Welcome

Welcome to the Fourth International Conference of Quantitative Genetics: Understanding Variation in Complex Traits - Edinburgh International Conference Centre, 17-22 June 2012

We invite you to join us in the beautiful city of Edinburgh for the Fourth International Conference of Quantitative Genetics, being held from 17 – 22 June 2012 at the Edinburgh International Conference Centre.

Variation in quantitative and other complex phenotypes underpins most important traits in human health and disease, agriculture, natural populations and evolution. The genomics revolution has provided the tools to start the dissection of such traits, enhancing both their understanding and exploitation. This has led to an explosion of interest and new studies across all of biology.

The aim will be to present and discuss state-of-the-art results, theoretical developments, understanding and methodology across the whole range of quantitative genetics - the genetic analysis and interpretation of data on complex traits - and to provide a stimulating conference in an attractive locale. Topics will include statistical methods for parameter estimation, including analysis of trait, genomic and functional genomic data; methods for QTL and gene identification; genetic control of complex traits; prediction of breeding value and individual risk, and interpretation of evolutionary change.

The meeting will pertain to and span results from, for example, humans, livestock, crops, micro-organisms and natural and experimental populations of all species.

The conference will be held at the Edinburgh International Conference Centre, and sponsored by the UK's Genetics Society. It follows previous successful meetings at Iowa State University in Ames, North Carolina State University in Raleigh, and Zhejiang University in Hangzhou, China.
Edinburgh has the best possible set up for a prestigious conference, and we are confident that the Fourth International Conference of Quantitative Genetics will meet the highest expectations. We look forward to welcoming you to Edinburgh.

**Bill Hill** (Chair), On Behalf of the Local Organising Committee.

---

**Conference Publicity**

We would be extremely grateful if you could assist us in publicising the conference. Please click on these links to download a [Poster](#), [PowerPoint slide](#) or [Flyer](#).

If you wish, please display them in a prominent place in your work place where it can be viewed by people who would find this conference of interest. We would also like to encourage you to use the PowerPoint slide during any presentations you may be giving over the coming year. We thank you in advance for helping us to promote this important conference.

**You may also be interested in the forthcoming conferences:**

- **Behavior Genetics Association**
- **British Society of Human Genetics**
- **European Society of Human Genetics**

---

**LOCAL ORGANISING COMMITTEE**

<table>
<thead>
<tr>
<th>Name</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bill Hill (Chair)</td>
<td></td>
</tr>
<tr>
<td>Lutz Bünger</td>
<td></td>
</tr>
<tr>
<td>Chris Haley</td>
<td></td>
</tr>
<tr>
<td>Mike Kearsey</td>
<td></td>
</tr>
<tr>
<td>DJ de Koning</td>
<td></td>
</tr>
<tr>
<td>Loeske Kruuk</td>
<td></td>
</tr>
<tr>
<td>Josephine Pemberton</td>
<td></td>
</tr>
<tr>
<td>Alan Wright</td>
<td></td>
</tr>
</tbody>
</table>

**CONFERENCE ORGANISERS**

<table>
<thead>
<tr>
<th>Company</th>
<th>Contact Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>In Conference Ltd</td>
<td>Tel: +44 (0)131 339 9235</td>
</tr>
<tr>
<td></td>
<td>Fax: + 44 (0)131 339 9798</td>
</tr>
<tr>
<td></td>
<td>E-Mail: <a href="mailto:icqg4@in-conference.org.uk">icqg4@in-conference.org.uk</a></td>
</tr>
</tbody>
</table>

---

**INTERNATIONAL ADVISORY COMMITTEE**

<table>
<thead>
<tr>
<th>Name</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>David Allison, USA</td>
<td>USA</td>
</tr>
<tr>
<td>David Balding, UK</td>
<td></td>
</tr>
<tr>
<td>Piter Bijma, Netherlands</td>
<td></td>
</tr>
<tr>
<td>Rachel Brem, USA</td>
<td></td>
</tr>
<tr>
<td>Ed Buckler, USA</td>
<td></td>
</tr>
<tr>
<td>Andrew Clark, USA</td>
<td></td>
</tr>
<tr>
<td>Mark Daly, USA</td>
<td></td>
</tr>
<tr>
<td>Rebecca Doerge, USA</td>
<td></td>
</tr>
<tr>
<td>Marie-Anne Felix, France</td>
<td></td>
</tr>
<tr>
<td>Jonathan Flint, UK</td>
<td>UK</td>
</tr>
<tr>
<td>Greg Gibson, USA</td>
<td></td>
</tr>
<tr>
<td>Mike Goddard, Australia</td>
<td></td>
</tr>
<tr>
<td>Ary Hoffman, Australia</td>
<td></td>
</tr>
<tr>
<td>Fred Hospital, France</td>
<td></td>
</tr>
<tr>
<td>Mark Lathrop, France</td>
<td></td>
</tr>
<tr>
<td>Trudy Mackay, USA</td>
<td></td>
</tr>
<tr>
<td>Albrecht Melchinger, Germany</td>
<td></td>
</tr>
<tr>
<td>Juha Merilä, Finland</td>
<td></td>
</tr>
<tr>
<td>Bill Muir, USA</td>
<td>USA</td>
</tr>
<tr>
<td>Patrick Phillips, USA</td>
<td></td>
</tr>
<tr>
<td>Daniel Pomp, USA</td>
<td></td>
</tr>
<tr>
<td>Pak Sham, China</td>
<td></td>
</tr>
<tr>
<td>Fred van Eeuwijk,</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Peter Visscher, Australia</td>
<td></td>
</tr>
<tr>
<td>Bruce Walsh, USA</td>
<td></td>
</tr>
<tr>
<td>Bruce Weir, USA</td>
<td></td>
</tr>
<tr>
<td>Qifa Zhang, China</td>
<td></td>
</tr>
</tbody>
</table>

---

Conference organisers: In Conference Ltd. [www.in-conference.org.uk](http://www.in-conference.org.uk) - Site by [Source](#)

---
PROFIT. IT’S IN OUR GENES.

Genetic excellence means Hy-Line layers have better liveability and lay superior quality eggs … while offering industry-leading feed efficiency. It all adds up to an unparalleled profit potential.

www.hyline.com  info@hyline.com

© 2012 Hy-Line International
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welcome/Committees</td>
<td>2</td>
</tr>
<tr>
<td>Scientific Programme</td>
<td>5</td>
</tr>
<tr>
<td>Industry Lunchtime Symposia</td>
<td>13</td>
</tr>
<tr>
<td>General Information</td>
<td>14</td>
</tr>
<tr>
<td>Social Programme</td>
<td>17</td>
</tr>
<tr>
<td>Poster Information</td>
<td>18</td>
</tr>
<tr>
<td>Conference Partners, Supporters &amp; Gold Sponsors</td>
<td>19</td>
</tr>
<tr>
<td>Exhibitors and Sponsors</td>
<td>22</td>
</tr>
<tr>
<td>Exhibition Floorplan</td>
<td>25</td>
</tr>
<tr>
<td>Edinburgh City Centre Map</td>
<td>26</td>
</tr>
<tr>
<td>EICC Floorplan</td>
<td>27</td>
</tr>
<tr>
<td>Oral Abstracts</td>
<td>29</td>
</tr>
<tr>
<td>Student Symposia Abstracts</td>
<td>60</td>
</tr>
<tr>
<td>Poster Abstracts</td>
<td>65</td>
</tr>
<tr>
<td>Author Index</td>
<td>261</td>
</tr>
</tbody>
</table>
WELCOME

Welcome to Edinburgh and the Fourth International Conference on Quantitative Genetics.

This is a time of rapid developments towards understanding the genetics of complex traits and in utilising that knowledge to improve health, food production and the natural environment. We have continuously improving technology for the collection and interpretation of molecular and cellular data and new ways to analyse these in combination with phenotypic data. Yet still we have much to learn and understand because many key traits are, indeed, very complex. We value this opportunity to bring you together to exchange ideas, discuss the latest results and consider the new challenges and opportunities.

We have organised a diverse programme which we hope you will find informative and stimulating. It was designed to cover most areas from genes to phenotypes, the molecular to the mathematical, and from the laboratory to farmed and wild populations. We shall be in plenary session throughout with the aim of maximising information exchange and interaction of ideas and knowledge across disciplines.

We have over 600 delegates registered from over 50 countries. In addition to the invited and contributed lectures we have around 400 posters to view and discuss, particularly in the early evening sessions on Monday and Tuesday. There will be considerable opportunity in the coffee and lunch breaks for further interactions as we shall be together the whole time in this fine facility, the Edinburgh International Conference Centre.

We hope also that many will participate in the student symposium and subsequent social activities on Wednesday.

Out of hours (only, of course) there is much else to occupy you in the beautiful and compact city of Edinburgh, with restaurants for all tastes and pockets, bars, museums and historic sights. In addition we start with the Sunday evening mixer and finish with a banquet at Dynamic Earth.

We very much hope you have an informative, productive and enjoyable time here. We hope all will go smoothly for you but if you have any problems please let one of us know.

With best wishes for a great week.

Lutz Bunter, DJ de Koning, Chris Haley, Bill Hill, Mike Kearsey, Loeske Kruuk, Josephine Pemberton and Alan Wright (Organising Committee), Kay Boulton (PG Symposium)
Conference Organisers

ICQG Secretariat
c/o In Conference Ltd, 4-6 Oak Lane, Edinburgh, EH12 6XH, Scotland, UK
Tel: +44 (0)131 339 9235  Fax: +44 (0)131 339 9798
Email: icqg4@in-conference.org.uk

COMMITTEES

LOCAL ORGANISING COMMITTEE
Bill Hill (Chair)
Lutz Bünger
Chris Haley
Mike Kearsey
DJ de Koning
Loeske Kruuk
Josephine Pemberton
Alan Wright
Kay Boulton (PG Symposium)

INTERNATIONAL ADVISORY COMMITTEE
David Allison, USA  Mark Lathrop, France
David Balding, UK  Trudy Mackay, USA
Piter Bijma, Netherlands  Albrecht Melchinger, Germany
Rachel Brem, USA  Juha Merilä, Finland
Ed Buckler, USA  Bill Muir, USA
Andrew Clark, USA  Patrick Phillips, USA
Mark Daly, USA  Daniel Pomp, USA
Rebecca Doerge, USA  Pak Sham, China
Marie-Anne Felix, France  Fred van Eeuwijk, Netherlands
Jonathan Flint, UK  Peter Visscher, Australia
Greg Gibson, USA  Bruce Walsh, USA
Mike Goddard, Australia  Bruce Weir, USA
Ary Hoffman, Australia  Qifa Zhang, China
Fred Hospital, France
Traditional style
Innovative content

Adapted to different housing systems
Persistent production for a prolonged laying cycle

LOHMANN SILVER
LOHMANN BROWN
LOHMANN TRADITION
LOHMANN LSL
LOHMANN SANDY

Excellent shell quality
Customer oriented after sales service

Outstanding breeding results and highest quality standards have safeguarded the success of our customers for years. High investments in the research and development division guarantee continuous breeding progress and yield constant innovations. www.ltz.de

Breeding for success...together
## SCIENTIFIC PROGRAMME

### Sunday 17th June

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.00 – 19.30</td>
<td>Registration Open</td>
</tr>
<tr>
<td>18.00 – 19.30</td>
<td>Welcome Reception Cromdale Hall, EICC</td>
</tr>
</tbody>
</table>

### Monday 18th June

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>08.00 – 19.30</td>
<td>Registration Open</td>
</tr>
<tr>
<td>09.00</td>
<td>Welcome from Bill Hill, Chair, Local Organising Committee</td>
</tr>
<tr>
<td>09.05 – 09.40</td>
<td>Session 1&lt;br&gt;The genetic architecture of quantitative traits&lt;br&gt;Chair: Bruce Walsh (University of Arizona, USA)</td>
</tr>
<tr>
<td>09.05 – 09.20</td>
<td>Introduction: The quantitative genetics landscape in 2012&lt;br&gt;Bruce Walsh</td>
</tr>
<tr>
<td>09.20 – 10.00</td>
<td>O-1&lt;br&gt;From Galton to GWAS: genetics of quantitative traits in human populations&lt;br&gt;Peter Visscher (University of Queensland, Australia)</td>
</tr>
<tr>
<td>10.00 – 10.40</td>
<td>O-2&lt;br&gt;Uniting the world’s maize diversity for dissection of complex traits and accelerated breeding&lt;br&gt;Ed Buckler (USDA, Ithaca)</td>
</tr>
<tr>
<td>10.40 – 11.10</td>
<td>Tea/Coffee</td>
</tr>
<tr>
<td>11.10 – 11.40</td>
<td>The genetic architecture of quantitative traits&lt;br&gt;Chair: Bruce Walsh (University of Arizona, USA)&lt;br&gt;O-3&lt;br&gt;An improved method for heritability estimation provides insights into the genetic architecture of epilepsy&lt;br&gt;Doug Speed (University College London, UK)</td>
</tr>
<tr>
<td>11.40 – 12.20</td>
<td>Genetics Society Mendel Lecture&lt;br&gt;Chair: Veronica van Heyningen (President, Genetics Society)&lt;br&gt;O-4&lt;br&gt;Secrets of the human genome&lt;br&gt;Eric Lander (Broad Institute of MIT &amp; Harvard, USA)</td>
</tr>
<tr>
<td>12.40 – 13.40</td>
<td>Lunch</td>
</tr>
<tr>
<td>13.40 – 15.20</td>
<td>Session 2&lt;br&gt;Evolutionary quantitative genetics&lt;br&gt;Chair: Derek Roff (University of California Riverside, USA)</td>
</tr>
<tr>
<td>13.40 – 14.20</td>
<td>O-5&lt;br&gt;Evolution of genetic variance under selection&lt;br&gt;Mark Blows (University of Queensland, Australia)</td>
</tr>
<tr>
<td>14.20 – 14.40</td>
<td>O-6&lt;br&gt;Plants in heterogeneous environments: determining when phenotypic plasticity is adaptive&lt;br&gt;Diane Byers (Illinois State University, USA)</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| 14.40 –    | **O-7**  
Discovery of cryptic genetic variation in *C. elegans* embryogenesis  |
| 15.00 –    | **O-8**  
Genomic comparisons between selected and relaxed chicken lines          |
| 15.20 –    | Tea/Coffee                                                              |
| **15.50 –**| **Evolutionary quantitative genetics**                                  |
| **17.20** | Chair: Derek Roff (University of California, USA)                       |
| 15.50 –    | **O-9**  
Gene interactions underlying the evolution of complex traits          |
| 16.30 –    | **O-10**  
What, me natural? Patterns of selection in wild systems and their     |
| 17.00 –    | theoretical implications                                               |
| 17.30 –    | **O-11**  
The deceit of monogamy: Quantitative genetic insights into the       |
| **19.30** | evolutionary ecology of polyandry in the wild                          |

**Tuesday 19th June**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>08.00 –</td>
<td>Registration Open</td>
</tr>
<tr>
<td><strong>09.00 –</strong></td>
<td><strong>Session 3</strong></td>
</tr>
<tr>
<td><strong>10.40</strong></td>
<td>Variation in the genome</td>
</tr>
</tbody>
</table>
| 09.00 –    | **O-12**  
Causes and consequences of new mutations                              |
| 09.40 –    | **O-13**  
The role of genetic variations on gene expression and splicing in |
| **10.00** | multiple regions of control in human post-mortem brain tissue          |
| **10.20:** | **O-14**  
Population genetics revealed by enormous structural variations         |
| **10.40** | discovered by population-scale sequencing                               |
| 10.40 –    | Tea/Coffee                                                              |
| **11.10 –**| **Variation in the genome**                                             |
| **12.40** | Chair: Alan Archibald (Roslin Institute, University of Edinburgh, UK)   |
| 11.10 –    | **O-16**  
Mapping the epigenetic basis of complex traits in Arabidopsis         |
| 11.50 –    | **O-17**  
The role of genetic variations on gene expression and splicing in |
<p>| <strong>12.10</strong> | multiple regions of control in human post-mortem brain tissue          |
| 12.10 –    | Tea/Coffee                                                              |</p>
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Title</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.50 – 12.20</td>
<td>O-17</td>
<td>The role of tandem repeats in mRNA and protein expression homeostasis</td>
<td>Sreenivas Chavali (MRC Lab Molecular Biology, UK)</td>
</tr>
<tr>
<td>12.20 – 12.40</td>
<td>O-18</td>
<td>Dissecting the contribution of regulatory sequence variation to quantitative haematological traits using maps of open chromatin in primary human blood cells</td>
<td>Cornelis Albers (University of Cambridge &amp; Sanger Institute, UK)</td>
</tr>
<tr>
<td>12.40 – 13.40</td>
<td></td>
<td>Lunch/Exhibition/Poster viewing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Affymetrix Lunchtime Symposium:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>How can you accelerate your genomics program from discovery to results?</strong></td>
<td>Professor David W. Burt, Chair of Comparative Genomics (The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, UK)</td>
</tr>
<tr>
<td>13.40 - 15.20</td>
<td></td>
<td><strong>Session 4</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Advances from numerical methods</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chair: David Balding (University College London, UK)</td>
<td></td>
</tr>
<tr>
<td>13.40 – 14.20</td>
<td>O-19</td>
<td>Statistical methods for the genetic analysis of arbitrarily structured populations</td>
<td>John Storey (Princeton University, USA)</td>
</tr>
<tr>
<td>14.20 – 14.40</td>
<td>O-20</td>
<td>CLIP Test: a new fast and simple method to distinguish between linked or pleiotropic quantitative trait loci in linkage disequilibrium analysis</td>
<td>Jean-Michel Elsen (INRA, France)</td>
</tr>
<tr>
<td>14.40 – 15.00</td>
<td>O-21</td>
<td>Sparse factor models for estimating the genetic architecture of gene expression traits</td>
<td>Daniel Runcie (Duke University, USA)</td>
</tr>
<tr>
<td>15.00 – 15.20</td>
<td>O-22</td>
<td>Big data, small cost – combining low-pass genome sequencing with long-range phasing and haplotype library imputation to create more powerful data for genomic prediction in plants and animals</td>
<td>John Hickey (University of New England, Australia)</td>
</tr>
<tr>
<td>15.20 –</td>
<td></td>
<td><strong>Tea/Coffee</strong></td>
<td></td>
</tr>
<tr>
<td>15.50 - 17.20</td>
<td></td>
<td><strong>Session 4</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Advances from numerical methods</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chair: David Balding (University College London, UK)</td>
<td></td>
</tr>
<tr>
<td>15.50 – 16.30</td>
<td>O-23</td>
<td>Haplotype phasing using next-generation sequencing reads</td>
<td>Jonathan Marchini (University of Oxford, UK)</td>
</tr>
<tr>
<td>16.30 – 17.00</td>
<td>O-24</td>
<td>Whole-genome regression and prediction of human complex traits using data from related and unrelated individuals</td>
<td>Gustavo de los Campos (University of Alabama, USA)</td>
</tr>
<tr>
<td>17.00 – 17.20</td>
<td>O-25</td>
<td>Relaxes the genetic model to identify quantitative trait loci having heterogeneous effects</td>
<td>Hugues Aschard (Harvard University, USA)</td>
</tr>
<tr>
<td>17.20 – 19.30</td>
<td></td>
<td><strong>Poster Session 2 – Even numbered posters</strong></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>08.00 – 16.00</td>
<td>Registration Open</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09.00 – 10.40</td>
<td><strong>Session 5</strong> &lt;br&gt;<strong>Technical advances and emerging areas</strong>&lt;br&gt;Chair: Marie-Anne Felix, (École Normale Supérieure, Paris, France)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09.00 – 09.40</td>
<td><strong>O-26</strong>&lt;br&gt;Rare variant patterns in 14,000 human samples and lessons for study design&lt;br&gt;Sebastian Zöllner (University of Michigan, USA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09.40 – 10.00</td>
<td><strong>O-27</strong>&lt;br&gt;Low depth, whole genome sequencing of Dai population demonstrates superiority over use of whole genome genotyping arrays in uncovering population structure, demographic history, selective pressures and phenotype associations in non-european populations&lt;br&gt;Lachlan Coin (BGI Shenzhen, China &amp; Imperial College London, UK)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.00 – 10.20</td>
<td><strong>O-28</strong>&lt;br&gt;The relationship between binary disease status and underlying heterogeneity in susceptibility and infectivity&lt;br&gt;Debby Lipschutz-Powell (Roslin Institute, University of Edinburgh, UK)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.20 – 10.40</td>
<td><strong>O-29</strong>&lt;br&gt;How do hosts shape their microbial communities?&lt;br&gt;Matthew Horton (University of Chicago, USA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.40 – 11.10</td>
<td><strong>Tea/Coffee</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.10 – 12.40</td>
<td><strong>Technical advances and emerging areas</strong>&lt;br&gt;Chair: Marie-Anne Felix, (École Normale Supérieure, Paris, France)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.10 – 11.50</td>
<td><strong>O-30</strong>&lt;br&gt;Quantitative association genetics of high dimensional cellular traits: extending beyond expression QTLs&lt;br&gt;Richard Durbin (Sanger Institute, UK)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.50 – 12.20</td>
<td><strong>O-31</strong>&lt;br&gt;Integrating gene expression into predictive health genomics&lt;br&gt;Greg Gibson (Georgia Tech, USA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.20 – 12.40</td>
<td><strong>O-32</strong>&lt;br&gt;Whole genome sequence analysis of a large Scottish family with bipolar disorder&lt;br&gt;Kathryn Evans (University of Edinburgh, UK)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Wednesday 20th June

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.00</td>
<td>Genetics Society Postgraduate Symposium</td>
<td></td>
</tr>
<tr>
<td>14.00</td>
<td>Welcome</td>
<td>Kay Boulton (University of Edinburgh, UK)</td>
</tr>
<tr>
<td>14.05</td>
<td>SO-1 / P-43</td>
<td>Markers as traits in multivariate BLUP: using REML for association testing and integration with breeding value prediction</td>
</tr>
<tr>
<td>14.20</td>
<td>SO-2 / P-46</td>
<td>Genetic markers as instruments for Mendelian randomization studies on vitamin D</td>
</tr>
<tr>
<td>14.20</td>
<td>SO-3 / P-41</td>
<td>Using KEGG pathways and expression studies and stuff for genomic partitioning</td>
</tr>
<tr>
<td>14.50</td>
<td>SO-4 / P-283</td>
<td>Genetic interactions in the human liver</td>
</tr>
<tr>
<td>15.05</td>
<td>SO-5 / P-265</td>
<td>Evidence of shared polygenic risk among smoking behaviors and body composition</td>
</tr>
<tr>
<td>15.20</td>
<td>SO-6 / P-214</td>
<td>Direct and indirect genetic effects for survival in purebred and crossbred laying hens</td>
</tr>
<tr>
<td>15.35</td>
<td>SO-7 / P-79</td>
<td>Maternal genetic effects set the potential for evolution in red squirrels</td>
</tr>
<tr>
<td>15.50</td>
<td>Concluding Remarks</td>
<td></td>
</tr>
<tr>
<td>20.00</td>
<td>Genetics Society Postgraduate Symposium Ceilidh</td>
<td></td>
</tr>
</tbody>
</table>

### Thursday 21st June

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>08.00</td>
<td>Registration Open</td>
<td></td>
</tr>
<tr>
<td>09.00</td>
<td>Session 6</td>
<td>Bridging the genotype-phenotype gap</td>
</tr>
<tr>
<td>09.00</td>
<td>O-33</td>
<td>Charting the genotype-phenotype map: lessons from <em>Drosophila</em></td>
</tr>
<tr>
<td>09.40</td>
<td>O-34</td>
<td>Heterosis as a systemic property emerging from the non-linearity of the genotype-phenotype relationship: evidence from metabolic models and test-tube genetics</td>
</tr>
<tr>
<td>10.00</td>
<td>O-35</td>
<td>Discovering early and late regulators of haematopoiesis through large-scale genomic analyses</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Thursday 21st June</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 10.20 – 10.40 | O-36  
A sexual ornament in chickens is determined by pleiotropic alleles at HAO1 and BMP2, selected during domestication  
Dominic Wright (Linköping University, Sweden) |
| 10.40 – 11.10 | Tea/Coffee  
Sponsored by Nature Communications |
| 11.10 – 11.40 | Bridging the genotype-phenotype gap  
Chair: Michel Georges (University of Liege, Belgium) |
| 11.10 – 11.50 | O-37  
Diamonds in the dirt: biological and translational insights into type 2 diabetes from large-scale genetic studies  
Mark McCarthy (University of Oxford, UK) |
| 11.50 – 12.20 | O-38  
The biologic relevance of eQTLs: a genome wide confirmation of eQTLs using two types of inbred populations in C. elegans  
Mark Sterken (Wageningen University, The Netherlands) |
| 12.20 – 12.40 | O-39  
Quantitative genetics of Drosophila life span  
Michael Magwire (North Carolina State University, USA) |
| 12.40 – 13.40 | Lunch/Exhibition/Poster viewing |
| 13.40 – 15.20 | Session 7  
The genetic architecture of quantitative traits 2  
Chair: Bruce Weir (University of Washington, USA) |
| 13.40 – 14.20 | The Genetics Society Balfour Lecture  
O-40  
Variance controlling genes and their role in the genetic architecture of complex traits  
Orjan Carlborg (Swedish University of Agricultural Sciences, Sweden) |
| 14.20 – 14.40 | O-41  
GWAS: 2D, or not 2D: that is the question  
Gibran Hemani (Roslin Institute, University of Edinburgh, UK & University of Queensland, Australia) |
| 14.40 – 15.00 | O-42  
Knowledge-driven analysis identified a gene-gene interaction affecting high-density lipoprotein cholesterol levels in multi-ethnic populations  
Alon Keinan (Cornell University, USA) |
| 15.00 – 15.20 | O-43  
The genetic architecture of quantitative traits: lessons from biochemical markers for disease  
Beben Benyamin (University of Queensland, Australia) |
| 15.20 – 15.50 | Tea/Coffee  
Sponsored by Nature Communications |
<table>
<thead>
<tr>
<th>Time</th>
<th>Session Description</th>
<th>Speaker/Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.50 - 17.20</td>
<td>The genetic architecture of quantitative traits 2 &lt;br&gt;Chair: Bruce Weir (University of Washington, USA)</td>
<td></td>
</tr>
<tr>
<td>15.50 – 16.30:</td>
<td>O-44 From Castle to the Collaborative Cross: evolution of the mouse in quantitative genetics research &lt;br&gt;Daniel Pomp (University of North Carolina, USA)</td>
<td></td>
</tr>
<tr>
<td>16.30 – 17.00</td>
<td>O-45 Genetic architecture and evolution of quantitative traits &lt;br&gt;Mike Goddard (University of Melbourne, Australia)</td>
<td></td>
</tr>
<tr>
<td>17.00 – 17.20</td>
<td>O-46 The quantitative genetics of phenotypic robustness &lt;br&gt;Hunter Fraser (Stanford University, USA)</td>
<td></td>
</tr>
</tbody>
</table>

**Friday 22nd June**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session Description</th>
<th>Speaker/Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>08.00 – 17.30</td>
<td>Registration Open</td>
<td></td>
</tr>
<tr>
<td>09.00 - 10.40</td>
<td>Session 8 Interactions among individuals and with the environment &lt;br&gt;Chair: Albrecht Melchinger (University of Hohenheim, Germany)</td>
<td></td>
</tr>
<tr>
<td>09.00 – 09.40</td>
<td>O-47 Studying the genotype-phenotype map in Arabidopsis &lt;br&gt;Magnus Nordborg (Gregor Mendel Institute, Austria)</td>
<td></td>
</tr>
<tr>
<td>09.40 – 10.00</td>
<td>O-48 Novel methods of GWAS for mapping genes of complex traits and their applications in crop breeding &lt;br&gt;Jun Zhu (Zhejiang University, China)</td>
<td></td>
</tr>
<tr>
<td>10.00 – 10.20</td>
<td>O-49 Nonlinear genotype x environment interaction &lt;br&gt;Rong-Cai Yang (University of Alberta, Canada)</td>
<td></td>
</tr>
<tr>
<td>10.20 – 10.40</td>
<td>O-50 Applying quantitative genetics to epidemics and disease resistance: implications of dynamic and noisy data &lt;br&gt;Steve Bishop (Roslin Institute and R(D)SVS, UK)</td>
<td></td>
</tr>
<tr>
<td>10.40 – 11.10</td>
<td>Tea/Coffee</td>
<td></td>
</tr>
<tr>
<td>11.10 - 12.40</td>
<td>Interactions among individuals and with the environment &lt;br&gt;Chair: Albrecht Melchinger (University of Hohenheim, Germany)</td>
<td></td>
</tr>
<tr>
<td>11.10 – 11.50</td>
<td>O-51 The consequences of indirect genetic effects for heritable variation and response to selection &lt;br&gt;Piter Bijma (Wageningen University, The Netherlands)</td>
<td></td>
</tr>
<tr>
<td>11.50 – 12.20</td>
<td>O-52 The genetics of social dominance: does competition cause constraint? &lt;br&gt;Alastair Wilson (University of Edinburgh, UK)</td>
<td></td>
</tr>
<tr>
<td>12.20 – 12.40</td>
<td>O-53 Simple and effective methods of addressing competitive effects in animal breeding programs &lt;br&gt;William Muir (Purdue University, USA)</td>
<td></td>
</tr>
<tr>
<td>12.40 – 13.40</td>
<td>Lunch/Exhibition/Poster viewing</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Title</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>13.40 - 15.20</td>
<td>Session 9</td>
<td>Genomic information in prediction</td>
</tr>
<tr>
<td>13.40 – 14.20</td>
<td>O-54</td>
<td>Towards genomic prediction from genome sequence data and the 1000 bull genomes project</td>
</tr>
<tr>
<td>14.20 – 14.40</td>
<td>O-55</td>
<td>Using whole genome sequence data to predict quantitative trait phenotypes in <em>Drosophila melanogaster</em></td>
</tr>
<tr>
<td>14.40 – 15.00</td>
<td>O-56</td>
<td>Next generation breeding using genotyping-by-sequencing</td>
</tr>
<tr>
<td>15.00 – 15.20</td>
<td>O-57</td>
<td>Genomic-BLUP decoded: a look into the black box</td>
</tr>
<tr>
<td>15.20 – 15.50</td>
<td></td>
<td>Tea/Coffee</td>
</tr>
<tr>
<td>15.50 - 17.30</td>
<td>Session 9</td>
<td>Genomic information in prediction</td>
</tr>
<tr>
<td>15.50 – 16.30</td>
<td>O-58</td>
<td>Predicting disease risk from marker genotypes under a polygenic model</td>
</tr>
<tr>
<td>16.30 – 17.00</td>
<td>O-59</td>
<td>Genome-based prediction in highly structured plant populations</td>
</tr>
<tr>
<td>17.00 – 17.20</td>
<td>O-60</td>
<td>Genomic prediction of colorectal cancer risk using GBLUP</td>
</tr>
<tr>
<td>17.20 – 17.30</td>
<td></td>
<td>Close of Conference - Thanks</td>
</tr>
<tr>
<td>19.30 – Midnight</td>
<td></td>
<td>Banquet at Dynamic Earth</td>
</tr>
</tbody>
</table>
INDUSTRY LUNCHTIME SYMPOSIUM

Affymetrix Lunchtime Symposium

Tuesday 19th June
12.55 – 12.25
Lomond Suite, located on the ground floor level of the Conference Centre
Bagged lunches will be available for collection to be taken into the Lomond Suite for the lunchtime symposium.

How can you accelerate your genomics program from discovery to results?

Abstract:
With so many genomics technologies available, knowing how to integrate them for maximum efficiency can be challenging. This is true whether the objective is to dissect complex traits, identify disease genes or improve breeding. The prevailing strategy is to integrate NGS and other technologies to quickly progress genomic markers from discovery into downstream studies and applications. For mid to high density marker sets, microarrays are widely preferred due to the high data quality, throughput, analytical efficiency and cost effectiveness they offer. Taking genomics in the chicken as an example, this scientific seminar will discuss best practice that is equally relevant across diverse species and applications.

Invited presentation:

Development and characterization of a high-density SNP genotyping assay for the chicken
Professor David W. Burt, Chair of Comparative Genomics,
The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh
GENERAL INFORMATION

Registration/Information Desks
All delegates will receive their name badge, meeting documents, ordered tickets and all relevant conference information upon arrival at the Conference Centre.

The Registration and Information Desks will be open at the following times:

- Sunday 17th June: 14.00 – 19.30
- Monday 18th June: 08.00 – 19.30
- Tuesday 19th June: 08.00 – 19.30
- Wednesday 20th June: 08.00 – 16.00
- Thursday 21st June: 08.00 – 17.30
- Friday 22nd June: 08.00 – 17.30

Exhibition Opening Hours
- Sunday 17th June: 18.00 – 19.30
- Monday 18th June: 08.00 – 19.30
- Tuesday 19th June: 08.00 – 19.30
- Wednesday 20th June: 08.00 – 16.00
- Thursday 21st June: 08.00 – 17.30
- Friday 22nd June: 08.00 – 15.50

Tea/Coffee Breaks and Lunch Arrangements
Catering points will be located with the exhibition in the Cromdale Hall which is located on the lower level of the Conference Centre.
On Tuesday lunchtime there will also be secondary catering points on the ground level of the Strathblane Hall for those attending the lunchtime symposium. Please follow the directions of the staff at the Conference Centre.
If you have requested a special diet at the time of registering (other than vegetarian), then your name badge will have a sticker on the back which you should show to the catering staff who will bring you your pre-ordered food.

Thank you to Nature Communications who have sponsored the tea and coffee breaks on Thursday.
Congress Etiquette
Delegates are advised that they are not allowed to take photographs of any posters or presentations without the author’s/presenter’s consent. Delegates should also obtain consent from an author before citing any of their work that was presented at the conference.
If you intend to ‘blog’ or ‘twitter’ results from this conference you must inform the conference organisers in advance and get the permission of the author or presenter.

Mobile phones should be switched off or placed on ‘silent’ during sessions. Thank you for your co-operation.

WiFi Access
The conference is providing WiFi access free to delegates who have their own laptops. Please see the registration notice board for password and log in details. Please note that there are printing facilities in the Business Centre although this will be chargeable to delegates.

Thank you to Genetics Society of America Journals (GENETICS and G3) who have sponsored the WiFi.

Certificate of Attendance
A certificate of attendance will be emailed to all participants following the conference.

Speaker Preview Room
This is located in Harris Room 1, on the first floor. All presenters are required to check in their presentation a minimum of 4 hours prior to their talk. The Speaker Preview Room will be open at the following times:

<table>
<thead>
<tr>
<th>Day</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sunday 17th June</td>
<td>14.00 – 19.30</td>
</tr>
<tr>
<td>Monday 18th June</td>
<td>08.00 – 17.30</td>
</tr>
<tr>
<td>Tuesday 19th June</td>
<td>08.00 – 17.30</td>
</tr>
<tr>
<td>Wednesday 20th June</td>
<td>08.00 – 16.00</td>
</tr>
<tr>
<td>Thursday 21st June</td>
<td>08.00 – 17.30</td>
</tr>
<tr>
<td>Friday 22nd June</td>
<td>08.00 – 15.50</td>
</tr>
</tbody>
</table>
Insurance
The Conference Organisers cannot accept any liability for personal injuries or for loss or damage to property belonging to delegates, either during, or as a result of the conference. Please check the validity of your own personal insurance before travelling.

Security
Your name badge must be worn at all times otherwise you will not be allowed entry to the conference centre. Exhibitor staff do not have access to the scientific sessions. Please leave any luggage or poster tubes with the staff at the cloakroom. This will be manned during the normal opening hours of the conference.

Social Programme

Welcome Reception
Sunday 17th June 18.00 – 19.30
Cromdale Hall, Edinburgh International Conference Centre
The Welcome Reception will take place at the Edinburgh International Conference Centre. The cost for this event is included in the registration fee.

Genetics Society Postgraduate Symposium Ceilidh
Wednesday 20th June 20.00 – Midnight
Edinburgh International Conference Centre
A Ceilidh will take place at the Edinburgh International Conference Centre following the Genetics Society Postgraduate Symposium. The cost for this event is £5 and tickets can be purchased at the registration desk. All delegates are welcome to attend.

Please note there will be no food served at this event and a cash bar will be available.
Conference Dinner  
Friday 22nd June 19.30 – Midnight  
Dynamic Earth, 112-116 Holyrood Road, Edinburgh, EH8 8AS  
The conference dinner will be held at Dynamic Earth on Friday evening.  
The evening will start with a pre-dinner drinks reception, followed by dinner and dancing.

Offering stunning 360 degree views across Edinburgh’s picturesque cityscape and Royal Park, the five star Our Dynamic Earth is a perfectly situated venue with a unique heritage twist. Located at the bottom of Holyrood Road, neighbouring the Scottish Parliament and the Palace of Holyrood House, getting to Our Dynamic Earth is easy and is approximately a 5 minute taxi ride from the EICC.

Tickets cost £50.00 and must have been pre-booked at the time of registration. Should you wish to attend and have not purchased a ticket please speak to the staff at the registration desk.

Limited shuttle buses will be available from 10.30pm – Midnight stopping at North Bridge, mid George Street, West End and the EICC.
Poster Information

Posters
All posters will be on display for the duration of the conference, however, there are two designated poster sessions to enable delegates to discuss the author’s work in more detail. We ask that you stand by your board during your allocated poster sessions so that delegates can ask you questions about your work.

<table>
<thead>
<tr>
<th>Poster Session 1</th>
<th>Poster Session 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday 18th June</td>
<td>Tuesday 19th June</td>
</tr>
<tr>
<td>17.30 – 19.30</td>
<td>17.20 – 19.30</td>
</tr>
</tbody>
</table>

Sponsored by

- Posters ending in an Odd-number: Stand by your board from 17.30 to 19.30
- Posters ending in an Even-number: Stand by board from 17.20 – 19.30

All posters are to be displayed from 18:00 Sunday and must be removed by 15:50 on Friday.
Conference Supporters

The Genetics Society
Contact: Linda Allardyce or Ivvah Chung
The Genetics Society, c/o Portland Customer Services,
Commerce Way, Colchester CO2 8HP
Tel: +44 (0)1206 796 351 / Fax: +44 (0)1206 799 331
Email: theteam@genetics.org.uk
Website:www.genetics.org.uk

The Genetics Society was founded by William Bateson in 1919 and is one of the oldest “learned societies” devoted to genetics in the world. Its membership of over 1700 comprises the UK’s active geneticists, including academics, researchers and students.
The Genetics Society organises meetings to promulgate genetics, supports students to attend meetings, sponsors research through fieldwork grants and student bursaries, and promotes the Public Understanding of Genetics.
It co-owns two leading journals in the field: Heredity and Genes and Development.
The Genetics Society is proud to be the major sponsor of the 4th International Conference on Quantitative Genetics.

UNIVERSITY OF EDINBURGH
The University of Edinburgh, Old College, South Bridge,
Edinburgh EH8 9YL
www.ed.ac.uk

The University is home to a vigorous complex trait genetics research community. There are particular strengths in this area in biomedical, evolutionary and livestock genetics. We welcome approaches from anyone interested in the many exciting opportunities available for research and education at all levels from MSc to senior scientist.
Over the past decade we have seen agriculture move from being regarded as a ‘sunset industry to one that is now core to contemporary political and societal thinking. Our food supplies, our environment, our water quality, and our renewable energy capability, all depend on the effective and efficient management of our land resources. SAC’s portfolio of expertise in research, in teaching and in consultancy, is being increasingly recognised as a core national and international resource that is competent to tackle many of these issues.

**Gold Sponsor**


**Aqua Gen AS**
Aqua Gen AS, PO Box 1240, Sluppen, N-7462 Trondheim  
Phone: +47 72 45 05 00  
Fax: +47 73 54 62 91  
E-mail: firmapost@aquagen.no  
Web: www.aquagen.no

Aqua Gen manages and further develops the world’s first family-based breeding programs of Atlantic salmon and rainbow trout. Today’s eggs from the Aqua Gen strain are selected for excellent performances on growth, disease resistance, robustness and filet quality by use of state-of-the-art methods within quantitative and molecular genetics.
Aviagen
Newbridge, Midlothian, EH28 8SZ, Scotland, UK
Tel: +44 131 333 1056
Fax: +44 131 333 3296
Email: infoworldwide@aviagen.com
www.aviagen.com

Aviagen is the world’s leading poultry breeding company, developing pedigree lines for the production of broiler chickens and turkeys under the Ross, Arbor Acres, Indian River, B.U.T. and Nicholas brand names. Aviagen has wholly-owned operations on five continents, employs almost 3,000 people and a distribution network that serves customers in 120 countries.

Hy-Line
Contact: Jerry Dreyer
Hy-Line International, PO Box 310,
Dallas Center IA 50063 USA
Tel: +1-515-992-3736
Email: jdreyer@hyline.com
www.hyline.com

Hy-Line International is the world’s oldest egg layer genetics company since being founded by the Wallace family as part of Pioneer Seed Corn in 1936. Hy-Line layers have become the largest selling layers across the globe and are known for their superior livability, feed conversion and egg quality.

LOHMANN TIERZUCHT GmbH
Am Seedeich 9 – 11, 27472 Cuxhaven
Phone: +49 4721 505 0
Fax: +49 4721 505 222
E-Mail: info@ltz.de

LOHMANN TIERZUCHT – Breeding for success … together!
LOHMANN TIERZUCHT is the specialist for layer breeding and as a worldwide leader in the market constantly improving the performance characteristics. The diagnostics conducted at LOHMANN’s veterinary laboratory and the resulting curative measures round up the range of products.
Exhibitors

AFFYMETRIX
Mercury Park, Wycombe Lane,
Wooburn Green, HP10 0HH,
United Kingdom
Tel: +44 (0)1628 552500
Email: Sales_Europe@affymetrix.com
Website: www.affymetrix.com

Affymetrix®, part of The Genome Generation, provides genomic solutions for agrigenomics and human health research. Our technologies for whole-genome analysis to single-gene validation enable sequencing discoveries to be confirmed and applied. Our solutions for expression, genotyping and cytogenetics enable your translational and applied research... we are all The Genome Generation.

BIOCOMPUTING PLATFORMS LTD
Elina Lamponen
BC Platforms Ltd., Innopoli 2, Tekniikantie 14,
Espoo, FI-02150, Finland
Tel: +358 9 2517 7340
Fax: +358 10 296 1288
Email: elina.lamponen@bcplatforms.com

BC Platforms offers data management solutions for genetic researchers. Our systems include SNP genotyping for GWAS, next-generation sequencing, CNV, MultiQTL, sample-tracking and management, pedigree and phenotype data into one powerful and secure data warehouse used by the world’s leading research institutes, giving multiple users easy access through a Web-based interface.
ILLUMINA, INC.
5200 Research Place, San Diego, CA 92122 USA
www.illumina.com

Illumina’s innovative DNA sequencing and genotyping products measure genetic variation on an unprecedented scale. HiSeq and MiSeq sequencers have been used by thousands of organizations to characterize DNA, RNA and functional elements in numerous species. Our iScan and BeadXpress® platforms are the standard for routine testing, powered by a customer-centric consortia product development model. We routinely support de novo and resequencing applications, transcriptomics, SNP discovery, parentage verification, and genomic selection.

LGC GENOMICS
Contact: Customer Services
Address: Ostendstr. 25, TGS Haus 8, 12459 Berlin, Germany
Email: service@lgcgenomics.com
Tel: +49(0) 30 5304 2200
Website: www.lgcgenomics.com

LGC Genomics offers a full range of genomics products and services including nucleic acid extraction, sequencing and genotyping. Our portfolio includes reagents and consumables in each of our core areas of expertise supported by a range of genotyping laboratory automation instrument solutions and service laboratories in Europe and the USA.
SOURCE BIOSCIENCE
Andrew Jones
Nottingham Business Park 1 Orchard Place,
Nottingham, NG8 6PX
Email: enquires@sourcebioscience.com
Tel: +44(0) 115 973 9012
Website: www.sourcebioscience.com

Source BioScience LifeSciences are European leaders in DNA sequencing, genomic services, and bioinformatics analyses and offers a comprehensive portfolio of genomic reagents and antibodies. Source BioScience LifeSciences offer a unique Sanger sequencing service; our local facilities in London, Cambridge, Oxford, Nottingham, Dublin and Berlin.

WISEPRESS MEDICAL BOOKSHOP
25 High Path, Merton Abbey,
London, SW19 2JL, UK
Phone: +44 20 8715 1812 / Fax: +44 20 8715 1722
bookshop@wisepress.com
www.wisepress.com

Wisepress.com, Europe’s leading conference bookseller, has a complete range of relevant books and journals which can be purchased at the stand or, if you would rather not carry them, posted to you – Wisepress will deliver worldwide. We also have a comprehensive medical and scientific online bookshop with great offers.
ICQG Exhibition and Poster Floorplan

Cromdale Hall

1  EW Group: AquaGen, Aviagen, Hy-Line & Lohmann Tierzucht
2  Affymetrix
3  Bio Computing
4  LGC Genomics
5  Cambridge University Press
6  Illumina
7  Source Bioscience
8  Wisepress
The main entrance to the EICC is on Morrison Street where there is a coach drop off point. Access to the EICC Loading Bay (deliveries only) is from the West Approach Road.
There are a number of car parks within walking distance of the EICC, they are marked on the map. For further details on car parking please visit our website.

Please note that there are a number of one-way streets in close proximity to the EICC. After 18.30 parking is allowed in certain areas.
Please be advised that these maps cannot be reproduced without the permission of PRM Marketing Ltd. www.prm.co.uk

Please see the appropriate copyright document.
EICC Room Layouts

LEVEL 3
PENTLAND
SIDLOW
FINTRY

LEVEL 1
GALLOWAY SUITE

LEVEL 0
LOMOND ROYAL
STRATHBLANE HALL

LEVEL -2
CROMDALE HALL

MAIN ENTRANCE
ASReml - R

ASReml-R fits the general linear mixed model to large data sets with complex variance models, using the Average Information (AI) algorithm and Residual Maximum Likelihood (REML) to estimate variance components, and sparse matrix methods for computational efficiency.

The R implementation offers a compact formula based syntax for specifying the linear model, and associated variance models for the random effects.

Types of analysis:

- (un)balanced longitudinal data
- repeated measures data (multivariate analysis of variance and spline type models)
- (un)balanced designed experiments
- multi-environment trials and meta analysis
- regular or irregular spatial data

ASReml - R is available on a variety of platforms including: Windows 32 bit / 64 bit, Macintosh, Linux (Fedora 12) 32 bit / 64 bit, Linux (Centos 5.5) 32 bit / 64 bit.

ASReml-R (Windows) is also made available FREE to educators worldwide and scientists in the developing world (qualifying conditions apply).

Contact us for further details at: support@vsni.co.uk
O-1
FROM GALTON TO GWAS: GENETICS OF QUANTITATIVE TRAITS IN HUMAN POPULATIONS

Peter Visscher
University of Queensland Diamantina Institute, University of Queensland, Princess Alexandra Hospital, Brisbane, Queensland, Australia

Variation in quantitative traits such as human height is caused by a combination of multiple genes and environmental effects. Traditionally, since Galton in the late 1800s, the genetics of such traits has been studied using concepts that refer to the combined effect of all genes (e.g., heritability), using the resemblance between relatives. For example, estimates of heritability for adult height and schizophrenia are ~0.8, so that 80% of differences between people in the trait is due to genetic factors. Genome-wide association studies (GWAS) facilitate the dissection of heritability into individual locus effect. They have been successful in finding many SNPs associated and have greatly increased the number of genes involved in complex trait variation. To date, for many complex traits, tens to hundreds of loci have been identified that explain in total up to 20% of narrow sense heritability. The proportion of additive genetic variation explained by all common SNPs together (not just the significant ones) is one-third to one-half for a range of quantitative traits and diseases. The variation is spread over all chromosomes in proportion to their length, implying that there are many more variants with effects sizes too small to be detected with sample sizes employed to date. The identification of tens to hundreds of associated variants facilitates the study of biology and function, in particular when genetic data is combined with data on gene expression. Empirical observations from the resemblance between relatives, from within-family segregation variance captured by markers and from population based studies are all converging to a highly polygenic model of complex traits, with a surprisingly large proportion of additive genetic variation due to variants that are in linkage disequilibrium with common SNPs. We will show results from empirical data on quantitative traits and disease in human populations to quantify and partition genetic variation.

O-2
UNITING THE WORLD’S MAIZE DIVERSITY FOR DISSECTION OF COMPLEX TRAITS AND ACCELERATING BREEDING

Edward Buckler
1USDA - ARS, Ithaca, NY, USA, 2Cornell University, Ithaca, NY, USA, 3Cold Spring Harbor Lab, Cold Spring, NY, USA, 4University of California, Irvine, Irvine, CA, USA, 5North Carolina State University, Raleigh, NC, USA, 6University of Missouri, Columbia, MO, USA, 7University of Wisconsin, Madison, WI, USA

Maize is one of the most diverse crops and species in the world. This diversity is a tremendous resource for understanding the genetic basis of complex traits and for plant breeding in general. However, it poses both a serious problem and substantial opportunity in relating these 10s of millions of variable sites to the complex traits they control. We have recently combined whole genome next generation sequencing with QTL mapping on the maize Nested Association Mapping (NAM) population of 5,000 recombinant inbred lines and thousands of other breeding lines. Through whole genome sequencing of maize and teosinte lines, we have identified more than 50 million variable regions in the genome. We have also combined this deeper sequencing with reduced representation genotyping (or Genotyping-By-Sequencing) that has brought genotyping costs down to $10-20/sample. By combining these two approaches to understand genome diversity, we were able to conduct GWAS that unites the world’s germplasm diversity and is providing insights into heterosis, central metabolism, development, and many other complex traits. We can now show that copy number and intergenic variation is very important for a series of traits. While in many cases the association and genes are very interpretable, there still remain numerous associations that remain complex. Despite this complexity, the ability to make predictions using diverse germplasm is high for many traits, which suggests great opportunities for merging, mining, and breeding with the world’s germplasm resources. Already mining diverse germplasm is advancing to make more nutritional maize for the developing world.
O-3
AN IMPROVED METHOD FOR HERITABILITY ESTIMATION PROVIDES INSIGHTS INTO THE GENETIC ARCHITECTURE OF EPILEPSY
Doug Speed, David Balding
UCL Genetics Institute, London, UK

The 2010 Nature Genetics paper of Yang, Visscher et. al. first promoted the idea of performing heritability analysis using SNP data for unrelated individuals. Their key insight was that by using unrelated individuals, the heritability estimates no longer correspond to the trait's overall heritability, but rather the proportion of variance which can be attributed to the causal variants among (or tagged by) the genotyped SNPs. Their work has provided strong evidence that the apparent missing heritability of many complex traits might be due to many causal variants of low influence; this would explain the failure of GWASs as most of the causal variants effect sizes would be too small to detect with standard sample sizes. For all the method's strengths, we discover that its results are highly sensitive to LD. Depending on the location of causal variants, the method's estimates can be out by up to a factor of two. This finding is pertinent for rare variant traits, as the method performs particularly poorly when the frequency of the causal alleles is low. In response, we propose a solution which corrects for the impact of LD, allowing unbiased heritability estimation regardless of the distribution of causal variants. We have applied our revised method to Susceptibility to Epilepsy, yet another trait known to be highly heritable, but for which GWASs have failed to explain much variation. We deduce that almost all the heritability can be explained by considering the contribution of all SNPs simultaneously. We show how this leads us to a prediction model with practical applications in determining personalised treatment, and by partitioning the genome provides insights into the areas most likely to harbour causal variants.

O-4
SECRETS OF THE HUMAN GENOME
Eric Lander
Broad Institute, MIT & Harvard, USA

No Abstract.

O-5
EVOLUTION OF GENETIC VARIANCE UNDER SELECTION
Mark Blows
University of Queensland, Brisbane, Australia

Determining how the genetic variance evolves under selection in the presence of widespread pleiotropy is a fundamental issue that is ultimately related to the question of how genetic variance is maintained in populations. The study of sexually-selected traits is particularly useful here as male sexually-selected traits are often relatively easy to identify and selection generated by mate choice can be readily quantified. At the same time, sexually-selected traits are often postulated to be subject to natural selection, either through pleiotropic associations with condition and other fitness components, or as a consequence of selection on mate recognition generated by the presence of other species. In a series of experiments using Drosophila serrata as a model, we have investigated the relationship between selection and genetic variance in a multivariate set of contact pheromones. While directional sexual selection is associated with very low levels of genetic variance in natural populations of this species, the genetic variance in these traits can rapidly evolve, increasing dramatically as males are selected for higher attractiveness.
However, it is the pleiotropic associations between the sexually-selected traits and other (unmeasured) traits under natural selection that have been revealed to play the key role in limiting the response to sexual selection, and the maintenance of multivariate genetic variance in traits under sexual selection.

O-6
PLANTS IN HETEROGENEOUS ENVIRONMENTS: DETERMINING WHEN PHENOTYPIC PLASTICITY IS ADAPTIVE

Diane L. Byers¹, Rebecca R. Miller¹,²
¹Illinois State University, Normal, IL, USA, ²Missouri Dept of Conservation, Blue Springs, MO, USA

The quality of natural environments is highly heterogeneous, imposing selection on the local populations. As plants are sessile for much of their life history, it has been proposed that phenotypic plasticity is a critical mechanism that allows plants to adapt to environmental heterogeneity. However, increased phenotypic plasticity does not necessarily lead to increased fitness. Furthermore, the evolution of increased phenotypic plasticity may be limited by strong cross-environment genetic correlations. We have been studying the response of wild Arabidopsis thaliana populations to the natural variation in soil nutrients. Using inbred lines from 18 populations located in USA, we tested and found extensive phenotypic plasticity in many fitness-related traits in response to different soil phosphorous treatments. We used a typical approach (modified phenotypic selection analysis, van Kleunen and Fischer 2001) for determining if the expressed phenotypic plasticity is adaptive. Using this approach we found very limited evidence for increasing phenotypic plasticity associated with greater seed production. We have done a simple mathematical alteration on this approach that uses a more realistic simulation of the environmental variation. Results from this analysis give different conclusions and insights for determining if and when phenotypic plasticity is adaptive. Even though phenotypic plasticity is adaptive under some environmental conditions, strong genetic correlations across the phosphorus treatments indicate that some of the traits will be constrained, as they will not evolve independently in the soil environments. Other traits, including root biomass (a trait likely critical for tolerance of lower nutrient soils), are not constrained. Our populations of A. thaliana show significant phenotypic plasticity in response to differences in availability of phosphorous, some of this plasticity is adaptive, while the evolution of plasticity will be constrained due to the cross-environment genetic correlations.

O-7
DISCOVERY OF CRYPTIC GENETIC VARIATION IN C. ELEGANS EMBRYOGENESIS

Annalise Paaby, Amelia White, Kris Gunsalus, Fabio Piano, Matthew Rockman
New York University, New York, NY, USA

In the evolution of complex traits, natural genetic variation may accumulate in pathways that determine robust phenotypes. Such variation is “cryptic” if it has no effect on phenotype under normal conditions, but an effect when the organism is perturbed, either environmentally or genetically. Thus cryptic genetic variation demonstrates conditional dynamics in the relationship between genotype and phenotype. Conditional effects are probably common in biological systems, but they pose barriers to the identification of causal alleles that underlie complex traits. Accumulation of cryptic genetic variation in developmental pathways may lead to evolution by developmental system drift, a phenomenon with consequences for speciation.

To explore cryptic genetic variation in C. elegans embryogenesis, we induced perturbations by feeding 57 worm strains a total of 41 RNAi vectors against early embryonic genes. We observed that embryonic lethality depended not only on the gene silenced but also on the worm strain, indicating that these animals harbor cryptic genetic variation for embryogenesis. We also observed substantial variation in responsiveness to germline RNAi. These results yield insight into the complex regulation of RNAi,
including the identification of a novel candidate locus for germline RNAi sensitivity. The cryptic genetic variation we observed is pervasive. Of the 41 germline genes we silenced, about half exhibited phenotypes that correlate significantly with one or more SNPs across the strains, indicating candidate loci for cryptic variants. However, of the significant SNPs we observed, few are shared across silenced genes. These results describe a genetic architecture of many variants of moderate effect, with low pleiotropy. Silenced genes that revealed particularly strong cryptic variation include par family members, perhaps a surprising result given their fundamental role in polarizing the early embryo. We are currently evaluating candidate loci for functional cryptic variants that may act in the polarity pathway.

O-8
GENOMIC COMPARISONS BETWEEN SELECTED AND RELAXED CHICKEN LINES
Mats Pettersson¹, Anna Johansson¹, Paul Siegel², Örjan Carlborg¹
¹Swedish University of Agricultural Sciences, Uppsala, Sweden, ²Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

In an earlier study, we genotyped individuals from generations 40 and 50 in a bi-directional selection experiment for body-weight in chickens, using a 60k SNP-chip to examine the genome wide effects of selection. The results showed that although selection had acted on many (likely hundreds) of genes, any one point in time selection was focused on a smaller set of genes.

Here, we describe results from a new analysis based on individuals from the selected lines at generation 53, as well as individuals from generation 9 of two relaxed selection lines spliced from the high- and low-selected lines at generation 45. We study genomic patterns on several spatial scales and compare over time, within the populations, as well as across the four different populations. Effects on e.g. expected heterozygosity, haplotype block size and allele frequencies were examined, to unravel the effects of applying, and subsequently relaxing, single trait directional selection over more than 50 generations. Initially we identified small stretches of markers with a consistent trend of allele frequency change over time that likely represent regions under current selection. Sliding windows covering a few hundred markers detected extended regions of near complete fixation, which were used to reconstruct a timeline of fixation based on the shared history of the studied lines. A particularly interesting example is a region on chromosome 1, which displays a drastic drop in diversity in late generations – likely due to a novel mutation that appeared in the high-weight selected line during the selection experiment. Finally, genome wide divergence was found to be higher between the low-weight selected and relaxed lines than in the high-weight selected pair. This is interesting, given that the low-weight selected line has reached a selection plateau, and suggests purging of deleterious alleles from the low-weight relaxed line.

O-9
GENE INTERACTIONS UNDERLYING THE EVOLUTION OF COMPLEX TRAITS
Patrick Phillips
University Oregon, Eugene

No Abstract
O-10
WHAT, ME NATURAL? PATTERNS OF SELECTION IN WILD SYSTEMS AND THEIR THEORETICAL IMPLICATIONS

Jarrod Hadfield1,2
1University of Oxford, Oxford, UK, 2University of Edinburgh, Edinburgh, UK

Most biologists believe that natural selection is stabilising and that trait means are close to their optimum. Contrary to this, evidence from wild populations suggests that selection is often directional. It has therefore been suggested that our estimates of natural selection are unrepresentative because incomplete fitness measures have been used or that short-term studies of selection fail to observe the temporal reversals in sign that would lead to no overall selection (i.e. fluctuating selection). However, studies that use complete fitness measures generally show stronger directional selection and reversals in the sign of selection are uncommon and largely due to measurement error. Given that abundant genetic variance segregates for most traits evolutionary change is expected and indeed has been documented using BLUP methodology. However, the observed response to selection usually falls far short of the expected cumulative response, and in many cases the evidence for a response to selection is not supported by more rigorous hypothesis tests. Measuring selection on the genetic component has been advocated in order to equate expectation with observation. However, because estimates of evolutionary change and estimates of ‘genetic selection’ use the same information in the data they will always be concordant even when the true values are not. Moreover, ‘genetic selection’ conflates selection and inheritance and therefore provides little insight into the two main remaining hypotheses regarding the lack of evolutionary change in natural populations: selection on unmeasured ‘traits’ resulting in the overestimation of i) the strength of selection on the focal trait or ii) the amount of useful genetic variance (through antagonistic genetic correlations). It is suggested that selection and inheritance should be kept distinct in empirical work and that both observational and experimental approaches need to be brought to bear on identifying the causal agents of fitness variation.

O-11
THE DECEIT OF MONOGAMY: QUANTITATIVE GENETIC INSIGHTS INTO THE EVOLUTIONARY ECOLOGY OF POLYANDRY IN THE WILD

Jane Reid
University of Aberdeen, Aberdeen, UK

Multiple mating by females, and hence genetic polyandry, are widespread across the animal kingdom. However, the evolutionary forces driving polyandry remain unclear and much debated. This is because polyandry cannot easily be experimentally manipulated in the wild, and evolutionary responses predicted from phenotypic selection analyses are likely to be biased by environmental covariances between polyandry and fitness components. Recent developments in quantitative genetic analysis and wild population datasets now provide new opportunities to explicitly test evolutionary theories explaining polyandry in the wild. I formulate expressions quantifying components of selection on polyandry in terms of additive genetic variances and covariances among polyandry and fitness components, and then estimate these (co)variances using 18 years of complete genetic pedigree data from free-living song sparrows (Melospiza melodia). First, I demonstrate substantial additive genetic variance and heritability in polyandry, measured as a female’s liability to produce offspring sired by an extra-pair male. Second, I demonstrate that indirect selection on female polyandry through increased additive genetic value of extra-pair offspring can be estimated as the additive genetic covariance between offspring fitness and male net paternity gain through polyandry. I show that, opposite to expectation, this covariance is significantly negative in song sparrows. Third, I demonstrate positive genetic covariance between female polyandry and seasonal reproductive success, predicting that polyandry will show a weak positive response to selection through this fitness component. These analyses provide the first rigorous estimates of key components of
selection on polyandry in the wild and predicted evolutionary responses, and demonstrate the potential for modern quantitative genetic approaches to provide major new insights into long-standing questions in evolutionary and behavioural ecology.

O-12
CAUSES AND CONSEQUENCES OF NEW MUTATIONS
Matt Hurles
Sanger Institute, Cambridge, UK

No Abstract

O-13
THE ROLE OF GENETIC VARIATIONS ON GENE EXPRESSION AND SPICLING IN MULTIPLE REGIONS OF CONTROL IN HUMAN POST-MORTEM BRAIN TISSUE

Daniah Trabzuni1,2, Adaikalavan Ramasamy4, Mina Ryten1, Colin Smith3, Robert Walker1, Michael Weale4, John Hardy1
1Institute of Neurology, University College London, London, UK, 2Department of Genetics, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia, 3Department of Neuropathology, University of Edinburgh, Edinburgh, UK, 4Department of Medical & Molecular Genetics, King’s College, London, UK

Introduction: Over the past few years there has been a realization of the importance of understanding the underlying molecular mechanism of complex neurodegenerative diseases. GWAS studies confirmed a significant association between SNPs with common complex neurodegenerative diseases. This highlights the importance of the genetic variations in individuals and how they contribute to increase susceptibility to neurodegenerative diseases by modulating the gene expression and splicing pattern in the relevant tissue. In this project we are assessing the correlation between genetic SNPs variations with gene expression and alternative splicing regulation in control post-mortem human brain tissues from 10 different regions, which support different functional roles. The regions presented here are frontal cortex, temporal cortex, occipital cortex, white matter, hippocampus, thalamus, putamen, substantia nigra, medulla and cerebellum.

Material and Methods: A total of 1231 RNA samples from 134 control human post-mortem brains were extracted. The QC assessed RNA samples were run on Affymetrix GeneChip® human Exon 1.0 ST Arrays. In parallel, DNAs from brains were run and analysed using the Illumina Omni 1M Beadchips. A regression correlation analysis was performed to identify gene expression and alternative exon splicing in different brain regions.

Results: We report highly significant region-specific exon and gene quantitative trait loci (xQTL and wQTLs respectively) with p-values <10⁻³⁵. Cerebellum and white matter show more unique QTLs in comparison with other brain regions. Furthermore, some QTLs have stronger effect on specific group of regions together than others such as cortical regions.

Conclusions: In addition to the novel regional specific QTLs, this study yields an insight into how different brain regions function through gene expression in normal brain. There are significant risk loci that have been confirmed in GWAS studies required to be taken to a further stage of functional studies. This dataset is a valuable resource for research into the complex genetics of neurodegenerative diseases.
O-14
POPULATION GENETICS REVEALED BY ENORMOUS STRUCTURAL VARIATIONS DISCOVERED BY POPULATION-SCALE SEQUENCING

Quan Long1, Dazhe Meng1,2, Fernando Rabanal3, Christian Huber4, Bjarni Vilhjalmsson1, Qingrun Zhang1, Alexander Platzer1, Viktoria Nizhynska1, Ashley Farlow1, Petar Forai1, Kai Ye4, Ines Hellmann1, Magnus Nordborg1,2
1Gregor Mendel Institute, Vienna, Austria, 2University of Southern California, Los Angeles, USA, 3University of Vienna, Vienna, Austria, 4Leiden University Medical Center, Leiden, The Netherlands

Objectives: To understand the population genetics of Arabidopsis thaliana when structural variations are considered.

Methods: We sequenced ~180 Arabidopsis thaliana inbred lines with Illumina at deep coverage (ranging from 20 to 60). By combining eight different methodologies based on read-pair, read-depth, split-reads, as well as de novo assembling, we discovered enormous number of indels and structural variants. We used Sanger sequencing and chip based genotyping data to validate the indels. Flow Cytometry data that measures the genome sizes is also generated.

Results: (1) Totally 38% of the variance of genome size can be explained by the structural variations. Ribosome repeats and centromere explains most of them, whereas insertions called from de novo assembly explains little variance. This indicates that the genome size variation is largely due to the repetitive region, instead of unique region. (2) We observed a large number of genes that have compensating frame-shift indels. Those pairs of indels are in LD and some of the genes shows footprint of selective sweep. (3) We find that most of the SNPs in long range LD are caused by structural variations, population structure, as well as TEs. In the retained ones after above filtering, we observed enrichment of regions under selection. (4) The recombination hotspots estimated from LD pattern (LDHat) are biased at the region of structural viable region, indicating the necessity of new methodology.

Conclusions: Sequence based whole genome analysis reveals important insight into population genetics that are not accessible by chip based genotyping platforms.

O-15
VARIATION IN TRANSCRIPTION FACTOR BINDING AMONG HUMANS

Maya Kasowski1, Fabian Grubert1, Chris Heffelfinger2, Manoj Hariharan1, Akwasi Asabere4, Sebastian Waszak3, Lukas Habegger1, Joel Rozowsky2, Minyi Shi1, Alexander Urban1, Mi-Young Hong2, Konrad Karczewski1, Wolfgang Huber3, Sherman Weissman2, Mark Gerstein2, Jan Korbel1, Mike Snyder1
1Stanford University, Stanford, CA, USA, 2Yale University, New Haven, CT, USA, 3EMBL, Heidelberg, Germany

Differences in gene expression may play a major role in speciation and phenotypic diversity. Although variations in gene expression among individuals have been documented, the origins of these differences are not clear, and studies that directly measure differences in transcription factor binding sites among humans have not been performed. We have examined genome-wide variation in transcription factor binding in different individuals and a chimpanzee using chromatin immunoprecipitation followed by massively-parallel sequencing (ChIP-Seq). The binding sites of RNA Polymerase II (Pol II) as well as a key regulator of immune responses, NFκB, have been mapped in ten lymphoblastoid cell lines derived from individuals of African, European, and Asian ancestry, including a parent-offspring trio. Using a stringent threshold, approximately 7.5% and 25% of the respective NFκB and Pol II binding regions exhibit differences between any two individuals. To understand the underlying basis of the variations, we examined the effect of SNPs and genomic structural variations (SVs) on binding differences among individuals. We find that many binding differences are associated with SNPs and SVs. Comparison of the binding data with gene expression data generated by RNA-Seq revealed that differences in binding often correlate with gene expression differences. Furthermore, comparison of the Pol II human sites with
Oral Abstracts

binding sites identified in the chimpanzee suggests a high level of divergence in binding relative to our closest evolutionary neighbor. Our results indicate that many differences in individuals occur at the level of TF binding and provide insight into the genetic events responsible for these differences.

**O-16**

**MAPPING THE EPIGENETIC BASIS OF COMPLEX TRAITS IN ARABIDOPSIS**

Frank Johannes  
University of Groningen, Groningen, The Netherlands

Inter-individual differences in DNA methylation states can provide a source of heritable phenotypic variation independent of DNA sequence changes. Accumulating evidence suggests that this mode of epigenetic inheritance is more widespread in plant populations than previously appreciated. Here I highlight our ongoing attempts to characterize the epigenetic basis of complex trait inheritance in an experimental system of so-called Epigenetic Recombinant Inbred Lines (EpiRILs) of the plant Arabidopsis. This population was derived from two parents with nearly identical DNA sequences but drastically divergent DNA methylation profiles. Employing a combination of population genetic approaches and genome-wide epigenomic profiling techniques, we are now obtaining first glimpses of the heritable epigenetic architecture underlying a spectrum of phenotypes in this system. Our work has relevance for understanding mechanisms of adaptation in this species.

**O-17**

**THE ROLE OF TANDEM REPEATS IN MRNA AND PROTEIN EXPRESSION HOMEOSTASIS**

Sreenivas Chavali, Madan Babu M.  
MRC Laboratory of Molecular Biology, Cambridge, UK

The information encoded in the DNA not only influences its own replication but also determines the abundance and function of mRNA, the encoded proteins, their interactions and ultimately the phenotype of a cell. Repeating sequences in the genome are an important source of genetic variation and are prevalent in all eukaryotes. For instance, ~10-20% of the eukaryotic genes contain continuous repeating sequences (e.g., CAG trinucleotide repeats in Huntingtin), referred as Tandem Repeats (TRs). Previous studies have linked variation in repeat length with phenotypes such as human diseases, canine skeletal morphology and have demonstrated few functional roles. However, an integrated understanding (i.e., systems view) of how TRs influence the flow of genetic information at the molecular level, determining phenotypes, under normal physiological conditions, has remained elusive. Through the investigation of multiple genome-wide datasets we provide insights into how TRs in the coding regions of genes and in proteins affect different regulatory steps involved in gene and protein-expression homeostasis in Saccharomyces cerevisiae. At the mRNA level, we show that genes with multiple TRs in the coding regions have higher transcript abundance, mRNA half-life with multiple polyadenylation sites and longer 5’ and 3’ UTRs compared to those without TRs. In contrast, proteins with aminoacid TRs are less abundant with lower ribosomal density and shorter half-life compared to those without TRs. Taken together, our results suggest that TRs in coding regions of genes are associated with enhanced mRNA stability while protein TRs bring about rapid turnover of proteins. Deriving such general principles allows the possibility of elucidating the effect of variations in repeat length on different molecular processes influencing the phenotypic outcomes, especially human diseases and in synthetic biology applications.
O-18
DISSECTING THE CONTRIBUTION OF REGULATORY SEQUENCE VARIATION TO QUANTITATIVE
HAEMATOLOGICAL TRAITS USING MAPS OF OPEN CHROMATIN IN PRIMARY HUMAN BLOOD
CELLS

Cornelis Albers1,2, Dirk Paul2, Augusto Rendon2, Katrin Voss1, Jonathan Stephens1, HaemGen Consortium1,2, Nicole Soranzo2, Panos Deloukas2, Willem H Ouwehand1
1University of Cambridge & NHS Blood and Transplant, Cambridge, UK, 2Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK

A major challenge in genome-wide association studies (GWASs) is to translate the associated sequence
variants into biological mechanisms and pathways, to identify relevant cell types and to characterize
small polygenic effects. We explored these issues using data from two meta-analyses of GWASs in over
60,000 individuals of two quantitative platelet traits (platelet count and volume) and six quantitative
red blood cell traits (including count, volume and haemoglobin). We generated genome-wide maps of
open chromatin using the FAIRE-Seq technique in three primary blood cell types of the myeloid lineage:
megakaryocytes, erythroblasts, and monocytes. Regions of open chromatin, i.e. nucleosome-depleted
regions (NDRs), may represent active gene regulatory elements harboring functional variants. We used
bootstrapped p-value distributions of sequence variants imputed from the 1000 Genomes project to
investigate enrichment of platelet and RBC associations in open chromatin in each cell type.

Considering only the subset of high-confidence NDRs, we found that red blood cell trait associations were
most enriched in NDRs from erythroblasts, the precursor cells for red blood cells, but weakly enriched in
megakaryocytes, the precursor cells of platelets, and monocytes. Platelet trait associations were enriched
in high-confidence megakaryocyte NDRs, and interestingly monocyte NDRs showed a similarly strong
enrichment. Considering the subset of NDRs present in two biological replicates of the same cell type,
which includes weak NDRs supported by relatively few sequencing reads, the strongest enrichment for
platelet traits was found in megakaryocyte NDRs, and for RBC traits in erythroblast NDRs. The strength
and relative orderings of enrichments in NDRs reveal interesting patterns across cell types indicating
distinct biological mechanisms underlying the traits. The enrichments are particularly clear for associations
not reaching genome-wide significance, suggesting that NDRs reveal polygenic effects.

We will report on ongoing eQTL analyses, and the application of variance-partitioning and genotype-
phenotype prediction methods incorporating NDRs as additional information.

O-19
STATISTICAL METHODS FOR THE GENETIC ANALYSIS OF ARBITRARILY STRUCTURED POPULATIONS

John Storey
Princeton University, US

No Abstract

O-20
CLIP TEST: A NEW FAST AND SIMPLE METHOD TO DISTINGUISH BETWEEN LINKED OR
PLEIOTROPIC QUANTITATIVE TRAIT LOCI IN LINKAGE DISEQUILIBRIUM ANALYSIS

Ingrid David1,2, Jean-Michel Elsen1, Didier Concordet1
1INRA UR631, Castanet Tolosan, France, 2INRA-ENVT UMR1331, Toulouse, France

An important question arises when mapping QTLs for genetically correlated traits: is the correlation due
to pleiotropy (a single QTL affecting more than one trait) and/or close linkage (different QTLs that are
physically close to each other and influence the traits)? Several methods have been previously reported to
distinguish between linked or pleiotropic QTLs especially in linkage analysis. None of them considers the
fact that, since the LD between a marker and the QTL is the same for the correlated traits studied under the pleiotropic assumption, the pattern of the SNP effects (i.e. the change in SNP effects when moving along the tested genomic region) should be similar for all these traits whereas it should be different under the close-linked QTL assumption. The Close Linkage versus Pleiotropism (CLIP) test that we propose is based on this consideration. It is a fast, simple and powerful method that compares the square of the observed correlation between a combination of apparent effects at the marker level, to the minimal value it can take under the pleiotropic assumption. A simulation study was performed to estimate the power and alpha risk of the CLIP test in a linkage disequilibrium analysis with dense SNP assays and complex pedigree structure. Performances of the test were compared to the ones of a test that evaluated whether the confidence intervals of the two QTLs overlapped or not (CI test). On average, the CLIP test showed a higher power (68%) to detect close-linked QTLs than the CI test (43%) and a same alpha risk (4%).

O-21
SPARSE FACTOR MODELS FOR ESTIMATING THE GENETIC ARCHITECTURE OF GENE EXPRESSION TRAITS

Daniel Runcie, Sayan Mukherjee
Duke University, Durham, NC, USA

Predicting responses to selection requires knowing the genetic architecture of the traits that underlie variation in fitness. Morphological or phenological traits may be closely associated with fitness and are often easy to measure, but their genetic architecture tends to be highly complex and it is rarely clear how particular genetic variants cause the phenotypic variation. Quantitative genetic studies of gene expression provide a way to indirectly measure cellular, developmental and physiological traits in organisms and thus can provide insights into mechanisms that link genotypes and phenotype. However, classical analytical techniques are poorly suited to quantitative genetics studies of gene expression, where numbers of traits assayed per individual can reach many thousand. Here, we derive a Bayesian sparse factor model for predicting responses to selection of gene expression traits given measures of fitness on pedigreed individuals. We assume that genes function within semi-independent modules (factors), eg. metabolic pathways, and that these modules control the variation in fitness. This implies a factor structure for the genetic covariance matrix (G) of the gene expression traits and fitness which is highly structured and much more parsimonious than classical estimators. We further simplify our estimates of G by forcing factors to influence a low number of genes through sparsity priors. The advantages of this approach are two-fold. First, the factors themselves can be explored to provide biological intuition into the evolutionarily important underlying traits by inspecting functional or pathway classifications of genes in the modules. Second, in many cases this combination of parsimonious modeling and shrinkage will lead to greatly improved prediction accuracy of selection responses, while the Bayesian approach automatically provides credible interval estimates of important parameters. We demonstrate our approach on an existing Drosophila gene expression data set from the Drosophila Genetic Reference Panel.
Low-pass sequencing (LPS) has been used for the past few years because it has the potential to deliver large amounts of genomic information at low cost as a potential alternative to current genome assaying technologies. However, LPS results in lots of missing genotypes in any given individual. In human populations Hidden Markov model imputation algorithms have been adapted to address this issue. Long-range phasing and haplotype library based imputation algorithms have been shown to give higher imputation accuracy compared to Hidden Markov models in animal populations because of the relationship structures that are pervasive, and the same is likely true for plant populations. The objective of this paper was firstly to describe how long-range phasing and haplotype library based imputation can be adapted to impute missing information from LPS in pedigreed and unpedigreed plant and animal populations. The second objective was to describe the construction of an incidence matrix, based on multiple allelic variants at each genomic position, which can be used for genomic prediction. These multi-allelic single locus variants are a novel by-product of this imputation procedure and are derived using information from nearby called, uncalled, and imputed alleles. In commercial Maize populations the method could impute between 2% and 10% of the missing data with strong evidence for the remainder being biologically missing and in cross-validation studies this imputation increased the accuracy of genomic prediction by up to 4%. Further small gains in accuracy of prediction were obtained using the multi-allelic approach. A pilot study is under way to test the procedure involving the use of LPS and SNP chip data in a commercial pig population with a pedigree structure that can be leveraged to enhance the power of the resulting genomic data.

O-23

HAPLOTYPE PHASING USING NEXT-GENERATION SEQUENCING READS

Jonathan Marchini
Oxford University, UK

Next-generation sequencing is now widely used in many studies of human disease and population genetics. Sequencing is still relatively expensive so many studies are collecting low-coverage data that necessitates the use of phasing and imputation methods to call genotypes and haplotypes at polymorphic sites. Sequencing reads carry valuable phase information when spanning two or more heterozygous sites. For example, in the Phase I 1000 Genomes Project data we have found ~33% of the heterozygous sites are covered by phase informative reads. This information is not used by current methods. We have extended SHAPEIT to incorporate the sequencing read information and boost performance. To compare methods and investigate parameters such as coverage, read length and insert size we have carried out an extensive simulation study of sequence read level data using the SFS_code. For example, with 5x paired-end 100bp reads (500bp insert size) using phase-informative reads increases the mean distance between phasing errors from 85.2kb to 94.0kb. The mean switch distance for BEAGLE was much lower at 63.5kb. By adjusting the distributions of read length and insert size performance can be substantially improved. These results indicate the likely benefit of future advances in sequencing technologies. On real data we
have achieved similar performance but have found necessary to carefully account for poorly mapped reads and poor calibration of base qualities. Overall our results highlight the gains that can be achieved by using phase-informative reads when estimating haplotypes from next-generation sequence data.

O-24
WHOLE-GENOME REGRESSION AND PREDICTION OF HUMAN COMPLEX TRAITS USING DATA FROM RELATED AND UNRELATED INDIVIDUALS

Gustavo de los Campos1, Ana Ines Vazquez1, Yann Klementidis1, Daniel Sorensen2
1University of Alabama at Birmingham, Birmingham, Alabama, USA, 2Aarhus University, Tjele, Denmark

In recent years GWAS identified unprecedented numbers of variants associated to complex human traits and diseases (CHT&D). However, these ‘hits’ usually explain a small fraction of genetic risk. Partly, this reflects low power of standard GWAS. In response, consortiums were formed and with increased sample size more variants were detected and prediction accuracy (PAC) increased (Lango Allen et al. Nature, 2010). Yet, PAC remains low and some authors (de los Campos, et al. Nature Reviews Genetics, 2010; Yang et al. Nature Genetics, 2010) proposed using Whole-Genome-Regressions (WGR, Meuwissen et al., Genetics 2001) for analysis and prediction of CHT&D. The Yang study showed that, with nominally-unrelated individuals, WGR on common SNPs explain 45% of the variance of height; a trait with 80% heritability. Makowsky et al. (PLoS Genetics 2011) applied WGR to human height and, using family data (FD), found no missing heritability and a PAC (R2) of 25%, but highly dependent on familial relationships. What explains the differences in the genomic heritability reported by these studies? Can WGR yield high PAC of CHT&D when applied to nominally-unrelated individuals? Using analysis, simulation and real data we show that when markers and QTL are in perfect linkage disequilibrium (LD), PAC is bounded by an index which is a weighted sum of squares of true genomic relationships (GRs) between training and validation samples. The expected value the GRs between unrelated individuals is zero and its variance is low if large numbers of loci are involved. Therefore, the contribution of unrelated individuals to PAC is minimal. Subsequently, we show that the negative impacts of any given level of imperfect LD between markers and QTL on PAC are larger with data from unrelated individuals. Simulation and data-analysis confirmed the analytical results and we propose avenues to improve the PAC of WGR applied to nominally-unrelated individuals.

O-25
RELAXING THE GENETIC MODEL TO IDENTIFY QUANTITATIVE TRAIT LOCI HAVING HETEROGENEOUS EFFECTS

Hugues Aschard1, Noah Zaitlen1, Rulla Tamini1,2, Sara Lindström1, Peter Kraft1
1Program in Molecular and Genetic Epidemiology and Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA, 2Channing Laboratory, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA

Objectives: Genome-wide association studies of quantitative phenotypes are usually conducted using linear regression assuming that risk alleles affecting natural variation in quantitative traits display a linear marginal effect. The utility of this assumption has been confirmed by the wide number of quantitative trait loci (QTL) identified using this approach. Still, summarizing the effect of a QTL with a single linear estimator may have limited statistical power when the magnitude and the direction of the effect depend on the genetic and environmental background of the individuals studied. To improve detection existing approaches have focused so far on strategies to allow for measured effect modifiers. We propose an alternative strategy which relies on a more flexible model for the marginal effect of QTLs.

Methods: We developed a non-parametric score that compares phenotypic distributions by genotypes. This score can capture any non-linear effects such as those displayed in the presence of interactions or
indirect association. We circumvented the non-trivial issue of identifying asymptotical null distributions of the score by using permutation, a solution that is now possible due to the tremendous gains in processing power of computers in the past decades.

**Results:** We show via simulation that such approach can be more powerful than the standard tests for marginal effects when the effect of the QTL is heterogeneous. We further applied our test in a discovery context of a genome-wide association study of mammographic density. We found a significant (\(p\)-value=2.10^{-3}) enrichment of genes differentially expressed in mammary epithelial cell or genes related to breast cancer among the top loci identified by our test, whereas no similar enrichment was observed for the marginal test (\(p\)-value=0.08).

**Conclusions:** Our study demonstrate that innovative statistical tests allowing for non-linear marginal effects and leveraging the recent improvement in processing power may help to discover new QTLs.

**O-26**

**RARE VARIANT PATTERNS IN 14,000 HUMAN SAMPLES AND LESSONS FOR STUDY DESIGN**

Sebastian Zöllner  
(University of Michigan, USA)

**O-27**

**LOW DEPTH, WHOLE GENOME SEQUENCING OF DAI POPULATION DEMONSTRATES SUPERIORITY OVER USE OF WHOLE GENOME GENOTYPING ARRAYS IN UNCOVERING POPULATION STRUCTURE, DEMOGRAPHIC HISTORY, SELECTIVE PRESSURES AND PHENOTYPE ASSOCIATIONS IN NON-EUROPEAN POPULATIONS**

Lachlan Coin\(^1\,\(^2\), Ruoyan Chen\(^1\,\(^3\), Zhang Fan\(^1\,\(^4\), Yingrui Li\(^1\), Yu Chang\(^1\), Yong Wang\(^5\), Rasmus Nielsen\(^1\,\(^5\)

\(^1\)BGI-Shenzhen, Shenzhen, China, \(^2\)Imperial College, London, UK, \(^3\)Wuhan University, Wuhan, China, \(^4\)South China University of Technology, Guangzhou, China, \(^5\)UC Berkeley, Berkeley, USA

The development of whole genome sequencing (WGS) technology provides an opportunity to overcome the limitations of previous approaches for calculating population genetic parameters. In particular, genome-wide inference using SNP genotyping arrays is largely based on a limited set of variants tagging common variation in a reference cohort, rather than all variation actually present in the study samples, which may limit resolution, particularly in the study of non-European populations.

In order to compare inferences made from genotyping arrays with those from low coverage WGS, we generated genotyping as well as low-coverage (4x) WGS data on 80 samples from the Dai population, one of the minority groups in Yunnan province, South China, whose demographic history has yet to be resolved. Genotype concordance between WGS and array derived genotypes was high (99%) after imputation. 31.2%, 35.8% of all Dai SNPs were tagged with \(r^2\) > 50% by the Illumina 610k, and Illumina Omni ‘China’ 1M chip respectively, compared to 31.2%, 37.3% in Han Chinese, and 39.4%, 43.0% in CEPH European samples, indicating bias towards European genomes in array design. WGS identified 3964 SNPs with allele frequency differences between Dai and Han greater than 80%, of which only 5 were visible to the array, indicating a vastly improved ability to identify positive selection with WGS.

WGS data we provided finer resolution of the internal structure of the Dai population than with array data, obtaining more IBD segments, and resolving relatedness within the sample more accurately. We also obtained more precise estimates of population genetic parameters relating to the history of Dai, and performed an evolutionary analysis of the Y and M chromosomes not possible with a 610k Illumina array. Our results suggest the Dai population arose initially from migration from South-Eastern Asia, but was later subject to substantial male-dominated migration from Han Chinese populations.
O-28
THE RELATIONSHIP BETWEEN BINARY DISEASE STATUS AND UNDERLYING HETEROGENEITY IN SUSCEPTIBILITY AND INFECTIVITY

Debby Lipschutz-Powell1, John A. Woolliams1, Piter Bijma2, Andrea B. Doeschl-Wilson1
1The Roslin Institute and R(D)SVS, Easter Bush, Midlothian, UK, 2Wageningen University, Wageningen, The Netherlands

Reducing disease prevalence through selection for host resistance offers a desirable alternative to chemical treatment. Selection for host resistance has proven difficult, however, due to low heritability estimates from field data. The hypothesis underlying our study was that these low estimates may partly be caused by a failure to capture all the relevant genetic variance in disease phenotypes, as standard genetic analyses currently cannot capture potential host genetic variation in infectivity. Host infectivity is the propensity of transmitting infection upon contact with a susceptible individual. There is strong epidemiological and evolutionary evidence for substantial host genetic variation in infectivity. However, infectivity is difficult to measure directly and may need to be inferred from more readily available binary disease data. We have previously shown that genetic variation in infectivity can be partly captured from binary data using a linear or standard generalized linear mixed model (GLMM) including indirect genetic effects. However these statistical models have severe shortcomings as they can’t account for the fact that expression of a host’s susceptibility and infectivity changes over the time course of an epidemic. In this study we combined epidemiological theory with quantitative genetics theory to derive a genetic-epidemiological link function that links binary disease data to the underlying parameters assumed to be under host genetic control (such as host susceptibility g and infectivity f). This function describes the time-dependent infection probabilities inherent in an epidemiological SIR model in terms of individuals’ infection status and underlying parameters (g and f), and thus captures the dynamic aspect of disease progression in the population. The link function was validated with simulated disease data generated by a stochastic genetic epidemiological SIR model. Implications for implementing this link function into genetic analyses and predicting selection response as well as the corresponding data requirements will be discussed.

O-29
HOW DO HOSTS SHAPE THEIR MICROBIAL COMMUNITIES?

Matthew Horton, Natacha Bodenhausen, Joy Bergelson
University of Chicago, Chicago Il, USA

Earlier results from greenhouse experiments suggest that individual accessions of Arabidopsis thaliana, the plant genetic model, host distinct microbial communities. To date, however, it is unclear whether the composition of a host’s microbial community is a heritable trait in a natural setting, or which (if any) host factors are responsible. We have conducted an experiment in which we allowed bacteria and fungi to colonize a mapping-panel of A. thaliana in a field setting. Characterizing each individual accession’s microbial community with next generation sequencing allows us unparalleled insights into the relationship between a host in its natural setting, and its associated microbial community. We find a strong and statistically significant association between host-genotype and the ordinations produced from complementary community ecology techniques (non-metric multidimensional scaling, correspondence analysis, and canonical correspondence analysis). Therefore, we scanned the genome of A. thaliana using genome-wide polymorphism data (~5 million SNPs) to identify the loci associated with the abundances of individual microbial species. I will discuss the most interesting candidate loci, as well as the gene ontology (GO) categories that show enrichment for these associated SNPs. Our results illustrate that host associated microbial communities are indeed heritable traits in ecologically relevant settings, and also highlight interesting candidate genes for follow-up studies.
O-30
QUANTITATIVE ASSOCIATION GENETICS OF HIGH DIMENSIONAL CELLULAR TRAITS: EXTENDING BEYOND EXPRESSION QTLS

Richard Durbin
Sanger Institute, Cambridge, UK

No Abstract

O-31
INTEGRATING GENE EXPRESSION INTO PREDICTIVE HEALTH GENOMICS

Greg Gibson
Center for Integrative Genomics, Georgia Institute of Technology Atlanta, USA

Whereas the current focus of personalized medicine is overwhelmingly exome sequence-oriented, explanation and prediction of ill health is constrained. Differential gene expression is a major component of disease risk, so it stands to reason that direct incorporation of the transcriptome into personalized genomics may improve on interpretation based solely on structural variants. We are taking two approaches to this effort. The first asks to what extent common variation in peripheral blood gene expression is associated with clinical attributes. Although the principal components of transcription are somewhat study specific, covariance of gene expression is pervasive, and we find that eight common axes of variation, each involving hundreds to thousands of genes, are consistently explain over half the variance in population samples. These axes relate to B and T cell signaling, metabolism, inflammation, immunity and anti-viral responses and collectively define a person’s position in immune space. We show how normal variation along the axes is influenced by the environment, and suggest that ill health arises when perturbation pushes a person toward extremes. Rather than a dichotomy of healthy/diseased, normality is itself highly structured and influences individualized risk profiles. Our second approach asks whether there is information at the extremes of the gene expression distribution that is equivalent to defining expression mutations, involving both loss and gain of function. Longitudinal RNASeq of healthy individuals in the Emory-Georgia Tech Center for Health Discovery and Well Being cohort reveals a few dozen divergently expressed genes in each person. In some cases these can be readily linked with clinical attributes. Sometimes these profiles may be traced to rare regulatory variants, sometimes to expression at the extremes of the aforementioned axes of variation. In this way, rare and common variants both contribute to the relationship between gene expression and wellness.
Bipolar disorder is a highly heritable illness that affects more than 1% of the population with substantial lifetime morbidity and societal cost. Despite its high heritability, the identification of causative variation for this illness has proven difficult. Even with the recent outpouring of GWAS data, which has identified variants that may act as risk factors, only a small proportion of the heritability has been explained, and the nature of specific functional variants remain elusive. In contrast, conventional linkage studies in relatively rare, large families afflicted with bipolar disorder and related conditions have previously demonstrated the existence of individual loci of high relative risk, but have lacked the mapping resolution to nominate candidate genes or mutations. The advent of next generation sequencing technology, however, provides the means to exploit the reduced phenotypic and genetic heterogeneity of such extended pedigrees and provides the opportunity to identify gene variants, and thus biological pathways, linked to illness.

Previously, we have described such a large family (family 22) in which bipolar disorder and the related condition, recurrent major depression, show linkage to a 20Mb region on chromosome 4p15-p16 (max LOD=4.4). We have carried out whole genome sequencing of three affected individuals and two unaffected, married-in members of family 22, allowing us to define 3672 SNPs that segregate with illness in family 22. We have assessed the candidacy of these SNPs by considering their predicted functional affects and their frequency in publicly available control data. This analysis has allowed us to prioritise a relatively small number of variants for further analysis. Our study demonstrates that next generation sequencing technology has fundamentally changed the way we can study the genetics of all human disorders, including the most complex, such as bipolar disorder, which have previously proved largely refractory.
O-34
HETEROSIS AS A SYSTEMIC PROPERTY EMERGING FROM THE NON-LINEARITY OF THE GENOTYPE-PHENOTYPE RELATIONSHIP: EVIDENCE FROM METABOLIC MODELS AND TEST-TUBE GENETICS

Dominique de Vienne1, Thibault Nidelet1, Christine Dillmann1, Julie Fiévet3
1University Paris XI, Gif-sur-Yvette, France, 2INRA, Montpellier, France, 3AgroParisTech, Paris, France

Described for more than two centuries, and largely used by breeders, the frequent superiority of hybrids over their parents for quantitative traits remains a debated phenomenon. The genetic and molecular approaches to understand heterosis usually do not rely on any model of the genotype–phenotype relationship. From a generalization of the Wright-Kacser metabolic model of dominance, we hypothesize that heterosis can arise from the non-linearity of the genotype-phenotype relationship, through the multi-dimensional hyperbolic-like response of the flux to variation of enzyme kinetic parameters and/or concentrations. To test this hypothesis, we modelled the carbon metabolism network of Saccharomyces cerevisiae, using a system of differential equations for seventeen reactions. We simulated genetic variability by randomly drawing parental enzyme concentrations with different constraints on total enzyme amounts. Virtual parents were “crossed” to obtain hybrids, and the kinetic model was run to compute the fluxes. We found that most hybrids exhibited either mid-parent or best-parent heterosis for metabolic fluxes. Heterosis was maximized when total enzyme concentration was constrained. The best predictive variables of heterosis proved to be the contrast between parental fluxes and the Euclidian distances between parental enzyme concentrations. Experimental validation was conducted by reconstituting in vitro the upper part of glycolysis. Varying enzyme concentrations in test tubes allowed us to mimic genetic variability. Mixing the content of “parental” tubes led to “hybrids”, the fluxes of which were heterotic under the conditions previously defined. Finally, the decomposition of the flux values into genetic effects (dominance and different types of epistasis) revealed that antagonistic additive-by-additive epistasis effects played the major role in determining the level of heterosis. This conclusion is consistent with observations in quantitative and evolutionary genetics, and provides a model unifying the genetic effects underlying heterosis.

O-35
DISCOVERING EARLY AND LATE REGULATORS OF HAEMATOPOIESIS THROUGH LARGE-SCALE GENOMIC ANALYSES

Nicole Soranzo1, Pim van der Harst2, Christian Gieger3, John Chambers4, Willem Ouwehand5
1Wellcome Trust Sanger Institute, Hinxton, UK, 2University of Groningen, Groningen, The Netherlands, 3Institute of Genetic Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany, 4Imperial College London, London, UK, 5University of Cambridge, Cambridge, UK

In a healthy adult person, approximately 10 billions new blood cells are produced daily in order to maintain steady state levels in the peripheral circulation. The process of blood cell formation from haematopoietic stem cells (HSCs) in the bone marrow and their subsequent survival once released in the circulation, is finely controlled to ensure maintenance of blood cell counts within physiologically ‘normal’ ranges. We have used large-scale meta-analyses of genome-wide association studies to discover genetic variants underlying variation in red and white blood cells, and in platelets identifying in excess of 150 genetic loci (Gieger et al, Nature 2011 and unpublished). We have further applied a host of follow-up analyses, including genome-wide transcription analyses, protein-protein interaction networks, in vitro differentiation of HSCs towards red cell and platelet precursors, and silencing experiments in model organisms (fly, zebrafish and mouse), to understand the function of genes and genetic variants underlying the observed associations. Our results support the notion that the regulation of the formation and survival of blood cells in healthy individuals is mediated through a host of previously unknown regulators. We further show that the newly identified genes are active in the late stages of lineage commitment, and that these genes affect blood cell formation in a prevalently lineage-specific manner. Finally, we show
widespread evidence for evolutionary conservation of gene function for the newly identified genes, and show how model organisms can further enhance our understanding of fate-determining events during human haematopoiesis.

O-36
A SEXUAL ORNAMENT IN CHICKENS IS DETERMINED BY PLEIOTROPIC ALLELES AT HAO1 AND BMP2, SELECTED DURING DOMESTICATION

Martin Johnsson¹, Olle Kämpe³, Leif Andersson², Per Jensen¹, Dominic Wright⁴
¹Linköping Universitet, Linköping, Sweden, ²Uppsala Universitet, Uppsala, Sweden, ³Uppsala University Hospital, Uppsala, Sweden

The genetic analysis of phenotypes and the identification of the causative underlying genes remains central to molecular and evolutionary biology. By utilizing the domestication process it is possible to exploit the large differences between domesticated animals and their wild counterparts to study both this and the mechanism of domestication itself. Domestication has been central to the advent of modern civilization and yet despite domesticated animals displaying similar adaptations in morphology, physiology and behaviour, the genetic basis of these changes are unknown. In addition, sexually selected traits, which have been ‘decoupled’ from natural selection in these animals, can also be analysed with this technique. Though sexual selection theory has been the subject of a vast amount of study, almost nothing is known about the underlying genetics of such traits. We have generated multiple intercrosses and advanced intercrosses based on wild derived and domestic chickens to fine-map genomic regions (or QTL) affecting a sexual ornament (one to less than 100kb in size). These regions have been over-laid with putative selective sweeps identified in domestic chickens, and found to be significantly associated with them. By using expression QTL analysis, we show that two genes in one region, HAO1 and BMP2, are controlling multiple aspects of the domestication phenotype, from a sexual ornament to multiple life history traits. This demonstrates the importance of both pleiotropy and close linkage in controlling these genetic changes.

O-37
DIAMONDS IN THE DIRT: BIOLOGICAL AND TRANSLATIONAL INSIGHTS INTO TYPE 2 DIABETES FROM LARGE-SCALE GENETIC STUDIES

Mark McCarthy
University of Oxford, Oxford, UK

Type 2 diabetes represents one of the major challenges for global health in the decades to come, yet a poor understanding of the biological basis of this disease compromises efforts to develop effective strategies for its treatment and prevention. As a result, T2D has also been one of the diseases most intensively targeted for genetic and genomic analysis in recent years, powered by the hope that these increasingly-powerful approaches will deliver insights into the fundamental defects involved in diabetes predisposition. Common variant-focused genome wide association studies have contributed to the identification of >60 loci influencing predisposition to T2D. This work has provided important clues to the relative roles of beta-cell function and insulin resistance to T2D predisposition; demonstrated that the epidemiological association between birthweight and later T2D is driven in part by genetic factors; and highlighted the contribution of defects in Wnt-signalling and cell cycle regulation (in the islet), and of KLF transcription factors signaling (in the adipocyte) to the development of disease. Current efforts in T2D genetic discovery have two major goals: the disclosure of novel loci and the characterization of causal variants within the association signals discovered. These efforts aim to further illuminate the mechanisms involved in disease pathogenesis and help explain a greater proportion of the genetic variance, enhancing the opportunities for stratified medicine. This work is increasingly reliant on the power of next-generation
sequencing to provide an exhaustive catalog of DNA sequence variation in large numbers of subjects with and without diabetes, and to define the architecture of genomic regulation in the cell types of particular interest (notably the pancreatic beta-cell). Efficient use of these data is dependent on substantial methodological developments related to the acquisition and analysis of rare allele data, some of which I will discuss.

Over the next five years, we can be confident of our ability to characterize the basis of genetic predisposition to T2D. Work in cellular and animal models, and in man, will have taken these genetic signals and defined the mechanisms through which they alter risk of diabetes. The challenges will lie increasingly in the translation of these findings into improvements in the capacity to prevent and/or treat diabetes.

O-38
THE BIOLOGIC RELEVANCE OF eQTLs: A GENOME WIDE CONFIRMATION OF eQTLs USING TWO TYPES OF INBRED POPULATIONS IN C. ELEGANS

Mark G. Sterken1, L. Basten Snoek1, Roel P.J. Bevers1, Rachel C. Brenchley2, Arjen E. van ‘t Hof2, Andrew R. Cossins2, Jan E. Kammenga1

1Wageningen University, Laboratory of Nematology, Wageningen, The Netherlands, 2University of Liverpool, Centre for Genomic Research, Institute of Integrative Biology, Liverpool, UK

With the advent of genetical genomics as a powerful tool for studying the genetic architecture of complex traits, questions arise with regard to the reproducibility and biological relevance of this method. Genetical genomics studies are usually conducted in recombinant inbred lines (RILs). A panel of RILs is a genetic mosaic of the parental genome and statistically very attractive for mapping of complex traits. But the biological effects of the genetic recombination can be hard to entangle. Therefore a population of near introgression lines (NILs), containing only a single introgressed part of one of the parental genomes, can be used to confirm the detected eQTLs.

Here we set out to biologically confirm genome wide eQTLs mapped in a RIL population with a NIL population. The two populations were created from the genetically most divergent C. elegans isolates, Bristol N2 and Hawaii CB4856 [Li et al. 2006; Doroszuk et al. 2009]. The genome wide transcription in L4 larvae was measured by microarray in both RILs and NILs. Quantitative trait locus mapping was applied to find the genomic loci that underlie the variation in transcriptional responses between the RILs. In the NILs significantly up or down regulated genes were identified independently from the RIL population.

Comparison of the outcomes show that the eQTLs detected using a RIL population can be confirmed to a high degree in a NIL population. Next to genome-wide confirmation, also trans-bands mapped in the RIL population were supported by the NIL population. These results provide a unique insight into the robustness of quantitative genetic analyses and exemplifies the biologic relevance of its outcomes.

Doroszuk et al. (2009) NAR 37: 16, e110
Li et al. (2006) Plos Genetics 2(12): e222

Funded by ERASysBio+ project GRAPPLE - Iterative modelling of gene regulatory interactions underlying stress, disease and ageing in C. elegans
**O-39**
SYSTEMS GENETICS OF LIFE SPAN IN THE DROSOPHILA GENETICS REFERENCE PANEL

Michael Magwire, Steven Vensko, Wen Huang, Brittny Calsbeek, Robert Anholt, Eric Stone, Trudy Mackay
NC State University, Raleigh, NC, USA

We are interested in understanding the underlying genetic architecture of various traits in *Drosophila melanogaster*. One of the primary resources we employ to this end is the *Drosophila* Genetics Reference Panel (DGRP). The DGRP consists of 192 inbred *D. melanogaster* lines derived from a natural population in Raleigh, NC USA. Complete genome sequences for these lines have been obtained using Illumina technology. In addition, expression data has been collected for these lines using Affymetrix tiling arrays (two replicates each for females and males). Using the sequencing data, we perform genome wide association (GWA) analyses to identify single nucleotide polymorphisms (SNPs) associated with many quantitative traits (quantitative trait loci, QTLs). We also use the sequencing data to adjust probe binding intensities on tiling arrays for probes containing SNPs relative to the reference genome. These adjusted probe intensities are then used to generate gene expression estimates for 14,637 genes in each of the DGRP lines, for both females and males. We then identify genetically variable transcripts, and perform GWA analyses to identify SNPs associated in cis and trans with each expression trait (expression QTLs, eQTLs). We also identify which of the genetically variable transcripts are also associated with each quantitative trait (quantitative trait transcripts, QTTs). Finally, we consider trios of QTLs, eQTLs and QTTs to derive causal models associating genetic, transcriptional and phenotypic variation. We illustrate the results of the systems genetics analysis using life span data from virgin flies.

**O-40**
VARIANCE CONTROLLING GENES AND THEIR ROLE IN THE GENETIC ARCHITECTURE OF COMPLEX TRAITS

Orjan Carlborg
Swedish University of Agricultural Sciences, Uppala, Sweden

The phenotypic effect of a gene is normally described by the mean difference between alternative genotypes. A gene may, however, also influence the phenotype by causing a difference in variance between genotypes. Here, I will introduce some of the recent work to develop theory and tools for genome-wide mapping of individual variance-controlling loci. Empirical findings will be used to illustrate the contribution of such loci to the genetic architecture of complex traits and the implications of the findings on our understanding of the genetic regulation of complex trait variation will be discussed.

**O-41**
GWAS: 2D, OR NOT 2D: THAT IS THE QUESTION

Gibran Hemani¹, Sara Knott², Chris Haley¹,²
¹The Roslin Institute, University of Edinburgh, Edinburgh, UK, ²MRC Human Genetics Unit at the MRC Institute of Genetics and Molecular Medicine at the University of Edinburgh, Edinburgh, UK, ³Diamantina Institute, University of Queensland, Brisbane, Australia, ⁴Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

A serious, empirical investigation into the prevalence of epistasis is long overdue. Here we show that from an evolutionary perspective the search for epistasis is theoretically justified, and we present software and analytical methods that can effectively perform these analyses.

First, we show that incomplete linkage disequilibrium (LD) between causal variants and observed SNPs is prone to causing epistatic patterns of variation to appear additive. In addition the decay of non-additive variance with decreasing LD proceeds much more rapidly in contrast to additive variance, so without
sequence or very dense genotype data or there is a strong bias towards the conclusion that most variants act additively.

Second, it is shown that when epistatic loci are under selection they stabilise at frequencies that release mostly non-additive variance, and that in fact they are unlikely to maintain much additive variance at all. Consequently, the most powerful method of detecting these variants is through exhaustive two-dimensional scans on dense SNP chips.

Third, software to perform such scans rapidly on dense SNP chips will be presented, and we use this to show that even with sequence data, the perception that significance thresholds for 2D searches will be too extreme is inaccurate, and that the multiple testing penalty for even sequence data is not prohibitive for exhaustive searches.

Finally, we present a new analytical method that further ameliorates the power limitations of epistatic searches. Decades of theory advocate the inclusion of epistasis in complex trait analysis, and now that sufficiently large datasets are available we conclude that searching for epistasis is painless, warranted, and long overdue.

O-42
KNOWLEDGE-DRIVEN ANALYSIS IDENTIFIED A GENE-GENE INTERACTION AFFECTING HIGH-DENSITY LIPOPROTEIN CHOLESTEROL LEVELS IN MULTI-ETHNIC POPULATIONS

Alon Keinan1, Li Ma1, Ariel Brautbar2, Eric Boerwinkle3, Charles Sing4, Andrew Clark1
1Cornell University, Ithaca, NY, USA, 2Baylor College of Medicine, Houston, TX, USA, 3University of Texas, Houston, TX, USA, 4University of Michigan, Ann Arbor, MI, USA

While the importance of epistasis is well appreciated, specific gene-gene interactions have been rarely identified and replicated in humans. We tested for interactions underlying four lipid levels that are important risk factors for coronary artery disease. To increase power, we considered only SNPs based on prior knowledge of each of established GWAS hits, protein-protein interactions, and pathway information. Testing for interactions using 9713 European Americans (EA) from the Atherosclerosis Risk in Communities (ARIC) study, we identified a significant interaction between HMGCR and LIPC in their effect on high-density lipoprotein cholesterol (HDL-C) level (Bonferroni corrected $P_c = 0.002$). Due to the decrease in the effect of tagged markers, gene-gene interactions are quadratically less likely to replicate than marginal associations. Hence, we employed an adaptive method that leverages signals from proximate linked SNPs. We successfully replicated this gene-gene interaction in EAs from the Framingham Heart Study ($P_c = 0.002$) and from the Multi-Ethnic Study of Atherosclerosis (MESA; $P_c = 0.006$). The interaction between the two loci is also significant in the ARIC African American sample ($P_c = 0.004$) and in the Hispanic American sample from MESA ($P_c = 0.04$), for a combined evidence for a gene-gene interaction of $P_c = 9.0 \times 10^{-8}$ following a conservative Bonferroni correction. Both these genes are involved in the metabolism of lipids, with previous GWAS reporting an association of LIPC with HDL-C, and an association of HMGCR with LDL-C (but not with HDL-C). Importantly, this novel gene-gene interaction explains 0.3% of the overall variation in HDL-C beyond the additive effect of the two genes themselves. For comparison, the whole set of associated variants in a recent meta-analysis of 46 GWAS explains 12% of the variation in HDL-C level (Teslovich et al. 2010). Our results suggest that it is possible to detect and replicate gene-gene interactions in current GWAS data. However, interactions affecting complex traits—much like main effects—can be of small magnitude, and hence might easily be masked by the trillions of tests required for a genome-wide study. The approach we undertook of focusing on a tiny fraction of SNPs between which epistasis signals are likely to be enriched, together with a locus-based replication, may hold the key to detecting interactions underlying complex human disease and traits.
O-43
THE GENETIC ARCHITECTURE OF QUANTITATIVE TRAITS: LESSONS FROM BIOCHEMICAL MARKERS FOR DISEASE

Beben Benyamin\textsuperscript{1,2}, Rita Middelberg\textsuperscript{2}, Grant Montgomery\textsuperscript{2}, Peter Visscher\textsuperscript{1}, Nicholas Martin\textsuperscript{2}, John Whitfield\textsuperscript{2}

\textsuperscript{1}University of Queensland, Brisbane, Australia, \textsuperscript{2}Queensland Institute of Medical Research, Brisbane, Australia

Biochemical markers measured in blood are widely used indicators of a current disease or predictors of disease risk. Genetic factors are significant sources of individual difference in biochemical markers with heritability ranging from 25 to 75%. Genome-wide association studies (GWAS) have identified a number of single nucleotide polymorphisms (SNPs) associated with biochemical markers. However, the genetic architecture of the biochemical phenotypes remains poorly understood. The aims of this study are to: 1) assess the spectrum of allelic effect sizes; 2) assess the contribution from SNPs whose associations do not reach genome-wide (GW) significance; 3) compare the sum of SNP-associated effects against heritability; 4) contrast the results across phenotypes. We analysed 16 biochemical markers measured in serum from ~11,000 individuals from the Queensland Institute of Medical Research (QIMR) twin and family studies, in which genome-wide (~2.4 million) SNPs data were available. These include markers for cardiovascular disease, such as triglycerides, LDL- and HDL- cholesterol. For most biochemical markers, GWAS found several GW-significant SNPs that in total account for <5% of the phenotypic variance. For some phenotypes, including bilirubin, cholinesterase, transferrin and uric acid, GW-significant SNPs explain between 10 to 20% of the phenotypic variance. On the other hand, no GW-significant SNPs were identified for alanine aminotransferase, total protein or albumin. By considering all SNPs simultaneously (Yang \textit{et al} Nat Genet 42:565-9, 2010), we found that common SNPs account for 20 - 80% of the heritability (except for albumin where common SNPs do not explain any of the heritability). For most phenotypes there is still some ‘heritability gap’ – presumably due to incomplete LD between causal variants and tagging SNPs.

O-44
FROM CASTLE TO THE COLLABORATIVE CROSS: EVOLUTION OF MOUSE MODELS IN QUANTITATIVE GENETICS RESEARCH

Daniel Pomp
University of North Carolina - Chapel Hill, Chapel Hill, NC, USA

Use of mouse models has been a driving force in understanding genetic architecture of quantitative traits. Commonly used inbred strains of mice constitute the primary mammalian model system. Within these lines and their derivatives, such as recombinant inbred lines, genomewide congenic strains, chromosome substitution lines, advanced intercross lines, long-term selection lines, and heterogeneous stocks, there exists a vast array of relevant genetic and phenotypic variation. Nearly 100 years of studying such variation has shed significant light on the genetics and genomics of a huge variety of traits. Despite these significant efforts, use of mouse models for analysis of complex traits is at a crossroad. Mouse crosses have been successful in populating predisposition maps with QTL, but few of these loci have been robustly characterized at the gene level. At the same time, GWAS in humans and livestock species are enabling direct identification of genes for complex traits with both biomedical and agricultural significance. Thus, new and improved mouse models and paradigms to use them are required. The Collaborative Cross (CC) is a recombinant inbred panel derived from eight strains that capture a level of genetic diversity unsurpassed by any existing mouse resource. The randomly distributed genetic variability and eternal reproducibility of the CC makes it a superior resource for systems genetics research, providing a platform for investigation into genetic architecture underlying complex interactions among diverse phenotypes and molecular networks. Current use of the CC includes the JAX Diversity Outbred
population derived from partially inbred Collaborative Cross strains, while future applications CC will exploit recombinant inbred intercrosses. Integration of rich phenotypic and genomic data over time and across a variety of fields will be vital to delivering on the key attribute of the CC, a common genetic reference platform for identifying causative variants and genetic networks controlling complex traits in mammals.

O-45
THE GENETIC ARCHITECTURE AND EVOLUTION OF QUANTITATIVE TRAITS

Michael Goddard1,2, Ben Hayes2
1University of Melbourne, Melbourne, Victoria, Australia, 2Department of Primary Industries, Victoria, Bundoora, Victoria, Australia

Complex traits are, by definition, affected by many genes but an important question is how many genes and what are their effects and allele frequencies. We have studied these questions for traits in 20,000 dairy and beef cattle. 700k SNP genotypes were used as explanatory variables in a linear model. A Bayesian model was used in which the effects of the SNP were distributed as a mixture of normal distributions. An MCMC procedure was used to estimate the mixing proportions. Most SNPs were found to have zero effect but thousands each explain 0.0001 of the genetic variance and only 11 each explain 0.01 of genetic variance. The alleles estimated to have effects by the analysis are slightly rare but not extremely so. Thus the genetic architecture of these traits is that they are controlled by 1000’s of genes with the number of genes falling in inverse proportion to the size of the effect and with a spectrum of allele frequencies biased towards rare alleles compared with the neutral model. This architecture is not expected from most models of the evolution of quantitative trait loci (QTL). A model that does predict the observed results includes: mutations at 600,000 sites within the genome affect a typical quantitative trait, mutations with very small effect are much more common than those of large effect, selection against these mutations increases more than linearly with the size of the effect. Under this model most of the mutation variance is due to genes of large effect but most of the standing genetic variance is due to genes of small effect.

O-46
THE QUANTITATIVE GENETICS OF PHENOTYPIC ROBUSTNESS

Hunter Fraser1, Eric Schadt2
1Stanford University, Stanford, CA, USA, 2Mount Sinai School of Medicine, New York, NY, USA

Phenotypic robustness, or canalization, has been extensively investigated both experimentally and theoretically. However, it remains unknown to what extent robustness varies between individuals, and whether factors buffering environmental variation also buffer genetic variation. Here we introduce a quantitative genetic approach to these issues, and apply this approach to data from three species. In mice, we find that for hundreds of gene expression traits, robustness is polymorphic and can be genetically mapped to discrete genomic loci. Moreover, we find that the polymorphisms buffering genetic variation are distinct from those buffering environmental variation. In fact, these two classes have quite distinct mechanistic bases: environmental buffers of gene expression are predominantly sex-specific and trans-acting, whereas genetic buffers are not sex-specific and often cis-acting. Data from studies of morphological and life-history traits in plants and yeast support the distinction between polymorphisms buffering genetic and environmental variation, and further suggest that loci buffering different types of environmental variation do overlap with one another. These results suggest that naturally occurring polymorphisms affecting phenotypic robustness are abundant, and that these polymorphisms generally buffer either genetic or environmental variation, but not both.
O-47
STUDYING THE GENOTYPE-PHENOTYPE MAP IN ARABIDOPSIS

Magnus Nordborg
Gregor Mendel Institute, Vienna

No Abstract

O-48
NOVEL METHODS OF GWAS FOR MAPPING GENES OF COMPLEX TRAITS AND THEIR APPLICATIONS IN CROP BREEDING

Jun Zhu
Agronomy Department, Hangzhou, China

It is a challenge to develop efficient statistical methods for mapping genes underlying complex traits. We have developed new mapping approaches that integrate the detection of gene-to-gene interaction and gene-to-environment interaction for quantitative trait nucleotides (QTNs) based on SNPs, and for quantitative trait transcripts (QTTs) based on variation in expression of transcripts. The genetic models can include cofactors and treatments, genetic main effects (additive A, dominance D), epistasis effects (additive by additive AA, additive by dominance AD, dominance by dominance DD), and gene-to-environment interaction (AE, DE, AAE, ADE, and DDE). Mixed linear model approaches are used for unbiased prediction of all the genetic main effects, epistasis effects, as well as gene-to-environment interaction effects. The variation contributing due to these effects can be estimated. Mapping software (QTLNetwork 3.0 and GMDR-GPU) have also been developed, which can be used under different operating systems (Windows, Mac, and Unix). By using the detected QTS and QTT information, we can then use our newly developed Web based knowledge-deliver tool (BioPubInfo) for searching gene-network and real gene functions. The applications of gene mapping and molecular selection in crop breeding for complex traits have been discussed.

O-49
NONLINEAR GENOTYPE × ENVIRONMENT INTERACTION

Rong-Cai Yang\textsuperscript{1,2}
\textsuperscript{1}University of Alberta, Edmonton/Alberta, Canada, \textsuperscript{2}Alberta Agriculture and Rural Development, Edmonton/Alberta, Canada

The presence of genotype × environment interaction (GEI) remains to be a major impediment to genetic improvement in crop plants and farmed animals. Voluminous literature has focused on the development and use of numerous linear-bilinear models for describing and characterizing the standing GEI variation but it has contributed little to our understanding of its genetic and physiological bases. In the classical joint regression analysis that uses a single environmental index as a surrogate to represent all of the complex and unobservable characteristics of the environment, the total GEI variation is partitioned into a portion due to differential responses of genotypes to changes in environmental gradient (linear regression) and the residual. However, linear regression often accounts for only a small portion of the total GEI variation and the majority of GEI variability remains unexplained. Moreover, in the joint regression analysis, both sub- and super-optimal environmental conditions would lead to a reduced performance but they are indistinguishable. In this presentation, we will explore the use of several nonlinear functions for capturing and recovering more GEI variation and for better characterizing genotypic responses to sub- and super-optimal environments, thereby providing more opportunities to understand the biological
basis of GEI. It will be shown that the nonlinear analysis is of value in matching ‘right’ genotypes with ‘right’ environments. An empirical analysis of barley variety trials over the Canadian Prairies will be also discussed.

**O-50**

**APPLYING QUANTITATIVE GENETICS TO EPIDEMICS AND DISEASE RESISTANCE: IMPLICATIONS OF DYNAMIC AND NOISY DATA.**

*Steve Bishop*

*The Roslin Institute & R(D)SVS, University of Edinburgh, Midlothian, UK*

Foremost amongst complex traits addressed by animal geneticists is resistance to infectious disease. Under controlled conditions considerable host genetic variation in resistance is invariably seen, but obtaining sufficient phenotypes for genomic studies usually requires field data; such data is noisy and complex. Issues include misclassification of animal disease state, incomplete exposure to infection and variable infection pressure as epidemics progress. Theoretical results and challenges are outlined here. First, assume static conditions, with probability of exposure to the pathogen (e), diagnostic test sensitivity and specificity S_p and S_e, and true and observed prevalence p and p'. Then, estimated allelic substitution effects (\(\alpha\)) from an association analysis are 

\[
\left(S_p + S_e - 1\right)\varepsilon \alpha
\]

and estimated heritabilities on the observed and liability scales are

\[
\varepsilon^2 \left(S_p + S_e - 1\right)^2 \phi \left(x_p\right)^2 p^{-1} \left(1 - p^{-1}\right)^{-1} h^2
\]

and

\[
\varepsilon^2 \left(S_p + S_e - 1\right)^2 \phi \left(x_p\right)^2 \phi \left(x_{p'}\right)^2 h^2,
\]

where \(h^2\) is the true liability heritability. Dynamic conditions increase complexity; force of infection varies spatially (between environments) and temporally (changing during epidemics), resulting in nonlinear changes in probability of infection. Consequently, the estimated mode of inheritance for resistance may vary from recessive at low infection pressure, through additive to dominant at high infection pressure. The quantitative genetic control of performance traits in the face of infection also requires redefinition. Performance phenotype when infected (\(P_I\)) may be expressed as 

\[
P_I = P_0 + T f(I),
\]

where \(P_0\) is performance when uninfected, \(T\) is tolerance of infection and \(f(I)\) represents some function of level of infection. Infection level, itself, is a joint function of an animal’s resistance and the force of infection, which in turn is influenced by the population mean resistance. Hence performance depressions occur as the product of tolerance and resistance, and are also influenced by associative effects, i.e. the mean resistance of cohorts. Challenges remain: to fully describe dynamic impacts of disease epidemics on genetic interpretations, to develop statistical methods to jointly estimate genetic and epidemiological parameters, and to optimise experimental designs utilising epidemic data.
O-51
THE CONSEQUENCES OF INDIRECT GENETIC EFFECTS FOR HERITABLE VARIATION AND RESPONSE TO SELECTION
Piter Bijma
Wageningen University, Wageningen, The Netherlands

In classical quantitative genetics, phenotypic variance is partitioned into a heritable component, additive genetic variance, and a residual. Additive genetic variance reflects both the contribution of heritable effects to phenotypic variance, and potential response to selection; Response is the product of strength of selection and additive genetic variance, $R = Gb$. The traditional QG-model, however, ignores the social organisation of life. With social interactions, trait values and fitness may depend on genes in others, a phenomenon known as indirect genetic effects. Moreover, some traits cannot be attributed to an individual, but are properties of groups or societies, such as the number of prey caught by a hunting pack. Classical QG-theory does not define heritable variation for such traits, and expressions for response do not distinguish selection from heritable variation. Kin-selection theory, for example, is poorly integrated into the standard QG-framework. Here I review recent developments in this field, focussing on heritable variation and response to selection, and discuss future research. I will present a general definition of the heritable variance that determines potential response to selection, which applies to any inheritance model, and reduces to additive genetic variance in the classical model. Relatedness between focal individual and the individuals affecting its fitness, and the level of selection are key determinants of the utilization of heritable variance for response. Future research questions include the importance and estimation of IGEs in natural populations, the identification of the traits causally underlying the IGEs, the genetic architecture of IGEs, the constraint of response to selection due to IGEs, the potential contribution of IGEs to long-term response, and the importance of drift vs. selection for traits affected by IGEs.

O-52
THE GENETICS OF SOCIAL DOMINANCE: DOES COMPETITION CAUSE CONSTRAINT?
Alastair Wilson
University of Edinburgh, UK

No Abstract

O-53
SIMPLE AND EFFECTIVE METHODS OF ADDRESSING COMPETITIVE EFFECTS IN ANIMAL BREEDING PROGRAMS
William Muir, Allan Schinckel
Purdue University, West Lafayette, Indiana, USA

Introduction: At the 2nd International Congress of Genetics, Griffing (1977, Proc 2nd Int. Conf. Quant. Genet., pages 413-434) generalized his multilevel selection methods to include an optimum index of direct and associative effects, with individuals housed in any family group structure, including random. Earlier Griffing advocated a suboptimal method for addressing competitive effects, termed group selection, or simply selection of entire families housed in groups. The power of group selection has been demonstrated both experimentally and commercially. Unfortunately group selection also results in a rapid increase in inbreeding due to selection of entire families. However, Griffing (1977), and also Bijma, et. al. (2007, Genetics, 175: 277-288), gave theoretical results showing that individual selection with individuals housed in family groups is usually superior to group selection. As such the rate of inbreeding should not be greater than that of traditional breeding programs. However, experimental verification of individual selection in family groups vs. random allocation was lacking.
Methods: Two groups of Japanese quail were mated, 1 male to 4 females, and housed in groups of 16, but assigned either at random (RG) or as half-sib groups (KIN). Birds were selected for increased 6 week weight based on traditional animal model BLUP for 25 cycles.

Results: Birds selected in RG resulted in a negative response to selection for body weight at 6 weeks of age (BW) (-.074±.26g/generation), in contrast, birds selected in KIN significantly increased BW (.749±.21g/generation). Similarly, mortality increased with RG (.3±.2/generation) but decreased with KIN (-.2±.2/generation).

Conclusions: These results show that competitive effects can be easily addressed in commercial breeding operations using traditional animal model BLUP, but with animals housed as families rather than at random.

O-54
TOWARDS GENOMIC PREDICTION FROM GENOME SEQUENCE DATA AND THE 1000 BULL GENOMES PROJECT

Ben Hayes1,2, Charlotte Anderson1,2, Hans Daetwyler1,2, Ruedi Fries4, Bernt Guldbrandtsen5, Mogens Lund6, Didier Boichard6, Paul Stothard5, Roel Veerkamp8, Ina Hulsegge8, Dominique Rocha6, Curt Van Tassell9, David Coote1,2, Mike Goddard1,2, The 1000 Bull Genomes Consortium0

1Biosciences Research Division, Department of Primary Industries, Victoria, Australia, 2Dairy Futures Cooperative Research Centre, Victoria, Australia, 3La Trobe University, Victoria, Australia, 4Lehrstuhl fuer Tierzucht, Technische Universitaet Muenchen, Freising, Australia, 5Aarhus University, Faculty of Science and Technology, Department of Genetics and Biotechnology, Tjele, Denmark, 6INRA, Jouy en Josas, France, 7Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, Canada, 8Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, Wageningen, The Netherlands, 9Bovine Functional Genomics Laboratory, Agricultural Research Service, USDA, Beltsville, USA

Using full genome sequence data in genomic prediction could be advantageous in at least three situations. If linkage disequilibrium between SNP on standard arrays and causative mutations affecting the quantitative trait is incomplete, accuracy of prediction could be improved by including the causative mutations in the data set. Secondly, if genomic predictions are made across breeds, using full sequence data is likely to be particularly advantageous, as there is no longer the need to rely on marker-associations which may not persist across breeds. Thirdly, persistence of accuracy of genomic predictions across generations should be improved with full sequence data. Unfortunately the cost of sequencing is such that it is unlikely that the entire reference population will be sequenced. An alternative strategy is to sequence key ancestors of the population, then impute the genotypes for the sequence variants into much larger reference sets with phenotypes and SNP panel genotypes. The 1000 Bull Genomes Project aims to build this database of sequenced key ancestor bulls for the bovine research community. Thus far the data set consists of 133 full genome sequences of Holstein and Fleckvieh bulls, sequenced at an average of 11.3 fold coverage. There were 17.4 million filtered sequence variants detected in the sequences, including 15.8 million SNP and 1.6 million insertion-deletions. Agreement of sequence genotypes to genotypes from a SNP array in the sequenced bulls was excellent at 98.8%. This increased to 99.7% when the genotypes were imputed given the probability of the genotypes from all sequences. This project will provide a strong opportunity to identify the most important causative variants. As a consequence, it will help to understand the biology, to differentiate shared causative variants across breeds from within breed variants. Finally, examples are given of genomic predictions for quantitative traits using imputed sequence data.
O-55
USING WHOLE GENOME SEQUENCE DATA TO PREDICT QUANTITATIVE TRAIT PHENOTYPES IN DROSOPHILA MELANOGASTER
Ulrike Ober1, Julien F. Ayroles2,3, Eric A. Stone2, Stephen Richards4, Diánhui Zhu2, Richard A. Gibbs4, Christian Stricker1, Daniel Gianola1, Martin Schlather1, Trudy F.C. Mackay2, Henner Simianer1
1Georg-August-University Göttingen, Animal Breeding and Genetics Group, 37075 Göttingen, Germany, 2North Carolina State University, Department of Genetics, Raleigh, NC 27695-7614, USA, 3Harvard University, Department of Organismic and Evolutionary Biology, Cambridge, MA 02138, USA, 4Baylor College of Medicine, Human Genome Sequencing Center, Houston, TX 77006, USA, 5agn Genetics GmbH, Davos 7260, Switzerland, 6University of Wisconsin-Madison, Department of Animal Sciences, Madison, WI 53706, USA, 7Georg-August-University Göttingen, Centre for Statistics, 37077 Göttingen, Germany

The ability to accurately predict values of complex phenotypes from genotype data will revolutionize plant and animal breeding, personalized medicine, and evolutionary biology. To date, genomic prediction has utilized high density single nucleotide polymorphism (SNP) genotyping arrays, but the availability of sequence data opens new frontiers for genomic prediction methods. Our study is the first application of genomic prediction using whole genome sequence data in a substantial sample of a higher eukaryote. We use ~ 2.5 million SNPs determined by sequencing the Drosophila Genetic Reference Panel population of inbred lines to predict phenotypes for the traits resistance to starvation stress, startle-induced locomotor behavior, and chill coma recovery. A genomic relationship matrix is constructed from the SNP data and used in a genomic best linear unbiased prediction (GBLUP) model. Predictive ability is assessed as the correlation between predicted genetic values and observed phenotypes via cross-validation (CV). The study systematically addresses prediction within vs. across sexes, GBLUP vs. a Bayesian approach, and the effect of SNP density. For starvation resistance and startle response, we find that (i) genomic prediction can be efficiently implemented with full genome sequence data via GBLUP; (ii) there is little gain in predictive ability if the number of SNPs is increased above 150,000; and (iii) neither implicit nor explicit marker selection substantially improves the predictive ability of the models. In a five-fold CV with GBLUP the predictive ability is 0.239 ± 0.008 (0.230 ± 0.012) for starvation resistance (startle response). In contrary, the GBLUP model completely fails for chill coma recovery (-0.038 ± 0.010). We itemize possible explanations for this, e.g. the bimodal distribution of phenotypes and the existence of many epistatic interactions between SNPs. Our results will help to assess the potential benefits of sequenced-based prediction applied to non-model organisms.

O-56
NEXT GENERATION BREEDING USING GENOTYPING-BY-SEQUENCING
Jose Crossa1, Paulino Perez6, John Hickey3, Juan Burgueño1, Gustavo de los Campos2, Ed Buckler4, Gary Atlin1, Joseph Beyene1, Kassa Semagn1, Sharon Mitchell5, Jeff Glaubitz5
1CIMMYT, Mexico DF, Mexico, 2University of Alabama, Birmingham, Alabama, USA, 3University of New England, Armidale, Australia, 4USDA-ARS, Ithaca, New York, USA, 5Cornell University, Ithaca, New York, USA, 6Colegio de Postgraduado, Edo. de Mexico, Mexico

Genomic Selection (GS) is being implemented at a fast pace in several plant and animal breeding programs. Most implementations so-far are based on genotypes derived from marker panels (e.g., SNP panels) discovered using reference samples. Recently, the cost of sequencing has decreased markedly and therefore, genotyping-by-sequencing (GBS) has become a good alternative. GBS has many attractive features: it can produce very detail characterization of genomes at a relatively low cost and the extent of ascertainment bias is expected to be less than that of genotyping using marker panels derived from reference populations. However, GBS brings important statistical and computational challenges. With low-coverage the proportion of ‘missing’ genotypes can be extremely high and it is not clear how to
incorporate the wealth of information that GBS can potentially produce into practical models for GS. In this study we used 504 doubled haploids CIMMYT maize lines (from the Drought Tolerance Maize for Africa project, DTMA) evaluated for grain yield, days to anthesis, and anthesis-silking interval (ASI) in four environments, and introduced methods for incorporating GBS information into parametric (GBLUP) and non-parametric (reproducing kernel Hilbert spaces, RKHS) models for GS. Some methods do not require imputing missing genotypes and others require imputation of missing genotypes as first step. We considered implementing these models using marker (imputed or not) or haplotype information. Also we evaluated the impact of combining GBS with pedigree data. Our results demonstrate the feasibility of incorporating GBS information for prediction of genetic values. We found (1) modest but consistent benefits of imputation, (2) superior accuracy of the non-parametric approach relative to the linear model, (3) modest benefits (but consistent) of combining pedigree and GBS data, and (4) slight superiority of methods based on haplotypes over those based on imputed (or not) SNPs.

O-57
GENOMIC-BLUP DECODED: A LOOK INTO THE BLACK BOX

David Habier, Rohan Fernando, Dorian Garrick
Department of Animal Science and Center for Integrated Animal Genomics, Iowa State University, Ames, Iowa, USA

Genomic-BLUP is used in animal and plant breeding to obtain genomic estimated breeding values (GEBVs) for selection. It replaces the numerator-relationship matrix in the traditional BLUP equations by a matrix \( G \) that is calculated entirely from genotypes of single nucleotide polymorphisms (SNPs). While the numerator-relationship matrix is known to explain additive-genetic relationships between individuals at quantitative trait loci (QTL), it is not obvious to what extend \( G \) captures relationships at the QTL, and how linkage disequilibrium (LD), co-segregation, and additive-genetic relationships contribute to the accuracy of GEBVs. Answering these questions helps to better predict this accuracy for complex pedigree relationships between training individuals and selection candidates, and may also contribute to the discussion of missing heritability in human genetics. We defined the three types of genetic-statistical information mentioned above, unified them in one coherent genetic model, and identified the contribution of each of these to the correlation between GEBVs and true breeding values, i.e., the accuracy of GEBVs. Simulations confirmed analytical results and showed how the contribution of LD, co-segregation and additive-genetic relationships depends on effective population size, genome length, training data size, and SNP density. When LD is not present, the accuracy of GEBVs can approach the accuracy of Pedigree-BLUP, and presumably that of linkage analyses. Genomic-BLUP captures more LD and less additive-genetic relationship information for a given training data size and high SNP density when the effective population size is low. When the training size is increased relative to the effective number of SNPs, Genomic-BLUP will not capture accurately additive-genetic relationships. This will lower the accuracy of GEBVs if LD between SNPs and QTL is incomplete. In this case, the genetic variability explained by SNPs will be lower than estimated from pedigree.

O-58
GENOME-BASED PREDICTION IN HIGHLY STRUCTURED PLANT POPULATIONS

Pak Sham
University of Hong Kong, Hong Kong

No Abstract
O-59
GENOME-BASED PREDICTION IN HIGHLY STRUCTURED PLANT POPULATIONS

Chris-Carolin Schoen
Technische Universitaet Muenchen, Freising, Germany

Most traits of importance in plant and livestock breeding follow a quantitative distribution. The assessment of these quantitative traits in performance tests on thousands of individuals is extremely resource consuming. Thus, prediction of the genetic potential of individuals from their DNA sequence becomes highly desirable.

We used cross validation and validation with independent samples to assess the performance of genome-based prediction of genetic values in experimental populations of maize (Zea mays L.), rice (Oryza sativa L.) and the model plant Arabidopsis thaliana (L.). The populations differed mainly with respect to effective population size, extent and distribution of linkage disequilibrium and mating system. All populations were genotyped with high density SNP markers and phenotyped for traits of different genetic architecture and heritability in multi-environment trials. We fitted different statistical models to investigate the effect of population substructure and assessed if variable selection or prior biological knowledge on model predictors could improve predictive abilities.

Average predictive abilities based on genomic data were high even for a complex trait like maize grain yield when the cross validation schemes allowed for a high degree of relatedness between the training and the validation set. Correlations between predicted and observed performance decreased substantially when training and validation sets were evaluated with different testers or in different years. For genome-based prediction of genetic values statistical models using variable selection and prior knowledge on genome structure did not outperform random regression genomic BLUP irrespective of the experimental population or trait under study.

O-60
GENOMIC PREDICTION OF COLORECTAL CANCER RISK USING GBLUP

Isobel Stewart1, Harry Campbell2, Ian Tomlinson3, John Hopper4, Steve Gallinger5, Graham Casey6, Richard Houlston7, Malcolm Dunlop8, Albert Tenesa1,7
1Roslin Institute, The University of Edinburgh, Edinburgh, UK, 2Centre for Population Health Sciences, The University of Edinburgh, Edinburgh, UK, 3Molecular and Population Genetics Laboratory, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK, 4Center for Molecular, Environmental, Genetic and Analytic Epidemiology, The University of Melbourne, Melbourne, Australia, 5Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Canada, 6Departments of Preventative Medicine and Biochemistry and Molecular Biology, University of Southern California, Los Angeles, USA, 7MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, The University of Edinburgh, Edinburgh, UK, 8Colon Cancer Genetics Group, Institute of Genetics and Molecular Medicine, University of Edinburgh and MRC Human Genetics Unit, Edinburgh, UK

Colorectal cancer is the fourth highest cause of cancer mortality and the third most common malignancy. Genetic risk variants, identified from candidate gene or genome wide association approaches, account for less than 2% of the phenotypic variance (liability scale). Around 170 common variants, of individually small effect, are thought to account for the remaining variance in “sporadic” CRC. The aim of this work was to assess the genomic prediction method GBLUP for CRC risk.

Genotyping data for 300K SNPs was available for 8999 individuals (3531 cases, 5468 controls), of mainly North European descent. A linear mixed model was implemented in ASReml (3.0 Gilmour et al., 2009), including gender as a fixed effect, additive genetic variance as a random effect, and using a genomic relationship matrix calculated in GCTA (Yang et al., 2011). The trait was assumed to be i) quantitative and ii) binomial (logistic). 10-fold cross validations were performed, and the predictive models assessed using ROC curves. These analyses were then repeated where 14 SNPs with previously shown associations with
CRC risk were excluded in the calculation of the GRM, and included as fixed effects in the linear model. The heritability estimate was 27.6% in the quantitative trait model and 5.3% (liability scale) in the logistic model (11.1% on the observed scale, after accounting for the case-control nature of the data (Lee et al., 2011) and a prevalence of 0.4%). The mean AUC was 0.56 and 0.55 respectively. In models including the fixed effects of SNPs, heritabilities were 22.4% and 4.7% (10.0%), and the AUC was 0.61 and 0.60 respectively.

GBLUP for CRC shows slight predictive power. The inclusion of SNPs with previously known associations with CRC improved the prediction model. Further methods of genomic risk prediction are being examined.
SO-1/P-43
MARKERS AS TRAITS IN MULTIVARIATE BLUP: USING REML FOR ASSOCIATION TESTING AND INTEGRATION WITH BREEDING VALUE PREDICTION

Jeremy Brawner1,2, Mark Dieters2
1CSIRO, St Lucia, QLD, Australia, 2University of Queensland, St Lucia, QLD, Australia

A pedigreed population of an outcrossing forest tree species that was established in multiple progeny trials and genotyped for an association study is used to demonstrate the concept of using markers as additional traits in a multivariate analysis. The population exhibited moderate to high narrow-sense heritability estimates for three economically important traits: stem diameter (0.28), wood density (0.44) and pulp yield and (0.36). For the 65 markers that were sequenced within 12 candidate genes, P-values were estimated for four tests of association using two different approaches: 1) two software packages (Tassel and Asreml) fit standard univariate mixed models that include each SNP as an additional fixed effect, and 2) bivariate or multivariate mixed models including each SNP as additional selection trait. The first approach determines if a significant proportion of phenotypic variance is attributable to the marker effect using an F-test and the second estimates the significance of the genetic correlation between additive genetic values and marker frequency using log-likelihood ratio tests. Within either the univariate or multivariate approach there was a significant correlation between the tests of significance. The two programs used for the first approach produced nearly identical results and the significance of marker-trait correlations generated with the bivariate and multivariate analyses were equally similar. However, there was little correspondence between the two approaches: only one of the 12 among-approach correlations was weakly positive (r=0.25 and p<0.05). This empirical study demonstrates that different SNP markers would be selected using multivariate analytical methods compared to the standard marker:trait association methods. In addition to demonstrating empirical results, the presentation provides an interpretation of genetic parameters estimated using REML derived variance components within a multivariate framework including two phenotypes and two markers as traits, including: 1) typical heritability of phenotypic traits, 2) heritability of genic traits, 3) typical trait:trait correlations, 4) maker:trait correlations as association tests, and 5) marker:marker correlations as LD estimates.

SO-2/P-46
GENETIC MARKERS AS INSTRUMENTS FOR MENDELIAN RANDOMIZATION STUDIES ON VITAMIN D

Diane Berry1, Karani Vimaleswaran1, John Whittaker1, Aroon Hingorani2, Elina Hypponen1
1UCL Institute of Child Health, London, UK, 2London School of Hygiene and Tropical Medicine, London, UK, 3Department of Epidemiology and Public Health and the Centre for Clinical Pharmacology University College London, London, UK

Observational studies on vitamin D and health maybe confounded by factors that cannot be fully controlled. Mendelian randomization (MR) makes use of genetic variants mimicking the influence of a modifiable exposure to overcome this problem. We examined genetic markers that index differences in 25-hydroxyvitamin D (25(OH)D) concentrations as instruments for MR studies of vitamin D. We used data from 4,572-6,877 participants in the 1958 British birth cohort with information on genetic markers and 25(OH)D. As potential instruments, we selected 20 single nucleotide polymorphisms (SNP) which are located in the vitamin D metabolism pathway or known to affect skin pigmentation/tanning, including 4 SNPs from recent genome-wide association (GWA) meta-analyses on 25(OH)D. We analyzed the SNP associations with 25(OH)D and evaluated the use of allele scores dividing genes to those affecting 25(OH)D synthesis (DHCR7 and CYP2R1) and metabolism (GC, CYP24A1, CYP27B1). In addition to the GWA SNPs, only two SNPs (in CYP27B1 and OCA2) showed evidence for association with 25(OH)D, with the OCA2 association abolished after lifestyle adjustment. Per allele differences varied between -0.02 and -0.08 nmol/L (P<0.02 for all), with a 6.1 nmol/L and a 10.2 nmol/L difference in 25(OH)D between individuals with highest compared lowest number of risk alleles in synthesis and metabolism allele scores,
respectively. Individual SNPs but not allele scores showed associations with lifestyle factors. An exception was region which was associated with synthesis score. Illustrative power calculations (80% power, 5% alpha) suggest that approximately 80,000 participants are required to establish a causal effect of vitamin D on blood pressure using the synthesis allele score. Combining SNPs into allele scores provides a more powerful instrument for MR analysis than a single SNP in isolation. Population stratification and the potential for pleiotropic effects need to be considered in MR studies on vitamin D.

SO-3/P-41
USING KEGG PATHWAYS AND EXPRESSION STUDIES AND STUFF FOR GENOMIC PARTITIONING

Stefan M. Edwards, Li Jiang, Per Madsen, Peter Sørensen
Department of Molecular Biology and Genetics, Faculty of Science and Technology, Aarhus University, Tjele, Denmark

A new approach building on linear mixed models was developed for identifying genetic variants in biological pathways that affect complex diseases. The method partitions the genome into sections based on prior information (e.g. KEGG Pathways or expression data), allowing the estimation of the variance components explained by the genetic variants found in each of these sections and variance components explained by non-genetic factors such as the environment.

A major advantage of our approach is, that it builds on a solid statistical modelling framework, that is already widely used and we can easily adjust for other known effects e.g. gender or complex pedigree structures. The methodology integrates several layers of information, but requires that there is a mapping between each layer, basically to place each genetic variant into each section. For well-studied organisms such as cattle or human, this information is readily available in public databases such as Ensembl.

Using this approach, we have used KEGG pathways to examine the variance components explained by each of the pathways as well as groups of pathways, such as Lipid Metabolism and Immune System, showing that genetic variants in genes associated to pathways involved in the immune system can explain a larger proportion of the genetic variance for bovine mastitis.

Currently we are evaluating bootstrapping and permutation methods to examine the significance of the findings.

SO-4/P-283
GENETIC INTERACTIONS IN THE HUMAN LIVER

Darren J. Fitzpatrick, Derek Greene, Denis Shields
University College Dublin, Dublin, Ireland

Quantitative phenotypes emerge as the result of multiple interactions, both biological and environmental. However, association studies tend not to take account of epistasis when mapping genotypic and phenotypic variation. The magnitude of the ‘interaction problem’ can be illustrated by considering pairwise interactions, the number of which increase combinatorially. The consequent problem of multiple testing requires that individual tests achieve minute p-value thresholds which most studies are insufficiently powered to detect. We present two strategies to detect epistatic loci associated with gene expression in the human liver. The first considers interactions between cis- and trans-acting SNPs known to exhibit a marginal effect (Schadt, et al, 2008). More than 9.3 million tests for interaction were conducted on a set of 2,433 SNPs associated with 5,764 transcripts and significance determined using a 5% false discovery rate (FDR) and Bonferroni correction. Results show a network of interactions with hubs centred on three genes. Interestingly, SLC6A12, a GABA transporter, is shown to interact with PITPNA, SLC16A2, NRG1, TSEN34, ADCY9, NPTXR, TULP3 and TNC all of which have a role in neuronal functions. The second strategy utilised a published network of protein-protein interactions (PPI) in the
human liver (Wang, et al, 2011) to define the search space. We conducted more than 100 million tests for interaction on 500,000 SNPs and 3098 transcripts for those genes whose proteins are known to physically interact. Considering SNPs in regions 2kb upstream of the gene, 3 interactions were significant after Bonferroni adjustment. We will also present results for downstream and intergenic SNPs. We contend that using a priori information is not only a sensible means of defining a search space but also aids interpretability. By informing the search space with functional information, e.g. PPIs, meaningful genetic interactions that effect the transcription of these proteins can be uncovered.

SO-5/P-265
EVIDENCE OF SHARED POLYGENIC RISK AMONG SMOKING BEHAVIORS AND BODY COMPOSITION

Roseann Peterson, Xiangning (Sam) Chen, Jingchun Chen, Bradley (Todd) Webb, Hermine Maes
Virginia Commonwealth University, Richmond, VA, USA

Obesity and nicotine dependence (ND) represent complex heterogeneous diseases which pose serious public health problems, affecting 33 and 20 percent of Americans, respectively. While cross-sectional studies of ND are typically supportive of a negative relationship between smoking and body mass index (BMI), a positive association is supported by the observations that, within smoking cohorts, heavy smokers tend to be of increased bodyweight compared to light smokers. A growing body of literature demonstrates the utility of genome-wide association studies (GWAS) for identifying single nucleotide polymorphisms (SNP) that contribute to disease risk. The GWAS approach has been applied to BMI and smoking behaviors (SB) using sample sizes in the tens of thousands and yielded several putative risk variants of small effects on individual traits. Many traits show comorbidity but most studies do not examine common versus specific variants. The purpose of this study was to investigate whether variants affecting BMI or SB were common to multiple behaviors or were trait specific. 75 BMI and 54 SB associated SNPs were catalogued from large-scale GWAS meta-analyses. These variants were tested for association in n=2,802 (41% African-American) older community-dwelling adults (68-80 years old) from the Health Aging and Body Composition study. Preliminary Results: Current smokers had significantly lower BMI and abdominal visceral fat than never or former smokers in males and females (p<0.0001). There were three BMI SNPs associated with both body composition variables and smoking traits: rs1900273 in STK33, rs2145270 near BMP2 and rs12127438 in the 1q42.2 locus. Additionally, there were three SB SNPs associated with multiple traits: rs11072774 in CHRNB4, rs2640732 in SCARA3 and rs6945244 in PDE1C. Results indicate shared genetic risk between smoking and body composition. Future research should confirm these associations and address mechanisms behind the common genetic architecture underlying these traits.

SO-6/P-214
DIRECT AND INDIRECT GENETIC EFFECTS FOR SURVIVAL IN PUREBRED AND CROSSBRED LAYING HENS

Katrijn Peeters1, Tineke T. Eppink1, Esther E. Ellen1, Jeroen Visscher2, Piter Bijma1
1Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands, 2Institut de Sélection Animale B.V., Hendrix Genetics, Boxmeer, The Netherlands

Through social interactions, individuals can affect each other’s phenotype. The heritable effect of an individual on the phenotype of a conspecific is known as an indirect genetic effect (IGE). IGEs contribute to heritable variation and can have a substantial impact on the magnitude and direction of response to selection. To study IGEs for survival in laying hens, we analyzed data on 6,227 W1 and 6,900 WB purebred individuals, as well as 7,668 W1xWB and 7,344 WBxW1 crossbred individuals. Birds of the same line were kept in groups of four. Feather pecking and cannibalism occurred, as beaks were kept intact. Survival was lower in crossbreds than in purebreds, indicating negative heterosis. A direct-
indirect animal model was used to estimate genetic parameters for each purebred and crossbred line. The contribution of IGEs to heritable variation was substantial, especially in crossbreds (65% in W1, 44% in WB, 87% in W1xWB and 72% in WBxW1). Moreover, the direct-indirect genetic correlation \( r_{DI} \) differed among purebreds and crossbreds (0.20 in W1, -0.28 in WB, -0.37 in W1xWB and -0.83 in WBxW1). As a result of the large IGEs and negative \( r_{DI} \), W1xWB would fail to respond to mass selection and WBxW1 would respond in the opposite direction. Overall, IGEs in crossbreds contributed more to heritable variation and had more impact on response to selection than in purebreds. In addition, direct and indirect genetic correlations between W1xWB and WBxW1 were estimated. The direct genetic correlation was high (0.95), whereas the indirect genetic correlation was moderate (0.42). Thus, for IGEs, it mattered which parental line provided the sire and which the dam. This phenomena is known as a parent-of-origin effect. The effect appeared to be paternally transmitted and is most likely Z-chromosome-linked.

SO-7/P-79
MATERNAL GENETIC EFFECTS SET THE POTENTIAL FOR EVOLUTION IN RED SQUIRRELS

S.E. McFarlane1, J.C. Gorrell2, D.W. Coltman2, M.M. Humphries3, S. Boutin2, A.G. McAdam1
1University of Guelph, Guelph, Ontario, Canada, 2University of Alberta, Edmonton, Alberta, Canada, 3McGill University, Montreal, Quebec, Canada

Microevolution has been notoriously difficult to document in the wild despite a perception that it is common. Many traits in wild populations seem to be static although they are both heritable and under strong selection. One explanation for this stasis is that evolutionary biologists typically focus on quantifying direct additive genetic variance and, thus, may not be measuring all of the available genetic variance in a trait. Genetic variance in fitness is necessary for evolution, as any trait must genetically co-vary with fitness in order to evolve. Indirect genetic variance, which results when the genotype of one individual affects the phenotype of another individual, could lead to unexpected rates and directions of evolution depending on how it correlates with direct genetic variance. We estimated both the additive genetic variance and maternal genetic variance of fitness (lifetime reproductive success) and fitness components, including longevity and annual reproductive success, in a long-term pedigreed population of red squirrels, near Kluane Lake, Yukon. We found that direct additive genetic variance explained only 0.03% - 0.05% of the phenotypic variance in fitness and had, at most, a coefficient of additive genetic variance of 3.73. In contrast, maternal genetic effects explained 2% - 34% of the phenotypic variance in both fitness and a variety of fitness components, and had a coefficient of maternal genetic variance of 16.27 – 85.79. High levels of maternal genetic effects, especially the significant levels found for fitness, suggest indirect genetic effects play an important role in the evolution of red squirrels. Incorporating indirect maternal genetic effects into predictions of evolution in a range of wild systems may lead to more precise estimates of microevolution.

SO-8/P-109
ANTAGONISTIC SELECTION PRESSURES ACTING ON HATCHING ASYNCHRONY IN A SIBLICIDAL SEABIRD

Emily Tompkins1, David J. Anderson1, Katharina Foerster2
1Wake Forest University, Winston-Salem, NC, USA, 2University of Tübingen, Tübingen, Baden-Württemberg, Germany

Directional selection is expected to reduce additive genetic variance for a given trait, but may instead maintain genetic diversity in a population by acting in opposing directions on different components of fitness. Early reproduction traits expressed in the mother may affect the fitness of both mother and offspring, providing an opportunity for antagonistic selection pressures acting on maternal and offspring fitness to impose an evolutionary constraint when the trait is considered part of the phenotype of both
individuals. We use a “de-lifing” approach to estimate fitness as an individuals’ relative survival and reproduction contributions to per year population growth, and examine selection acting on hatching asynchrony (HA), a maternal trait, in a wild population of a long-lived, obligately siblicidal, seabird, the Nazca booby (*Sula granti*). Asynchronous hatching occurs when incubation is initiated before laying is complete, resulting in a competitive advantage for earlier-hatching siblings. Selection on hatching asynchrony in obligately siblicidal species has been related to HAs facilitation of brood reduction. Nazca boobies lay two eggs and from the maternal, paternal, and first-hatched chicks’ perspective a longer hatching asynchrony would facilitate speedy siblicide, reducing the duration of nestling competition. Earlier lay dates increase a chick’s probability of surviving to fledge and from the perspective of the second hatched chick, a shorter hatching asynchrony may increase its survival prospects as well as lengthening its time in the nest. Analyses of direct and correlated selection acting on maternal, A-chick, and B-chick components of fitness for hatching asynchrony may show antagonistic selection pressures capable of maintaining genetic diversity in this trait. Future work will examine heritability of this trait to further assess its evolutionary potential.
P-1
QTL IciMAPPING v3.2: INTEGRATED SOFTWARE FOR BUILDING LINKAGE MAPS AND MAPPING QUANTITATIVE TRAIT GENES

Huihui Li, Luyan Zhang, Lei Meng, Jiankang Wang
Institute of Crop Science and CIMMYT China, Chinese Academy of Agricultural Sciences, Beijing, China

Objectives: QTL IciMapping is freely-available public software, capable of building high-density linkage genetic maps and mapping quantitative trait loci (QTL).

Methods and Results: There are seven functionalities fully implemented in version 3.2. 1. MAP: Construction of Genetic Linkage Maps in Biparental Populations. Grouping can be based on (i) anchored marker information, (ii) a threshold of LOD score, and (iii) a threshold of marker distance. Three ordering algorithms are (i) SER: SERiation, (ii) RECORD: REcombination Counting and ORDERing, and (iii) nnTwoOpt: the nearest neighbor algorithm was used to construct an initial map, and two-opt algorithm was used to improve the initial map. Four rippling criteria are (i) SARF (Sum of Adjacent Recombination Frequencies), (ii) SAD (Sum of Adjacent Distances), (iii) SALOD (Sum of Adjacent LOD scores), and (iv) COUNT (number of recombination events). 2. BIP: Mapping of Additive and Digenic Epistasis Genes in Biparental Populations. Five mapping methods are (i) SMA, Single Marker Analysis, (ii) SIM: Simple Interval Mapping, (iii) ICIM-ADD: Inclusive Composite Interval Mapping of ADDitive (and dominant) QTL, (iv) ICIM-EPI: Inclusive Composite Interval Mapping of digenic EPistatic QTL, and (v) SGM: Selective Genotyping Mapping. 3. CSL: Mapping of Additive and Digenic Epistasis Genes with Chromosome Segment Substitution Lines. Three mapping methods are (i) SMA, and (ii) RSTEP-LRT-ADD: Stepwise regression based likelihood ratio tests of additive QTL, and (iii) RSTEP-LRT-EPI: Stepwise regression based likelihood ratio tests of digenic epistasis QTL. 4. MET: QTL by Environment Interaction in Biparental Populations. Two mapping methods are (i) ICIM-ADD: ICIM of additive QTL by environment interaction, and (ii) ICIM-EPI: ICIM of digenic epistatic QTL by environment interaction. 5. NAM: QTL Mapping in NAM Populations. One mapping method is JICIM: Joint Inclusive Composite Interval Mapping of additive QTL. 6. SDL: Mapping of Segregation Distortion Loci in Biparental Populations. Two mapping methods are (i) SMA, and (ii) SIM. 7. iMAP: Integration of multiple genetic linkage maps sharing common markers.

Conclusions: QTL IciMapping is user friendly software to facilitate the complex traits dissection.

P-2
INBREEDING EFFECTIVE POPULATION SIZE IN BREEDS WITH NUCLEUS AND MULTIPLIER BREEDING UNITS

Fabricio Bongiorno1, Rodolfo J. C. Cantet2
1Universidad Nacional del Centro de la Provincia de Buenos Aires, Azul, Buenos Aires, Argentina, 2Universidad de Buenos Aires - CONICET, Ciudad Autónoma de Buenos Aires, Argentina

Reduction of genetic variability in newly formed breeds of cattle poses a threat to their successful commercial establishment. If the inbreeding effective population size ($N_e$) is low, appropriate modifications of recording policies and bull recruiting permits correcting for the reduction in genetic variability. Determination of average inbreeding ($F$) in these populations is uncertain as these breeds have poor pedigree information, especially from the dam side. Most formulae used to calculate $N_e$ are inefficient in these populations as selection, overlapping generations, drift, subdivision and high variance of family size, take place simultaneously. On the other hand, characterization of the mating system in the breed by means of demographic parameters is possible due to the extensive data recording used in setting genetic evaluations. We propose an estimator of inbreeding $N_e$ for a breed composed from nucleus and multiplier herds. Males are selected within the nucleus and have progeny in both strata. Therefore, $F$ is only possible within the nucleus. A system of recursive equations comprising the following parameters was set: 1) $F$ in the nucleus; and average coancestry between two individuals: 2) from different nucleus herds ($\theta_L$); 3) from the same nucleus herd ($\theta_W$); 4) one in the nucleus and the other in the multiplier ($\theta_{NM}$);
and 5) from multiplier herds \((\theta_m)\). The solution for \(F\) was obtained from equations 1 to 3, and was set equal to one over 2 times \(N_e\), and solved for the last parameter. This inbreeding \(N_e\) measures genetic variability in the nucleus and is relevant to sustained response to selection. A second expression \((N_{eM})\) was obtained by expressing the solutions for the two remaining parameters as a function of \(F\). \(N_{eM}\) is relevant to the variability of multiplier males used in commercial populations. \(N_{eN}\) and \(N_{eM}\) were estimated for the Brangus breed of Argentina.

P-3
GENETIC COMPONENT OF MILK FT-IR SPECTRA USED TO PREDICT BREEDING VALUES (EBV) FOR MILK COMPOSITION AND QUALITY TRAITS

Binyam Dagnachew, Theo Meuwissen, Tormod Ådnøy
Norwegian University of Life Sciences, Aas, Norway

Genetic analysis of milk FT-IR spectral variables may lead to a better estimation of genetic parameters for milk composition and quality traits. Instead of first predicting phenotypes from the spectra and then estimating heritability and breeding values using the phenotypes (indirect prediction – IP), the genetic components of all wavelengths of the spectra and their predicted breeding values need to be calculated for the direct prediction (DP). Breeding values of derived milk traits (fat percentage, coagulation, etc.) can then be predicted as a trait correlated to the genetic information of the spectra. The aim of the study is to compare the effects of IP and DP of milk composition and quality traits on prediction error variance (PEV) and genetic gain. A dataset containing 28,000 milk FT-IR spectral observations which belong to 14,869 goats of 271 herds was used for training and evaluating models. REML was used to estimate variance components of the spectral variables after principal component analysis (PCA) was applied to reduce the spectral dimension. EBVs were predicted for fat, lactose and protein percentages using DP and IP methods. The DP approach reduced the mean PEV by 3.19%, 4.07%, and 6.14% for fat, lactose and protein percentages respectively compared to IP method. Given the reduction in PEV, relative genetic gains were 2.56%, 2.78% and 3.72% in fat, lactose and protein percentages respectively. It is concluded that more accurate EBV could be found using genetic component of milk FT-IR spectra compared to single trait animal models analyses on phenotypes predicted from the spectra. Lowly heritable milk quality traits could benefit more from the proposed method as it evaluates animals through genetic correlations of many spectral variables.

P-4
A COMPUTATIONALLY EFFICIENT BAYESIAN MODEL FOR INCORPORATING A VAST NUMBER OF MARKERS

Daniel Sorensen, Luc Janss
Aarhus University, Aarhus, Denmark

Objectives: To develop a computationally efficient Bayesian genomic model that incorporates a vast number of markers and simultaneously partitions the total genomic variance into independent contributions from unobserved sources of population substructure. Two partitions are possible: one is trait-specific and the other depends on the distribution of genetic markers. The methodology is applied to the analysis of height (HT), systolic blood pressure (SYS) and cholesterol (HDL) from 3,000 individuals from the 1958 British cohort study genotyped for 1 million SNP markers. An efficient back transformation is described to obtain SNP effects on the original scale. The predictive ability of the model is studied using a 10-fold cross-validation.

Methods: An eigen-value (EV) decomposition of the genomic relationship matrix is performed that leads to a probabilistically equivalent form of the classical genomic model. An Markov chain Monte Carlo (McMC) algorithm is developed to implement the EV model where maker effects are updated simultaneously.
Results and Conclusions: It takes approximately 30 minutes to generate chain samples of adequate size in an 3.5 GHz 64 bit intel linux cluster. Estimates of the proportion of genomic variance based on posterior means (posterior s.d. values) were 0.40 (0.11), 0.15 (0.09) and 0.21 (0.10) for HT, SYS and HDK respectively. The results of partitioning the total genomic variance into independent components indicate that very little substructure is present in the British cohort 58 data. The proposed method yields a numerically stable partition of the variance as a function of the number of eigenvectors, in marked contrast with an advocated approach, based on including the largest eigenvectors as fixed covariates in a linear model. The 10-fold cross-validation study shows that the markers explain only 1.7% of the total variance in phenotype for HT, and less than 1% for SYS and HDL.

P-5
HIERARCHICAL GENERALIZED LINEAR MODELS WITH RANDOM EFFECTS AND VARIANCE HETEROGENEITY FOR GWAS AND BEYOND

Lars Rönnegård1,2
1Dalarna University, Borlänge, Sweden, 2Swedish University of Agricultural Sciences, Uppsala, Sweden

Statistical methods in animal breeding and genome-wide association studies (GWAS) are heavily based on linear models with normal random effects. A challenge to this day is to fit non-normally distributed data in a computationally fast and unbiased manner. Several methods have been proposed but unlike most frequentist and Bayesian ones, hierarchical generalized linear models, implemented in the hglm package, combine modelling flexibility with deterministic computations and are based on Lee & Nelder’s h-likelihood theory. hglm can be used for linear mixed models and generalized linear mixed models with random effects for a variety of links and a variety of non-normal distributions for both the outcomes and the random effects. Correlated random effects, including genomic relationship matrices, can be fitted as well. Fixed effects can also be fitted in the dispersion part of the model to estimate variance heterogeneity. Due to its computational speed the main function hglm() has been incorporated into the well-established GenABEL package for analysis of GWAS data. The hglm package is available on CRAN and has been applied by several research groups on a number of problems varying from e.g. oyster breeding value estimation to analysis of diallel experimental cross data having non-normal outcomes. The methodology has recently also been used for fitting: genomic selection models on data sets having large number of markers (>100,000), models including random genetic heterogeneity in the residual variance on a large dairy data set, and variance-controlling QTL (vQTL) models. Methods to detect vQTL have developed rapidly over the past two years and play an important role in search for loci regulating homeostasis and epistasis.

P-6
PREDICTION OF DISEASE CAUSING NON-SYNONYMOUS SNPS BY THE ARTIFICIAL NEURAL NETWORK PREDICTOR NETDISEASESNP

Morten Johansen, Thomas Petersen, Søren Brunak, Ramneek Gupta
Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark

Non-synonymous single nucleotide polymorphisms (nsSNPs) may affect protein structure and function and several are disease-associated. We have developed a sequence homology based artificial neural network predictor called NetDiseaseSNP which classifies nsSNPs as disease or not disease causing. NetDiseaseSNP uses the excellent alignment generation algorithm of SIFT to extract sequence homology data which is augmented and converted to an extensive input feature space containing 31 features including surface accessibility prediction which is used as input to NetDiseaseSNP. NetDiseaseSNP produces a single output score which can be used to rank nsSNPs based on their disease causing
potential and NetDiseaseSNP has a performance in terms of Matthews correlation coefficient of up to 0.70 depending on the quality of the sequence homology data. When SIFT is unable to process a protein sequence, BLOSUM62 scores are used to generate the input features for NetDiseaseSNP. We show that NetDiseaseSNP has the ability to discriminate cancer driver mutations from passenger mutations. NetDiseaseSNP is shown to outperform SIFT on disease/non-disease datasets as well as on cancer driver/passenger mutation datasets. NetDiseaseSNP can be regarded as an improvement of SIFT which has the potential to use all information in the alignment generated by SIFT. NetDiseaseSNP can thus be used to pinpoint and prioritize plausible disease candidates among nsSNPs for further investigation. The method will be made available as an online tool as well as a web service: http://www.cbs.dtu.dk/services/NetDiseaseSNP/

P-7 – ABSTRACT WITHDRAWN

P-8
POOLING ESTIMATES OF COVARIANCE COMPONENTS USING A PENALIZED MAXIMUM LIKELIHOOD APPROACH

Karin Meyer
Animal Genetics and Breeding Unit, University of New England, Armidale, NSW 2351, Australia

Estimates of large genetic covariance matrices are commonly obtained by pooling results from a series of analyses of small subsets of traits. Procedures available to pool the part-estimates differ in their efficacy in accounting for unequal accuracies of estimates and sampling correlations, and ensuring that pooled matrices are within the parameter space. We propose a maximum likelihood (ML) approach to combine estimates, treating sets from individual part-analyses as matrices of mean squares and cross-products from independent families. This facilitates simultaneous pooling of estimates for all sources of variation considered, readily allows for weighted estimation or a given structure of the pooled matrices, and provides a framework for regularized estimation by penalizing the likelihood.

A simulation study is presented, comparing the quality of combined estimates for several procedures, including truncation or shrinkage of either canonical or individual matrix eigen-values, iterative summation of expanded part matrices, and the ML approach, considering a range of penalties. Shrinking eigen-values of individual matrices towards their mean reduced losses in the pooled estimates, but substantially increased proportional losses in their phenotypic counterparts and thus yielded estimates differing most from corresponding full multivariate analyses of all traits. Assuming a simple pseudo-pedigree structure when combining estimates for all random effects simultaneously using ML allowed sampling correlations between estimates of different components from the same part-analysis to be approximated sufficiently to yield pooled matrices closest to full multivariate results, with little change in phenotypic components. Imposing a mild penalty to shrink matrices for random effects towards their sum proved highly advantageous, markedly reducing losses in estimates and more than compensating for the reduction in efficiency of using the data inherent in analyses by parts. Penalized ML provides a flexible alternative to current methods for pooling estimates from part-analyses with good sampling properties, and should be adopted more widely.
P-9
POWERFUL FAST APPROXIMATION OF THE VARIANCE COMPONENTS MODEL FOR WHOLE-GENOME ASSOCIATION ANALYSIS

Gulnara Svishcheva¹, Tatiana Axenovich¹, Nadezhda Belonogova¹, Cornelia van Duijn², Yurii Aulchenko¹
¹Institute of Cytology & Genetics SD RAS, Novosibirsk, Russia, ²Department of Epidemiology, Erasmus MC Rotterdam, Rotterdam, The Netherlands

Objective: The use of variance component (VC) model for genome-wide association studies (GWAS) of large samples become computationally exhaustive when the number of genetic markers is over a few hundreds of thousands.

Method: We propose a new two-step VC score-test-based method GRAMMAR-Gamma. In the first step, the heritability model and GRAMMAR-Gamma correction factor, γ, are estimated. Fast estimation of the heritability model was enabled by optimization of the matrix operations based on the eigendecomposition of the relationship matrix and analytical estimation of model parameters. In the second step, the SNP effects and statistics are estimated without explicit use of the relationship matrix and then corrected by the γ-factor. The time complexity of this step is linear on the number of individuals and markers and is equivalent to the time complexity of the methods, which do not account for the genetic structure of the sample.

Results: Statistical properties of the method and running time of our implementation were estimated using simulated data. Type 1 error, power, and the SNP effect estimates were very close to those of the most accurate two-step methods, such as mmscore and EMMAX. We compared the running time of different methods and demonstrated that GRAMMAR-Gamma provides the fastest means to run GWAS using mixed models. It achieved speed-up of up to 10 times in comparison with the fastest existing method. As one of the current challenges in statistical genomics is the analysis of whole-genome re-sequencing data, we estimated that GWAS performed by GRAMMAR-Gamma on re-sequencing data for 3,000 individuals was completed in 38 min.

Conclusions: GRAMMAR-Gamma provides a practical, fast, and powerful tool for the analysis of human GWAS scans. The role of the method may increase in the future with the availability of larger sample sizes and increased number of markers.

P-10 – ABSTRACT WITHDRAWN

P-11
OPTIMISATION OF GENOME-WIDE COMPLEX TRAIT ANALYSIS

Alan Gray, Isobel Stewart, Albert Tenesa
The University of Edinburgh, Edinburgh, UK

Genome-Wide Complex Trait Analysis (GCTA) is a freely-available software package for quantifying the contribution of genetic variation to phenotypic variation for complex traits. This comprises two stages: estimation of the “Genetic Relationship Matrix” (GRM) containing identity by descent (IBD) coefficients calculated using SNP data, and the use of Restricted Maximum Likelihood (REML) to maximise the likelihood of the phenotypes given these IBD coefficients. We have re-developed GCTA to dramatically improve efficiency on modern multi-core CPU hardware. We adapt the GRM stage to be expressed in a form suitable for execution using the highly optimised Basic Linear Algebra Subroutines (BLAS) library. We introduce the Linear Algebra Package (LAPACK) library to perform those matrix inversions that dominate the REML stage. We use parallel versions of these libraries, and introduce OpenMP parallel computing directives throughout the rest of the software, to enable efficient utilisation of modern multi-core processors. We perform a range of serial optimisations to further boost performance. We report benchmark results for our new package, which we call Advanced Complex Trait Analysis (ACTA), using a compute
node comprising two Intel Nehalem multi-core CPUs. The performance improvement, for our test case with 8999 individuals and 279,435 SNPs, of ACTA over GCTA is a factor of 161 (GRM estimation) and a factor of 13 (REML analysis), giving an overall reduction in runtime from around 17 hours to around 10 minutes. The REML stage is of ACTA was measured to be a factor of 70 faster than the alternative ASReml software. ACTA promises to make practically feasible the analysis of those large numbers of people and SNPs associated with sequence data. Furthermore, this work allows us to employ new increasingly sophisticated techniques: we describe the automation of the recursive analysis of multiple localised SNP windows to capture genetic variation at functional genomic regions.

P-12
USING SNP DOSAGE DATA FOR LINKAGE ANALYSIS AND QTL MAPPING IN AUTOTETRAPLOID SPECIES

Christine Hackett¹, Karen McLean², Glenn Bryan²
¹Biomathematics and Statistics Scotland, Dundee, UK, ²James Hutton Institute, Dundee, UK

SNP genotyping of autotetraploid mapping populations not only provides large numbers of markers for map construction, but also additional information about SNP dosage, i.e. whether the genotype is AAAA, AAAC, AACC, ACCC or CCCC. Using such dosage information, the recombination frequency can be estimated more precisely than with presence/absence data alone. For measuring dosage on a large chip, for example the Potato SolCap Infinium Chip, which assays 8300 SNPs, it is necessary to automate the dosage calling. We have developed an approach fitting normal mixture models to the allele intensity ratio theta from the BeadStudio software to identify the dosage for parents and offspring at each SNP locus. Previous methods for linkage analysis in tetraploid species have been extended to calculate recombination frequencies and most likely phases from allele dosage data using an EM algorithm. The JoinMap 4 software (Van Ooijen, 2006) can be used to order the markers using the pairwise recombination frequencies and lod scores, taking the map after two rounds of ordering as a basic map. The remaining markers have been placed in bins neighbouring the mapped SNPs. A check of the linkage map can be made by mapping the theta scores as quantitative traits to see whether the position and phase corresponds to that of the corresponding marker. These methods have been applied to an established cross between potato genotypes Stirling and 12601ab1, to produce SNP maps for each chromosome with 82-168 mapped SNPs, together with up to 200 further SNPs placed into bins neighbouring each mapped SNP. Reference


P-13 – ABSTRACT WITHDRAWN
P-14
DETECTION OF IMAGING GENETICS ASSOCIATIONS IN HEALTHY NORWEGIAN INDIVIDUALS USING SPARSE REDUCED-RANK REGRESSION

Eva Janousova1,2, Maria Vounou3, Lars T. Westlye4,5, Andrea Christoforou1,5, Astri J. Lundervold6,7, Ivar Reinvang2,5, Vidar M. Steen1,3, Thomas Espeseth5,8, Giovanni Montana3, Stephanie Le Hellard1,5

1Dr E. Martens Research Group for Biological Psychiatry, Department of Clinical Medicine, University of Bergen, Bergen, Norway, 2Institute of Biostatistics and Analyses, Masaryk University, Brno, Czech Republic, 3Department of Mathematics, Imperial College London, London, UK, 4Center for the Study of Human Cognition, Department of Psychology, University of Oslo, Oslo, Norway, 5Centre for Advanced Study, Oslo, Norway, 6Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway, 7Kavli Research Centre for Aging and Dementia, Haraldsplass Deaconess Hospital, Bergen, Norway, 8Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway

Traditional genome-wide association studies (GWASs), based on univariate linear modelling, do not account for correlations in genotype and phenotype data as each genetic marker is independently tested for association with one phenotypic trait at a time. To overcome these limitations, our group has developed a multivariate regression technique known as sparse reduced-rank regression (sRRR). The sRRR technique simultaneously models all genotype markers and phenotypic traits, thus accounting for the multivariate nature of GWAS data. By adopting the Lasso penalty, the sRRR model achieves variable selection in the genotypic domain. The sRRR model is combined with a data resampling technique, where the model is fitted repeatedly in data subsamples, and the genetic markers are then ranked according to their frequency of selection (selection probability). This approach has been previously validated on simulated data and on the Alzheimer’s Disease Neuroimaging Initiative (ADNI) dataset. Here, we applied the sRRR model to 636,668 SNPs and 10,914 imaging traits, representing brain cortical thickness and area, in 440 healthy volunteers in the Norwegian Cognitive NeuroGenetics (NCNG) sample. A total of 29 SNPs with selection probabilities greater than 0.5 were identified by the sRRR model. Using our LD-based gene-binning tool, LDsnpR, 14 of these SNPs were assigned to 13 Ensembl-defined genes, some of which have already been shown to play important roles in neuronal development. Further work following up these findings is warranted. In conclusion, using the sRRR multivariate model, we were able to analyse high-dimensional brain imaging and GWAS data from healthy Norwegian subjects and identify interesting candidate genes worthy of further investigation.

P-15
STATISTICAL IMPLICATIONS OF UNOBSERVED PARENTAL GENOTYPES IN ESTIMATING LINKAGE DISEQUILIBRIUM

Emma Huang1, Colin Cavanagh2

1CSIRO Mathematics, Informatics and Statistics and Food Futures National Research Flagship, Dutton Park, QLD, Australia, 2CSIRO Plant Industry and Food Futures National Research Flagship, Acton, ACT, Australia

Estimation of linkage disequilibrium (LD) plays important roles both in characterizing the genetic history of populations and in designing genomewide association studies. Numerous measures have been devised to estimate LD between biallelic loci, and have been extended to multiallelic markers. However, not all of these estimators perform equally well for both types of data. We consider the behaviour of different estimators in the situation where observed genetic data are biallelic surrogates for unobserved multiallelic genotypes which identify a limited number of genetic founders. This occurs in practice for complex experimental crosses. We estimate LD in two ways: first, based directly on the biallelic data; and second, based on imputing the fully informative multiallelic markers which identify each founder. We compare the performance of common LD measures for these two options through simulations and practically relevant multiparent experimental cross data. Characterizing the relationship between the two approaches as the number of parents increases (and approximates a typical association mapping population) will
demonstrate the danger in ignoring the unobserved parental alleles for certain LD measures. We will see that standard methods of estimating linkage disequilibrium need to be given greater attention, particularly in comparisons of values between populations.

P-16

OPTIMISED DENDROGRAM: EXTRACTING POPULATION INFORMATION USING SERIATION TO ENHANCE HIERARCHICAL CLUSTERING

Vivi Arief1, Ian DeLacy1, Kaye Basford1,2, Mark Dieters1

1The University of Queensland, School of Agriculture and Food Sciences, Brisbane, QLD, Australia, 2Australian Centre for Plant Functional Genomics (ACPFG), Brisbane, QLD, Australia

Hierarchical clustering or classification procedure is used to investigate structure in any two-way (rows and columns) table of data. Structure refers to patterns among rows (cases or categories) or among columns (variables). The data table or matrix is two-mode if the rows and columns refer to different classes of entities (say genotypes and traits) or one-mode if the entities are the same (as in a correlation matrix). Investigation of population structure is a prerequisite to genome wide association studies. Structure from hierarchical clustering is visualised using a dendrogram where the order of entities (leaves or branches) is arbitrary, being restricted only by the avoidance of crossovers. Re-ordering of the leaves of a dendrogram based on a seriation method produces an ‘optimised dendrogram’ order, which provides information on both the grouping and the ordering of the entities. This order is a minimum distance walk in the $N$ dimensions defined by the data (where $N$ is the rank of the data matrix) and has been shown to be a powerful solution to the travelling salesman problem. The procedure has been used in fields such as archaeology, psychology, and ecology, but to our knowledge its use in genetics has been limited to the analysis of microarray data. We provide a number of examples relevant to quantitative genetics to demonstrate the advantages of this optimisation for: i) two-mode data – genotype by marker, genotype by trait and marker by trait (association analysis) tables; and ii) one-mode data – coefficient of parentage, recombination fraction, and gametic phase disequilibrium matrices. We emphasise that the procedures are completely general and argue that optimised dendrograms should be used where ever clustering is an appropriate analytical procedure, particularly in the analysis of genetic data where dendrogram order should ideally reflect genetic similarity.

P-17

A NEW APPROACH FOR QTL MAPPING IN MULTI-PARENT RECOMBINANT INBRED LINES

Rohan Shah2,1, Andrew George2,1

1Food Futures National Research Flagship, Brisbane, Queensland, Australia, 2CSIRO Mathematics, Informatics and Statistics, Brisbane, Queensland, Australia

QTL mapping is the process of locating important regions of DNA, and has traditionally been performed using two-parent experimental designs. These designs are very good at detecting the presence of QTL, but not so good at estimating the locations of the detected QTL. Multi-parent designs have been developed with the aim of improving the precision of these location estimates. They allow for greater genetic diversity, but involve a more complicated experimental setup and are new enough that QTL mapping techniques are still being developed. In this talk we introduce a new QTL mapping method for a specific class of multi-parent designs. Our technique is an extension of the well accepted approach of Multiple Interval Mapping (MIM), which is based on the EM algorithm. We will discuss how the method has been extended to multi-parent designs, and some of its strengths and weaknesses compared to competing techniques. As this method uses the EM algorithm an efficient implementation is crucial; we give details of our current implementation, which is highly scalable and suitable for use on a computing cluster. We also present results from the analysis of a real wheat data set.
P-18

BENEFITS OF USING LOCAL WHOLE EXOME REFERENCE PANELS FOR THE IMPUTATION OF RARE VARIANTS IN TWO EUROPEAN POPULATIONS

Peter Joshi, Ross Fraser, Caroline Hayward, Igor Rudan, Alan F Wright, Harry Campbell, Pau Navarro, Jennifer Huffman, Véronique Vitart, Chris Haley, James F Wilson
University of Edinburgh, Edinburgh, UK

Increasing focus is being put on the potential for rare variants to explain a significant portion of the heritability of complex diseases, which are a substantial health burden in western countries. While whole genome sequencing remains unduly expensive, imputation provides a cost effective method of inferring rare variants on the basis of existing high density genome-wide SNP genotyping. This study examines whether the use of a local reference panel of whole exome sequences can increase the accuracy of rare variant imputation compared with the use of the 1000 Genomes reference panel alone, in the ORCADES and CROATIA-KORCZ ZULA genetic isolate studies.

889 Orcadian and 898 Korculan samples were genotyped using the Illumina Hap300 or 370CNV arrays, and ∼300k autosomal SNPs were used for imputation. The samples were phased using SHAPE-IT and then three separate imputations were conducted using IMPUTE2. The imputations used the 1000 Genomes reference panel (500 Europeans and 1221 non-Europeans), local panels of 100 exomes in each population and both panels together. The accuracy of the three imputations was measured against independent genotyping of rare and low frequency variants (MAF 0.5-10%) using the iSelect custom “metabochip” array. We examine the differences in accuracy by allele frequency and population. Where discrepancies between the imputations arose, we attempt to infer the cause, for example whether the smaller reference panel fails to capture enough variation, or whether identification of local long haplotypes and population-specific rare variants improve imputation.

This study gives an indication of the value and source of improvement in accuracy of using local genome sequences to supplement a general population reference panel, when imputing genotypes for rare variants using genome-wide SNP data.

P-19

GENETIC MAPPING USING SEQUENTIAL MONTE CARLO ALGORITHMS

Christian Schäfer1,3, Willem Kruijer2, Marco Bink2, Nicolas Chopin1, Cajo Ter Braak2, Marcos Malosetti2, Fred van Eeuwijk2
1Université Paris Dauphine, Paris, France, 2Biometris, Wageningen University and Research Centre, Wageningen, The Netherlands, 3CREST (ENSAE), Paris, France

Bayesian variable selection (BVS) methods are well established in quantitative genetics. Considering a linear model for a quantitative trait and a large number of genetic predictors, the quantity of interest is the posterior mean on the binary model space (probability that a marker is included or not). In genomic selection BVS is used for prediction, whereas in other contexts (e.g. Guan and Stephens, 2010) the goal is variable selection as such.

Despite various successful applications, most BVS methods suffer from two important drawbacks. First, they typically rely on MCMC algorithms with local transitions, which do not easily allow to profit from parallel computing environments, while competing non-Bayesian methods based on permutations or subsamples are straightforward to parallelise. Secondly, Markov chains driven by local transition kernels are known to mix slowly in the presence of strong local modes. Due to this slow mixing, MCMC algorithms may converge extremely slow and produce poor estimates of the actual posterior mean. This becomes very predominant when the predictors are highly correlated (e.g. high marker density) and even more when epistatic interactions are included in the model.

Although these problems are somewhat alleviated by recent MCMC-improvements (Bottolo and Richardson, 2010), we advocate Sequential Monte Carlo (SMC) algorithms (Del Moral et al., 2006).
These were adapted to BVS by Schäfer and Chopin (2011). They constructed global, fast-mixing adaptive transition kernels with independent proposals, drawn from a suitable parametric family. This allows for reliable sampling from highly multi-modal distributions and for massive parallelisation. Using SMC we analyse real and simulated data from natural- and experimental populations, such as flowering time measured in the arabidopsis hapmap-collection. SMC produced much more accurate estimates of QTL-locations than existing MCMC-based implementations of BVS such as Rqtl/bim, and also performed better than non-Bayesian methods such as composite interval mapping.

P-20
MIXED MODELLING WITH WHOLE GENOME DATA

Jing Hua Zhao, Jian’an Luan
MRC Epidemiology Unit & Institute of Metabolic Science, Cambridge, UK

Objective: We consider a unified modelling framework that is able to account for correlations among individuals and various sources of variations. These correlations could be relationships known among relatives in families or uncertain among unrelated individuals in a general population; both could be refined or derived with whole genome data. Variations can include oligogenes, polygenes, single nucleotide polymorphism (SNP) and covariates.

Methods: We describe mixed models as a coherent theoretical framework to accommodate correlations for various types of outcomes in relation to many sources of variations. The framework also extends to consortium meta-analysis involving both population-based and family-based studies.

Results: Through examples we show that the framework can be furnished with general statistical packages as viable computational tools whose great advantage lies in simplicity and flexibility to study both genetic and environmental effects. Areas which require further work are also indicated.

Conclusion: Mixed models will play an important role in practical analysis of data on both families and unrelated individuals when whole genome information is available.

P-21
REGIONAL HERITABILITY MAPPING FOR THE DISSECTION OF COMPLEX TRAITS

Yoshitaka Nagamine2,6, Ricardo Pong-Wong1, Pau Navarro1, Masoud Shirali2,5, Yoshinobu Uemoto2,7, Jennifer Huffman1, Veronique Vitart1, Caroline Hayward1, Valentina Riggio1, Oswald Matika1, Steve Bishop1, Igor Rudan6,4, Harry Campbell1, Nick Hastie1, Alan Wright1, Chris Haley1

1MRC Human Genetics Unit, MRC IGMM, University of Edinburgh, Edinburgh, UK, 2Roslin Institute and R(D)VS, University of Edinburgh, Edinburgh, UK, 3Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK, 4Centre for Global Health, Faculty of Medicine, University of Split, Split, Croatia, 5Department of Animal Science, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran, 6National Institute of Livestock and Grassland Science, Tsukuba, Japan, 7National Livestock Breeding Centre, Nishi-shirakawa, Fukushima, Japan

Genome-wide association studies (GWAS) can be effective for identifying common variants responsible for trait variation. However, standard GWAS analyses have limited power to find the clusters of potentially rare variants at a single locus that may characterise much trait variation. Using high density marker data it is possible to estimate genetic relationships between individuals in isolated and other populations. Relationships can be estimated for all of the genome or for shorter regions down to the level of an individual locus. These relationships can then be used to estimate the trait heritability for the chosen region that integrates the variance contributed by rare and common variants into a single estimate. This regional heritability analysis can thus be used to estimate the summed additive variance at individual loci contributing to trait variation. Because the regional heritability combines different sources of variation within the region, it takes into account genetic heterogeneity, thus allowing for the identification of new loci that
cannot be found by standard single-SNP GWAS analyses. We show by simulation that a regional heritability based genome scan has substantially greater power than a standard GWAS to map loci that contribute variance due to the segregation of several common or rare variants whilst retaining similar power to map loci segregating for a single common variant. We present analyses of real data from both human and livestock populations that demonstrate that the approach can find those loci identified by standard GWAS analyses; additionally it will detect new loci that were not previously identified but which can be subsequently confirmed by large meta-analyses. Regional heritability mapping provides a valuable complement to standard GWAS analysis, helping to identify some of the loci contributing missing variance. Importantly, it also provides the framework for uniting results from linkage and association analysis.

P-22
GENETIC CORRELATION BETWEEN FEED EFFICIENCY TRAITS WITH PRODUCTION AND CONFORMATION TRAITS IN DAIRY CATTLE

Ghader Manafiazar, Laki Goonewardene, Thomas McFadden, Erasmus Okine, Zhiquan Wang
Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, Alberta, Canada

Although feed cost is important in the dairy industry, little attention has been paid to selection for improved feed efficiency due to cost associated with individual feed intake measurements. Therefore, finding indicator traits such as conformation might be more applicable industry-wide. Individual daily feed intake (FI) of 191 first lactation dairy cows from 1 to 305 days in milk were recorded at the University of Alberta Dairy Research and Technology Centre; individual milk yields, milk composition, and confirmation traits were obtained from Canadian Dairy Network. The calculated efficiency traits were residual feed intake (RFI), feed conversion ratio (FCR), and gross energy efficiency (GEE). The conformation traits were considered in the analysis included rump angle, pin width, udder depth, rear attachment height, rear attachment width, teat length, dairy strength, final score, feet& leg score, and mammary system. Multiple-trait analysis was performed using ASREML software to estimate heritability and genetic correlation for the traits. Estimated heritabilities were, 0.24, 0.31, 0.17 and 0.38, for RFI, FCR, GEE and FI, respectively. RFI had low phenotypic correlation with milk production (0.07), and three top genetically correlated conformation traits were dairy strength (0.21), rear attachment width (0.19), and rump angle (0.17). FCR had a high phenotypic (0.5) and genetic correlation (0.4) with milk production. GEE had a high phenotypic (0.5) and moderate genetic correlation (0.2) with milk production. Overall, RFI had near zero correlation with production traits, while FCR and GEE had strong correlation with production. Therefore, RFI can be referred as net feed efficiency and the combined conformation traits might be used as indicator traits for RFI in dairy industry to increase profitability by selecting animals that are genetically superior in energy efficiency.

P-23
CASE-CONTROL MAXIMUM WEIGHTED BIPARTITE MATCHING IN GENOME WIDE ASSOCIATION STUDIES

Andre Lacour1, Tim Becker1,2
1German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany, 2Institute for Medical Biometry, Informatics and Epidemiology, University of Bonn, Bonn, Germany

Population stratification in samples of genome wide association studies give rise to large obliterations in the results of statistical tests. In order to correct for stratification effects we have implemented a pairwise case-control matching that is based on the identity-by-state matrix. We obtain a ‘maximum weighted bipartite matching’ by making use of an improved Kuhn-Munkres “hungarian” algorithm which solves the assignment problem of weighted bypartite graphs in polynomial time.
A quality control on the matched pairs as well as a rematching of residual sample elements makes sure that we do not lose power due to reducing the sample. In this way, the pairwise matching is extended to tiny clusters with at least one case and one control. Association P-values are obtained by within cluster case-control permutation. The matching can be performed both genome-wide and window-wise (‘localized matching’). The latter will be applied to the analyses of rare variants, where one would expect that the amount of stratification vary according to genomic location.

As it turns out from simulation studies the statistical niveau is maintained: local window sizes of a few thousand SNPs are enough to guarantee identification of strata. Thus, our method leads to an increase of power and simultaneously to a reduced false-positive rate in simulations compared to unstratified analyses. As a byproduct, our implementation strongly outperforms common covariate approaches based on multidimensional scaling in runtime, and makes genome-wide application possible. Our method for stratified analyses is implemented in the genome-wide interaction analysis software INTERSNP.

P-24
ONEQTL: AN R PACKAGE FOR SINGLE AND MULTIPLE TRAITS MULTIPLE INTERVAL MAPPING OF QUANTITATIVE TRAIT LOCI FROM BIPARENTAL INBRED AND OUTBRED CROSSES

Luciano E Silva, Gabriel Margarido, Antônio Garcia
University of São Paulo, Piracicaba, SP, Brazil

Oneqtl is an R package under development (beta version) that implements the statistical methods for multiple interval mapping of quantitative trait loci (QTL) on single (Kao, 1999 Genetics 152:1203-1216.) and multiple traits (L.D.C. E Silva, S. Wang and Z-B. Zeng, submitted for review) from biparental inbred line crosses: backcross, F2 intercross, and recombinant inbred lines derived by selfing or sib-mating. Oneqtl also implements our newly developed statistical method for single and multiple traits multiple interval mapping of QTL from biparental outbred cross, i.e. full-sibs (being written up for submission to a peer reviewed journal). Oneqtl contains many functions that allow fine-tuning of a linear model from normally distributed data via either mixture model (Lander and Botstein, 1989 Genetics, 121, 185–199) or Haley-Knott regression (Haley and Knott, 1992 Heredity, 69, 315–324.). For example, there are functions to estimate genome-wide empirical threshold via permutation (Churchill and Doerge, 1994 Genetics 138:963-971) and resampled score statistics (Zou et al., 2004 Genetics 168:2307-2316), to search for QTL effects (main and interaction), to refine QTL positions, and to test QTL effects. The search for QTL interaction can be done between QTL already in the model (0-D scan), between QTL already in the model and QTL not yet in the model (1-D scan), and between QTL not in the model (2-D scan). Since our package is fully integrated with the OneMap R package (Margarido et al. 2007, Hereditas 144: 78-79) for map estimation from biparental inbred and outbred crosses, we expect that Oneqtl will not only have broad application in dissecting the genetic architecture of complex traits from biparental inbred line crosses, but also from biparental outbred cross, due to the lack of specialized statistical tools for this type of cross. We exemplify the usefulness of our package through the analyses of real data.

P-25
A METHOD FOR GENETIC ANALYSIS OF ENVIRONMENTAL VARIATION

Majbritt Felleki1,2, Dongwhan Lee3, Youngjo Lee3, Arthur Gilmour4, Lars Rönnegård1,2
1Dalarna University, Falun, Sweden, 2Swedish University of Agricultural Sciences, Uppsala, Sweden, 3Seoul National University, Seoul, Republic of Korea, 4University of Wollongong, Wollongong, Australia

Standard tools in animal breeding like ASReml can be used for estimating parameters of environmental variation, for instance of a genetic heterogeneity model such as the exponential model (Hill & Mulder 2010) of SanChristobal-Gáudy et al. (1998). It is possible also to include a genetic correlation between
effects of mean and variance model. The proposed method is much faster than Markov Chain Monte Carlo method using minutes rather than days for the same analysis (Rönnegård et al. 2010). Simulation studies gave unbiased estimates of genetic effects. For permanent environmental effects some bias was present generally moving variance mass from the mean model to the variance model. A comparison study using data from Sorensen and Waagepetersen (2003) gave similar results to the MCMC method.

We call the method an iterative reweighted least square (IRWLS) approximation of DHGLM, because it builds on the theory of Double Hierarchical generalized linear models (DHGLM) (Lee & Nelder 2006). It can be intuitively explained, the main idea is simultaneous fitting and iteratively updating mean and variance. It will be implemented in ASReml 4, which will simplify analysis.

P-26
DERIVATION OF A NEW LAMB SURVIVAL TRAIT FOR IMPLEMENTATION IN THE NEW ZEALAND SHEEP INDUSTRY

Sylvie Vanderick1, Benoit Auvray2, Sheryl-Anne Newman2, Ken Dodds2, Julie Everett-Hincks2
1Gembloux Agro-Bio Tech. University of Liège, Gembloux, Belgium, 2AgResearch Ltd., Mosgiel, New Zealand

Lambing percentage is one of the most significant factors affecting profitability on New Zealand sheep farms. Since the early 1990s, lambing percentage has increased at about 1% per year from a relatively stable level of approximately 100%, and top performing sheep farms are now consistently achieving 150% or more. As lambing percentage increases, the proportion of ewes bearing twins and triplets increases accordingly. Lamb mortality rate in these multiples is higher than in singles, with triplets being particularly susceptible. Consequently, lamb survival has become increasingly important to the New Zealand sheep industry. Sheep Improvement Ltd (SIL, New Zealand’s national sheep genetic evaluation system provided by Beef & Lamb NZ) records lamb survival to weaning but genetic improvement has been limited due to the low heritability of the trait and the current method of recording.

This poster will present how we reviewed the existing SIL survival to weaning and developed an improved survival to weaning trait for industry implementation. This included a comparison of current and proposed methods to record and score lamb survival, decision rules for data inclusion/exclusion, development of a new genetic evaluation model and estimation of the associated variance components, comparisons of genetic parameters and breeding values for the current and new traits, and identification of flocks by birth year combinations with unusual relationships between the two traits or where significant re-ranking occurred.

P-27
GENETIC ANALYSIS OF GROWTH CURVE PARAMETERS IN BALUCHI SHEEP

Mohammad Reza Bahreini Behzadi
Yasouj University, Department of Animal Sciences, Yasouj, Iran

Weight – age data were fitted to a Brody function to estimate parameters of growth curve and their genetic and phenotypic parameters. Genetic and phenotypic relationships were also estimated between growth curve parameters and weight at birth (BW), weaning (WW), six-month (W6), nine-month (W9) and yearling (YW). The data was collected from total of 16650 body weight records of 2071 animals at Abbas-Abad Sheep Breeding Station, from 2002 to 2009. The fixed effects of the model were birth year, age of dam, type of birth and sex. Additive direct genetic effect was considered as random effect. Heritability estimates of growth curve parameters A (asymptotic mature weight), B (proportion of mature weight attained after birth) and k (the rate of maturity) were estimated 0.06 ± 0.03, 0.16 ± 0.04 and 0.15 ± 0.05, respectively. Genetic correlations between A and B, A and k and B and k were -0.46±0.28, -0.69±0.09
and 0.15±0.18, respectively. The negative association between mature weight and rate of maturing indicate that animals maturing early are less likely to obtain large mature weights than those individuals growing slowly in early life. Genetic correlations of A and k with BW, WW, W6, W9 and YW were 0.62 and −0.04, 0.46 and -0.04, 0.57 and -0.05, 0.39 and -0.08 and 0.89 and -0.09, respectively. Medium to high genetic correlations between mature weight and immature weights indicated that selection for these weights will result in a correlated response in mature weight.

P-28
THE USE OF ARTIFICIAL NEURAL NETWORK TO PREDICT SOME CARCASS TRAITS TO BE USED IN SELECTION PROGRAM

Mohammad Reza Bahreini Behzadi1, Ali Asghar Aslaminejad2, Mohammad Ali Norouzian3
1Yasouj University, Department of Animal Sciences, Yasouj, Iran, 2Ferdowsi University of Mashhad, Department of Animal Sciences, Mashhad, Iran, 3University of Tehran, Abourayhan Pardis, Department of Animal Sciences, Tehran, Iran

The aim of this study was to test the usefulness of artificial neural network (ANN) for prediction of some carcass traits in Baluchi sheep based on body weights data. Recently, attempts have been made to measure carcass quality in the live animal before slaughter. ANN has proved exceptionally useful for solving many problems related to data prediction and classification. The investigated traits were hot carcass weight (HCW), cold carcass weight (CCW), empty body weight (EBW), omental and mesentery fat (OMF) and fat-tail weight (FTW). The study was carried out on 70 lambs weaned at 42 days of age and fattened from 63 to about 150 days of age in the Research Farm of Ferdowsi University of Mashhad, Iran. Each lamb was weighed weekly from birth to the end of the period. At the end of the experiment, 35 lambs were weighed after 12h feed deprivation and slaughtered. Data were analyzed using NeuroSolutions for Excel Release 5 software. The accuracy of ANN predictions was evaluated by using the correlation coefficient between the actual values of traits and the corresponding ANN predictions. It was found that a multilayer perceptron with a backpropagation network with 4 neurons in hidden layers gives the best fitting with the experimental data, which made it possible to predict HCW, CCW, EBW, OMF and FTW with acceptable correlation coefficient (0.87, 0.55, 0.88, 0.81 and 0.73, respectively). These results show that ANN could potentially be used to predict sheep carcass traits. Due to difficulties in recording of the carcass traits and the cost, there is no selection for these traits in Baluchi sheep right now. Therefore, it would be helpful using the predicted data by ANN in the selection program of this breed.

P-29
PRACTICAL CONSIDERATIONS FOR REGIONAL HERITABILITY MAPPING

Yoshinobu Uemoto1,2, Ricardo Pong-Wong1, Pau Navarro3, Chris Haley1,3
1Roslin Institute and Royal (Dick) School of Veterinary Studies University of Edinburgh, Edinburgh, Midlothian, UK, 2National Livestock Breeding Centre, Nishi-Shirakawa-gun, Fukushima, Japan, 3MRC Human Genetics Unit at the MRC Institute of Genetics and Molecular Medicine at the University of Edinburgh, Edinburgh, Midlothian, UK

Genome-wide association studies (GWAS) have provided valuable insights into the genetic basis of complex traits. However, they have explained relatively little of the heritability. Recently, a new variance component approach called regional heritability mapping (RH) has been suggested to capture more of the missing genetic variation. In RH analyses, the genome is divided into windows and two genetic effects (one regional and one genomic) are fitted for every window, requiring a large number of calculations. Computational needs depend on the analysis window-size. We investigated the window-size (number of SNPs) needed to capture a QTL effectively, and tested the implementation of RH using a GRAMMAR-type approach, where the data is preadjusted for the genomic background, decreasing the computational load.
Simulations based on real data from a human population of 1851 individuals genotyped with 300,000 SNPs were used to assess the optimum analysis window-size. We simulated regions spanning 100 SNPs (approximately 1 megabase of DNA on average) explaining 5% of the total variance due to the combined effect of 1, 5 or 10 QTL. These regions were analysed using RH with windows containing 10, 20, 50 or 100 SNPs. Results were compared with single-SNP GWAS analysis in terms of power of detection. Across all situations varying the number of simulated QTL or the window size during the analysis, RH has a greater power than GWAS for some or all windows. Using a larger window size was more powerful when the effect was due to 5 or 10 QTL. Implementing the GRAMMAR-type approach speeded the analysis but led to underestimation of p-values relative to standard RH, however, results from the two approaches were very highly correlated. Therefore, we suggest that the optimum practical approach would be to perform GRAMMAR-type RH, and then, re-analyse the top ranking regions using the standard RH method.

P-30
RANDOM SURVIVAL FOREST HELPS IDENTIFY A VARIANT OF THE OBESITY GENE THAT AFFECTS B-CELL PROGRESSION

Emmanouela Repapi1,2, Jorge Zeron-Medina1, Anton Parker3, Jonathan Strefford4, David Oscier5, Nicolai Meinshausen2, Gareth Bond1
1Ludwig Institute for Cancer Research, University of Oxford, Nuffield Department of Clinical Medicine, Oxford, UK, 2Department of Statistics, University of Oxford, Oxford, UK, 3Royal Bournemouth Hospital, Bournemouth, UK, 4Cancer Molecular Genetics, University of Southampton, Southampton, UK, 5Department of Haematology, Royal Bournemouth Hospital, Bournemouth, UK

The identification of genetic variants, such as single nucleotide polymorphisms (SNPs), that affect cancer progression and overall survival, could provide a significant aid in the betterment of treatment strategies. The genome-wide association study (GWAS) is a powerful method for detecting SNPs associated with disease susceptibility and quantitative traits. However, the GWAS is not as well suited for time-to-event phenotypes, such as cancer progression and survival, especially in low-incidence diseases, which comprise most types of cancer. Specifically, commonly used survival models are only valid under particular assumptions that are difficult to test in a genome-wide setting. In addition, the large sample sizes required for GWAS pose a considerable obstacle. To address these limitations, our approach focuses on SNPs in genes whose protein products are clearly involved in cancer in order to reduce the number of SNPs tested in each patient cohort and increase our power. Secondly, our approach utilizes a combination of two non-parametric survival methods, the log-rank test and the random survival forest, both of which are model-free and therefore do not require diagnostic tests. We have utilized simulations to examine whether the variable selection method of the random survival forest improves the power of the log-rank test to identify genetic variants associated with survival phenotypes and to test the power of our methodology. These studies have resulted in the identification and validation of a high frequency SNP in the obesity gene that associates with progression of B-cell chronic lymphocytic leukaemia, a common leukaemia for which biomarkers for, and a deeper understanding of, cancer progression is greatly needed.

P-31
ADDITIVE GENETIC VARIANCE OF FITNESS-RELATED TRAITS – ESTIMATION BASED ON BAYESIAN ANIMAL MODELS AND INTEGRATED NESTED LAPLACE APPROXIMATIONS (INLA)

Anna Marie Holand1, Ingelin Steinsland2, Sara Martino1, Henrik Jensen1
1Centre for Conservation Biology, Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway, 2Centre for Conservation Biology, Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway, 3Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway
Estimates of the additive genetic variance of fitness-related traits (heritability) are vital in evolutionary biology as they give insight into the potential for evolutionary change and enable us to predict evolutionary response to selection. The animal model is a general linear mixed model (GLMM) that enables us to identify the genetic part of quantitative traits. To do inference for Bayesian animal models we use the recently introduced Integrated Nested Laplace Approximation (INLA) methodology. We demonstrate that the INLA methodology can be used for many versions of a Bayesian animal model using both simulated data and real data from natural house sparrow (Passer domesticus) populations. Additive genetic variances of Gaussian, Binomial and Poisson likelihoods / traits were estimated, showing low heritability of traits closely related to fitness. These results were compared with results using MCMC (Markov Chain Monte Carlo) methods. The two methods for inference gave similar results, but inference using INLA was significantly faster.

We also introduce an R package, AnimalINLA, which will enable evolutionary biologists and animal breeders to easily do fast inference for Bayesian Animal models using INLA. We expect that this will aid our understanding of evolutionary processes in natural populations and in animal breeding.

P-32
A SIMPLE METHOD TO CALCULATE IBD PROBABILITIES

Brian Kinghorn
University of New England, Armidale, Australia

This pedigree-free method predicts the IBD probability for each locus in a pair of haplotypes or gametes. A Bayesian approach is used, with a kernel that considers a proband locus, whose IBD probability is calculated, and an adjacent supporting locus, whose state (IBS or not) is used in this calculation. This procedure is applied from left-to-right, treating each locus as the proband and the locus to its right as the supporting locus, then right-to-left, and this cycle is iterated till convergence in IBD probabilities. The probability of IBS at the supporting locus is initially calculated, independently from its known genotype, but conditional on probability of crossover between the two loci since a common ancestral gamete (CAG), assuming known recombination fraction and generations since the CAG. The latter assumption is of little consequence, as in later iterations the probability of crossover since the CAG is derived from prob(IBD) at the supporting locus, conditional on IBD at the proband locus. The method has not been extensively tested and some further development is probably warranted. Simple testing by simulation involved unique tagging of foundation SNPs through 100 generations of burn-in (Ne=100) plus four generations of a real pedigree structure (n=1091). True IBD was thus diagnosed for all 1000 loci in 2.3 million gamete pair contrasts, as well as IBS, and IBD probabilities from the method described. The correlation between true IBD and IBS averaged 0.556, and that between true IBD and IBD probability averaged 0.804 after four iterations (8 minutes CPU time). IBD probabilities at the locus level should lead to much greater linkage disequilibrium with QTL, compared to IBS. It is proposed that IBD probability can be used instead of IBS to build gametic relationship matrices, and other such components/applications that do not already infer IBD in a competitive manner.

P-33
PROGENO: AN INTEGRATED SOFTWARE SYSTEM FOR PHENOTYPIC AND GENOMIC SELECTION

Steven Maenhout, Bernard De Baets, Geert Haesaert
1Ghent University, Ghent, Belgium, 2University College Ghent, Ghent, Belgium

As the price tag of high-density genotyping comes down, genomic selection gradually finds its way into breeding programs of various plant and animal species. Widespread adoption of this selection approach is, however, hindered by a lack of user-friendly software that allows practical breeders to distill reliable genomic breeding values from their available phenotypic and genotypic data collections. The increasing
density of commercial SNP chips and the computational challenges this imposes on existing software solutions forms a second hurdle which is generally difficult to overcome without considerable concessions with respect to the dimensionality of the initial problem. These two obstacles have been the key motivation for the development of the Progeno software system. In its core, Progeno is a linear mixed model computing engine which was written from scratch to solve problems involving millions of unbalanced phenotypic observations in combination with dense molecular marker profiles containing tens, if not hundreds of thousands of individual marker scores. It offers ample model flexibility including a wide range of possibilities for imposing a predefined structure on the variance of random effects and residuals including unstructured, compound symmetry, autoregressive, anisotropic and many other variance and correlation structures. Pedigree and molecular marker information can be integrated in the variance structure of random effects using various approaches ranging from a classic numerator relationship matrix to more recent advances such as Reproducing Kernel Hilbert Space Regression. The computational workload can be spread over multiple processors and flexible out-of-core storage techniques avoid having to compromise with respect to data quantity or model complexity due to computer memory limitations. The resulting pedigree- or genomic-based prediction models can be valorized by breeders directly as a user-friendly web application allows them to generate estimated breeding values, cross-predictions and even parental cross advice.

P-34
EFFECTS OF CAUSAL STRUCTURES BETWEEN TRAITS FOR PREDICTION OF BREEDING VALUES

Toshihiro Okamura, Motohide Nishio, Akiko Nishiura, Osamu Sasaki, Masahiro Satoh
NARO Institute of Livestock and Grassland Science, Tsukuba, Japan

Estimating structure between phenotypic traits may be effective for more accurate prediction of breeding values. Valente et al., (2010, Genetics) have developed techniques to estimate the causal structure using structural equation models (SEMs) and inductive causation (IC) algorithm. The objective of this study was to investigate the effects of causal structure between traits for predicting breeding values using computer simulation. Data sets including 1,500 subjects were generated from two parameters sets (genetic (co)-variances, residual (co)-variances and causal structure); one set had true causal structure that can be estimated using IC algorithm (D1), and the other set had that cannot be estimated (D2). These subjects were consisted of 150 full-sib families and a full-sib family has 10 full-sibs. After being confirmed the feasibility of estimation of causal structure using IC algorithm two kind of breeding values using BLUP procedure taking effects of causal structure into account in the model (PBV1) or not (PBV2). For instance, if a causal effect of trait 1 on trait 2 exists, the prediction of breeding values for trait 2 is conducted including trait 1 as a covariate. The correlation between true and predicted breeding values was defined as the accuracy of predicted breeding values and the accuracies of PBV1 and PBV2 were compared. The accuracy of PBV1 was larger than PBV2 in D1. However, the accuracy of PBV1 was almost the same as PBV2 in D2. The result suggested that if causal structure could be estimated with IC algorithm, taking effects of causal structure between traits into account in the model was effective for predicting more accurate breeding values.
P-35
DEVELOPMENT OF A METHOD FOR HAPLOTYP-E-BASED ASSOCIATION ANALYSIS OF BINARY
TRAITS IN STRUCTURED POPULATIONS

François Guillaume, Zhiyan Zhang, Arnaud Sartelet, Carole Charlier, Michel Georges, Frédéric Farnir, Tom
Druet
University of Liège, Liège, Belgium

Population or family structure (stratification) can cause spurious associations. Mapping methods can help
identifying underlying causal variants only if true signals of functional association can be separated from
false signals due to stratification. Linear mixed models approaches have been proposed to account for
stratification by including a genomic relationship matrix. We herein extend this approach to Generalized
Linear Mixed Models (GLMM) more appropriate for non normally distributed traits such as disease status for
instance. With a previously developed model relying on a hidden Markov model (Druet & Georges, 2010),
reconstructed haplotypes are assigned to ancestral haplotypes that are shown to correspond to clusters of
genealogically related chromosomes. In the original study, it was shown that these cluster states can directly
be used to fine map QTL. We developed a method to solve GLMM and compute score tests with such
ancestral haplotypes. To study properties of the newly developed method we simulated data with several
levels of stratification (population and/or polygenic effects) in addition to single SNP effects. Haplotypes
presented high linkage disequilibrium with underlying SNPs and higher power was achieved if they were
used instead of SNPs. It was shown that the method correctly accounted for stratification. The method was
also tested on real data from genetic defects in Belgian Blue cattle. In conclusion, the proposed GLMM
method successfully accounted for stratification and should prove useful since it is flexible (it is possible
to include additional effects) and does not rely on a particular biological model. It was shown to correctly
map trait loci with simple genetic architecture (monogeneic, recessive or dominant) but can also handle
incomplete penetrance or complex traits. The model can also be applied with other link functions for
survival data or count traits. The method is implemented in GLASCOW, a freely available software (http://

P-36
A SUPERVISED LEARNING APPROACH TO GENETIC CLASSIFICATION

Carlos Pinto1, Michael Bridges2, Darach Golden1, Jimmy Tang1, Iria Rodriguez Perez3
1University of Dublin, Dublin, Ireland, 2University of Cambridge, Cambridge, UK, 3University of Santiago
de Compostela, Santiago de Compostela, Spain

We address the problem of detecting and quantifying systematic genetic differences between pre-defined
groups of individuals.
This problem arises in many contexts. An obvious theoretical application is the study of geographically
separated populations. From a more applied standpoint, such systematic differences are a critical
consideration in case-control studies of traits with a genetic component.
In such instances, the candidate groups are known in advance. However most approaches to this
problem, such as principal components analysis (PCA) do not take exploit this prior knowledge and do
not, therefore, achieve optimal sensitivity.
We have adopted a supervised learning approach in which prior knowledge of group membership is
explicitly taken into account.
An initial feasibility study confirmed that supervised learning is significantly more sensitive than PCA in
situations where group membership is known.
We now apply supervised learning techniques to the control datasets genotyped by the Wellcome Trust
Case Control Consortium (WTCCC)[2]. These controls are widely used in case control studies by the
genetics community. They consist of two large groups of individuals each typed on two different arrays.
This poster will report on our analysis of the WTCCC control data data and compare it with more
traditional approaches to quality control. 
This poster will report on our analysis of the WTCCC control data and compare it with more traditional approaches to quality control.

P-37
MIXED MODEL WITH SOCIAL EFFECTS IN A DUROCK PIG BREED DATABASE

Mónica Santos-Cristal de Sivak¹, Rodolfo J.C. Cantet¹,²
¹Departamento de Producción Animal, Facultad de Agronomía, Universidad de Buenos Aires, Buenos Aires, Argentina, ²Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

In animal breeding, more specifically for pigs and laying hens, social effects have an increased interest for animal welfare and production. A problem with the animal model used to estimate direct and social effects (SAM) with a real set of data, is the lack of identifiability of the (co)variance components. This results from the confounding between the pen and the social effects. Our objective in this study is to find a solution for this lack of identifiability in a Durock pig data set, bred in pens located at stables. In several pens, different feeding groups of animals spent part of the period together; and most of the pens have only one feeding group. The SAM equation is $\mathbf{y} = \mathbf{X}\mathbf{\beta} + \mathbf{Z}_d\mathbf{a}_d + \mathbf{Z}_c\mathbf{a}_c + \mathbf{e}$, where $\mathbf{y}$ is the data vector; $\mathbf{a}_d$ and $\mathbf{a}_c$ are random vectors with direct (DBV) and social (SBV) breeding values; $\mathbf{X}$, $\mathbf{Z}_d$ and $\mathbf{Z}_c$ are the incidence matrices that relate the data with fixed ($\mathbf{\beta}$) and random effects (DBV, SBV); and $\mathbf{e}$ is the random error vector. The phenotypic value of animal $i$ is affected by the additive social genetic components of other animals ($j$) in the pen: $f_{ij}a_{c_i} + f_{i2}a_{c_2} + \ldots + f_{im}a_{c_m} = \sum_{j=1}^{m} f_{ij}a_{c_j}$. The values of $f_{ij}$ are the intensity of competition (IC) that the SBV of 1, 2, … $m$ animals exert over the phenotype of $i$, and are the non-zero elements in the $i$-th row of $\mathbf{Z}_c$. In previous papers, the restriction $\sum_{j=1}^{m} f_{ij}a_{c_j} = 1$, was proposed to standardize the variation of the SBV. To avoid collinearity between $\mathbf{X}$ and $\mathbf{Z}_c$, the objective of the current work is to define the ICs as a function of two factors: the time that animals spent together in the same pen, and the relationship between them as social effects would have been different between related and non related animals. The $\mathbf{Z}_c$ matrix proposed allows for the identifiability of the (co)variance components, as shown in the examples presented. The $\mathbf{Z}_c$ of Muir is a special case of the $\mathbf{Z}_c$ matrix proposed here.

Key words: social breeding values, mixed models, identifiability, (co)variance components.

P-38 – ABSTRACT WITHDRAWN

P-39
VARIANCE CONTROLLING GENES AND EPISTASIS: NEW OPPORTUNITIES IN GENETICAL GENOMICS

Ronald Nelson¹, Xidan Li¹, Xia Shen², Örjan Carlbord¹
¹Swedish University of Agricultural Sciences, Uppsala, Sweden, ²Uppsala University, Uppsala, Sweden

Combining genome-wide genotypic data and microarray expression data make it possible to associate polymorphisms on the DNA level to biochemical pathways. The usefulness of this approach has been
shown in earlier association studies, where large effects of individual genes have been implied. The nature of biochemical pathways is, however, a complex interplay of molecular components that are likely to act in both an additive and a non-additive fashion. Although the traditional association and linkage studies are powerful in identifying additive effects, they are less suitable for identifying complex interactions involving multiple loci. Using a newly developed method (vGWAS) where the variance of the phenotype (here, gene expression) is analysed instead of the mean, we are able to find loci that regulate the variability in expression phenotypes rather than the mean. Using expression data from a cross of two yeast strains (BY4716 and RM11-1a), we show that the vGWAS is useful for identifying gene-gene interactions on both expression and DNA levels. In this way, it presents an approach to fill the gaps between the genotype and expression phenotype in populations by assigning new functions to known biochemical pathways and expand our understanding of incomplete gene networks and biochemical pathways. Also, once the markers or genes that significantly affect the phenotype are found, the differences in response between crossed strains can be investigated. vGWAS is thus emerging as a useful tool to supplement current analysis methods, that use genomic and expression data to understand the role of variance and epistasis in biochemical pathways.

P-40

HOST RESISTANCE TO A GASTROINTESTINAL PARASITE REGULATED BY A SNP MODULATING SMAD3 BINDING TO A TGFB-REGULATED TRANSCRIPTIONAL ENHANCER OF MINA, ENCODING AN IMMUNOREGULATORY MEMBER OF THE JMJC PROTEIN FAMILY

Shangli Lian, Meenu Pillai, Mark Bix
St. Jude Children’s Research Hospital, Memphis, TN, USA

In a hunt for Th2-bias regulatory loci in mice, we have used interval-specific congenic mapping to positionally clone a gene called Mina, encoding an immunoregulatory member of the JmjC protein family. Th2-bias is a genetic trait defined as the propensity of naïve T helper cells to develop in vitro under neutral conditions into T helper 2 (Th2) effector cells. Th2 cells play a critical role in host protection from multicellular parasites such as intestinal helminths and when disregulated contribute to allergic diseases such as asthma. The Th2-bias trait is known to vary across different inbred mouse strains and correlates inversely with the relative abundance of intracellular Mina protein. Adding further support for a causal link between Mina protein level and Th2-bias, Mina was shown to act as a dose-dependent transcriptional corepressor of the gene encoding interleukin-4 (IL4), a key regulator of Th2 development. We have now identified a SNP that explains differential Mina expression level in high and low Th2-biased strains such as BALB/c and C57BL/6. Most interestingly, this SNP acts by modulating SMAD3 binding to a TGFB-regulated intronic Mina enhancer. We are currently tying this into its physiological role in pulmonary inflammation and intestinal helminth infection using Mina conditional KO mice, which exhibit dramatic phenotypes in both models.

P-41

USING KEGG PATHWAYS AND EXPRESSION STUDIES AND STUFF FOR GENOMIC PARTITIONING

Stefan M. Edwards, Li Jiang, Per Madsen, Peter Sørensen
Department of Molecular Biology and Genetics, Faculty of Science and Technology, Aarhus University, Tjele, Denmark

Abstract can be found in the student symposium abstract section on page 61.
Quantitative trait locus (QTL) mapping has been valuable in characterizing the “genetic architecture” of quantitative traits. By regarding gene expression as a quantitative trait and monitoring gene expression of thousands of genes, new inferences involving concepts from systems biology are possible. “Genetical genomics” uses the genetic variation in segregating progenies to identify eQTLs (expression QTLs), then from co-localized eQTLs, we can reconstruct networks of gene expression; in certain cases, the direction of causality can be inferred (a “directed graph” constructed). Most interestingly, the roles of a gene network in causing or reacting to a particular trait of interest, such as a disease, can be inferred. Genetical genomics can augment, and even advance, our knowledge about the functional biology of genes. We have been conducting both theoretical and empirical studies in this area. Distant pair mating designs, factorial crossing designs, and statistical power to infer networks and directionality are discussed. For empirical studies, we have been using spruce. We have assayed a 21832-member cDNA microarray and SNPs to infer gene networks using progeny of a 2x2 factorial cross in Interior spruce, and a 3x2 factorial cross in Sitka spruce, both involving ca. 60 progeny per family. For example, patterns of QTL co-expression revealed an inferred network of phenylpropanoid genes and MYB transcription factors which was previously identified by functional studies. We have also evaluated the causative roles of these expressed genes in growth and insect resistance. In a third study of Interior spruce (in progress), we are using next-generation sequencing to simultaneously genotype and measure gene expression in segregating progenies.
of purpose-built pig model for obesity to conduct this systems genetics investigation. The human study design consists of 4,600 Indian and 1,500 Danish obesity case-control cohorts for EWAS and the pig (F2 intercross) design consists of genomic data from 60k porcine SNPchip, transcriptomic data (from RNAseq) and a range of DEXA scanned obesity traits and traits related to metabolic syndrome. With these study designs as a background, we present analytical models for within- and between-populations (meta) analyses to detect potential causal DNA variants, whole genome predictions for obesity and metabolic disorders in Danish and Indian populations and in pig models. Further, we present analytical framework for construction of weighted gene co-expression networks and systems genetic investigations to build causal, regulatory and transcriptional networks leading to development of obesity and diabetes or variation in quantitative metabolic traits. Some preliminary results on quantitative genetic characterization (heritabilities and genetic correlations) of obesity and metabolic traits in pig F2 intercross populations will be presented. We also show concepts behind how these ‘animal model’ results will be linked with human EWAS hits, via translational approaches, to underpin the causality and systems biology of childhood obesity in two human populations.

P-45

COMBINING GENETIC VARIATION WITH TARGETED KNOCK-DOWNS/OUTS TO UNDERSTAND COMPLEX HUMAN DISEASE PATHWAYS IN C. ELEGANS

Jan Kammenga
Wageningen University, Wageningen, The Netherlands

Objectives: The nematode worm C. elegans has intensively been studied for complex human disease pathways. These analyses are rooted in forward or reverse genetic screens in a single standard laboratory wild type. Yet, induced mutations or RNAi-screens in one genotype do not reveal the allelic effects that segregate in natural populations and contribute to phenotypic variation. Our objective is a quantitative genetic analysis of natural variation in complex disease signalling in C. elegans to identify novel alleles and their regulators.

Methods: We have performed a quantitative pathway analysis of natural variation in complex disease signalling in C. elegans and present a robust approach for selecting candidate genes associated with the Notch, Wnt and RAS pathway. Using a powerful combination of sequenced recombinant inbred lines (RILs), mutant and RNAi screens, we focused on the transcriptome, proteome and cellular development.

Results: We highlight a case of cryptic genetic variation in genome wide gene expression regulation (eQTL) underlying a gld-1 knock-down. Furthermore we will show how mutations of the RAS and Wnt pathways in different genetic backgrounds lead to the discovery of hidden modifiers affecting germline-development, an important phenotypic readout of cancer signalling pathways. We will illustrate how (transcriptomics, proteomics and phenotypic) data will feed into a predictive model of germline development which can be used to investigate the relative contributions of various recombinants on the phenotypic output.

Conclusions: The combination of genetic variation with forward and reverse genetic screens is a powerful approach to identify novel regulators in complex disease signalling which allows for predictive systems biology modelling.

This work was supported by the EU-FP7 project PANACEA, contract nr. 222936.
P-46
GENETIC MARKERS AS INSTRUMENTS FOR MENDELIAN RANDOMIZATION STUDIES ON VITAMIN D
Diane Berry1, Karani Vimaleswaran1, John Whittaker2, Aroon Hingorani3, Elina Hypponen1
1UCL Institute of Child Health, London, UK, 2London School of Hygiene and Tropical Medicine, London, UK, 3Department of Epidemiology and Public Health and the Centre for Clinical Pharmacology University College London, London, UK

Abstract can be found in the student symposium abstract section on page 60

P-47
VARIATION IN THE HYPOTHALAMIC-PITUITARY-ADRENAL AXIS: ASSESSMENT OF ADRENAL STEROIDS TO IMPROVE STRESS MANAGEMENT AND ROBUSTNESS
Denise Hough1, Schalk Cloete1,2, Pieter Swart1
1Stellenbosch University, Stellenbosch, Western Cape, South Africa, 2Institute for Animal Production: Elsenburg, Elsenburg, Western Cape, South Africa

The analysis of endogenous steroids in humans and livestock has become the focus of endocrinologists and physiologists that investigate these important hormones. The quantification of these steroids is imperative to our understanding of reproductive fitness and stress management. Stress responses are mediated by the hypothalamic-pituitary-adrenal (HPA) axis in combination with the autonomic nervous system and behavioural adaptation. Stimulation of the HPA axis results in the secretion of adrenocorticotropic hormone from the pituitary gland, which in turn stimulates the release of glucocorticoids, such as cortisol (sheep, cows, pigs, fish) or corticosterone (birds, rats, mice), from the adrenal gland. The HPA axis is a pathway with large individual variation and holds potential for genetic selection for robustness and related traits. Animals with superior HPA axis responses to stressors (such as environmental stress, general husbandry procedures, starvation, etc.) are likely to adapt more rapidly to stressful conditions than animals with inferior HPA axis responses. A comprehensive analysis of all the adrenal steroid hormones, as well their intermediary precursors, is more informative pertaining to the evaluation of HPA axis activity than the analysis of one or two steroid hormones (such as only cortisol or corticosterone) with immunoassays. The comprehensive analysis allows for the detection of variation within the HPA axis responses of animals, as well as the identification of animals for breeding purposes. This study presents the utilization of UPLC-MS/MS analysis for the simultaneous quantification of adrenal steroids. This technique allows for a holistic evaluation of the adrenal response to HPA axis stimuli. Such an evaluation contributes to understanding the adrenal steroidogenesis pathway and may be used to identify genetic factors that affect robustness and related traits. The system was tested in Merinos divergently selected for reproduction, and seems to hold potential for identifying animals with superior HPA axis activity (and presumably improved fitness).

P-48
GENE PRIORITIZATION FOR QUANTITATIVE TRAITS BY DATA INTEGRATION
Li Jiang, Luc Janss, Stefan Edwards, Axel Skarman, Bo Thomsen, Peter Sørensen
Department of Molecular Biology and Genetics, Faculty of Science and Technology, Aarhus University, Tjele, Denmark

Quantitative traits such as milk production in dairy cattle have complex modes of inheritance. Genome-wide association studies for such traits typically show many genetic variants with small effects. However, it appears that often multiple independently associated variants are located in the same genes and that genetic variants tend to occur in genes whose products are connected in biological pathways or are enriched for likely functions. These findings provide valuable insight into the design of network-based
approaches for prioritizing candidate genes associated with quantitative traits. They support the idea that mutations in the same gene (allelic heterogeneity) or mutations in different members of a gene complex (pathway) may lead to identical or similar trait phenotypes. Thus, we implemented a network-based gene prioritization approach for ranking of genes associated with a quantitative trait. The approach integrates information from trait phenotypes, SNP data and protein-protein associations and uses a Bayesian mixture model for quantifying the variance explained by genetic variants linked to each candidate gene complex. The approach is illustrated on milk production traits in dairy cattle. It provides a general framework for prioritizing genes associated with various quantitative traits in different species.

P-49
DETECTING AND ANALYZING VARIANCE-CONTROLLING GENES IN STRUCTURED POPULATIONS

Marcin Kierczak, Örjan Carlborg
Computational Genetics Group, Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden

Traditionally, the statistical tests used to detect genotype-phenotype associations in genome-wide association studies (GWAS) aim to detect significant differences in phenotypic means between genotypes. However, polymorphisms in some genes cause a shift in the phenotypic variance, rather than the mean, between genotype classes. Such genes will be missed in analyses based on traditional GWAS methods. Recent studies show that such variance controlling genes are not uncommon and that they may play an important role both in explaining the phenotypic variance and the evolution of traits. For example, strong selection for mean phenotype may also result in the increase of variance, therefore preserving phenotypic diversity.

Here, we present an analytical approach that can be suitable to detect and analyze variance-controlling genes in highly stratified populations, e.g. populations including many varieties of the same plant species or several breeds of a domesticated animal. The proposed methodology, based on the combination of statistical tests for detecting difference between variances and various genomic distance-based approaches for detecting and correcting for population stratification, allow us compare the effects of variance-controlling genes in different environments and genetic backgrounds. Our research focuses on both methods development as well as analyses of three different species: 1) subpopulations of Arabidopsis thaliana collected from different locations across the globe; 2) different varieties of rice Oryza sativa collected across the globe as well as 3) several canine Canis lupus familiaris breeds and subpopulations that represent European population.
GENOME-WIDE ASSOCIATION STUDIES (GWAS) ARE CRITICALLY DEPENDENT ON DETAILED KNOWLEDGE OF THE PATTERN OF LINKAGE DISEQUILIBRIUM (LD) IN THE HUMAN GENOME, AS CHARACTERISED, FOR EXAMPLE, BY THE HAPMAP PROJECT. A GWAS GENERATES A LIST OF VARIANTS, USUALLY SINGLE-NUCLEOTIDE POLYMORPHISMS (SNPs), RANKED ACCORDING TO THE SIGNIFICANCE OF THEIR ASSOCIATION TO THE PHENOTYPE OF INTEREST. DOWNSTREAM ANALYSES FOCUS ON THE GENE(S) THAT ARE PHYSICALLY CLOSEST TO THE AFOREMENTIONED SNPS, FAILING TO ACCOUNT FOR THEIR LD PROFILE WITH OTHER SNPS. WE HAVE DEVELOPED A FLEXIBLE R PACKAGE, LDsnpR, WHICH EFFICIENTLY Assigns SNPs TO GENES BASED BOTH ON THEIR PHYSICAL POSITION AND ON THEIR PAIRWISE LD WITH OTHER SNPs.

Genome-wide association studies (GWASs) are critically dependent on detailed knowledge of the pattern of linkage disequilibrium (LD) in the human genome, as characterised, for example, by the HapMap project. A GWAS generates a list of variants, usually single-nucleotide polymorphisms (SNPs), ranked according to the significance of their association to the phenotype of interest. Downstream analyses generally focus on the gene(s) that are physically closest to the aforementioned SNPs, failing to account for their LD profile with other SNPs. We have developed a flexible R package, LDsnpR, which efficiently assigns SNPs to genes based both on their physical position and on their pairwise LD with other SNPs.

As proof of principle, we used the traditional positional binning and LD-based binning approaches to investigate whether including these “LD-based” SNPs affects the interpretation of three published (unimputed) GWASs of bipolar affective disorder (BP). We show that including LD can be critical for interpretation and comparison of GWASs. LD-based binning increased the overall number of SNPs involved in the binning process and effectively “recovered” 6.1-8.3% of Ensembl-defined genes in these datasets. It also altered the scores and subsequent ranks of the genes, and resulted in a >27% difference between the lists of the top 2,000 genes emerging from the two binning approaches. Most interestingly, LD-based binning improved the overall gene-based concordance between independent BP studies. The method was also applied to two corresponding imputed datasets. While the increase in coverage was more modest (0.4 and 2.9%), even greater concordance between datasets was observed, highlighting the benefit of LD-based binning on imputed datasets as well. Thus, our study shows how ignoring LD risks the misinterpretation of the GWAS findings and can have non-trivial consequences on subsequent genetic and functional studies. These results call for a re-assessment of GWASs, putting LD back into the picture.
P-52
QTL MAPPING OF BEHAVIOUR AND GENE EXPRESSION IN AN ADVANCED INTERCROSS OF WILD AND DOMESTIC CHICKENS

Martin Johnsson1, Ida Gustafsson1, Leif Andersson2, Per Jensen1, Dominic Wright1
1Department of Physics, Chemistry and Biology, Linköping University, Linköping, Sweden, 2Department of Medical biochemistry and microbiology, Uppsala University, Uppsala, Sweden

Animal domestication provides a dramatic example of phenotypic evolution on a short time scale. Domestic chickens and the Red Junglefowl, the main progenitor of domestic chickens, differ in both morphological and behavioural traits, including fearfulness. Along with a compact genome and high recombination rates, this makes the chicken a useful study animal for the genetics of domestication.

We performed quantitative trait locus (QTL) mapping of fear-related behaviours, as measured by a tonic immobility and a social reinstatement test, in an advanced intercross of Red Junglefowl and White Leghorn layers. 575 individuals were genotyped for 650 informative SNP markers. We detected eight QTL for social reinstatement, and two for tonic immobility (average lod score 6 and average 4 Mb support interval).

To identify potential causal variants acting by means of gene expression differences, we performed expression QTL (eQTL) mapping using 129 hypothalamus mRNA samples and NimbleGen expression arrays covering all known Ensembl and RefSeq transcripts, as well as sequences from a hypothalamus cDNA library. We detected eQTL for 652 probes, derived from 543 genes. eQTL are non-uniformly distributed across the genome, forming clusters on chromosomes 3, 7, 9, and 14. This modular distribution could be caused either by pleiotropic regulatory genes or modules of linked eQTL. The social reinstatement and tonic immobility QTL have on average five overlapping eQTL. A social reinstatement QTL on chromosome 2 overlaps an eQTL for TTRAP, encoding a transcription factor and DNA repair related protein. TTRAP expression correlates with time spent in the start zone, even when including QTL genotype in the model, suggesting a causal role.

Our eQTL results suggest a modular genetic architecture of transcript abundance traits under chicken domestication. We also implicate candidate genes, including TTRAP, for the difference in social reinstatement behaviour between wild and domestic chickens.

P-53
CIRCULAR GENOMIC PERMUTATIONS: EVALUATING GENE-SETS ASSOCIATED WITH COMPLEX TRAIT VARIATION

Claudia P Cabrera1, Pau Navarro1, Jennifer E Huffman1, Alan F Wright1, Caroline Hayward1, Harry Campbell2, James F Wilson1, Igor Rudan1, Veronique Vitart1, Chris S Haley1,2
1Human Genetics Unit, MRC, IGMM, University of Edinburgh, Edinburgh, UK, 2The Roslin Institute, University of Edinburgh, Edinburgh, UK, 3Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK

In genome-wide association studies (GWAS) only a few single nucleotide polymorphisms (SNPs) pass the stringent significant thresholds. Furthermore, the variants identified explain very little of the trait variation. Gene-set approaches applied to GWAS aim to identify sets of SNPs, which together have a detectable effect on a trait although they may not reach genome-wide significance individually. However, there is no standard way of dealing with the many variables (i.e. SNP-to-gene annotation window size, representation of a single association value for a group of SNPs, test of significance). Most methodologies select the most significant SNP to represent a gene (hence overestimating the association) and/or apply random permutation to assess the results which does not account for linkage disequilibrium or functional clustering of genes. Circular permutation accounts for genomic structure and utilizes individual SNP association p-values obtained from the GWAS analysis directly, instead of combining them into a single gene p-value. The SNPs are ordered according to their genomic locations, where the genome is
considered as circular (ordered from chromosome 1 to X, and restarting at 1 again). Then, the p-values of association are permuted by rotation with respect to their genomic locations. They therefore retain the same position with respect to each other but at each permutation gain a new random position. The methodology was applied to a GWAS data with ~700 individuals, measured for 51 traits. We used KEGG pathways as the predefined gene-sets, or in this case SNP-sets. We show that the evaluation of gene-set analyses is very variable with standard methods. Circular permutation provides a more realistic and believable evaluation of gene set involvement in complex trait variation.

**P-54 – ABSTRACT WITHDRAWN**

**P-55**
INFERRING CAUSAL PHENOTYPE NETWORKS USING STRUCTURAL EQUATION MODELS AND GENOMIC INFORMATION

*Guilherme J. M. Rosa*, Bruno D. Valente, Aniek C. Bouwman, Francisco Peñagaricano, Raphael R. Wenceslau

1University of Wisconsin, Madison, WI, USA, 2Wageningen University, Wageningen, The Netherlands

Phenotypic traits may relate to each other through complex networks of causal effects. For example, high yield in dairy cows may increase the liability to certain diseases and metabolic disorders and, conversely, the incidence of a disease may affect yield negatively. Likewise, transcription levels may be a function of the reproductive status or developmental stage in plants and animals, which may depend on other physiological variables or gene activity. To study such phenotypic networks in the context of quantitative genetics, mixed effects structural equation models (SEM) have been suggested as a flexible tool to express functional relationships between traits. The most challenging component when modeling phenotypic networks refers to inferring which pairs of traits present direct causal relationships. In this context, there are algorithms that search for plausible causal network structures, driven by observed conditional dependencies and independencies between variables. Such algorithms rely on specific assumptions, from which the causal sufficiency seems to be the strongest one, especially as multiple traits are often affected by pleiotropic genes. Nonetheless, genomic information such as DNA polymorphisms and QTL signal can be used to aid phenotypic network inference. In this presentation we will discuss how molecular marker data can be used to attenuate confounding of the search due to genetic factors, as well as how QTL can be utilized as instrumental variables, capitalizing on the concept of Mendelian randomization. Simulated data and real examples will be used to illustrate the methods and demonstrate how SEM coupled with genomic information can be used for further learning regarding relationships between phenotypic traits, providing valuable information for drug development and disease treatment, and for optimal management practices and decision making in agriculture.

**P-56 – ABSTRACT WITHDRAWN**
P-57
PATHWAY ANALYSIS OF MULTIPLE BRAIN PHENOTYPES MEASURING ACTIVATION WHEN VIEWING FACES IMPLICATES VARIANTS IN CALCIUM/CALMODULIN RESPONSIVE PATHWAYS


1Wellcome Trust Sanger Institute, Cambridge, UK, 2University of Oslo, Oslo, Norway, 3Oslo University Hospital, Oslo, Norway, 4Sørlandet Hospital HF, Kristiansand, Norway, 5Charité Universitätsmedizin, Berlin, Germany

Functional magnetic resonance blood-oxygen level dependent imaging provides an indirect measure of regional brain activity during task or rest. The spatial resolution allows quantitative measures of activity to be inferred at hundreds of thousands of locations. Several lines of evidence suggest that the variation in this activity has a genetic component, but the exact variants responsible have proved elusive. One reason could be the logistic challenges of scanning large numbers of people limit sample size. In the context of data from an experiment where participants matched faces displaying strong negative emotions (anger and fear), we attempted to combine information across both genes within pathways and multiple fMRI phenotypes to discover pathways enriched with variants affecting the induced global patterns of activation. Firstly, whole brain phenotypes were created by applying principal components analysis to the contrast maps of activation. Then GWAS scans were performed, and a list of variants moderately associated (p<1e-4) with at least one whole brain phenotype was produced. Variants were annotated to genes, and a score for a pathway was produced as a sum of the number of genes within a pathway with evidence of association, weighted by the eigenvalue of the associated principal component. We found significant evidence that variants within pathways related to post NMDA receptor activation events were enriched, chiefly genetic variants in calcium/calmodulin-dependent protein kinase II (CAMK2G, CAMK2D) and a calcium-regulated nucleotide exchange factor (RASGRF2), all of which are activated by intracellular Ca2+/calmodulin. By mapping these variants back to the normalized brain template, the strongest associations were localized to the left inferior frontal gyrus (p=1.03e-9), a region primarily involved in semantic processing but also involved in processing negative emotions.

P-58
GENETIC VARIATION IN RENAL EXPRESSION OF FOLATE RECEPTOR 1 (FOLR1) AS A RISK FACTOR FOR FEATURES OF THE METABOLIC SYNDROME

Michal Pravenec, Viktor Kozich, Jakub Krijt, Vaclav Zidek, Vladimir Landa, Ludmila Kazdova, Theodore Kurtz

1Institute of Physiology AS CR, Prague, Czech Republic, 2Institute of Inherited Metabolic Disorders, 1st Medical Faculty, Charles University, Prague, Czech Republic, 3Institute for Clinical and Experimental Medicine, Prague, Czech Republic, 4University of California, San Francisco, USA

Background: Metabolism of homocysteine and other sulfur amino acids is closely related to the metabolism of folates. It is possible that mild hyperhomocysteinemia observed in hypertension and metabolic syndrome could reflect genetic disturbances in folate homeostasis.

Methods: In the current study, we analyzed the possible role of folates/sulfur amino acids in the development of features of the metabolic syndrome in the spontaneously hypertensive rat (SHR) using linkage and correlation analyses in the BXH/HXB recombinant inbred (RI) strains derived from the SHR and Brown Norway (BN-Lx) progenitors and in vivo functional studies in SHR.BN-chr.1 congenic and SHR-Folr1 transgenic lines.

Results: We demonstrated that the SHR strain compared to the normotensive BN-Lx strain exhibits significantly lower plasma folate concentration and increased plasma total homocysteine and cysteine. In the BXH/HXB RI strains, we mapped a QTL (quantitative trait locus) for cysteine concentrations to a
region of chromosome 1 that contains a cis-acting expression QTL regulating mRNA levels of Folr1 (folate receptor 1) in the kidney. Sequence analysis revealed a 5.7 kb insertion variant in the Folr1 promoter region of the BN versus SHR strains. Transfection studies demonstrated that the SHR promoter region of Folr1 is less effective in driving luciferase reporter gene expression than the BN promoter region of Folr1. Results in the SHR.BN-chr.1 congenic strain with the BN Folr1 allele confirmed that the SHR variant in Folr1 cosegregates with reduced renal expression of Folr1, reduced renal folate reabsorption, decreased serum levels of folate, increased levels of plasma total cysteine and homocysteine, reduced muscle insulin sensitivity, and increased blood pressure. Transgenic rescue experiments in the SHR with increased expression of the Folr1 transgene confirmed these results.

Conclusions: These findings are consistent with the hypothesis that inherited variation in renal expression of Folr1 is associated with some components of the metabolic syndrome.

P-59
GENETIC PARAMETERS FOR FATNESS-RELATED TRAITS IN PIGS: ANIMAL MODEL FOR HUMAN OBESITY

Lisette J.A. Kogelman, Thomas Mark, Peter Karlskov-Mortensen, Camilla S. Bruun, Susanna Cirera, Claus B. Jørgensen, Merete Fredholm, Haja N. Kadarmideen
Faculty of Health and Medical Sciences, University of Copenhagen, 1870 Frederiksberg C, Denmark

Background & Objectives: Obesity is an epidemic worldwide public health problem leading to high morbidity and mortality rates. The recent increase in the worldwide prevalence has increased the need for knowledge about the genetic and biological background of obesity. The lack of comprehensive phenotype information and influencing environmental factors limit the detection of the genetic background for human obesity in human research. These problems can be reduced by using the pig as an animal model where several traits related to obesity can be measured in an experimental setting. The main objective of this study is to estimate genetic parameters for fatness-related traits in a pig animal model.

Methods: A total of 450 F2 intercross pigs resulting from crossing obesity-predisposed Göttinger minipigs and lean production pigs were phenotyped in respect to weight, conformation, dual-energy x-ray absorptiometry (DXA) scans, and fatness traits at slaughter. Heritabilities and genetic correlations were estimated for a total of 27 obesity-related traits using a BLUP-animal model fitting age, batch and sex as fixed effects. All pigs will be genotyped using the 60K SNP chip and 30 pigs will be selected for selective transcriptive profiling by RNA-Seq methods. Data will be analyzed using an integrative systems genetics approach.

Results: Heritabilities for weight ranged between 0.21 and 0.93 (±0.20) at different ages, for conformation traits between 0.09 and 0.38 (±0.15), for DXA-traits between 0.10 and 0.38 (±0.15), for growth rate between 0.38 and 0.66 (±0.20) and for slaughtering characteristics between 0.01 and 0.28 (±0.35). Low to high genetic and phenotypic correlations were found between the different measurements.

Conclusions: High heritabilities and genetic correlations in the experimental population suggest a strong genetic background for fatness-related traits and give a good rationale for further genomic and systems genetics approaches for detecting DNA variants, pathways and gene networks affecting human obesity.
P-60
THE INFLUENCE OF POLYGENIC RISK FOR BIPOLAR DISORDER ON NEURAL ACTIVATION ASSESSED USING fMRI

Heather Whalley\textsuperscript{1}, Martina Papmeyer\textsuperscript{1}, Emma Sprooten\textsuperscript{1}, Liana Romaniuk\textsuperscript{1}, Douglas Blackwood\textsuperscript{1}, David Glahn\textsuperscript{0,2}, Jeremy Hall\textsuperscript{1}, Stephen Lawrie\textsuperscript{1}, Jessika Sussmann\textsuperscript{1}, Andrew McIntosh\textsuperscript{1}

\textsuperscript{1}Division of Psychiatry, University of Edinburgh, Edinburgh, UK, \textsuperscript{2}Department of Psychiatry, Yale University School of Medicine, CT, USA

Objectives: Genome-wide association studies have demonstrated a significant polygenic contribution to bipolar disorder (BD) where disease risk is determined by the summation of many alleles of small individual magnitude. Modelling polygenic risk scores may be a powerful way of identifying disrupted brain regions whose genetic architecture is related to that of BD.

Methods: We determined the extent to which common genetic variation underlying risk to BD affected neural activation during an executive processing/language task in individuals at familial risk of bipolar disorder and healthy controls. Polygenic risk scores were calculated for each individual based on GWAS data from the Psychiatric GWAS Consortium Bipolar Disorder Working Group (PGC-BD) of over 16 000 subjects.

Results: The familial group had a significantly higher polygene score than the control group (p=0.04). There were no significant group by polygene interaction effects in terms of it’s associations with brain activation. However, we did find that an increasing polygenic risk allele load for BD was associated with increased activation in limbic regions previously implicated in BD, including the anterior cingulate cortex and amygdala, across both groups.

Conclusion: The findings suggest that this novel polygenic approach to examining brain imaging data may be a useful means of identifying genetically-mediated traits mechanistically linked to the aetiology of BD.

P-61
WHITE MATTER INTEGRITY IN INDIVIDUALS AT HIGH RISK OF BIPOLAR DISORDER AND HEALTHY CONTROLS: A GENOME-WIDE ASSOCIATION STUDY

Emma Sprooten, Katy Fleming, Pippa Thomson, Heather Whalley, Jeremy Hall, Jess Sussman, James McKirdy, Douglas Blackwood, Stephen Lawrie, Andrew McIntosh

Division of Psychiatry, University of Edinburgh, Edinburgh, Scotland, UK

White matter integrity, as measured using diffusion tensor imaging (DTI), is reduced in individuals with bipolar disorder (BD), their unaffected relatives and in carriers of specific risk-alleles. Fractional anisotropy (FA), one of the most common indices of white matter integrity, is highly heritable but the genetic architecture of this trait has, until now, received little investigation. In the current study we performed a genome-wide association study of FA as a quantitative phenotype, in unaffected relatives of patients with BD and a matched control group. Amongst our top results were SNPs located in genes known to be involved in cell adhesion, white matter development and neuronal plasticity. These functions were supported by pathway analysis of the top associated polymorphisms and genes which confirmed the enrichment of processes relevant to BD, white matter development and plasticity including axon guidance, ErbB-signalling, neurotrophin signalling, phosphatidylinositol signalling, and cell adhesion. The majority of genes implicated in these pathways were differentially associated with FA in individuals at high familial risk for BD, suggesting interactions with genetic background or environmental factors secondary to familial risk for BD. Although the present findings require independent replication, the results encourage the use of global FA as a quantitative phenotype in future large-scale studies that may also help to identify the biological processes underlying reduced FA in BD and other psychiatric disorders.
P-62
COPY NUMBER VARIATION IS ASSOCIATED WITH CELL SPECIFIC EFFECTS ON GENE EXPRESSION

*Esther Ng, Seiko Makino, Ben Fairfax, Julian Knight*
*University of Oxford, Oxford, UK*

**Objectives:** Although 12-15% of the human genome is subject to variation in copy number, the effect of copy number variants (CNVs) on gene expression remains underexplored. Previous analysis of small, mixed population cohorts using immortalised tissue supports a role for CNVs in the regulation of gene expression. The reproducibility of this in primary cells is unclear however, as is the degree to which CNV regulation of gene expression is constant across different cell types. Here we investigate the relationship between CNVs and gene expression in B cells and monocytes from healthy Caucasian individuals.

**Methods:** CD19+ B cells and CD14+ monocytes were positively selected from peripheral blood mononuclear cell fractions prepared from 288 healthy Caucasians. Genome wide gene expression profiling and genotyping were performed with HumanHT-12v4 BeadChip and HumanOmniExpress-12v1.0 (Illumina) respectively. CNVs were identified from the genotyping data using PennCNV and QuantiSNP. Spearman correlation was performed between copy number and robust spline normalised probe expression. HLA status was inferred by imputation using the HLA*Imp software package.

**Results:** After quality control, a total of 550 unique CNV segments were identified from 278 individuals. Probes corresponding to 13 genes displayed significant correlation between expression level and copy number (false discovery rate of <0.1). The set of CNVs that were associated with gene expression showed significant enrichment of large CNVs with this effect. Notably, 5 of these genes; SULT1A, NAIP, LILRA-6, GOLGA8B and SOCS7, demonstrated cell-specific association (correlation present in one of the cell types but not the other). Furthermore, some of these associations are to genes linked to health and disease: SULT1A activity has previously been associated with tamoxifen metabolism and breast cancer outcome whilst the murine ortholog of NAIP has been shown to confer resistance to Legionella infection. A further CNV was identified at Chr6:31384588-31393270 which showed significant negative correlation with HLA-C expression in both cell types. This CNV demonstrated linkage disequilibrium with the classical HLA alleles 0602 and 1302.

**Conclusions:** Cell-specific CNV-gene associations can be identified through a combination of high-resolution genotyping and gene expression microarrays. The functional effects of these associations in the context of health and disease warrants further attention.

---

P-63
THE INFLUENCE OF GENETIC VARIATION ON PROTEIN TRANSLATION

*Frank Albert1, Dale Muzzey2, Jonathan Weissman2,3, Leonid Kruglyak1,3*
*1Princeton University, Princeton, NJ, USA, 2University of California, San Francisco, CA, USA, 3Howard Hughes Medical Institute, Princeton, NJ, USA*

DNA variation between individuals is the substrate of evolution and influences susceptibility to many human diseases. Because some DNA variants exercise their effects on the phenotype by influencing gene expression, a common research strategy is to search for DNA variants that affect the expression level of one or several genes (expression quantitative trait loci, “eQTL”). Although coding genes exercise their functions through their protein products, most eQTL studies so far have focused on mRNA levels, rather than on the proteins themselves. To bridge the gap between studies of mRNA and protein variation, we study the influence of genetic variation on protein translation.

We have profiled genome-wide translation in two genetically different strains of the yeast *S. cerevisiae* using ribosomal footprinting. We measured the abundance of mRNA fragments that are bound to ribosomes by massively parallel sequencing. Comparisons to matched total mRNA levels provide a highly detailed view of gene expression differences between the strains that specifically involve translation, rather than mRNA abundance. By comparing the strain differences to allele-specific expression in the diploid hybrid of the
same strains, we are determining the fraction of translational variation that is due to allele-specific (“cis”) vs. distant (“trans”) genetic loci. We identify individual genes whose translation, but not mRNA abundance, is almost entirely regulated in cis or in trans. These data provide a starting point for further investigation of genetic influences on variation in posttranscriptional gene expression.

P-64
POLYMORPHISMS IN GYOXALASE I GENE CONSTITUTE TO ALTERED ENZYME ACTIVITY
Ilze Konrade1, Raitis Peculis2, Elina Skapare1-2, Janis Klovins2, Maija Dambrova1-4
1Riga East Clinical University Hospital, Riga, Latvia, 2Latvian Biomedical Research and Study Centre, Riga, Latvia, 3Latvian Institute of Organic Synthesis, Laboratory of Pharmaceutical Pharmacology, Riga, Latvia, 4Riga Stradins University, Faculty of Pharmacy, Riga, Latvia

Objectives: The glyoxalase system and mainly its bottleneck enzyme glyoxalase I protects cells from advanced glycation endproducts (AGEs), such as methylglyoxal (MG) and other reactive dicarbonyls. AGEs are partly responsible for harmful protein alteration in living cells, notably in neurons leading to their dysfunction and death. Therefore neuronal dysfunctions are so common diabetes complications due to elevated blood sugar that lead to high levels of AGES.

The aim of our study was to determine wether single nucleotide polymorphisms (SNP) or their haplotypes in gene of glyoxalase I (GLO1) enzyme influences its detoxification capability of MG and wether these SNP have impact on development and/or rate of diabetes neurological complications.

Methods: Total of 326 individuals were genotyped for 3 common SNPs (rs2736654 A111E, rs1130534 G124G and rs1049346 5’UTR) in GLO1 using minisequencing and subsequent MALDI-TOF mass spectrometry analysis. Glyoxalase I activity was determined in blood samples by spectrophotometry method.

Results: Our data revealed significant association of rs1130534 G124G and rs1049346 5’UTR with glyoxalase I activity decreased in minor allele homozygotes to 25.2±10.0 U/g Hb (p=0.001) (wild type 31.6±8.5 U/g Hb) and 27.2±8.0 U/g Hb (p=2.6*10^-5) (wild type 32.8±7.1 U/g Hb), respectively. Haplotype analysis supported this trend: carriers of both minor alleles in these markers had average enzyme activity 21.8±4.3 U/g Hb while homozygote wild type allele carriers had 33.0±7.0 U/g Hb. With other haplotypes showing result between these values depending on minor allele count in individuals.

Conclusion: SNPs in GLO1 can be used to predict enzyme activity and its detoxifying capability as our research provides evidence that the presence of the minor alleles at rs1049346 and rs1130534 are significantly associated with reduced enzyme activity.

P-65
NEW GENOMIC APPROACHES TO UNDERSTAND THE MYSTERY CHILDHOOD KAWASAKI DISEASE
Hariklia Eleftherohorinou1-2, Victoria Wright1, Lachlan J. Coin1, Michael Levin2
1Centre of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK, 2Section of Paediatrics, Division of Infectious Diseases, Department of Medicine, Imperial College London, London, UK

Background: Kawasaki disease (KD) is an inflammatory disorder of unknown aetiology that occurs almost exclusively in children and is the leading cause of acquired childhood heart disease in developed countries. Intravenous immunoglobulin is the only evidence-based treatment available, yet 25% do not respond (and 40% of untreated) and develop coronary artery aneurysms (CAA), which can lead to coronary artery thrombosis and death in 3% of cases.

Scope: A major challenge in KD research is to identify the basis of genetic susceptibility to CAA development and understand the pathogenesis of arterial wall damage that will ultimately enable targeted therapy for children at greatest risk.

Methods: We analyzed a GWAS on 600K SNPs of 1,000 cases/ 2,000 controls using our pathway-driven
methodology that assesses the effect of the cumulative genetic variation along a pathway on disease outcome. Over 2,500 pathways were examined, as acquired by public databases. The top ~100 were further explored using a novel gene stability selection methodology that identifies the most influential genes within associated pathways.

**Results:** Of initially >8,000 genes our method selected 25, the functional role of which was further examined using a series of bioinformatics tools. Interestingly, all 25 are involved in Cell signaling ($P<1 \times 10^{-38}$), with 16/25 having enriched toxicity function Cardiac arteriopathy ($P<3.3 \times 10^{-3}$) and 19/25 having a role in Cardiovascular disease ($P<3 \times 10^{-10}$). Microarray analysis of 110 paired acute and convalescent KD samples showed coordinated regulation at the gene expression level too. Further replication is underway using custom SNP typing of the identified genes on an independent study of 500 KD cases.

**Conclusion:** This is the first study to identify functionally interacting genetic factors that potentially control CAA formation. Our results demonstrate that our pathway-driven analysis of GWAS is a powerful tool likely to shed light in the mechanism of CAA development of the mysterious KD.

**P-66**

**THE IMPACT OF EPIGENOMIC VARIATION IN THE PREDICTION OF PHENOTYPES AND GENETIC MERIT**

*Oscar Gonzalez-Recio¹, Luis Alberto García Cortés¹, Beatriz Villanueva², Eva Ugarte²*

¹INIA, Madrid, Spain, ²NEIKER, Vitoria-Gasteiz, Alava, Spain

Today, at the dawn of the 21st century, another aspect of complex traits inheritance is being unveiled: epigenomic variation and epigenetic control mechanisms of DNA. The modern definition of epigenetics is the study of possibly heritable changes in gene expression that are not caused by changes in DNA sequence. One of the most known epigenetic markers is the methylation of DNA (DNAm). Recent advances in genome sequencing techniques have paved the way to obtain genome-wide epigenetic information. Genomic regions with large amount of DNAm tend to be less transcriptionally active, possibly having some impact on the phenotypes (e.g. productive traits or metabolic diseases).

In this work we evaluate the consequences of ignoring DNAm information in the genetic merit prediction of complex traits, and explore new methods to incorporate epigenomic information at predicting genetic merit (or risk) and phenotypes.

The DNA methylation status of 2104 bovine individuals genotyped with the Illumina Bovine50K chip were simulated along with their respective phenotypes. Twenty percent of phenotypes were masked to mimic a genome-wide prediction scenario with yet to be observed data. Common methods applied to genetic and genomic evaluations were compared with ad-hoc methods incorporating DNA methylation information. Twenty replicates were analyzed. The results showed that traditional genetic evaluations methods, such as BLUP, achieve 10% lower accuracy when epigenomic variability exists, whereas the predictive ability of genomic methods, such as GBLUP or Bayesian LASSO, decreased 12% in the presence of epigenomic variance. Ad-hoc methods incorporating DNAm information were used to capture epigenomic variation in the phenotype decomposition models. In the presence of epigenomic effects, these methods were more accurate at predicting genetic merit (0.69-0.75) and phenotypes (0.52-0.81) in both the training and validation samples than genomic models ignoring information from DNAm (0.67-0.74 in the training sample and 0.45-0.72 in the validation sample).
P-67
HIGH YIELDING RED RICE: GENES FROM THE WILD RICE TO CULTIVAR

Md. Atiqur Rahman Bhuiyan\textsuperscript{1}, Narimah Md. Kairudin\textsuperscript{1}, Abdul Rahim Harun\textsuperscript{2}, Kalluvettankuzhy K Sabu\textsuperscript{4}, Abdullah Md. Zain\textsuperscript{3}, Tilakavati Karupaiah\textsuperscript{1}, Wickneswari Ratnam\textsuperscript{1}
\textsuperscript{1}Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia, \textsuperscript{2}Malaysian Nuclear Agency, Bangi, Selangor, Malaysia, \textsuperscript{3}Universiti Malaysia Terengganu, Kuala Terengganu, Malaysia, \textsuperscript{4}Jawaharlal Nehru Tropical Botanic Garden & Research Institute (JNTBGRI), Palode, Thiruvananthapuram, India

A red rice transgressive variant was developed from a cross between \textit{Oryza rufipogon} Griff.(IRGC105491) and a Malaysian high yielding cultivar MR219. The field trials revealed the yield potentiality of the variant and showed significantly (p<0.05) higher (5.20 \text{tha}^{-1}) than the control, MR219 (4.53 \text{tha}^{-1}). Quantitative trait loci (QTLs) for agronomic traits were validated in the advanced generation. The yield of this variant was influenced by several QTLs viz. qGPL-1, qSPL-1-2, qSPL-8 and qYLD-4. The red pericarp colour from the wild rice was transferred to the transgressive variant and the pericarp colour stability was confirmed. The variant was screened and showed resistance against two pathotypes of blast disease (\textit{Magnaporthe oryzae}). Chromosome segment analysis confirmed the introgression of wild genes. In the evaluation of Glycaemic and Insuline response, the variant produced the lowest postprandial glycaemia (GI: 51) and insulinaemia (II: 39), which were comparatively lower than the control (white pericarp rice, GI: 86 and II: 68). It has been filed for IPR under the NPV Act in Malaysia with the preferred name of UKMRC9 (PVBT039/09).

P-68 – ABSTRACT WITHDRAWN

P-69
INVESTIGATION OF HUMAN BRAIN TISSUE PHENOTYPES AND THEIR RELATIONSHIP WITH NEUROLOGICAL DISORDERS

Amy Rowlatt\textsuperscript{1}, Chris Haley\textsuperscript{1,2}, Albert Tenesa\textsuperscript{1,2}
\textsuperscript{1}Roslin Institute at the University of Edinburgh, Edinburgh, Midlothian, UK, \textsuperscript{2}MRC Human Genetics Unit at the MRC Institute of Genetics and Molecular Medicine at the University of Edinburgh, Edinburgh, Midlothian, UK

Traditional genome-wide association studies, which use population based methods to investigate the correlation of single nucleotide polymorphisms (SNPs) with complex traits, have been successful in identifying multiple risk alleles for disease. However, for the majority of traits, this approach has explained only a small proportion of the estimated heritable variation and the underlying causative pathways and interactions. One approach that may provide insight into the biological interactions underlying disease is the use of intermediate phenotypes measured in the relevant tissue. In this study we select locations of the genome known to influence neurological health and use a regional heritability approach to partition the variance of two intermediate phenotypes, CpG methylation and mRNA abundance, into variance attributable to cis and trans acting genetic factors. For regions of the genome showing a significant contribution of cis acting SNPs to both methylation and mRNA abundance, we calculate the phenotypic, genetic and environmental correlations between the two traits. Additionally, we compute the extent of sharing of cis and trans acting regions within traits but across tissues. We conduct this analysis on samples taken post-mortem from four brain regions for 148 human samples that died from various causes. Our investigation provides both insight into the relationship between mRNA levels and methylation and an indication of the extent of shared and independent control of gene expression and methylation for genomic regions of medical importance.
NEW METHOD FOR EFFICIENT PHASING OF SMALL POPULATIONS USING HIDDEN MARKOV MODELS WITH A SCALAR PHASE VARIABLE

Carl Nettelblad
Uppsala University, Uppsala, Sweden

Objectives: Reconstructing phase can help in diverse genetic analysis settings, both in itself, and as a method for imputation of missing genotypes. One popular approach is to use Markov Chain Monte Carlo methodology to iteratively sample the haplotype distribution in each individual as mosaics of all other individuals, treating the phase in each position as a binary parameter. We make such methods deterministic and improve accuracy by analysing local as well as global convergence.

Methods: We propose a model where the phasing of individual genotypes is treated as a scalar ordering probability, with no use of pedigree data. This allows traditional, mostly deterministic, algorithms for Hidden Markov Model training to be used. Based on analysing the structure of the likelihood function, we introduce a specific gradient-descent based approach rather than general expectation-maximization, and explicitly included inversion operations, reducing the number of incorrect phase switches.

Results: In comparisons against existing state-of-the-art algorithms, including Beagle and MaCH, our implementation performs favourably. Specifically, we look at data from pre-existing simulated datasets for crosses with known phase (our method reduces the number of incorrect haplotype switches, sometimes by more than 90%), experimental data sampled from natural populations (our method increases average LD block length by more than a factor of 2), as well as simulated data based on human populations (similar results).

Conclusions: The use of a scalar variable for representing phase can improve phasing accuracy. In MCMC methods with a binary representation of phase, the presence of multiple highly genetically similar individuals (e.g. full and half siblings, or highly preserved haplotype blocks among more distantly related individuals), can result in optimization being caught in detrimental local minima. The introduction of a scalar variable improves rectifies this. Our proposed method also better lends itself to analysis regarding convergence.

ESTIMATION OF GENETIC PARAMETERS FOR METHANE INDICATOR TRAITS BASED ON MILK FATTY ACIDS IN DUAL PURPOSE BELGIAN BLUE CATTLE

Purna Bhadra Kandel1,2, Nicolas Gengler1, Helene Soyeurt1,3
1University of Liege, Gembloux Agro Bio-Tech, Animal Science Unit, Animal Breeding and Genetics Group, Gembloux, Belgium, 2ITN Marie Curie GreenHouseMilk Project, Edinburgh, UK, 3National Fund for Scientific Research (FNRS), Brussels, Belgium

Dairy production stands out for its large methane emission. Therefore, specific nutritional strategies are being applied to abate methane emission but very little information is available about the genetic variability of methane emission. Methane indicators using traits indirectly related to methane and easily recorded like the mid-infrared (MIR) prediction of fatty acid could be used to conduct genetic studies. MIR methane indicators used in this study were derived from published fatty acid based methane indicators using 597 calibration samples. Genetic parameters of these MIR indicators were estimated by single trait random regression test-day models from 13,389 records collected on 1602 Dual Purpose Belgium Blue cows in their first 3 lactations. For the published indicator showing the highest relationship ($R^2 = 0.88$) with Sulphur hexafluoride ($SF_6$) methane emission data, the average daily heritability was $0.25 \pm 0.06$, $0.25 \pm 0.07$ and $0.18 \pm 0.09$ for the first three lactations, respectively. Similarly, the heritability of lactation yield was $0.45 \pm 0.09$, $0.46 \pm 0.11$ and $0.24 \pm 0.14$. The sire genetic variability was $3.60$, $4.08$, $1.19$ kg² of methane for the first three lactations, respectively. The genetic difference between sires having daughters eructing the highest and the lowest methane content was $11.62$, $13.01$ and $5.98$ kg per lactation for
the first three parities. This study suggests that methane indicator traits can be predicted by MIR and the genetic variability of these traits seems to exist. Therefore, it also indicates there is genetic variability of methane content eructed by dairy cows. These first findings might open new opportunities for animal selection programs on methane emission.

P-72
ANALYSIS OF A LONG-TERM DIVERGENT SELECTION FOR BODY WEIGHT IN CHICKENS WITH AN EXPONENTIAL MODEL
Tina Flisar, Špela Malovrh, Dušan Tercic, Antonija Holcman, Milena Kovac
University of Ljubljana, Biotechnical Faculty, Deparment of Animal Science, Ljubljana, Slovenia

The aim of this study is to estimate genetic parameters for body weight in divergently selected lines of chickens with non-linear model. Chickens were divergently selected over 34 generations for high and low body weight (BW) at 8-wk of age. Estimates of body weight for each generation were provided by mixed model. For fitting generation means against generation or cumulative selection differential exponential model was used. Estimates of realized heritability over generations were derived from regression of response to cumulative selection differential. Data were analyzed by SAS, using GLM and NLIN procedures. Fitting the generation means in non-linear model was better for the low line. Coefficient of determination was 90.96% for females and 88.99% for males. In high line, R² was 58.10% for females and 63.00% for males. Estimates for selection limit in high line were 2598.4 g and 2144.1 g, for males and females, respectively. Despite good fitting, selection limit have not been accomplished in low line. In high line, half of the selection response was obtained after about 6 to 8 generations, and in low line, about 20 to 28 generations. The estimated realized heritability decreased over generations. Analyses revealed differences between lines. During experiment realized heritability declined from 0.24 to 0.02 in males and from 0.55 to 0.31 in females of the high line. The use of non-linear model is recommended for modeling the response of long-term selection in poultry because of the well fitting and informative parameter estimates.

P-73 – ABSTRACT WITHDRAWN

P-74
RETAINED PLACENTA IN MRIJ CATTLE IS AFFECTED BY HERITABLE AND NON-HERITABLE FACTORS
Lindert Benedictus¹, Ad Koets¹, Peter Eldik, van¹, Henri Heuven¹²
¹Utrecht University, Utrecht, The Netherlands, ²Wageningen University, Wageningen, The Netherlands

Failure of the timely expulsion of the fetal membranes, called retained placenta (RP), leads to reduced fertility, increased veterinary costs and reduced milk yields. Immune-mediated rejection of the fetal membranes by the maternal immune system plays an important role in the breakdown of the fetal–maternal attachment and RP can be linked to failure of immune-mediated detachment of the fetal membranes. Immune MHC 1 (Major Histocompatibility Complex type I) compatibility between calf and dam depends on whether the paternal MHC I haplotype inherited by the calf is compatible to the MHC I haplotypes of the dam. The objectives of this study were to concurrently look at the heritable and non-heritable genetic aspects of retained placenta and test the hypothesis that a higher coefficient of relationship between dam and calf increases the risk of RP in the dam. The average incidence of RP in 43661 calvings was 4.5%, ranging from 0% to 29.6% among half-sib groups in MRIJ cattle. The average pedigree based relationship between the sire and the maternal grandsire (CRs*mgs) was 0.05 and ranged from 0 to 1.04. Using a sire-maternal grandsire model the h² was estimated at 0.22(± 0.07) which is
comparable with estimates for other dual purpose breeds. There was a tendency for an effect of \( \text{CR}_{\text{v,mg}} \) on \( \text{RP} \) (\( p<0.1 \)). The \( \text{CR}_{\text{v,mg}} \) was used as an estimate for the CR between dam and calf, which is correlated with the probability of MHC I compatibility between dam and calf. The study shows that selection against \( \text{RP} \) is possible and that mating of related parents should be prevented.

P-75

MACHINE-VISION APPROACHES TO PHENOTYPING ARABIDOPSIS

Anyela V Camargo\(^1\), Dimitra Papadopoulou\(^2\), Zoi Spyropoulou\(^2\), Konstantinos Vlachonasios\(^2\), Alan P. Gay\(^1\), John H. Doonan\(^1\)

\(^1\)Aberystwyth University, Aberystwyth, UK, \(^2\)Aristotle University of Thessaloniki, Thessaloniki, Greece

With recent advances in genomics, phenotyping has become the major bottleneck in exploiting this information both in medicine and agriculture. Rosette shape and size is an attractive trait on which to develop and evaluate automated approaches for objective machine-assisted plant phenotyping. Shape and size describe how the plant covers the ground surface, affecting both photosynthetic potential and canopy closure. The latter affects a crop’s ability to suppress competition and, in dry climates, affects soil water conservation. We used a machine-vision based method to compare the rosette stage of 19 genetically variable natural accessions of \textit{Arabidopsis} as a pilot study. This group of 19 lines comprise the parents of a Multiparent Advanced Generation Inter-Cross (MAGIC) mapping population (Kover, Valdar et al. 2009). Images of the plants were taken at different time steps to capture their patterns of growth. Images were then segmented to separate rosettes from background and 20 features describing rosette shapes, such as Area, Compactness and Roundness, were extracted and analysed. We show that shape descriptors can be used to parameterise and determine significant differences between rosette shapes of different accessions of \textit{Arabidopsis}, follow their development through time and we discuss the scalable potential of this approach in relating phenotype to genotype in large mapping populations.

References


P-76

TOWARDS IN-VITRO GENETICS; A CASE STUDY OF SOCS2

Fernando Lopes Pinto\(^1\), Chloe Pass\(^2\), Muhammad Muddassir Ali\(^3\), Vicky MacRae\(^2\), Chris Haley\(^2\)\(^4\), Simon Horvat\(^5\), Colin Farquharson\(^2\), Dirk-Jan de Koning\(^1\)\(^2\)

\(^1\)Swedish University of Agricultural Sciences, Uppsala, Sweden, \(^2\)The Roslin Institute, Edinburgh, UK, \(^3\)University of Veterinary and Animal Sciences, Lahore, Pakistan, \(^4\)MRC Institute of Genetics and Molecular Medicine, Edinburgh, UK, \(^5\)University of Ljubljana, Domžale, Slovenia

The \textit{in vitro} study of gene expression, using both natural variation and induced perturbations, might provide a rapid and effective method for complex trait dissection. To further explore the potential of this approach we must compare gene expression signatures across various scenarios, such as \textit{in vivo} signatures of a known mutation versus the signature of the same mutation in an \textit{in vitro} sample.

In this study, we compared \textit{in vivo} and \textit{in vitro} gene expression data by both wild-type and Suppressor of Cytokine Signalling-2 knockout (\textit{Socs2 KO}) mice. \textit{Socs2} is a negative regulator of growth hormone (GH) signalling and endochondral bone growth via inhibition of the JAK/STAT pathway. This has been classically demonstrated by the overgrowth phenotype of the \textit{Socs2} KO mouse, which has normal systemic IGF-1 levels.

We bred ten litters of mice (1-day old): five litters with the wild-type allele and five KO litters. From
each littler one pooled sample of primary growth plate chondrocytes was used for direct RNA extraction while another sample was cultured for 7 days. Chondrocytes were chosen because of the role of Socs2 role in endochondral bone growth. The resulting 20 samples were tested for differential gene expression on Affymetrix gene chips. After detection of differentially expressed genes using GENSTAT, further, bioinformatic analyses were performed using BioLayout and DAVID. The gene expression signature of the Socs2 KO is dwarfed by the signature of the culturing process: At a threshold of $P < 0.001$ the Socs2 KO gives 52 (in vitro) and 42 (in vivo) differentially expressed genes, compared to nearly 14000 at the same threshold for the effect of culturing. Direct comparison of gene lists showed a very limited overlap: out of the 1500 most significant ($P < 0.05$) genes, only 128 show overlap between the in vitro and in vivo results. While the study was quite small and may have limited power to detect the full gene signature of the Socs2 KO, the results suggest that we have to be careful when substituting in vivo with in vitro approaches.

P-77 – ABSTRACT WITHDRAWN

P-78
PREDICTING THE FITNESS DECLINE AFTER POPULATION SHRINKAGE

Aurora García-Dorado
Universidad Complutense, Madrid, Spain

The evolution of mean fitness expected after population shrinkage is relevant to evolutionary genetics, as well as to the conservation of endangered populations. Here I study this process considering the deleterious alleles that segregate at low frequency in the ancestral population or arise by mutation. A main source of fitness decline is inbreeding depression. It is due to the expression of the load initially concealed in heterozygosis, i.e., to the ancestral inbreeding depression rate $\delta$. However, after shrinkage, inbreeding induces an increase of natural selection against the ancestral (partially) recessive deleterious alleles, which is known as “purge” and is due to the extra fitness disadvantage ($2\delta$) of homozygotes.

I define a purged inbreeding coefficient $g$ that is weighed by the reduction of the frequency of deleterious alleles caused by purging. This coefficient can be used instead of the classical inbreeding coefficient $f$ in order to account for the consequences of purging when predicting inbreeding depression. When the shrunk population has an effective size $N$, assuming $Nd > 1$, the purged inbreeding coefficient after $t$ generations can be predicted as $g_t \approx \left((1-1/2N)g_{t-1} + 1/2N\right)(1-2\delta f_t)$, or can be computed from genealogical information. This method improves a previous heuristic approximation published by the author, and shows how purging acts upon previously accumulated inbreeding, leading to an early fitness decay followed by some recovery. Furthermore, after shrinkage, the population transits towards a mutation selection drift (MSD) balance characterized by a smaller inbreeding depression rate ($\delta^*$). During this transition, standard (i.e., non-purging) selection cancels out the inbreeding depression ascribed to $\delta^*$, just as in the new MSD balance that will be eventually attained. Therefore, purge and inbreeding only operate upon the remaining $\delta-\delta^*$. Simulation results show that this approximation is remarkably accurate under a wide range of mutational models and population sizes.
P-79
MATERNAL GENETIC EFFECTS SET THE POTENTIAL FOR EVOLUTION IN RED SQUIRRELS

S.E. McFarlane1, J.C. Gorrell2, D.W. Coltman2, M.M. Humphries1, S. Boutin2, A.G. McAdam1
1University of Guelph, Guelph, Ontario, Canada, 2University of Alberta, Edmonton, Alberta, Canada, 3McGill University, Montreal, Quebec, Canada

Abstract can be found in the student symposium abstract section on page 63

P-80
USE OF PURGING FOR POPULATION MANAGEMENT IN CONSERVATION PROGRAMMES

Ángeles de Cara1, Beatriz Villanueva1, Miguel Toro2, Jesús Fernández1
1INIA (Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria), Madrid, Spain, 2Universidad Politécnica de Madrid, Madrid, Spain

Conservation programmes aim at maximising the probability of survival of the populations of interest. The goals in these programmes are to minimise the loss of genetic diversity, which allows populations to adapt to changes, and to reduce the increase in inbreeding. An optimal strategy to achieve these goals is using optimal contributions, that is, calculating the number of offspring every individual should contribute to the next generation to minimise global coancestry. However, this strategy may allow deleterious mutations to be maintained in the population, compromising the long-term viability of the population. In order to avoid this, optimal contributions can be combined with inbred matings, to expose and eliminate recessive deleterious mutations by natural selection in a process known as purging. Whether purging by inbred matings is efficient in conservation programmes depends on the balance between the loss of diversity, the increase in fitness and the reduction in mutational load. Although some studies have concluded that purged populations experience reduced inbreeding depression, others have found that purging had negative consequences in small populations because it increased the probability of extinction. In this study, we perform computer simulations to determine whether management of a population with inbred matings can help to improve its long-term viability while keeping reasonable levels of diversity. Our results are strongly dependent on the genetic architecture and the mutational model assumed. Using molecular coancestry in the management of the population maintains a larger genetic diversity but leads to a lower fitness than using genealogical coancestry.

P-81
FLAGELLIN PERCEPTION VARIES QUANTITATIVELY IN ARABIDOPSIS THALIANA AND ITS RELATIVES

M. Madlen Vetter1,2, Fei He1,6, Ilkka Kronholm1,4, Heidrun Häweker5, Matthieu Reymond1,3, Joy Bergelson2, Silke Robatzek1,5, Juliette de Meaux1,6
1Max-Planck-Institute for Plant Breeding Research, Cologne, Germany, 2University of Chicago, Department of Ecology and Evolution, Chicago, USA, 3INRA-Centre de Versailles-Grignon, Laboratoire de Génétique, Versailles, Cedex, 4Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, UK 5The Sainsbury Laboratory, Norwich, UK 6Institute for Evolution and Biodiversity, Münster, Germany

Much is known about the evolution of plant immunity components directed against specific pathogen strains: they show pervasive functional variation and have the potential to coevolve with pathogen populations. However, plants are effectively protected against most microbes by generalist immunity components that detect conserved pathogen-associated molecular patterns (PAMPs) and control the onset of PAMP-triggered immunity. In Arabidopsis thaliana, the receptor kinase FLAGELLIN SENSITIVE2 (FLS2) confers recognition of bacterial flagellin (flg22) and activates a manifold defense response. To decipher the evolution of this system, we performed functional assays across a large set of A. thaliana genotypes and Brassicaceae relatives. We reveal extensive variation in flg22 perception, most of which results from
changes in protein abundance. The observed variation correlates with both the severity of elicited defense responses and bacterial proliferation. We analyzed nucleotide variation segregating at FLS2 in *A. thaliana* and detected a pattern of variation suggestive of the rapid fixation of a novel adaptive allele. However, our study also shows that evolution at the receptor locus alone does not explain the evolution of flagellin perception; instead, components common to pathways downstream of PAMP perception likely contribute to the observed quantitative variation. Within and among close relatives, PAMP perception evolves quantitatively, which contrasts with the changes in recognition typically associated with the evolution of *R* genes.

P-82
SEXUAL CONFLICT OVER POST-REPRODUCTIVE LIFESPAN IN PRE-INDUSTRIAL HUMANS

*Sandra Bouwhuis*1,2, Jenni E. Pettay2, Virpi Lummaa2
1University of Oxford, Oxford, UK, 2University of Sheffield, Sheffield, UK

Nearly all contemporary human populations exhibit sexual dimorphism in lifespan, with women outliving men by on average 5 years. Such dimorphism may have resulted from sexually antagonistic selection, as recent studies have reported positive selection on female lifespan, but stabilising or negative selection on male lifespan in historical human populations. Here we assess whether selection on lifespan differed between men and women in a pre-industrial human population, and estimate sex-specific heritabilities and genetic correlations for lifespan (LS), post-reproductive lifespan (PRL) and lifetime reproductive success (LRS). Our analyses confirm sexually antagonistic selection for LS and PRL. At the genetic level, there was significant heritability for all traits in both sexes, as well as a strikingly similar and positive correlation between LS and LRS. The genetic correlation between PRL and LRS, however, differed between the sexes, being negative in men, but slightly positive in women. The intersexual genetic correlation for PRL was positive, and constrained the predicted response to selection by c. 500% in women, while allowing a near-perfect predicted response to selection in men. These results provide evidence for strong sexual conflict over PRL in humans, which should have genetically constrained the evolution of sexual dimorphism.

P-83
TEMPERATURE ADAPTATION IN RHYNCHOSPORIUM COMMUNE

*Tryggvi Stefansson*1, Yvonne Willi2, Bruce McDonald1
1ETH Zurich, Zurich, Switzerland, 2University of Neuchâtel, Neuchâtel, Switzerland

*Rhynchosporium commune* is a haploid ascomycete fungus and a major pathogen of barley (*Hordeum vulgare*). Despite the global economic importance of *R. commune*, little is known about its evolutionary ecology, including adaptation to different thermal regimes across populations from climatically diverse locations. We conducted common garden experiments with 126 genetically distinct isolates from 9 field populations to measure variation in growth rates at 12°C, 18°C and 22°C. Populations from colder climates with higher temperature variation grew faster at all three temperatures compared to populations from warmer climates. There was a strong positive correlation between the variance in mean annual temperature for the 9 locations and the mean population growth rates ($r^2 = 0.86$, $p<0.003$). We found a significant genotype-by-environment interaction between the two lower and relatively more benign temperatures and the more stressful temperature of 22°C. Population differentiation for growth rates at 18°C ($Q_{ST}$) was significantly lower than population differentiation at neutral microsatellite loci ($C_{ST}$) and not significantly different at 12°C, but at 22°C the $Q_{ST}$ was significantly higher, consistent with local adaptation for growth at higher temperatures.

This is one of the first studies in plant pathogenic fungi to use $Q_{ST}$ / $C_{ST}$ comparisons to evaluate the potential for local adaptation to different temperatures. We found that *R. commune* has a high potential to rapidly adapt its growth rate to different thermal environments. There was high genetic variation for
plasticity in thermal adaptation, especially in response to stress (22°C). We found that this globally distributed pathogen has adapted to local climatic conditions, not through a shift in temperature optimum, but rather by acquiring generally fast growth in cooler/variable climates or slow growth in warmer/constant climates. This latter finding implies that there may be costs associated with fast growth under warm/constant climates.

**P-84**
PARENT CONFLICT IN PLANTS REVEALED BY DIVERGENT INHERITANCE OF SEED SIZE IN SELFING AND OUTCROSSING POPULATIONS

Yvonne Willi
University of Neuchâtel, Neuchâtel, Switzerland

Outcrossing opens up a venue for several kinds of parental conflict. For example, when one sex provides parental care to offspring fertilized by several partners, the non-providing sex is under selection to maximally exploit the caring sex. Theory then predicts a co-evolutionary arms race between the sexes at loci affecting offspring provisioning. In plants, this conflict has been suggested to affect seed size. If true, there should be strong paternal genetic and higher-order genetic effects on seed size in outcrossing plant populations, whereas these effects should be absent in selfing populations. This hypothesis was tested by measuring seed size from within-population crosses (diallele cross) and between-population crosses including 4 outcrossing, 4 selfing and 1 mixed mating population of *Arabidopsis lyrata*. Paternal genetic and higher-order genetic interaction effects explained more phenotypic variance in outcrossing than in selfing populations, supporting the hypothesis that the conflict has been reduced under selfing. This conclusion was further supported by an increase in seed size when pollen came from an outcrossing population, compared to when the pollen came from a selfing or the same population. Still, paternal genetic and higher-order genetic effects were present in selfing populations, suggesting that genetic variation has not disappeared in the time since the transition to selfing.

**P-85**
USING GLMMS TO ESTIMATE GENETIC VARIABILITY OF LIFE-HISTORY REPRODUCTIVE TRAITS IN FOREST TREES. THE CASE OF TWO MEDITERRANEAN CONIFERS, PINUS PINASTER AND PINUS HALEPENSIS

Luis Santos-del-Blanco1,2, Ricardo Alía1,2, José Climent1,2
1CIFOR-INIA, Madrid, Spain, 2Sustainable Forest Management Research Institute, Palencia, Spain

Research of genetic variability of reproductive life-history traits of trees at the species level posits several research challenges related to the size and longevity of these organisms. Forest tree provenance and provenance-progeny trials planted by many forest services and institutes years ago, can now serve as a privileged tool to answer relevant questions in evolutionary ecology, provided that suitable statistical tools are used.

Here we focus on Maritime pine (*Pinus pinaster*) and Aleppo pine (*Pinus halepensis*), two monoecious conifer species of great ecological importance in the Western Mediterranean basin. We use an extensive dataset from multi-site common garden provenance-progeny experiments with range-wide populations for which male and female reproductive data are available for a series of years since reproductive onset. Animal and standard Generalised Linear Mixed Models are fitted by Bayesian methods (MCMCglmm package in R) in order to describe basic life history traits in these relevant species. Threshold size for male or female reproduction is modelled using the binomial family applied to data of presence/absence of reproductive structures, including vegetative size as a covariate. Reproductive allometry, indicative of reproductive allocation is modelled using the Poisson family applied to count data, including tree size as a covariate. Vegetative growth traits are analysed by Linear Mixed Models.
Our results support life history theory expectations, as within both species, early and heavy female reproduction is coupled with reduced growth across environmental gradients reflecting different growth conditions. On the other hand, male reproduction shows lower population differentiation and heritability compared to female reproduction.

P-86
ESTIMATES OF (CO)VARIANCE COMPONENTS FOR DIRECT AND MATERNAL EFFECTS ON SOME MEAT PRODUCTION TRAITS IN A PALAS MERINO LINE SHEEP

Cristina Lazar1, Mihail Alexandru Gras1, Rodica Stefania Pelmus2, Horia Grosu1, Gabriel Vicovan1,2
1National Research Development Institute for Animal Biology and Nutrition, Balotesti, Ilfov, Romania, 2Sheeps and Goats Breeding Research and Development Institute, Palas, Constanta, Romania

Genetic correlations are of the most interest in animal breeding field because of their magnitude depends the genetic improvement. The aim of this paper was to estimate the genetic parameters for some meat production traits in a Romanian meat Merino sheep line. The study was carried out on a number of 142 progenies of a Palas Merino line, belonging to 16 rams and 125 ewes. The following traits were analyzed: 1) body weight at lambing (BWL), at 30 (BW30), 70 (BW70), 150 days (BW150); 2) total weight gain between: 30 days and lambing (TWG L-30), 70 and 30 days (TWG 30-70), 150 and 70 days (TWG 70-150), 70 days and lambing (TWG L-70), 150 days and lambing (TWG L-150); 3) average daily weight gain between: 30 days and lambing (DWG L-30), 70 and 30 days (DWG 30-70), 150 and 70 days (DWG 70-150), 70 days and lambing (DWG L-70), 150 days and lambing (DWG L-150). Covariance components were estimated with REML method, using the Maternal Animal Model (Mrode, 2005) and were analyzed with MATLAB student program, (version 5.3, 1999). The heritability for these traits was: 0.28 for BWL, 0.16 for BW30, 0.26 for BW70, 0.34 for BW150, 0.19 for TWG and 0.16 DWG. Extremely high correlations were found between: BW 30 and DWG L-30 (0.9471); BW 70 and TWG L-70 (0.9897); BW 70 and DWG L-70 (0.9892); BW 150 and TWG L-150 (0.9917); BW 150 and DWG L-150 (0.9907).

P-87
EVOLUTION OF EVOLVABILITY UNDER FLUCTUATING SELECTION WITH MULTILINEAR EPISTASIS

Arnaud Le Rouzic1, José Álvarez-Castro2, Thomas F. Hansen2
1LEGS, CNRS, Gif sur Yvette, France, 2University of Santiago de Compostela - Department of Genetics, Lugo - Galiza, Spain, 2CEES, University of Oslo, Oslo, Norway

Theoretical and experimental quantitative genetics have evidenced the importance of genetic variation and variability to understanding and predicting evolution. The puzzling issue of the evolutionary properties of this genetic variance arose more recently. Indeed, the potential influence of natural selection on evolvability is still poorly understood, and involves mechanisms that were largely disregarded previously in evolutionary quantitative genetics, including epistasis. We investigated how evolvability changes based on a simple quantitative genetics model (including drift, selection, and mutations) in which epistasis results in a linear combination of marginal genetic effects -- the multilinear model. Both individual-based simulations and analytical results show that genetic architectures may evolve in a deterministic way under selection; stabilizing selection generates moderate canalization, while directional selection promotes canalization or decanalization, depending on the genetic architecture. Since long-term directional selection is unrealistic, the possibility of an adaptive increase in evolvability thus remains hypothetical. Nevertheless, we showed that all kinds of changes in genetic architectures, including deceased or increased evolvability, can be retrieved if selection is fluctuating. Our results suggest that the evolution of genetic architectures may depend on both the frequency and the size of evolutionary challenges encountered by a species in its recent history, and offer new keys to understanding why evolvability differs across traits, populations, and species.
P-88

CONTEXT-DEPENDENT GENETIC ARCHITECTURE IN WILD BIRD POPULATIONS

Anne Charmantier
CNRS, Montpellier, France

Any heritability estimate is specific to one trait measured in a particular population and a given environment. However, because of data limitations, quantitative genetics applied to wild populations has often underestimated how the genetic architecture of traits can change across space and time, but also can vary with the sex, age or life-stage of individuals. Although estimating such variation in genetic (co)variances and heritabilities in the wild is very challenging, it provides valuable insight into past selection processes that have shaped the traits and their variances. Using data collected in wild bird populations, I will show how long term studies in natural populations can address the question of context-specific genetic architecture, and the consequences of this variation on our perception of evolutionary potential. First, I will illustrate how age-dependent variances in a reproductive trait in the Mute swan Cygnus olor can contribute to testing evolutionary theories of senescence. Second, I will discuss how the genetic matrix of a set of morphological and life-history traits varies across habitats in Mediterranean Blue tit Cyanistes caeruleus populations. This case study will highlight how the stability of the genetic architecture can impose constraints to evolution. Finally, I will illustrate how canalization has shaped contrasting phenotypic and genetic variation in spring arrival date of common terns Sterna hirundo across their different reproductive stages. The declining additive genetic variance displayed by the phenological trait across life stages might here limit the evolutionary potential of migration timing in the context of climate change. Overall, these studies illustrate how differences in genetic architecture are shaped by selection pressures, and how they can impact future evolutionary responses.

P-89

ESTIMATION OF (CO)VARIANCE COMPONENTS FOR DIRECT AND MATERNAL EFFECTS ON BODY WEIGHT IN ROMANIAN TELEORMAN BLACK HEAD TSIGAI LAMBS

Rodica Stefania Pelmus, Gina Cecilia Pistol, Cristina Lazar, Horia Grosu, Mihail Alexandru Gras, Elena Chita, Ionelia Taranu
National Research-Development Institute for Animal Biology and Nutrition, Balotesti, Ilfov, Romania

The objective of this study was to estimate the variances as well as the covariances between direct and maternal additive genetic effects and to investigate the importance of these effects on body weight of lambs belonging to a Romanian local breed Teleorman Black Head Tsigai at different ages. Records of 162 lambs from 11 rams and 89 ewes were used in the study. Estimates of co(variance) components were obtained for the body weight at birth (BW), 30 days (BW30), weaning (BW60) and the average daily gain (ADG) from birth to 60 days. Genetic parameters were estimated for these traits using Restricted Maximum Likelihood (REML) method, with a model including maternal effects (Mrode 2005). Our results showed that the direct heritability for BW, BW30, BW60 and ADG was 0.248, 0.189, 0.280 and 0.451, respectively and the maternal heritability for BW, BW30, BW60 and ADG was 0.150, 0.115, 0.168, 0.274. It was concluded that for the increasing efficiency of selection for this local breed, inclusion of maternal effects as well as direct-maternal genetic covariance especially for early growth traits has to be taken into account.
P-90

POPULATION GENOMICS OF PARALLEL EVOLUTION IN THE MARINE SNAIL LITTORINA SAXATILIS

María Saura1,2, María José Rivas1, Emilio Rolán-Álvarez2, Humberto Quesada2
1Inia, Madrid, Spain, 2Universidad de Vigo, Vigo, Spain

Formation of new species has been widely studied in a variety of organisms with biological properties of special interest that give information on the ecological and geographical aspects leading to speciation. Traditionally, the best candidate organisms for this type of studies were those whose genomes are poorly known (non-model organisms).

The recent emergence of high throughput sequencing technologies has generated an important impact in the development of genomic approximation in non-model species. The most recent hypotheses on sympatric speciation have been focused on the genomic architecture of divergence, trying to elucidate aspects including the amount of variation shared among incipient species, how this variation is distributed along the genomes, and the effect of selection fingerprint along this process. The development of this kind of studies supposes an urgent need in order to elucidate the genetic architecture of reproductive isolation and the mechanism by which selection generates new species.

The marine snail Littorina saxatilis represent an excellent opportunity for the study of the genetic architecture of speciation. On the exposed marine rocky shore from Galicia (northwest Spain) two ecotypes coexist in sympatry, adapted to different environmental conditions and showing a certain degree of reproductive isolation. Previous studies have suggested that these ecotypes have been originated independently and in parallel.

We used the technology of microarray hybridization to elucidate the genomic patterns of variability and the differences in gene expression between ecotypes, using DNA and cDNA respectively. We sampled three different localities along the Galician coast.

Our preliminary results showed that differences among localities are higher than differences between ecotypes, thus supporting parallel evolution. The next step include to determine whether the loci differing between ecotypes at the genomic level are correlated to those differing at the expression level.

P-91

MODELS OF GENETIC EFFECTS WITH MULTIPLE ALLELES AND APPLICATIONS TO STUDY THE ACTION OF SELECTION

José M Álvarez-Castro1, Arnaud Le Rouzic2, Rong-Cai Yang3,4
1University of Santiago de Compostela, Lugo, Galiza, Spain, 2CNRS, Gif sur Yvette-Île de France, France, 3Alberta Agriculture and Rural Development, Edmonton-Alberta, Canada, 4University of Alberta, Edmonton-Alberta, Canada

We here present recent progress made in theoretical models of Quantitative Genetics and their applications to the field of Evolutionary Biology. First, we show our latest improvements in models of genetic effects—new extensions of the models for the decomposition of the genetic variance to more general situations, including multiple alleles with arbitrary departures from the Hardy-Weinberg equilibrium. Second, we show an application of these models to understanding the multiallelic human red cell acid phosphatase locus 1 (ACP1) polymorphism. Although a recent communication claims that one allele is deleterious and would thus eventually be removed by directional selection, we find the data available to support the hypothesis that this polymorphism is maintained by stabilizing selection. Finally, we perform simulations to show that our theoretical developments can be used in a broader range of situations than previous models to obtain correct predictions of the output of selection using the breeder’s equation.
P-92
SEXUALLY ANTAGONISTIC SELECTION ON THE TIMING AND RATE OF REPRODUCTION IN HUMANS

Elisabeth Bolund, Sandra Bouwhuis, Jenni Pettay, Virpi Lummaa
1Sheffield University, Sheffield, South Yorkshire, UK, 2University of Oxford, Oxford, Oxfordshire, UK, 3University of Turku, Turku, Finland

A trait can be beneficial in one sex but detrimental in the other, resulting in a sexual conflict and hence different selection pressures in the two sexes. Such sexually antagonistic selection has received tremendous attention in recent years and is now recognised as a significant evolutionary force across taxa. However, its role in human evolutionary studies has been largely ignored despite theoretical grounds for conflict over reproductive timing in men and women. We investigate if sexual conflict on the genetic level over key life history traits has been present historically in humans. We use extensive (n > 80,000) demographic pedigree data collected from church registers of 8 parishes in Finland. The dataset spans up to 10 generations across the demographic transition from natural high (5-6 offspring) to low (< 2 offspring) reproductive rate. The life-history events of all born individuals and pedigree members were followed, and accurate fitness data for both sexes recorded. Previous studies have established significant heritabilities as well as sex-specific phenotypic selection patterns of both the timing of reproduction during life and of age-specific rates of reproduction. To determine whether this reflects sexual conflict on the genetic level, we use the ‘animal model’ framework to estimate genetic variances of these traits and fitness, as well as cross-sex genetic correlations and sex-specific genetic selection pressures. We also investigate whether the conflict has changed with the transition to low reproductive rates in both sexes over the past 3-4 generations. Finally, we will highlight analytical obstacles in these types of studies along with possible solutions. This research has both field-specific and cross-disciplinary implications, furthering our understanding of both past and present evolutionary patterns and of sex differences in humans.

P-93
EVOLUTION OF BODY MASS CURVE IN FLOUR BEETLES: A FUNCTION-VALUED APPROACH

Patrick Carter, Kristen Irwin
Washington State University, Pullman, WA, USA

Objectives: Many traits of central importance in biology are function-valued (FV); that is, they can be described as a mathematical function of an independent index. We use FV methods to study the evolution of growth curves in larval red flour beetles Tribolium castaneum, an organism that can increase body mass 300 fold within 21 days. Previous studies demonstrated that the additive genetic variance-covariance (G) function explained most of the phenotypic variance in Tribolium body mass curves, and identified non-zero additive genetic covariances between mass and life history traits. Our objective in this study were to apply artificial selection along the entire body mass curve to test fundamental predictions from FV theory about evolution of FV traits under maximal selection, and to test for correlated responses in life-history traits.

Methods: Four replicate selection lines and 4 randomly bred replicate control lines were derived from the base population. A “selection index function” w_i = \int z_i(a) \beta(a) da
(where z_i(a) is mass as a function of age for individual i and \beta(a) is the selection gradient function of age) was used to identify breeders in the replicate selection lines. A maximum \beta(a) was used for 4 generations, and body mass and life history traits measured.

Results: A rapid response to selection of body mass at all ages was measured, and some life history traits showed evidence of a correlated response. However by generation 4 very high mortality rates among selected individuals was also measured.

Conclusions: Body mass curves responded to selection as predicted by FV theory; however, increased
mortality among selected lines was unexpected. A strong negative additive genetic covariance between body size and the life history trait of larval period is hypothesized to be one possible cause of this mortality.

P-94
BREEDING FOR EXTENDED LACTATION IN AUSTRALIAN DAIRY CATTLE

Mary Abdelsayed, Herman Raadsma, Peter Thomson
University of Sydney, Sydney, NSW, Australia

Introduction: Numerous changes in the dairy industry have taken place such as introduction of robotic milking and increased milk production which has led dairy producers to reassess if such traditional seasonal production systems are in fact the most optimal. An ongoing trend in the dairy industry is a shift to non-seasonal calving patterns providing scope to extending lactation beyond the traditional 305 days. Extending lactation has attested to increasing production and lactation efficiency through increased reproductive performance, decreased health issues associated with metabolic stresses around calving and early lactation, increased productive life of the cow, and increased profitability for the dairy producer. This project is aimed at examining the genetic and environmental variance components and obtaining estimated breeding values (EBVs) for extended lactation traits. In addition the project will assess the genetic covariance between extended lactation traits, persistency traits and other important milk and cow traits (survival, fertility indicator traits) which have not been estimated to date. Such findings will enable producers to select cows better suited to longer lactations and whether extended lactation traits should be in a breeding objective on its own or perhaps there needs to be a modification in the selection index to help producers maximise their profit from breeding.

Materials and Methods: Data obtained from ADHIS which includes ~158 million test day records from 1985 to 2010 from ~7 million cows and the extended lactation traits that will be looked at include milk yield, fat, protein, lactose percentage, Australian Selection Index (fat + protein-volume) and energy outflow of fat, protein and lactose as a measure of energy (MJ) per lactation. Extended lactation curves will be modelled using two methods, Wood model and random regression model to derive persistency and extended lactation traits that will be used in the genetic analysis. Genetic parameter, covariance estimates and estimated breeding values for these traits will be derived using linear mixed animal models using ASReml-R program.

P-95 – ABSTRACT WITHDRAWN

P-96
ALLELIC DIFFERENTIATION AS A DETERMINANT OF LONG TERM RESPONSE TO SELECTION FOR QUANTITATIVE TRAITS UNDER DIVERGENT SELECTION IN STRUCTURED POPULATIONS

Armando Caballero
Universidad de Vigo, Vigo, Spain

It is known from long time ago that whereas the short-term response to selection depends primarily on genetic variance and gene diversity (expected heterozygosity), the long-term response to selection depends on allelic diversity (the number of alleles available for selection in the population). It may be hypothesized that in a subdivided population subject to divergent selection for a given quantitative trait and different selection optima, allelic differentiation is also a main determinant for long-term response to selection. To show this, computer simulations were run assuming a subdivided population with a number of demes connected by migration. A quantitative trait subject to stabilising selection with different optima for the different subpopulations was assumed to be controlled by multiallelic loci with variable
gene effects. After the population reached a mutation-selection-drift equilibrium, a change in the optima was assumed to occur for the different subpopulations and the short (10 generations), medium (10 to 50 generations) and long-term (50 to 100 generations) response to selection ( adaptation to the new optima) was computed. By using a wide range of parameters in the simulations (variable subpopulation size, mutation and migration rate, and strength of stabilising selection), the relationship between the response to selection and the initial population diversity parameters could be established. A multiple regression analysis showed that the main determinant of long-term response to selection is the allelic differentiation between subpopulations ($A_{ST}$), rather than the quantitative genetic differentiation ($Q_{ST}$) or the gene frequency differentiation ($G_{ST}$).

**P-97**

**GENETIC CONSTRAINTS ON ADAPTATION TO A CHANGING ENVIRONMENT**

*Luis-Miguel Chevin*

CNRS, Montpellier, France

Genetic correlations between traits can constrain responses to natural selection. To what extent this reduces the rate of adaptation depends on the pattern of directional selection. I will present results for the average rate of fitness increase (or evolvability) under randomly changing selection gradients, allowing some phenotypic directions to be selected more often, or more strongly, than others. With an arbitrary covariance matrix of directional selection gradients, the average rate of adaptation depends on genetic correlations between traits, contrary to the isotropic case investigated in previous studies. Adaptation may be faster on average with more genetic correlation between traits, if these traits are selected to change jointly more often than the average pair of traits. However the rate of adaptation is not what is maximized by natural selection. Instead, the genetic covariance $G$ matrix that is favored in the long run is the one that maximizes long-term fitness. I will provide the average lag load caused by deviation of the mean phenotype from an optimum, under several forms of environmental changes typically experienced by natural populations, both stochastic and deterministic. This produces simple formulas for how the $G$ matrix affects long-term fitness, and I will discuss how their parameters can be estimated empirically.

**P-98 – ABSTRACT WITHDRAWN**

**P-99**

**GENETIC CORRELATIONS AMONG ANTAGONISTIC TRAITS IN FIGHTING COWS**

*Cristina Sartori, Roberto Mantovani*

University of Padua, Legnaro (PD), Italy

A dominant status in a female is often accompanied by an over-expression of some masculine traits, mainly due to changes in hormone balance. Moving from the phenotypic level to the genotype, it is likely to assume that “good genes” for fighting and thus useful to acquire a dominant status may possibly be accompanied by “bad genes” for feminine traits and therefore for fitness. Following these assumptions, the fighting ability, i.e., the capability to win a contest and important in achieving a dominant status, was assessed and linked to some traits either indicators of a “masculine” or a “feminine” phenotype. Fighting ability was measured in Aosta Chestnut and Aosta Black Pied breeds ($Bos taurus$), two Alpine rustic cattle traditionally involved in bloodless tournaments in which pairs of cows fight to assess dominance. The trait was scored as the number of victories achieved by a cow after fought against various opponents during one day of battles, and adjusted for tournament size. Information of more than 25,000 fighting results got by 12,000 cows in ten years of battles were related to the respective milk lactation yields and to individual morphological traits scored linearly after first calving. Data were analysed via REML method in a series
of bi-variate animal models to estimate the genetic correlations between traits. Fighting ability revealed medium-low negative genetic correlations of about -22% with some feminine traits as the milk, fat and protein lactation yields, the udder width, and the thinness. Conversely, low positive genetic correlations, around 12%, were obtained with more masculine traits as the front muscularity, and the thorax depth. These results suggest that selective pressures for female competitive success and dominant position, lead to an evolutionary trade-off because of the contemporarily action on antagonistic traits.

P-100
USING RNA-SEQ DATA TO IDENTIFY GENETIC VARIATION ASSOCIATED WITH LIFE HISTORY TRAITS IN THE GLANVILLE FRITILLARY BUTTERFLY

Virpi Ahola1, Patrik Koskinen2, Jouni Kvist2, Marjo Saastamoinen1, Petri Auvinen1, Lars Paulin2, Rainer Lehtonen1, Ilkka Hanski1
1University of Helsinki, Helsinki, Finland, 2Institute of Biotechnology, Helsinki, Finland

We have used next-generation RNA sequencing to identify candidate genes under putative selection in the Glanville fritillary butterfly (Melitaea cinxia), an ecological model species for which genomic resources are still limited. We sequenced pooled samples from three regions, the Åland Islands in Finland, China and France. The well-studied Åland Islands metapopulation was colonized probably 200 to 400 years ago, and the population is assumed to have become adapted to new climatic conditions and a highly fragmented habitat. We searched for genes in which variation has been significantly reduced in the Åland Islands in comparison with the other regions. We used gene set enrichment analysis for gene ontologies (GO) to assess whether genes in particular functional categories were overrepresented in the set of genes with reduced variation in the Åland Islands; the genes belonging to the overrepresented GO groups were hypothesized to have been under selection. Among the overrepresented GO groups, genes belonging to serine proteases, serine protease inhibitors and cuticular proteins were selected as candidates for an association study involving larval development in the Åland Islands metapopulation. We used selective genotyping to increase the power to detect small effects and effects associated with uncommon alleles. The data were analysed using single SNP and Bayesian multiple SNP approaches. The results suggest association of SNP variation in three serine proteases with larval growth, and reveal interactions with sex and rearing temperature.

P-101
QUANTIFYING SELECTION IN THE EVOLUTION OF INFLUENZA

Chris Illingworth, Ville Mustonen
Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK

Quantifying the selection acting on observed mutations is a central question in scenarios from the progression of healthy cells towards cancer to the development of drug resistance in pathogens. The rate of genetic changes in these scenarios can be rapid, with multiple interactions between genetic loci, such that the background upon which a mutation arises can be critical for its eventual fate. We here describe a multi-locus model of evolution, allowing for the inference of selection effects against a complex genetic background. We apply our method to time-resolved data describing the evolution of the haemagglutinin gene of H3N2 influenza since 1968. We demonstrate the effects of clonal interference on the fate of mutations, quantify the effect of deleterious mutations on the evolution of the virus, and discuss the potential to predict future changes in the viral population on the basis of genetic data alone.
P-102
SOME EVOLUTIONARY CONSEQUENCES OF EPISTASIS

James Cheverud
Washington University, St. Louis, Missouri, USA

Epistasis occurs when the genotypic values at one locus vary depending on the genotypes present at another locus. The importance of epistasis for evolution has sometimes been controversial. We will review models of epistasis’ role in populations undergoing genetic drift and directional selection along with experimental analyses of how epistasis interacts with these population processes through its role in the evolution of gene effects themselves. Specific examples include the evolution of genetic background through drift, altering both the magnitude and direction of genetic effects caused by newly introduced genetic variants, and its role in re-shaping heritable variation patterns under directional selection to facilitate adaptation. Under genetic drift, epistasis causes subpopulations to differentiate not only in their average values but also in their genetic backgrounds. New genotypes introduced into these populations can have drastically altered effects. These effects are demonstrated experimentally using subpopulations formed by inbreeding from the intercross of two inbred mouse strains, LG/J and SM/J. The LGXSM quasi-inbred recombinant lines were outcrossed to four standard inbred mouse strains and the relative effects of the outcross examined for variation. Up to 30% of the variance in body size can be attributed to these genetic interactions.

With epistasis, directional selection can alter patterns of heritable variation so that they display increased variation in the dimension being selected. Because of epistasis, there is genetic variation in the relationships (variances/covariances) between traits. These variations map to so-called relationship QTLs (rQTL). There may or may not be any differences in genotypic means for the traits at these locations but the regression of one trait on the other is altered. Under directional selection, the genotype-specific regression slope most closely aligned with the direction of selection is favored and increases in frequency. Examples from skeletal biology illustrating this model will be presented.

P-103
ANIMAL MODELS CONSTRUCTED WITH HYBRID RELATEDNESS MATRIX VS INCOMPLETE PEDIGREE: AN EXPERIMENTAL AND SIMULATION APPROACH

Julie Gauzere1, Sylvie Oddou-Muratorio1, Etienne Klein1, Laurène Gay2
1INRA, Avignon, France, 2INRA, Montpellier, France

The use of animal model to estimate quantitative genetics parameters in wild populations is often hampered by the lack of pedigree information on sampled individual. Molecular markers provide an alternative way to estimate the genetic relationships among individuals and to obtain a relatedness matrix to construct the animal model. However, for some species phenotypic information are still easier to obtain than molecular information. This study aims at investigating the efforts that should be allocated to obtain respectively phenotypic and molecular information in quantitative genetic studies in wild population.

In a common garden experiment of 60 maternal progenies (100 individuals per family) from 3 wild populations of Fagus sylvatica in the South East of France, we measured several putatively adaptative and performance traits on 6000 seedlings: height and diameter increment and budburst phenology. Furthermore we genotyped at 13 microsatellite markers a subset of 625 offspring and all their putative fathers, and we used paternity analysis to identify full-sibs within progenies and paternal half-sibs from different progenies. The additive variance (VA) and heritability (h²) of the traits were then estimated: i) using the 6000 phenotyped seedlings, and assuming that they were all half-sibs, ii) using only the sub-sample of 625 individuals, still assuming that they were all half-sibs, iii) using the same sub-sample and accounting for the paternal relatedness between individuals.

Accounting for complex kinship structure markedly changed VA and h² estimates: for the different studied traits, the estimates of h² varied between 0.22 and 0.45 when using paternal information, versus
h² between 0.39 and 0.63 assuming only maternal half-sibs. A simulation analysis was developed to investigate the relevance of constructing a « hybrid » relatedness matrix incorporating several sources of available information (incomplete pedigree, paternity assignments, relatedness estimated from markers) instead of using simplifying assumptions.

P-104 – ABSTRACT WITHDRAWN

P-105
EVALUATING UNCERTAINTY IN PREDICTED RESPONSE TO SELECTION FOR MULTIVARIATE TRAITS

John Stinchcombe1, Anna Simonsen1, Mark Blows2
1University of Toronto, Toronto, ON, Canada, 2University of Queensland, St. Lucia, QLD, Australia

Objectives: Estimating uncertainty in predicted responses to selection for multivariate traits remains a major challenge for two reasons: first, hypothesis testing about predicted responses to selection are difficult, and second, environmental covariances between traits and fitness may inflate our estimates of selection. Here we develop and test methods for addressing these challenges.

Methods: We analyze a quantitative genetic field experiment with Bayesian methods, to compare predicted responses to selection between treatments and test for environmental covariances between traits and fitness.

Results and Conclusions: Environmental covariances between traits and fitness deflated estimates of natural selection, tending to underestimate its strength in this experiment. Predicted responses to selection did not differ between experimental treatments, although utilizing phenotypic data instead of genetic data resulted in markedly different estimates. More generally, the methods we develop can be used to test hypotheses about G, Beta, and delta-Z.

P-106
THE GENETICS OF HYBRID INCOMPATIBILTY IN A DIVERSE MOUSE PANEL

David Aylor1, Fanny Odet1, Timothy Bell1, Darla Miller1, Elissa Chesler2, Deborah O’Brien1, Fernando Pardo-Manuel de Villena1
1University of North Carolina, Chapel Hill, NC, USA, 2The Jackson Laboratory, Bar Harbor, MA, USA

The Collaborative Cross project is creating a large multi-parental recombinant inbred panel of laboratory mice. Up to three-quarters of incipient lines cease to produce offspring during inbreeding. Since these mice are descended from highly diverse inbred strains, we expect this observation to be explained by Dobzhansky-Muller epistasis. To test this hypothesis and assess the mechanism of these incompatibilities, we measured fertility and reproductive parameters in male mice from over 350 independent extinct lines. Our preliminary studies indicate that male infertility is the mechanism for a majority of this extinction. We will present an analysis of these phenotypes, genotype-phenotype associations, and the genome architecture of the extinct lines.
P-108
MAXIMIZING GENETIC DIVERSITY UNDER STABILIZING SELECTION IN SUBDIVIDED POPULATIONS
Enrique Sánchez-Molano¹, Armando Caballero Rúa², Jesús Fernández Martín³
¹The Roslin Institute, Edinburgh, UK, ²Universidad de Vigo, Vigo, Spain, ³INIA, Madrid, Spain
In conservation programs, endangered species tend to be subdivided into several subpopulations maintained with few individuals. Their management implies the control of genetic diversity and differentiation among subpopulations, as well as the avoidance of inbreeding within demes through the maintenance of a certain degree of gene flow. A one-migrant-per-generation strategy has been commonly accepted for this purpose, but a method recently proposed by Fernández et al. (2008) has been proven to be more efficient. This dynamic method accounts for the genetic structure of the population, restricting inbreeding by determining both the optimal contributions for each individual and the number of migrants and subpopulations involved in the exchanges each generation. However, individuals may be adapted to different environmental conditions in each deme and, therefore, their survival probability could be reduced if they are moved to a new one. This consideration raises doubts about the performance of the dynamic method, as the death of migrants would imply a small or null gene flow and, consequently, would impede the method to effectively control inbreeding. Computer simulations have been carried out to test the performance of the dynamic method considering the existence of a quantitative trait under stabilizing selection with a different optimum in each subpopulation. Different selection intensities and heritabilities for the trait have been tested in a wide range of reasonable scenarios, comparing the obtained results with other procedures such as one-migrant-per-generation or maintaining isolated subpopulations. As expected, the stronger the selection pressure, the worse the performance of the dynamic method. Notwithstanding, the dynamic method is still more efficient than the one-migrant-per-generation strategy except for very strong selection intensities, which are highly unlikely to be found.

P-109
ANTAGONISTIC SELECTION PRESSURES ACTING ON HATCHING ASYNCHRONY IN A SIBLICIDAL SEABIRD
Emily Tompkins¹, David J. Anderson¹, Katharina Foerster²
¹Wake Forest University, Winston-Salem, NC, USA, ²University of Tübingen, Tübingen, Baden-Württemberg, Germany
Abstract can be found in the student symposium abstract section on page 63

P-110
FITNESS COMPONENTS DIFFER BETWEEN DROSOPHILA SUBOBSCURA INTER- AND INTRAPOPULATION HYBRIDS UNDER LEAD POLLUTION
Marina Stamenkovic-Radak¹, Marija Tanaskovic¹, Bojan Kenig², Marko Andjelkovic³
¹University of Belgrade, Faculty of Biology, Belgrade, Serbia, ²University of Belgrade, Institute of Biological Research, Belgrade, Serbia, ³Serbian Academy of Sciences and Arts, Belgrade, Serbia
Lead is one of the most present environmental contaminants and populations of different organisms respond differently to such stress. If combined with genomic stress, such as inbreeding and/or hybridization, lead as environmental stress may act synergistically which may cause significant decrease in fitness components and deterioration of population survival. Thus, the consequences of heavy metal pollution and habitat fragmentation on genetic structure should preferably be studied
in combination rather than separately. The possible effect of interpopulation hybridization is either outbreeding depression, as a consequence of breakdown of coadapted gene complexes, or the increased heterozygosity, which can reduce the negative effect.

In the present study, *Drosophila subobscura* originating from two ecologically different habitats, were used. One is sampled from highly polluted urban part of Belgrade, Serbia (Botanical garden) with a specific microclimate and under constant anthropogenic influence. Another is sampled from Sicevo gorge (Central Balkan), a natural unpolluted resort with rather stable microclimatic conditions. Intra and interpopulation crossings between isofemale lines of these populations were made. The progeny was reared at lead concentration of 200 μg/mL and fitness components (viability and developmental time) were monitored.

Results suggest that response to lead contamination is population specific. The increased adaptive value of interpopulation hybrids under stressful environment would suggest that heterozygosity hypothesis prevails as a mechanism for maintaining genetic homeostasis in populations under stress.

P-111
AN EMPIRICAL ESTIMATE OF THE EFFECTIVE PURGING COEFFICIENT

*Diego Bersabé, Aurora García-Dorado*
*Universidad Complutense de Madrid, Madrid, Spain*

The consequences of inbreeding depression depend upon the strength of purging selection against deleterious recessive alleles. Recently, a theoretical model has been developed to predict the combined effects of inbreeding and purging on fitness depression, in which the consequences of purging rely upon an “effective purging coefficient ($d$)”. We have studied the effectiveness of purging upon egg-to-pupae viability ($EPV$) after a drastic reduction in size starting from a natural population of *Drosophila melanogaster*. For this purpose, we first obtained estimates of the inbreeding depression rate ($\delta$) ascribed to recessive lethals and to non-lethal deleterious alleles for $EPV$ in the base population. Then, two different sets of lines were founded from this base population and were maintained through single mating with different constant size ($N=6; N=12$) throughout the rest of the experiment. Mean $EPV$ was assayed in both reduced populations at different generations and was used to obtain empirical estimates of $d$.

Our results show that the observed $EPV$ declines in these inbred lines were considerable smaller than the corresponding neutral predictions, which should be ascribed to purging. About 36% of $\delta$ was due to segregating recessive lethals with $d_r \approx 0.5$, and the remaining 64% was attributable to non-lethal deleterious alleles with $d_e \approx 0.05$. We also obtained an estimate of $d_e$ of about 0.11 which roughly accounts for the pooled inbreeding depression from lethal and non-lethal deleterious alleles. Finally, we discuss a number of potential sources of bias, which suggest that the actual values of $d_e$ might be larger than our current estimates, thus implying an even greater efficiency of purging.

P-112
THE HISTORY OF MAIZE DOMESTICATION AND ADAPTATION AS REVEALED BY A GENOME-WIDE SURVEY OF SNP AND HAPLOTYPE VARIATION

*Xiaohong Yang1, Huihui Li2, Lin Li1, Warburton Marilyn3, Suketoshi Taba4, Weiwei Wen4, Jiansheng Li1, Jianbing Yan1*

1China Agricultural University, Beijing, China, 2Chinese Academy of Agricultural Sciences, Beijing, China, 3USDA ARS Corn Host Plant Resistance Research Unit, Mississippi, USA, 4International Maize and Wheat Improvement Center, Mexico, Mexico

Modern cultivated maize (*Zea mays* L. ssp. *mays*) has been heavily selected during domestication and adaptation. To better understand these processes, we conducted a genome-wide survey of 982 maize inbred lines and 190 teosinte accessions using over 40,000 SNP markers. Population structure, principal
component analysis, and phylogenetic trees all consistently reflected historical evolutionary relationships among Zea species and subspecies. Shared haplotype analysis showed similar high levels of gene flow from Z. mays ssp. parviglumis and ssp. mexicana, confirming the critical contribution of ssp. mexicana to the maize gene pool. Scans for selection signatures identified 5.0% of the SNPs that have experienced selection during domestication, and 2.5% of the SNPs that have experienced selection during adaptation; these include hundreds of genomic regions in strong linkage disequilibrium (LD) due to probable selection. Four patterns of LD formation and decay were observed during maize domestication and adaptation. The characterization and relative magnitude of the selection pattern reflected by each of the four LD block types in the maize genome are consistent with maize domestication and adaptation history.

P-113
THE ACTION OF STABILIZING SELECTION, MUTATION AND DRIFT ON EPISTATIC QUANTITATIVE TRAITS

Victoria Avila1, Andrés Pérez-Figueroa1, Armando Caballero1, Aurora García-Dorado2, Carlos López-Fanjul3
1Universidad de Vigo, Vigo, Spain, 2Universidad Complutense, Madrid, Spain, 3University of Edinburgh, Edinburgh, UK

We simulated the effect of successive bottlenecks on the redistribution of the genetic variance of a quantitative trait in lines derived from populations at the mutation-selection-drift balance. We assumed 0.01 new mutations affecting the trait per gamete and generation and that same rate for neutral marker loci. Mutational effects on the trait followed a reflected exponential distribution accounting for a mutational variance about $10^{-3}$ (relative to the environmental variance), and they were assigned a random degree of dominance whose expected value exponentially decreased as the corresponding homozygous effect increased. Epistatic effects were assigned to a fraction of pairs of segregating loci. Cases of weak and strong epistasis (reinforcing or diminishing) or no epistasis were considered. Real stabilizing selection at different intensities (or no selection) was imposed upon the metric trait, and bottlenecking at different sizes was reiterated until an inbreeding coefficient $F$ very close to one was reached. To obtain accurate predictions for the evolution of the variance components of a neutral trait, $F_{st}$ estimates from neutral markers should be preferred to $F$ values expected from population size. For a neutral trait, the three components of the genetic variance (additive, dominant and epistatic) initially increased until a critical $F_{st}$ value was reached, then subsequently decreased to an equilibrium value. Due to non-additive gene action, $F_{st}$ exceeded the quantitative index of population divergence $Q_{st}$ for a wide range of $F_{st}$ values. On the whole, stabilizing selection did not essentially modify the neutral results obtained for the change of the genetic variance components after bottlenecks. Moreover, convergent selection blocked the change of the between-line variance with inbreeding and, consequently, $Q_{st}$ was again substantially smaller than $F_{st}$.

P-114
MAPPING THE GENETIC BASIS OF ECOLOGICALLY IMPORTANT TRAITS IN NICOTIANA ATTENUATA (SOLANACEAE) WITH ADVANCED INTERCROSS-RECOMBINANT INBRED LINES

Shuqing Xu, Celia Diezel, Ian Baldwin
Department of Molecular Ecology, Max-Planck Institute of Chemical Ecology, Jena, Germany

Identifying genetic basis of ecologically important traits – traits that are associated with organisms’ fitness in natural environments – is a key step for understanding how adaptation proceeds. Nicotiana attenuata, a diploid annual tobacco plant ($2n=2x=24$) native to Western North America, initiates growth in a post-fire environment in response to water-soluble chemical cues in wood smoke. Because of this ‘fire-chasing’ behavior, plant-insect community has to re-establish itself with new populations, generating highly unpredictable biotic environmental selective pressures. In order to deal with such dynamic selective pressures, N. attenuata evolved sophisticated herbivore induced defense strategies and remarkable
To better understand the genetic basis underlying these ecologically important traits, we are establishing advanced intercross-recombinant inbred lines (AI-RIL) by using two accessions that differ in many phenotypic traits, such as seeds germination, herbivore induced defense, green leaf volatiles, floral nectar volume. To increase the mapping resolution, we will perform four generations of intercrossing followed by six generations of self-fertilization to create the AI-RILs. Since one of parental accessions has been chosen for whole genome sequencing, we will perform low coverage genome re-sequencing on the other parental line to obtain high-density markers. The established AI-RIL population will be genotyped by restriction site associated DNA sequencing (RAD-seq). Traits that vary between parental lines will be phenotyped in AI-RIL population for genetic mapping. This AI-RIL population will provide an excellent resource for high precision mapping of ecologically important traits in the ecological model plant *N. attenuata*.

**P-115**

**CHARACTERIZATION OF GENETIC VARIATION AMONG SIX VIETNAMESE INDIGENOUS CATTLE POPULATIONS**

Doan Lan Pham¹, Ngoc Duy Do¹², Trong Binh Nguyen¹, Van Ba Nguyen¹, Quang Nam Le¹, Thi Thu Thuy Tran¹, Xuan Hoan Tran¹, Chi Cuong Vu¹, N Kadarmideen Haja²

¹National Institute of Animal Science, Hanoi, Viet Nam, ²University of Copenhagen, Faculty of Life Sciences, Department of Animal and Veterinary Basic Science, Division of Genetics & Bioinformatics, Copenhagen, Denmark

Cattle play very important role in agriculture and food security in developing world. It is important to assess cattle genetic diversity for optimal breeding and conservation strategies. In Vietnam, cattle have not yet been molecularly characterized. The aim of this study was to evaluate genetic diversity and genetic structure of Vietnamese indigenous cattle populations, using 27 microsatellite markers. For this purpose, a total of 410 individuals from six indigenous cattle populations (Ha Giang, Lang Son, U Dau Rui, Nghe An, Phu Yen and Thanh Hoa) and an exotic breed (Brahman) was characterized. A total of 362 alleles was detected and an average number of alleles per locus ranged from 8 (INRA005 and ILSTS005) to 17 (ETH185). High level of gene diversity (He) indicated by mean number of expected heterozygote (He) among seven cattle populations across 27 loci was 0.73. These populations expressed moderate level of inbreeding (mean Fis = 0.05) and genetic differentiations (mean Fst = 0.04). Phylogenic tree based on Nei’s (1972) genetic distances among seven populations reflected geographic distances. Individual assignment based on the software, Structure 2.1, indicated five possible homogeneous clusters. The Brahman, Lang Son, Ha Giang and U Dau Rui cattle were assigned independent clusters while Nghe An, Thanh Hoa and Phu Yen cattle were grouped in a cluster. We conclude that Vietnamese indigenous cattle have high level of genetic diversity and distinct genetic structures. Based on these results, we recommend that conservation program could group homogenous populations (Nghe An, Thanh Hoa and Phu Yen) to reduce costs and a cross between Lang Son and Ha Giang populations be made to increase heterozygosity.
MAINTENANCE OF GENETIC VARIATION IN HUMAN PERSONALITY: TESTING EVOLUTIONARY MODELS BY ESTIMATING HERITABILITY DUE TO COMMON CAUSAL VARIANTS AND INVESTIGATING THE EFFECT OF DISTANT INBREEDING

Karin Verweij¹,², Brendan Zietsch¹,²
¹Queensland Institute of Medical Research, Brisbane, Queensland, Australia, ²University of Queensland, Brisbane, Queensland, Australia

Personality traits are basic dimensions of behavioural variation, and twin, family, and adoption studies show that around 30% of the between-individual variation is due to genetic variation. There is rapidly-growing interest in understanding the evolutionary basis of this genetic variation. Several evolutionary mechanisms could explain how genetic variation is maintained in traits, and each of these makes predictions in terms of the relative contribution of rare and common genetic variants to personality variation, the magnitude of nonadditive genetic influences, and whether personality is affected by inbreeding. Using genome-wide SNP data from >8,000 individuals, we estimated that little variation in the Cloninger personality dimensions (7.2% on average) is due to the combined effect of common, additive genetic variants across the genome, suggesting that most heritable variation in personality is due to rare variant effects and/or a combination of dominance and epistasis. Furthermore, higher levels of inbreeding were associated with less socially-desirable personality trait levels in three of the four personality dimensions. These findings are consistent with genetic variation in personality traits having been maintained by mutation-selection balance.

USING SELECTIVE SWEEPS TO MAP QTLS

Leopold Parts¹,², David Knowles³,⁴, Gianni Liti⁵, John Winn⁴, Richard Durbin¹
¹Wellcome Trust Sanger Institute, Hinxton, UK, ²University of Toronto, Toronto, Canada, ³University of Cambridge, Cambridge, UK, ⁴Microsoft Research, Cambridge, UK, ⁵University of Nice, Nice, France

When a large, genetically diverse population is subject to selection, genetic variants that are beneficial for survival and propagation rise in frequency. This idea can be used to map quantitative trait loci (QTLs) by measuring the changes in allele frequencies from pooled population DNA before and during selection, and detecting selective sweeps. However, even if only a single sequence variant provides the fitness advantage, linked alleles will also increase in frequency, “hitchhiking” together with the driver variant. While QTLs can be mapped fairly easily by finding the regions under selection, narrowing them down to the underlying genes and alleles responsible for the sweep is not straightforward. Here, we present a Bayesian approach to prioritise candidate causative genes in the selected regions by combining data from allele frequency changes with prior information on the functional impact of each genetic variant. We demonstrate on simulations that our approach is more accurate in inferring population allele frequencies compared to common smoothing approaches, is able to narrow down the QTL intervals, and identify the driver alleles. We then apply our approach on data from a very large segregating yeast population grown under high temperature stress, and obtain candidate causative genes and alleles for the detected QTL regions. Finally, we compare this mapping approach with traditional linkage and association methods, and derive guidelines for experimental design.
P-118
IDENTIFICATION OF X-LINKED MOLECULAR VARIANTS AFFECTING COLD STRESS RESISTANCE IN DROSOPHILA MELANOGASTER BY MEANS OF A COMBINED QUANTITATIVE TRAIT LOCI AND SELECTIVE SWEEP MAPPING APPROACHES

Ricardo Wilches, Pablo Duchen, Sussane Voigt, Wolfgang Stephan
University of Munich, Department of Evolutionary Biology, Planegg-Martinsried, Germany

With the aim of understanding the molecular basis of cold adaptation in Drosophila melanogaster, we initiated the identification of the genes/functional elements affecting the fly’s chill coma recovery time (CCRT). CCRT is a proxy for cold stress tolerance. Based on quantitative trait loci (QTL) mapping between inbred lines from temperate and tropical origin, three X-linked QTL influencing CCRT were identified at 17, 18 and 24 cM. The two former QTL were resolved to fragments of approximately 150kb each by means of quantitative deficiency mapping. We achieved further resolution, to the gene/nucleotide level, for the locus at 18cM by subjecting the targeted chromosome fragment to population genetic analysis. In the light of selective sweep theory we expect the alleles responsible for cold tolerance to have been positively selected in the temperate population, therefore regions in the QTL that show footprints of positive selection may contain the candidate molecular elements behind cold stress tolerance. However, selective sweep mapping overlooks sites with other modes of evolution. In order to cover a broader range of potential sites influencing cold stress tolerance we conducted a parallel search with a phenotype-genotype approach suitable for fine scale mapping of QTL. Within the QTL defined at 18 cM, a 3-kb long transcription enhancer element approximately 7-kb upstream of gene brinker, shows footprints of positive selection and is strongly associated CCRT. We are currently studying the role of this enhancer element in the phenotype of interest, by measuring expression-level differences of the gene brinker and its immediate neighbor genes, both in wild type tropical/temperate flies and transgenic tropical flies with the putatively selected temperate version of the enhancer.

P-119
WHY DO SEYCHELLES WARBLERS HELP? FITNESS CONSEQUENCES AND HERITABILITY OF HELPING BEHAVIOUR

Hannah L Dugdale1,2, David S Richardson3, Jan Komdeur2, Terry Burke1
1University of Sheffield, Sheffield, UK, 2University of Groningen, Groningen, The Netherlands, 3University of East Anglia, Norwich, UK

Cooperative breeding is a social system in which individuals help to raise offspring that are not their own. It is a widespread system occurring in birds, mammals, fish and insects. As individuals are selected to maximise their fitness, this raises the question: why do some individuals help others to breed, rather than focusing on their own reproduction? Studies of cooperative breeding have tended to focus on the environmental factors that lead to the evolution of helping behaviour, with few studies quantifying the strength of selection on helping and whether helping is heritable. Knowledge of these factors is vital for a complete understanding of the evolution of cooperative breeding. This requires long-term study of individual behaviour, a multigenerational pedigree, and molecular and statistical techniques only recently adapted for natural systems. We evaluate selection for and heritability of helping using an 18-year dataset of 1,566 individually marked Seychelles warblers Acrocephalus sechellensis. We monitored individuals in this contained wild-living island-population from birth to death, presenting a rare opportunity to measure fitness accurately. Using Bayesian methods, 30 microsatellite loci and phenotypic data we constructed a multi-generation pedigree. We show that there are direct fitness benefits to helping, due to increased fecundity but not survival, and that there is individual variation in helping behaviour. We assess the environmental and genetic basis of helping and combine our findings to improve our understanding of the adaptive basis of reproductive decisions in the cooperatively breeding Seychelles warbler.
**P-120**

**HETEROZYGOSITY-FITNESS CORRELATIONS IN MARITIME PINE**

*Isabel Rodríguez-Quilón, Luis Santos-del-Blanco, Juan Pablo Jaramillo, Santiago C. González-Martínez, Ricardo Alía*

*Center of Forest Research (CIFOR-INIA), Madrid, Spain*

Correlations between molecular-marker heterozygosity and fitness traits, also known as heterozygosity-fitness correlations (HFCs), have been examined for more than three decades in numerous organisms. These studies, in general, indicate a positive relationship between heterozygosity and fitness traits. Nowadays, the use of a high number of SNP marker data in the analysis of HFCs may help us disentangle the direct and indirect effects of polymorphisms, shedding light on this classic issue. Provided that most economically important forest trees, and particularly conifers, are outcrossing wind-pollinated species, often suffering strong inbreeding depression, a positive relationship between molecular-marker heterozygosity and fitness is also expected in these species. A maritime pine (*Pinus pinaster* Ait.) common garden experiment was established at four sites, covering distinct environmental conditions (from Mediterranean to Atlantic). Each trial comprised 650 clones from 37 populations (20,800 trees were phenotyped). At each clonal trial, we studied the correlation between observed heterozygosity at 266 candidate gene SNPs and survival and total height, as fitness proxies. Variance components and clone BLUPs for both variables were estimated by restricted maximum likelihood (REML), using a linear mixed model approach for height, and a generalized linear mixed model for survival. Clonal heritability ranged between 0.26±0.03 and 0.47±0.07 for height, and between 0.05±0.03 and 0.14±0.04 for survival, at different locations. Observed heterozygosity ranged from 0.174 to 0.335, depending on population. Correlations between SNP heterozygosity and clone BLUPs for survival or total height were non-significant at the four sites. HFCs were neither significant when clones were pooled in metapopulations according to their genetic structure based on the 266 genotyped SNPs. Possible causes of our results are discussed based on the estimate both of the inbreeding load and the amount of variance in fitness explained by inbreeding.

**P-121**

**THE GENETIC BASIS OF PARASITE RESISTANCE IN A NATURAL POPULATION OF RED GROUSE**

*Marianne James1, Alex Douglas1, Gill Murray-Dickson1, Lucy Webster1, Pablo Vergara1,2, Jesus Martinez-Padilla1,2, Stuart Piertney1*

1Institute of Biological and Environmental Science, University of Aberdeen, Aberdeen, UK, 2Museo Nacional de Ciencias Naturales, Department of Evolutionary Ecology, Madrid, Spain

Understanding the genetic basis of parasite burden among individuals in natural populations is of major importance in ecology and evolution. We have been examining the host-parasite relationship between the red grouse *Lagopus l. scoticus* and its gastrointestinal nematode parasite *Trichostrongylus tenuis*. It is known that the vast majority of individuals are parasitized to differing levels, and that high parasite burden can affect grouse condition, survival and fecundity, which can have population dynamic consequences. The aims of our work is to define an interacting gene network of quantitative trait loci (QTL) associated with high worm burden, and expression QTLs that influence variation in transcript abundance of parasite-upregulated genes. The population genetic structure of these markers can then be examined in a landscape genomics context across a broad geographic range of populations exhibiting different levels of parasite burden, isolation, climate and density.

We have developed a grouse genomic resource by preparing two restriction site associated DNA (RAD) sequencing libraries. Over 256 million reads were sequenced which have been cleaned and filtered. A subset of these will be assayed using BeadXpress to genotype 960 individuals. Our SNP selection criteria was based upon the sequences containing a BLAST result from the chicken genome (for annotation), having low e-values, being well spaced through the chicken genome (for linkage analysis),
and elimination of exceptionally high or low coverage (to avoid repetitive elements or sequencing error respectively). The resulting genotypic information will be the basis of a genome-wide association study.

P-122
THE TRANSCRIPTIONAL ARCHITECTURE OF THE GENETIC VARIANCE FOR MALE AND FEMALE FITNESS IN A NATURAL DROSOPHILA POPULATION

Scott Allen, Steve Chenoweth
School of Biological Sciences, The University of Queensland, St Lucia, Australia

The genetic basis of reproductive fitness is central to our understanding of evolution. While additive genetic variance for fitness is known to be present in wild populations, the dissection of its genetic architecture presents some fundamental challenges. We have taken a transcriptional approach to the identification of genes associated with the genetic variance for fitness in the fruit fly Drosophila serrata. A panel of 41 inbred lines was established from a natural population and used to estimate the genetic variance for fitness in both sexes. Genetic variance for fitness was significantly greater than zero in males and females with broad-sense heritabilities of 8% and 14% respectively. The intersexual genetic correlation for fitness was positive (0.58) but significantly less than one, indicating a degree of independence between the genetic basis of fitness in each sex but also the potential existence of axes of standing variance summarising sexually concordant, sex-limited and also sexually antagonistic fitness effects. We then used custom designed 12 x 135K Nimblegen expression arrays to perform whole-genome transcriptional profiling of male and female adults within this mapping panel. We then associated genetic variance in transcript abundance with genetic variance in fitness. This allowed us to identify both individual transcripts and sets of co-expressed networks genetically correlated with fitness and to estimate how these molecular phenotypes influence fitness in each sex. Our analyses provide critical insights into the genetic basis of fitness and how it differs between males and females.

P-123
GENETICS OF ADAPTIVE VARIATION IN ARABIDOPSIS LYTARA

Tuomas Toivainen1, Saana Remula2, Timo Vesimäki2, Helmi Kuittinen1,2, David Remington3, Outi Savolainen1,2
1Biocenter Oulu, University of Oulu, Oulu, Finland, 2Department of Biology, University of Oulu, Oulu, Finland, 3Department of Biology, University of North Carolina at Greensboro, Greensboro, North Carolina, USA

Variation in quantitative traits can be maintained by a balance between mutation and selection, or by different forms of balancing selection. The underlying loci could be in mutation-selection balance, or spatially or temporally varying selection could maintain the variation. Overdominance, can also maintain quantitative genetic variation, especially in outcrossing species. Our aim was to measure the levels of adaptive genetic variation within a northern Arabidopsis lyrata population. We aimed at examining the genetic basis of variation: how much additive genetic variance, which loci govern the variation, are there any loci with large allelic effect? Do different alleles confer selective advantage in different environments? What is the frequency spectrum of the variants? We planted 1100 plants, consisting of half-and full-sib families (North Carolina II design), in two environments: the native environment, Spiterstulen (Norway, lat 61°N, long 08°E altitude 1100 m), and the experimental field in Oulu (Finland lat 65°N, long 25°E altitude 7 m). We collected phenotypic data on different morphological and phenological traits over two years from both environments. Variance components were estimated for several fitness related traits. Those traits had low to moderate heritabilities. To examine the role of flowering time loci on the observed phenotypic variation, we designed 70 SNP-markers in 20 well-known flowering time genes and 16 reference loci. Association mapping was used to identify the loci governing this variation and estimate
their allelic effects. Preliminary results showed that the following loci from the flowering time pathway had an effect on seed production in both years in plants grown in Oulu: CONSTANS, PHYB, FT, VRN1 and VRN2.

P-124
PEDIGREE RECONSTRUCTION FOR COSEXUAL SPECIES USING SIMULATED ANNEALING: CASE STUDY OF OIL PALM (ELAIS GUINEENSIS)

David Cros1, Jesus Fernandez2, Jean-Marc Bouvet1, Leopoldo Sanchez3

1CIRAD, Montpellier, France, 2INIA, Madrid, Spain, 3INRA, Orléans, France

The approach of Fernandez and Toro (2006) to reconstruct pedigrees from molecular data of contemporaneous individuals using a simulated annealing algorithm was extended to hermaphroditic and monoecious species. New features also include the possibility of selfings, accounting for a predefined coancestry matrix between founders and specifying different number of individuals per generation. The new method was validated using 16 individuals from the last generation of the Yangambi breeding population of oil palm. Their pedigree was known for 5 generations and they were genotyped with 166 SSR. In the study, the number of used SSR varied from 6 to 166 and the percentage of unknown parentages from 20% to 100%. The Pearson correlation between the pedigree-based coancestries calculated on the true and on the reconstructed genealogies ranged from 0.74 to 0.99. The RMSE ranged from 0.02 to 0.12. When pedigree was assumed completely unknown, reliable reconstruction required at least 38 SSR. Using 100 SSR or more, the Pearson correlation was very high (0.98) and the RMSE very low (0.06). The new method was also applied to 104 individuals from the last generation of a key breeding population (Deli) originated from 4 oil palms. The individuals were genotyped with 160 SSR. Records of their pedigree only existed for the recent past. Results of pedigree reconstruction detected a family coming from old selfings looking as outliers, with pedigree-based coancestries much higher than molecular coancestries, indicating old selfings were erroneous. After correcting the recorded pedigree, pedigree-based coancestries calculated on the reconstructed genealogy and molecular coancestries were highly correlated (> 0.9) when using 80 markers or more. In conclusion, this method gave likely pedigrees with satisfactory reliability for cosexual species, using a realistic number of polymorphic markers. Also, it seems very helpful to correct historical pedigrees. The methodology has been implemented in the software MOL_COANC_v2.

P-125
SCREENING PRIMATE GENOMES FOR ENDOGENOUS RETROVIRUSES.

Katherine Brown, Richard Emes, Ed Louis, Rachael Tarlinton
University of Nottingham, Nottingham, UK

Endogenous retroviruses (ERVs) are retroviruses which have integrated into the DNA of the germline cells of a host and are transmitted horizontally from generation to generation. Once a retrovirus has become endogenous, it behaves like any other genetic element, so it is subject to selection, mutation and genetic drift and can spread through the host population to fixation, or be eliminated from the population entirely. By analysing the ERV content of a genome it is possible to gain insight into the evolutionary history of both the virus and its hosts. Several large-scale genome screening projects have been carried out on mammalian genomes to identify ERVs and the human and chimpanzee genomes have been comprehensively screened. However, although all primate genomes have basic RepeatMasker annotation there is limited information about their retroviral content.

This project is based on use of the Exonerate algorithm (Slater and Birney, 2005) to screen the 10 available primate genomes for regions resembling retroviral genes. An input database has been created with comprehensive coverage of known retroviral diversity and these sequences are used as queries to locate
similar regions in primates. It is then possible to generate phylogenetic trees to investigate the likely origin of these sequences. Where multiple related sequences are available, the genetic structure of the different sequences can be examined to see how selection has affected ERV regions. It is also possible to date these sequences and compare this information to the host phylogenetic tree, providing insight into the possible evolutionary and demographic history of the host. Preliminary results suggest that this method is suitable for locating endogenous retroviral regions in primate genomes and our analysis of the human genome has generated results consistent with previous studies. This technique is now being applied to other primate genomes.


P-126
DECIPHERING THE MATING STRUCTURE OF AN ENDANGERED WILD SALMON POPULATION WITH DNA BASED SIBSHIP RECONSTRUCTION AND A NOVEL METHOD OF SEX INFERENCE

Louise de Mestral Bezanson1, Christophe Herbinger2, Patrick O’Reilly1
1Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada, 2Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada

The recent development of methods for sibship inference among individuals based on DNA marker data has allowed insight into mating systems in the wild when no information on putative parents is available. The microsatellite locus Ssa202 has recently been shown to be closely linked to the sex-determining region of the Atlantic salmon (Salmo salar) Y chromosome. Capitalizing on this finding, we show here a novel method to infer the sex of the unsampled, unseen, wild-spawing parents that produced some of the last remaining wild salmon in the endangered Stewiacke River, Nova Scotia, Canada. We inferred that a large proportion of the sampled juveniles belonged to a limited number of full-sib nested within half-sib families. This was corroborated in a second step by introducing individuals from a known and similarly complex pedigree into the original dataset and showing that the methods reconstructed nearly perfectly the family structure of these known individuals. Using the sex linked locus information, we show that the full sib parents were male and were generally monogamous while the half sib parents were female and polygamous. There are very few large anadromous salmon returning to spawn in this river and we conclude that the putative male parents were likely mature male parr (juvenile) rather than anadromous males. These results suggest that salmonid populations undergoing extreme declines maybe composed of a large number of related individuals from a few sib groups, and that mature male parr may provide an important genetic and demographic buffer to population decline. The methods used here allowed precise insight in the wild mating system and the maintenance of genetic variation in that population. These results had important consequences for the development of a conservation genetic strategy.

P-127
DNA MARKER-BASED INFERENCE OF REPRODUCTIVE SUCCESS BASED ON PROBABILISTIC PEDIGREES

Phillipe Fullsack1, Christophe Herbinger2, Kathryn MacEachern2, Patrick O’Reilly3
1Department of Mathematics and Statistics, Dalhousie University, Halifax, Nova Scotia, Canada, 2Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada, 3Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada

The propensity of individuals to transmit their genes to the next generation is often measured by their reproductive success, a key element of individual fitness. When the parental filiation between offspring and their parents cannot be observed, or is only partially known, sampling the genotypes of candidate parents and offspring via DNA markers can help reconstructing this filiation. Many variations on this
theme (no parental sample available, rich or poor sib structure, availability of complementary non-genetic information, spatial correlation structure of gene flow, marker type, fractional vs categorical allocation) have been explored in a number of numerical tools (Parentage, Colony, Pedigree, MasterBayes, Nest). In this study, we propose new inference techniques based on MCMC simulations and subsampling, and highlight the need to explore a continuous scale between categorical and probabilistic pedigrees (i.e. fractional allocation) to properly take into account the ambiguity in parentage allocation. We calibrate our method on a ‘gold standard solution’ constructed from genotype data from a known Atlantic salmon pedigree. Simulations are used to quantify a range of scenarios, and the degree of robustness or bias of various estimators. Finally, we explore some consequences of the use of probabilistic pedigrees on the estimation of reproductive success and individual inbreeding in a reconstructed multigenerational probabilistic pedigree of sea trout.

P-128
CONTRASTING GEOGRAPHIC PATTERNS OF GENETIC VARIATION FOR MOLECULAR MARKERS VS PHENOTYPIC TRAITS IN THE ENERGY GRASS Miscanthus sinensis

Gancho Slavov¹, Paul Robson¹, Elaine Jensen¹, Edward Hodgson¹, Gordon Allison¹, Sarah Hawkins¹, Sian Thomas-Jones¹, Xue-Feng Ma², Nickolai Alexandrov², Maxim Troukhman², Liping Zhang², Kerrie Farrar¹, Timothy Swaller², John Valentine¹, John Clifton-Brown¹, Richard Flavell², Iain Donnison¹

¹Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, Ceredigion, Wales, UK, ²Ceres, Inc., Thousand Oaks, California, USA

Because of their high productivity, C4 grasses from the tropical genus Miscanthus are believed to have great potential as a bioenergy crop. However, Miscanthus species are essentially undomesticated, and their accelerated breeding is hampered by their primarily outcrossing mating systems, perennial life cycles, and unreliable establishment from seeds in temperate climates. To overcome these challenges, we have assembled a large (>1500 accessions) and genetically diverse germplasm collection for several species of Miscanthus, obtained high-quality data for a large number of phenotypic traits and are using genotyping by sequencing approaches to generate dense molecular marker data for genomewide association studies (GWAS) and genomic selection (GS). Comparative multivariate analyses of molecular marker and phenotypic trait data for M. sinensis (i.e., one of the key species in our breeding programme) revealed several significant geographic patterns of genetic variation. A longitudinal cline accounts for a highly significant proportion of putatively neutral genetic variation, as revealed by single-nucleotide polymorphism and microsatellite markers ($R^2 = 0.60, P = 3E-15$). In contrast, genetic variation for 17 phenotypic traits related to biomass productivity, phenology and cell wall composition appears to follow primarily latitudinal and altitudinal gradients. Interestingly, $Q_{ST}$ vs $F_{ST}$ comparisons suggest that all three of these sets of traits may have been affected by divergent selection. Our results have several implications. First, the genetic clines we identified will need to be reflected in the design of populations for GWAS and GS, as well as in the analyses of the resulting data. Second, the contrasting geographic trends for neutral markers vs potentially adaptive phenotypic traits should be a favourable factor for the detection of molecular signatures of natural selection. Finally, the spatial patterns of genetic variation that we detected will directly inform our breeding programme and provide important clues about the evolutionary history of M. sinensis.
P-129
TWO-LOCUS EPISTATIC INCOMPATIBILITIES IN THE CONSTRUCTION OF RECOMBINANT INBRED LINES

Mohammad Shojaei Arani, Maria Colomé-Tatché, Frank Johannes
University of Groningen, Groningen Bioinformatics Centre, Faculty of Mathematics and Natural Sciences, Groningen, The Netherlands

In the construction of Recombinant Inbred Lines (RILs) certain combinations of parental genotypes at two (or more) loci may be genetically “incompatible”, thus resulting in lower fertility. Selection against these epistatic combinations during inbreeding causes significant segregation distortion and complicated patterns of long-range linkage disequilibrium (LD) within or across chromosomes. These undesirable effects have important implications for subsequent quantitative trait locus mapping experiments in the final RIL population. Here we study several two-locus epistatic selection models and quantify their impact on the genetic composition of the genomes of RILs during intermediate inbreeding generations and at fixation. Using our theoretical results, we develop maximum likelihood (ML) tests which can be used to infer the model that is most consistent with observed epistatic patterns in RIL genotype data. We illustrate this ML approach in the context of several RIL panels which show strong evidence for long-range LD and segregation distortion.

P-130
ADAPTATION TO URBAN ENVIRONMENTS IN ARABIDOPSIS THALIANA

Angela Hancock1,2, Felice Sperone2, Joy Bergelson2
1University of Vienna, Vienna, Austria, 2University of Chicago, Chicago, IL, USA

Plants growing in urban environments face unique challenges, and their abilities to thrive in these conditions may be mediated by natural selection. We scanned the Arabidopsis thaliana genome to identify genetic variants with higher frequencies in urban compared to non-urban environments. We found that truncating amino acid changes were significantly over-represented among the variants most strongly associated with urban environments. Furthermore, when we asked which gene ontology categories were enriched among strongly associated variants, we found categories related to heme biosynthesis, electron transport chain functioning and photosynthesis, suggesting that high levels of ozone may be an important selective pressure. Our results are consistent with the idea that urban environments pose novel challenges and lend insight into the physiological impacts of these environments.

P-131 – ABSTRACT WITHDRAWN

P-132
BALANCING SELECTION AND ASSOCIATIVE OVERDOMINANCE SIMULATIONS RECOVER UNEXPECTEDLY HIGH DIVERSITY OBSERVED IN C. ELEGANS EXPERIMENTAL POPULATIONS

Ivo Chelo, Henrique Teotónio
Instituto Gulbenkian de Ciência, Oeiras, Portugal

How can high diversity be maintained in populations in face of directional selection and drift is a recurrent question in evolutionary genetics. In this study we subjected Caenorhabditis elegans populations, that either reproduced exclusively by outcrossing (dioecy) or by partial selfing and outcrossing (androdioecy), to laboratory adaptation for 100 generations. Surprisingly, if we assume that loci evolve independently, we find similar fitness effect distributions in both mating systems. This is only possible because androdioecious populations maintain much lower inbreeding than expected for
their observed degree of assortative mating. In order to explain this excess in heterozygosity, numerical simulations accounting for drift, recombination and mating system were done to test for selection against partially-recessive loci or for balancing selection. We find that associative overdominance can explain diversity changes during the initial stages of evolution, but balancing selection is needed to explain later stages.

P-133
X-LINKED ADDITIVE VARIANCE REDUCES THE BETWEEN SEX GENETIC CORRELATION
Matthew E. Wolak1, Alastair J. Wilson2, Loeske E.B. Kruuk2, Derek A. Roff1, Daphne J. Fairbairn1
1University of California Riverside, Riverside, USA, 2University of Edinburgh, Edinburgh, UK

In organisms with separate sexes, the response to selection for each sex is influenced by the amount of shared additive genetic variation between the sexes. If each sex is undergoing selection for a different optimal trait value then shared variation, measured by the between sex genetic correlation, will act as a constraint on the response to selection. Evolutionary theory predicts genes with sex-specific effects to be linked to or translocated on the sex chromosomes (i.e., X in XX/XY or Z in ZZ/ZW systems). However, estimates of between sex correlations usually ignore any sex-linked genetic variation. It has long been known, but rarely considered, that even when X (or Z) linked effects are perfectly correlated between the sexes, males and females will differ in the amount of X (Z) linked variance because of sex-differences in the number of these chromosomes. Here, we show that whenever additive genetic variance is located on the X (Z) chromosome, the between sex genetic correlation (autosomal + sex-linked) is always less than one. This result holds even when the homogametic sex (e.g., XX female) randomly inactivates one copy of the sex chromosomes, as is the case in many mammals (i.e., lyonization). We use analytical and simulation approaches to demonstrate the impact of sex-linked variance on estimates of the between sex genetic correlation under three scenarios: 1) perfect correlation between the sexes in both autosomal and sex-linked additive genetic effects, 2) perfect correlation between the sexes in autosomal additive genetic effects but no correlation in sex-linked additive genetic effects, and 3) perfect correlation between the sexes in autosomal additive genetic effects but a complete negative correlation in sex-linked additive genetic effects. Our results highlight the potential importance of sex-linked variance in population responses to sex-specific selection pressures, despite both sexes sharing the same genes.

P-134
GENETIC INTEGRATION OF DISPERSAL AND EXPLORATION BEHAVIOUR IN A WILD BIRD
Peter Korsten1, Thijs van Overveld1, Frank Adriaensen1, Erik Matthysen1
1University of Antwerp, Antwerp, Belgium, 2University of Groningen, Groningen, The Netherlands

Dispersal – the movement of individuals from their place of birth to the place of breeding – is an important ecological parameter influencing population dynamics and gene flow. In wild animal populations, there is often large variation in the distance of dispersal. Part of this variation may have a genetic basis, but this has rarely been tested. Recently, it has been suggested that consistent and heritable individual differences in basal behavioural traits, such as exploration, boldness, sociability, and aggression – often described as variation in ‘animal personality’ – may be functionally linked to individual life-history and dispersal strategies. While there is good empirical evidence for phenotypic correlations between dispersal and personality traits, it is unclear whether dispersal and personality are also genetically integrated. Here, we show dispersal and exploration behaviour to be both heritable, and genetically coupled, in a common wild songbird, the great tit (Parus major). Using pedigree and phenotypic data from a wild great tit population, we found quantitative genetic variation both for the distance of dispersal within our study area and the rate at which individuals explore a novel environment measured in a standard behavioural test. Moreover, we found a positive genetic correlation between dispersal distance
Poster Abstracts

and exploration rate, thereby providing evidence for the genetic integration of animal personality and dispersal. These findings indicate a potentially important behavioural mechanism underlying between-individual differences in dispersal, show that dispersing individuals may form a genetically non-random subset of the population, and highlight the potential for dispersal to evolve in response to selection.

P-135
PIONEERING THE DISCOVERY OF THOUSANDS OF SNPS IN UNREFERENCED SPECIES - THE CASE OF EXTANT SIGNATURES OF SELECTION AMONG THE ATLANTIC HERRING POPULATIONS

Alvaro Martinez Barrio1,2, Sangeet Lamichhaney1,2, Nima Rafati1,2, Görel Sundström1,2, Carl-Johan Rubin1,2, Manfred Grabherr1,2, Nils Ryman1, Leif Andersson1,2
1Science for Life Laboratory, Uppsala, Sweden, 2Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden

The Atlantic herring (Clupea harengus) has been an important food resource in Northern Europe since ancient times. The Atlantic herring was one of the first marine fishes to be used for population studies. Previous studies based on a limited number of genetic markers revealed that most loci show no or minute genetic differentiation between populations, even including the herring adapted to the Baltic Sea with its low salinity. These data imply that the effective population size in the Atlantic herring is enormous which creates an opportunity to study the consequences of natural selection when genetic drift is of minor importance.

In this study we have combined whole genome sequencing and transcriptome analysis and performed a population study based on more than 400,000 SNP loci. We first assembled a muscle transcriptome from a single fish using the Trinity software. We then mapped genomic reads to the transcriptome and produced an exome assembly. We then carried out whole genome sequencing of pooled samples of herring (50 fish in each pool) from different localities in the Baltic Sea, Skagerak, Kattegat, North Sea and the Atlantic Ocean. The genomic reads were aligned to the exome assembly that resulted in the detection of more than 400,000 SNPs. Population genetic analysis of these data revealed that the great majority of loci showed very little genetic differentiation, and the average Fst value was 2-3%. However, a few percent of the loci showed highly significant allele frequency differences between populations, and the distribution deviated significantly from the one expected under neutrality. The study is a breakthrough as regards our ability to distinguish genetically distinct subpopulations of the Atlantic herring and establish this fish as a unique model for evolutionary studies of natural selection and genetic adaptation.

P-136
INVESTIGATION OF APOE GENE IN IRANIAN PATIENTS AFFECTED BY AGE-RELATED MACULAR DEGENERATION (AMD) AND ASSOCIATION BETWEEN THIS GENE AND SEVERITY OF DISEASE

Golnar Aryan1, Mohamad Hossein Sanati0,2, Hajar Aryan1, Alireza Irani0,3
1Fazeli-Sanati Genetic Laboratory, Tehran, Iran, 2National Institute for Genetic Engineering and Biotechnology, Tehran, Iran, 3Rasool-e-Akram Hospital, Tehran, Iran

Background: Based on previous studies, the E4 allele of APOE gene is significantly associated with reduced risk of age-related macular degeneration (AMD). On the other hands The E2 allele of this gene has been reported possibly to have an association with increased risk of the disease.

Methods: The study consisted of 240 subjects including 120 patients with AMD and 120 age and sex matched healthy controls.

Result: In total of AMD patients and healthy control we had 71 (29.59%) individuals with mutations in the APOE gene and 169 (70.41%) individuals without mutations (E3/E3 genotype). E3/E3 genotype had the highest frequency in two groups. The genotype E2/E2 with p-value=0.64, odds Ratio=7.37 and genotype E2/E3 with p-value=0.018, odds Ratio=2.60 were the most important genotypes significantly associated
with Increased risk of AMD. The E2/E3 genotype with p-value=0.02 and positive familial history with p-value=0.016 appears to be highly associated with early onset and greater risk of AMD progression. In the AMD patients we had 20% smokers and 6.6% with positive family history and in healthy controls we had 13.3% smokers (p-value=0.2) and no one had a positive family history (p-value=0.007) in this group. This data showed a significant association between positive familial histories and increasing the risk of AMD. On the other hand cigarette smoking showed no significant association with AMD.

**Conclusion:** The information obtained from this study showed that the E2 allele is associated with increased risk of AMD. Also no evidence was found to support an association between AMD and the protective effect of E4 allele because of low number of this allele among our subjects.

**P-138 – ABSTRACT WITHDRAWN**

**P-138**

ESTIMATION OF HERITABILITY FROM LIMITED FAMILY DATA USING GENOMIC IDENTITY-BY-DESVENT RELATIONSHIPS

Jørgen Ødegård1–2, Theo Meuwissen2

1Nofima, Aas, Norway, 2Norwegian University of Life Sciences, Aas, Norway

In classical quantitative genetic analyses, genetic variance is estimated from between-family variation based on pedigreed populations with numerous families, which is often complicated by close relatives sharing environmental factors. Genomic information may revolutionize how data can be utilized in analyses of quantitative traits. For human height it has been shown that genetic variance can be inferred solely from within-family variation using genomic identity-by-descent (IBD) relationships. However, this requires large datasets with genotyped individuals. However, for highly fecund species (e.g., fish, insects and poultry) population structures are very different from humans. In this simulation study we tested whether genomic IBD could be used to estimate additive genetic variance from within-family variation using restricted family material (1-10 full-sib families) with family sizes typical for highly fecund species. The analysis assumed that genomic data can be used to accurately trace inheritance of genomic blocks from parents to offspring. The results showed that the genomic IBD can be used to estimate genetic variance of polygenic traits from a limited family material. Furthermore, for a given number of observations, the most accurate estimates were obtained when data was restricted to a single large family only. Compared with classical pedigree-based methods the proposed method is more robust to selection among parents and confounding of additive genetic and environmental effects. Furthermore, information on individual breeding values is utilized, improving precision of the estimated genetic variance. The proposed testing method is of particular relevance for estimating additive genetic variance in highly fecund species using data from populations with shallow pedigrees and/or few available parents. The method may also be extended to situations were only one of the sexes have high fecundity (e.g., livestock). Due to the limited number of recombinations occurring from parent to offspring, genomic IBD relationships can likely be calculated with relatively sparse marker panels.
P-139
GENETIC EVALUATION OF ANIMALS BASED ON SIRES’ GENOTYPES FOR DENSE MARKERS USING COMPUTER SIMULATION

Majid Khansefid
University of Tehran, Tehran, Karaj, Iran

In this research the effect of selecting individuals for genotyping in the reference and validation population on the accuracy of estimated genomic breeding values (EGBV) was investigated. A program was designed to perform the investigations and the markers effects were estimated by BLUP characteristics by solving the equations with Gauss-Seidel method. As in the current research the classic estimated breeding values of individuals in the reference population were used to estimate the genomic breeding values of the validation individuals and due to the fact that the accuracy of estimated breeding values for conventionally proven sires are higher than the other individuals in their contemporary groups and most of the next generation individuals are the progenies of them, it was tested whether genotyping proven sires could lead to the higher EGBVs accuracies. The results indicated that ascending the number of individuals in the reference population, increased the EGBVs accuracies in the validation population and setting sires in the reference population because of their higher accuracies in classic estimated breeding values in comparison with the dams, caused higher EGBVs accuracies in the validation set. In all tests focused on using all or top individuals as the validation population, the results indicated that setting all of the individuals yields more reliable EGBVs. However, while selection was made among sires and dams based on their EBVs, the differences between accuracies in the validation where all or top individuals were chosen as the test population, reduced. Testing top or randomly chosen individuals in the reference population showed whenever the reference group was of top individuals, the accuracy of EGBVs descended. It might happen due to the closer relationship of top individuals in the reference population.

P-140
STATISTICAL MODELS TO COMBINE INFINITESIMAL AND MARKED GENETIC EFFECTS

Eduardo Manfredi
INRA, Toulouse, France

The infinitesimal (or polygenic) approach exploits successfully phenotypes and pedigree data to estimate additive genetic values. In its simplest version, the model is

\[ y = \mu + u + e \]

with \( \mathbb{E}[u] = 0 \) and \( e \sim N(0, \sigma_e^2) \). The “blackbox” vector \( u \) is supposed to capture the homogeneous additive effects of many unknown genes, and \( A \), the additive relationship matrix among individuals, is supposed to give the genetic covariances among individuals as fractions of the additive genetic variance,

\[ \text{cov}(u_i, u_j) = a_{ij} \sigma_u^2. \]

SNP and other DNA markers data can be used to replace the “blackbox” \( u \) by \( Mm \) where \( m \) is a vector of marked genetic effects and \( M \) is a centered matrix of marker genotypes, or to replace \( A \) by a genomic relationship matrix equal to \( MM' \). Alternatively, some authors included both \( u \) and \( Xm \) in their model while other tried to found optimum contributions from both sources of genetic variation.

When DNA data is available, it is convenient to represent genetic effects by including the marked effects \( Mm \) and also \( v \), a vector representing the sum of nonmarked additive genetic values. Moreover, \( v \) can be partitioned into \([v_b \quad v_d]\) for base and descendant individuals, and \( v_b \) can be estimated as a function of \( v_b \) and the molecular data \( M \). Models of this form can be also used with \( M \) selected in previous steps. These models will be illustrated by examples.
P-141
GENOME PREDICTION OF GENETIC VALUES USING ADDITIVE × ENVIRONMENT AND ADDITIVE × ADDITIVE × ENVIRONMENT INTERACTIONS

Juan Burgueno1, Jose Crossa1, Gustavo de los Campos2, Susanne Dreisigacker1, Paulino Perez3
1International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, Mexico, 2University of Alabama, Birmingham, Alabama, USA, 3Colegio de Postgraduados, Montecillo, Texcoco, Mexico

The availability of high density panels of molecular markers has prompted the adoption of genome-enable prediction and selection (GS) in animal and plant breeding. Most GS applications use a single-environment model. Genotype × Environment (G×E) interactions are pervasive in plants and evidence supports the existence of various types of epistatic interactions (G×G). The single-trait additive models commonly used in GS do not accommodate G×G interactions nor can they exploit correlated performance across environment for prediction of genetic values. Drawing on quantitative genetic theory of infinitesimal models under G×E and G×G we propose a multivariate mixed models for incorporating pedigree and molecular marker information in models for multi environment data which accommodate additive (A) and additive × additive interactions (AA) described using pedigree and/or marker-derived relationships. We illustrate and evaluate the proposed methods by comparing the prediction accuracy of models for pedigree, markers and pedigree plus markers which can consider A or A+AA effects in the context of single-environment and multi-environment analyses. The evaluations of this study were carried out using two validation designs: one representing the prediction of performance of newly developed lines (i.e., those with no field evaluation) and one representing the prediction of performance of lines that have phenotypic performance evaluated in incomplete field trials using. Data was from an historical set of 599 wheat lines evaluated in four environments. We found that: (1) combining markers and pedigree information yields higher accuracy than the use of markers or pedigree data alone, (2) modeling G×E can boost prediction accuracy, especially in incomplete field trials, and (3) in some environments modeling of AA effects can increase prediction accuracy. In conclusion, in most instances GS can benefit from modeling G×E and G×G.

P-142
ASSESSMENT OF SINGLE-STEP GENOMIC EVALUATION OF A MIXED-BREED DAIRY CATTLE POPULATION IN NEW ZEALAND

Bevin Harris, Anne Winkelman, David Johnson
LIC, Hamilton, New Zealand

The single-step procedure of combining genomic and pedigree information in the mixed model equations was used to obtain genomically-enhanced breeding values for production traits in the New Zealand dairy cattle population. The breed composition was predominantly Friesians (F), Jerseys (J) and their crosses (FJX). Genomic information was available for 5425 sires, either through genotyping (50K chip, 38,296 genotypes) or imputation from a lower density (18K) chip. Analyses were done using the same multiple-trait animal model used for the national evaluation, which was modified to include the genomic information. Genomic information was incorporated as either the genomic relationship matrix (GRM) or the Euclidean-squared (E2) matrix. Phenotypic records for production traits (milk volume (M), fat yield (FY) and protein yield (PY)) from seasons 1987 to 2008 were used. Evaluations for season 2011 were predicted for 601 sires whose first-lactation daughters calved in 2009 or 2010. The accuracies (R²) of the within-breed regressions of the daughter-proven evaluations on the genomically-enhanced predictions ranged from 0.49 to 0.64, 0.34 to 0.41, and 0.39 to 0.60 for M, FY and PY, respectively. Differences in R² obtained from using the GRM and E2 matrices were small (≤0.02) for all traits in the F and J breeds but using the E2 matrix resulted in higher values in the FJX breed (0.41 vs. 0.34 for FY, 0.60 vs. 0.53 for PY). The inflation estimates (regression slopes) ranged from 0.80 to 1.06, 0.97 to 1.03 and 0.80 to 1.03 for M,
FY and PY, respectively. The lowest values (0.80 to 0.85) were found for the J breed for M and PY. The single-step method of calculating genomically-enhanced breeding values resulted in higher accuracies and lower inflation than an alternative multiple-step method that is currently in use.

P-143
INCORPORATING PRIOR KNOWLEDGE INTO THE PREDICTION OF GENOTYPIC VALUES

Robin Wellmann, Jörn Bennewitz
University of Hohenheim, Stuttgart, Germany

The number of phenotyped individuals with marker data is increasing rapidly in many breeds. This marker data can be used to recover high density marker genotypes or even whole genome sequences. Thus, massive marker data is becoming available that can be used not only for the prediction of genomic breeding values, but also for the prediction of genotypic values. This enables breeders to find the optimal mating partners for their animals. However, an accurate prediction of genotypic values is challenging and the incorporation of prior knowledge about the genetic architecture into Bayesian models can increase the accuracy considerably. We demonstrate how Bayesian models can be extended to account for genetic architectures of quantitative traits that are suggested in the literature. The proposed models allow modelling the joint distribution of additive effects and dominance effects dependent on the allele frequencies. A dependency between additive and dominance effects can be modelled such that the additive variance at most loci is small even when the additive effect may be not small. This is likely to hold for populations under selection since selection has likely shifted allele frequencies of QTL away from values for which they have a large contribution to the additive variance. The models assume that absolute additive effects of markers come from a mixture of two folded Student t distributions with different variance, so they generalize BayesA and BayesC and have a similar computation time per random effect. Different so-called BayesD models with dominance effects are compared with respect to their ability to predict genotypic values, dominance deviations, and breeding values for simulated populations.

P-144
RIDGE REGRESSION FOR RISK PREDICTION WITH APPLICATIONS TO GENETIC DATA

Erika Cule1, Maria De Iorio2
1Imperial College London, London, UK, 2University College London, London, UK

We address the challenge of using genetic data to predict disease risk. Standard regression techniques traditionally used to fit prediction models cannot be applied to contemporary genetic data sets, due to the high dimensionality of the data and the correlation structure among genetic variants. Penalised regression methods are a family of regression techniques that can be used with such data. Among penalised regression methods, ridge regression has been demonstrated to offer the best predictive performance. Ridge regression requires the specification of a penalty parameter, which controls the amount of shrinkage of the regression fit. Several methods have been proposed in the literature to choose the ridge parameter from the data. However, previously proposed methods fail when the data comprise more predictors than observations, as is typically the case in current genetic data sets.

Here, we propose a semi-automatic method to guide the choice of ridge parameter from the data when the data comprise many more predictors than observations. We propose choosing the penalty parameter such that the degrees of freedom for variance is the same as that of a principal components regression with a specified number of principal components. We discuss ways to choose the number of components to use, with the aim of good predictive performance. Using simulation studies, we demonstrate that when the number of causal variables is large and effect sizes are small, a plausible situation in the case of complex diseases, our method offers improved predictive performance over other penalized regression methods. We apply our method to out-of-sample prediction using two Bipolar Disorder genome-wide...
association studies, with data from the Wellcome Trust Case-Control Consortium and the Genetic Association Information Network.

**P-145**

**USING REALIZED GENOMIC COVARIANCE TO CAPTURE MENDELIAN SAMPLING EFFECT IN BREEDING: AN EMPIRICAL STUDY FROM LOBLOLLY PINE**

Jaime Zapata-Valenzuela\(^1\), Christian Maltecca\(^1\), Ross Whetten\(^1\), Steve McKeand\(^1\), Fikret Isik\(^2\)

\(^1\)North Carolina State University, Raleigh, NC, USA, \(^2\)Unidad Bioinformática, Conception, Chile

Accuracies of genomic estimated breeding values based on realized genomic covariance matrix (\(G\) matrix) derived from SNP markers were compared with estimates based on the average additive genetic relationship matrix (\(A\) matrix) derived from pedigree. We used two methods described by vanRaden (2008), one based on frequency of alleles shared by individuals and another based on regression, to generate \(G\) matrices from 3461 biallelic SNP markers genotyped in 165 individuals. In a BLUP framework, the \(A\) matrix derived from pedigree was substituted by the inverse of \(G\) matrices for prediction of genomic estimated breeding values of individuals. Genomic estimated breeding values had higher accuracy values compared to predictions based on the \(A\) matrix. On average, predictions based on the \(A\) matrix had accuracy of 0.60 across four cross-validation methods. The range of accuracies of genomic estimated breeding values was 0.65 to 0.76 across different validation methods. The two methods for deriving \(G\) matrices from SNP markers gave comparable accuracies of prediction. Compared to standard genetic evaluation based on pedigree and phenotype, the similar or higher accuracies of genomic estimated breeding values could be explained by the power of markers capturing the Mendelian segregation effect in full-sib families, which was not the case for average relationships. Realized genomic covariances derived from markers in a BLUP framework could be a powerful tool in breeding to increase genetic gains by reducing breeding efforts. In such a framework, fixed effect covariates and genotype by environmental interactions can be easily modeled.

**P-146**

**WHOLE-GENOME ENABLED PREDICTION IN MODERATELY HERITABLE COMPLEX HUMAN TRAITS**

Ana I. Vazquez, Gustavo de los Campos, Yann C. Klementidis, David B. Allison

University of Alabama at Birmingham, Birmingham, USA

In recent years Genome Wide Association Studies identified unprecedented numbers of variants significantly associated to important complex human traits and diseases. However, usually these variants explain such a small fractions of inter-individual differences in genetic risk (i.e. ‘the missing heritability’) that the practical usefulness of models based on significant associations is limited. Recently, much of the missing heritability has been recovered using models that incorporate hundreds of thousands of the genome-wide SNPs (WGP=Whole Genome Prediction) as opposed to just the ‘top hits’. Additionally, a study recently showed moderate to high prediction accuracy of WGP for height. WGP was originally developed for continuous traits; however, most traits of public health relevance are either binary or continuous but subject to censoring. Here, we generalized WGP to accommodate this types of outcomes and evaluate the predictive performance of these models with two traits of great relevance, expected years of life (YL) and skin cancer (SC), (data from the Framingham Heart Study, N=5,117). Our baseline models accounted for basic risk co-variates. These models were extended by adding 40,000 evenly spaced markers. Our WGP models accounted for, and were able to predict, substantial fractions of interindividual differences in risk which were un-explained by the baseline model. The cross-validation R-square for the baseline and WGP models were 0.11 and 0.21 for YL and 0.07 and 0.18 for SC, respectively. Results of area under the curve showed similar patterns. Furthermore, we evaluated the impact of population structure and of familial relationships on the prediction accuracy of WGP and concluded that the prediction accuracy of WGP can be explained partially, but not fully, by such factors. We conclude
that WGP offers opportunities to advance our ability to predict complex human traits. This has major implications for practical applications, including enhancing the potential of personalized medicine.

P-147
LOW-COST GENOTYPING STRATEGIES TO MAXIMIZE IMPUTATION ACCURACY IN PEDIGREED POPULATIONS

Matthew Cleveland, Yijian Huang, Christian Maltecca, John Hickey
1Genus plc, Hendersonville, TN, USA, 2North Carolina State University, Raleigh, NC, USA, 3University of New England, Armidale, NSW, Australia

The goal of commercial breeding programs is to maximize the rate of genetic gain at minimum cost. The use of genomic information has the potential to increase rates of genetic gain but the cost of genotyping may be a constraint. Genus has implemented a genomic evaluation pipeline that combines a low-cost genotyping strategy with genotype imputation to dramatically reduce the cost of generating genomic information. The objective of this project was to evaluate the cost and imputation accuracy of alternative genotyping strategies in a pedigreed population. Pedigree and genotype data from a commercial pig population were used and several genotyping strategies were explored. Each strategy varied the density of genotyping for individuals to be imputed and their ancestors. Parents, grandparents and other ancestors of a set of test animals were genotyped at high density, low density or very low density, and both cost and imputation accuracy were evaluated for each scenario. As expected, the costs and imputation accuracy were sensitive to the level of genotyping in alternative scenarios. When test animals were genotyped at very low density (384 SNP) the imputation accuracy was dependent on the density of genotyping in the sire and dam, and less sensitive to genotyping in other ancestors. Given the disparate numbers of offspring produced by males and females, an optimized low-cost genotyping strategy for a commercial pig population could involve genotyping male parents at high density, female parents at low density (e.g. 3000 SNP), and selection candidates at very low density. In this case, such a strategy yielded 95% genotype imputation accuracy for selection candidates, at a cost per selection candidate that was 21% of the cost required to genotype all individuals at high density. Reducing the density for female parents decreased the costs to 17% while still yielding 93% imputation accuracy.

P-148
A GENETIC CORRELATION MODEL FOR MULTIPLE BREED GENOMIC PREDICTION. APPLICATIONS IN DAIRY CATTLE AND SHEEP

Sofiene Karoui2, Andres Legarra1, Clara Diaz2, Maria Jesus Carabano2
1INRA, SAGA, Toulouse, France, 2INIA, Madrid, Spain

Multiple breed genomic predictions are of interest to increase the size of the reference population but also for genomic predictions in crossbred animals. Models used up to now have considered several breeds as the same one. However, in most cases animals are selected within and not across breeds; and there is no reason to expect similar LD patterns and SNP effects to be consistent across breeds. We posit the (genomic) breeding value in two different breeds as correlated, but not identical, traits. Each animal expresses, in its phenotype, the genotype of one or several breeds. If two breeds have the same genetic architecture then the genetic correlation \( r_g \) will be close to 1; otherwise it will be close to 0. This model was tested in dairy cattle: Holstein, Montbeliarde and Normande considering Milk Yield, Fat Content and Fertility. Training populations included 2976, 950 and 970 bulls; validation ones, 964, 222 and 248. Estimated genetic correlations across breeds within traits oscillated between 0.79 and 0. Later, three models were tested for prediction: \( r_g = 0 \); estimated \( r_g \); \( r_g =1 \)Holstein or Normande did not benefit from the multiple breed prediction, because either the data set was already large (Holstein) or the
r, with other breeds, low (Normande). However, Montbeliarde did improve in prediction (R² for milk yield increased from 0.27 with r = 0 to R² = 0.33 with estimated r or r = 1). It is known that some introgression of Holstein into Montbeliarde occurred some decades ago.

The GENOMIA project (www.genomia.info) works in analogous uses in dairy sheep. Results are on the way. Correlation between breeds helps assessing possible “boosting” in accuracy of genomic predictions from combining specific populations, and describes their divergence for a quantitative trait, much like the Qst fixation index.

P-149
ACCURACY OF GENOMIC SELECTION FOR FERTILITY TRAITS IN AUSTRALIAN BRAHMAN CATTLE

Antonio Reverter1,2, Marina Fortes1,2, Bolormaa Sunduimijid1,3, Yuandan Zhang1,4, Sigrid Lehnert1,2
1Cooperative Research Centre for Beef Genetic Technologies, Armidale, Australia, 2CSIRO Livestock Industries, Brisbane, Australia, 3Victoria Department of Primary Industries, Melbourne, Australia, 4University of New England, Armidale, Australia

Genotypes for 729,254 SNP on Brahman bulls (N = 1,115) and related heifers (N = 1,014) were used to ascertain the accuracy (ACC) of genomic selection (GS) for fertility traits including scrotal circumference at ~12 months of age (SC12; 21.24 ± 0.08 cm) and percentage of normal sperm at ~24 months of age (PNS24; 73.55 ± 0.71 %) in bulls; as well as age at first corpus luteum (ACL; 750.60 ± 4.48 d) and first post-partum anoestrum interval (PPAI; 180.11 ± 4.33 d) in cows. The available data was halved into calibration and validation population based on sires and in a way such that no paternal halfsibs were represented across the two populations. The ACC of GS was explored using various amounts of SNP to build the marker-based relationship matrix according to the significance of their association to phenotypes. These genomic-based ACC were further compared with those obtained from traditional pedigree-based BLUP approaches after tracing back three generations of ancestors to build the numerator relationship matrix. We observed a monotonic increase in the genomic-based ACC as the number of SNP increased and reaching an asymptote at around 0.282, 0.312, 0.445 and 0.241 respectively for SWC12, PNS24, ACL and PPAI. These values compared favorably against those obtained from pedigree-based BLUP approaches which equated 0.117, 0.167, 0.332 and 0.234 for the same set of traits.

P-150 – ABSTRACT WITHDRAWN

P-151
FREQUENCY AND REPEATABILITY OF RECOMBINATION RATES IN BEEF CATTLE

Bruce Tier
AGBU/UNE, Armidale, NSW, Australia

Research at the Cooperative Research Centre for Beef Technologies over the last 18 years has resulted in genotyping over 11000 cattle from a variety of breeds using a mixture of high and low density SNP tests. After a stringent quality assurance process, sire haplotypes were inferred from progeny genotypes and sire genotypes where available, and haplotypes determined for progeny. Numbers of recombinations in each progeny’s sire gamete were calculated for each chromosome and the total for each individual. Over 250 small to medium sized half-sib families were available. Total recombinations per individual were analysed with a sire model to ascertain the repeatability of recombination rate. There were an average of 0.4 to 1.7 recombinations per chromosome, proportional to the length of the chromosome. The frequency of inheritance of a complete chromosome, without recombination, ranged from 0.2 to 0.6 and was inversely proportional to its length. The average number of recombinations per gamete was 21, with modest variation among breeds. The repeatability of recombination was estimated from a subset of
the data with families of 10 (5) half-sibs as 0.25 (0.30). The low frequency of recombinations per gamete means that accurate haplotypes for progeny can be imputed from low density genotypes if the haplotypes of one parent are available.

P-152
INCREASING ACCURACY OF GENOMIC PREDICTION COMBINING COW AND BULL REFERENCE POPULATIONS

Mario Calus, Yvette de Haas, Roel Veerkamp
Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, Wageningen, The Netherlands

Genomic selection might enable selection for expensive or difficult to measure traits, like feed efficiency and methane emission in dairy cattle. However, only a few thousand phenotypic records are likely to be collected for these expensive traits. Therefore, we suggest that an optimal strategy may be to combine this genotypic and phenotypic information, with information of predictor traits that are available from national recording schemes. To combine this information optimally we investigated a Bayesian genomic prediction model that allows us to analyse two traits while animals only have a phenotype on either of the traits. The impact of adding the bull information was evaluated for accuracy of prediction, and for posterior QTL probabilities to assess effects on power for a genome-wide association study (GWAS). The model was tested on a data set with 1,609 cows and 296 bulls with phenotypes for fat and protein yield and with genotypes for 36,346 SNPs. All bulls had highly accurate daughter yield deviations (DYD) for fat and protein yield from the Irish national evaluations. Estimated genetic correlations between the cow and bull traits were either low (0.22-0.26) or moderate (0.55-0.56). Prediction accuracies were calculated via cross-validation, while the whole data set, including or excluding one of the bull traits, was analysed to investigate effects on GWAS. Results indicated that adding information of just a few hundred bulls did not significantly increase prediction accuracy, despite the high accuracy of their DYD. To achieve higher accuracy for the genomic predictions, apparently much larger national bull reference populations need to be added. Adding the bull information did however increase the power for GWAS. So, indicator traits used in a genomic prediction model improve power to identify genomic regions associated with traits of interest, even when the accuracy of genomic prediction with similar models remains unaffected.

P-153
A BAYESIAN FRAMEWORK TO COMPARE TRIAL ENRICHMENT PROGRAMMES: APPLICATION TO CHOROIDAL NEOVASCULARISATION (CNV) PREVENTION DESIGNS

Paul Newcombe1, Linda McCarthy1, John Whittaker1,4, John Wurzelmann1, Michael Fries2, Nancy Burnham2, Gengqian Cai3, Sandy Stinnett1, Trupti Trivedi2, Chun-Fang Xu1
1GlaxoSmithKline, Stevenage, UK, 2GlaxoSmithKline, Renaissance Center, Pennsylvania, USA, 3GlaxoSmithKline, Research Triangle Park, North Carolina, USA, 4London School of Hygiene and Tropical Medicine, London, UK, 5GlaxoSmithKline, Upper Merion, Pennsylvania, USA

There has been a lot of interest in the performance of genetics for predicting disease risk. A related question is whether genetics can be used to aid in the selection of patients for high risk clinical trials. Our aim was to assess the most cost effective criteria for recruiting high risk Age-Related Maculopathy patients into prevention trials for conversion to CNV and, in particular, whether the generation of genetic data for use in screening programmes would be cost effective.

Using publically available data from the Age-Related Eye Disease Study (AREDS), we fitted a Bayesian model regressing CNV on all clinical risk factors of interest and variants from 8 candidate genes. Then, we developed a novel Bayesian trial simulation framework which allowed estimation of parameters such as incident rate, size and cost under various screening programmes. Plausible per subject screening and
recruitment costs were used and trial lengths up to 4 years were investigated. Screening criteria explored included predicted risk from logistic regression models both with, and without, genetic factors, and simpler filters according to severity scales commonly used in the clinic. There was very strong evidence for an overall genetic effect on progression (p < 1E-7). However, genetics only demonstrated a marginal improvement in positive predictive value on top of well documented clinical risk factors. Consequently, across all trial lengths it was estimated that including genetics in an enrichment design would result in a marginal difference to trial costs.

Our trial simulation framework allowed us to infer that the benefit of including genetic factors to screen patients for recruitment to CNV prevention studies of less than 4 years is likely to be marginal. Furthermore, the framework supported identification of cost effective simple clinical filters which combine good predictive performance with ease of application in the clinic.

P-154
COMPARING PREDICTIONS OF RESPONSE TO GENOMIC SELECTION USING SELECTION INDEX THEORY AND GENOTYPE SIMULATIONS

Alban Bouquet1, Louise Pedersen2, Mark Henryon3, Anders Christian Sørensen2, Jarmo Juga1
1Department of Agricultural Sciences - University of Helsinki, Helsinki, Finland, 2Centre for Quantitative Genetics & Genomics, Aarhus University, Tjele, Denmark, 3Danish Agriculture and Food Council, Pig Research Centre, Copenhagen, Denmark

We compared rates of genetic gain (ΔG) and inbreeding (ΔF) realised by genomic selection (GS) in simulated breeding schemes where full genotypings (FG) were simulated or direct-genomic breeding values were sampled using selection index theory (SIT). Breeding programmes were simulated for a 25-year period and comprised 5000 breeding females. Selection was undertaken on a single trait with heritability of 0.4 and only expressed in females. A conventional progeny-testing (PT) scheme without GS was simulated as reference. Each year, 200 1-year-old males were chosen for PT using mid-parental breeding values (EBV) and 20 5-year-old males were selected as active sires based on PT results. Two GS breeding programmes were simulated, namely the pre-selection (PS) and juvenile schemes (JS), assuming that 500 males were genotyped each year using mid-parental EBV. No females were genotyped. In PS, the best 200 genotyped males, ranked on genomic EBV (GEBV), were progeny-tested to finally select the 20 best 5-year-old sires for reproduction. In JS, 20 1-year-old males were selected among genotyped candidates using GEBV and used as active sires. To have reference populations of sufficient size, 180 additional males were selected each year and used outside the breeding population. All scenarios were successively simulated using FG and SIT with identical GEBV accuracies. ΔG predicted over the last 10 years using SIT was higher than FG prediction by 9%, 15% and 18% for PT, PS and JS schemes, respectively. In FG simulations, true genomic ΔF were underestimated by using pedigree-based estimates by 9%, 10% and 18% in PT, PS and JS schemes, respectively. Pedigree-based ΔF estimated in SIT simulations were close to those obtained with FG approach for PT (+6%) and PS (-7%) schemes but higher (+38%) for JS scenario. Hence, SIT approach overestimated ΔG realised in both GS schemes and overestimated ΔF specifically in JS scenario.
**P-155**
**IMPACT OF SEQUENCING DESIGN ON MISSING MARKER IMPUTATION AND GENOMIC SELECTION IN CATTLE**

*Tom Druet¹, Iona MacLeod²,³, Ben Hayes²*

¹Unit of Animal Genomics, University of Liège, Liège, Belgium, ²Department of primary industries, Victoria, Australia, ³Melbourne School of Land and Environment, University of Melbourne, Victoria, Australia

Rapid advances in genotyping technologies have allowed implementation of genomic selection in livestock species with outstanding results, particularly in dairy cattle. The accuracy of genomic predictions, and the persistence of accuracy of these predictions over time, could potentially be improved with whole genome sequencing to identify causative mutations affecting the target traits. At present the cost of whole genome sequencing is such that only a subset of individuals from a population can be sequenced. Here we study the impact of selection of animals for sequencing, and depth of sequencing, on marker imputation and genomic selection in a Belgian Blue cattle population. A set of ~1000 bulls was selected and data was simulated with Fregene to match a realistic evolution of efficient population size, and the Belgian Blue pedigree. Individuals were all genotyped at a marker density of 20 SNPs / Mb. Three different methods to select bulls for sequencing were compared, maximizing either 1) the expected average sequenced portion of the genome of individuals from the population (selecting founders with important contribution to the population), 2) the expected number of independent genomes sequenced and 3) the coverage of different haplotypes observed in a population. Designs with equal total cover (number of sequenced bulls x cover) were compared. Imputation accuracy, of the SNP from the sequence, was measured in the whole genotyped population or only in sequenced individuals. Then, genetic values were simulated by selecting a subset of SNP having an effect sampled from a double exponential distribution. A second scenario was also simulated where only SNPs with a MAF < 0.10 had a phenotypic effect. The results showed that selecting individuals having a broad range of different haplotypes improves accuracy and that the optimal cover was between 6 and 12x. Studies on the precision of genomic selection are still ongoing.

**P-156**
**REML IMPLEMENTATIONS FOR SINGLE AND MULTI-TRAIT GENOMIC PREDICTION**

*Marcos Malosetti, Martin Boer, Fred Van Eeuwijk*

Wageningen University, Biometris, Wageningen, The Netherlands

The advent of molecular marker technology has equipped breeders with information that can improve breeding decisions. Typically, markers have been employed in detecting QTLs, which in turn can be used to predict the genetic value of breeding lines. Originally, the main statistical problem in relation to QTL mapping was model selection: select a handful of markers from a larger set, and use those as predictors in (usually) a linear model. However, the increasing number of markers has made this task increasingly difficult.

Genomic prediction appears as an interesting alternative because it avoids the problem of model selection by using all markers in the model. The animal breeding field has been particularly active in this area, e.g. the seminal work by Meuwissen et al. (2001). However, the much larger number of predictors in relation to the number of observations calls for statistical methods that impose some penalty on the model parameters. The Bayesian framework has been particularly useful in this respect with methods as Bayes A and Bayes B as popular examples. More recently, semi-parametric models, reproducing kernel Hilbert space regression models have been proposed (De los Campos et al., 2009). In this paper, we discuss non-Bayesian implementations for genomic prediction, within a mixed model formulation in a similar spirit to Piepho (2009). Essentially, the models discussed here use a function of a marker based relationship matrix to constrain the distribution of random genotypic effects. We discuss how the models can be extended to multiple traits (or multiple environments), in which case also genetic correlations
between traits (or environments) should be modelled. We illustrate the case using a published wheat multi-environment data set (Cossa et al. 2010).

**P-157**

**GENETIC ADMIXTURE AND RISK OF COLORECTAL SPORADIC ADENOMAS: EVIDENCE OF THE ASSOCIATION OF EUROPEAN ANCESTRY AND CANCER PRECURSOR'S LESION IN LATINO POPULATIONS**

Gustavo Hernandez-Suarez1,2, Carolina Sanabria1,4, Martha Serrano4,1, Patria Rojas1, Adriana Umana4, Jesus Perez5, Oscar Herran1, Myriam Sanchez1, Luis E. Bravo5, Jovanny Zabaleta3, Albert Tenesa2

1National Cancer Institute, Bogota, Colombia, 2Roslin Institute, Edinburgh, UK, 3Louisiana State University, Louisiana, NO, USA, 4Universidad Nacional, Bogota, Colombia, 5Universidad Industrial de Santander, Bucaramanga, Colombia, 6Universidad Libre, Barranquilla, Colombia, 7Universidad del Valle, Cali, Colombia

**Objective:** To evaluate the association of genetic ancestry and the main precursor lesion to most colorectal cancers, sporadic adenoma, in an admixed population with European, African and Native American origins.

**Methods:** A multicentre case-control study recruited 130 patients with colorectal adenomas and 264 healthy controls attending GP practices in five Colombian cities. Participants responded a complete standardized epidemiologic questionnaire (including height and weight measurements), a food frequency questionnaire and donated a blood sample. For SNP analysis we used a commercial microarray (Illumina Cancer Panel) that consists of 1,421 screened and validated SNPs from 503 genes (covering all chromosomes) known to be involved in cancer etiology according to the SNP500Cancer project. Fractions of ancestral origin for each individual were estimated using the STRUCTURE software assuming three distinct population origins, and using available genotyping information from HapMap reference populations: African (LWK) European (CEU) and Asian (CHB). Quality control and conditional Multinomial Liner regression analysis was done using PLINK and STATA software.

**Results:** 1,196 SNPs (84%) and 353 (89%) individuals analyzed did pass our quality control protocol. 678 SNPs cross-matched with the HapMap dataset, were not in LD, and were used to estimate the proportions of different ancestries for each individual. The mean proportion of European ancestry was significantly different between cases and controls (Controls 0.39, SE=.007; Cases mean=0.44, SE=0.012, p=0.0001). After controlling for sex, age, city of enrolment, education level, BMI and energy intake, the net European ancestry proportion (i.e.: European minus African minus Asian Ancestry proportion) was positively associated with adenoma (OR=5.61, CI95%=1.65-19.03). In the same model, Education level showed a significant positive trend association with adenoma (p for trend<0.01).

**Conclusion:** We show for the first time that the proportion of genetic European ancestry is a strong predictor of colorectal adenomas in an admixed population as Colombians and this is independent of socioeconomic status.

**P-158**

**USE OF HIGH DENSITY SNP IN GENOMIC EVALUATION IN A DUTCH HOLSTEIN-FRIESIAN POPULATION**

Chyslaine Schopen, Chris Schrooten

CRV BV, Arnhem, Gelderland, The Netherlands

The objective of this study was to reduce the number of HD SNP data to reduce computer requirements and to see the effect on the reliability of genomic breeding values using the reduced HD dataset. DNA was isolated from semen samples of 548 bulls of the Eurogenomics consortium. These bulls were genotyped with the BovineHD Beadchip (777k SNP) and used to impute all genotyped Holstein-Friesian animals from 50k to HD, using the Beagle software package. The final dataset consisted of 30,483 animals
and 603,145 SNP genotypes. For each locus, a haplotype score was obtained from Beagle. Three subsets of variable sizes (38,355, 115,690 and 322,360 loci) were made based on deleting obsolete loci (loci that do not give extra information compared to the neighbouring locus). Error of imputation from 50k to HD was assessed by masking genotypes for SNP that are not on the BovineSNP50 in a subset of 60 animals, and impute those based on the remaining HD genotypes. Subsequently, a validation study using the haplotypes was performed to see the effect on reliability of genomic breeding values for nine traits (production, conformation and functional traits).

Average imputation error was 0.784%. The dataset of 115,690 loci gave the highest increase (0.10% to 0.60% compared to other subsets) in reliability based on all validation bulls, averaged across all traits. Using only old validation bulls, the increase was 0.90% to 1.70% compared to other subsets, on average. Eliminating obsolete loci will enormously decrease computation time to estimate genomic breeding values based on HD data. The more HD loci used the higher the increase in reliability of genomic breeding value. The reduced dataset resulted in an increase in reliability of genomic EBV of 1 to 2% compared to the routinely used SNP dataset (~38k SNP).

P-159

LOW PREDICTIVE ABILITY OF LOW PENETRATION GENETIC VARIANTS FOR BLADDER CANCER RISK AMONG NON-SMOKERS

Evangelina López de Maturana1, Douglas Mata2, Jesús Herranz1, Nathaniel Rothman3, M Luz Calle4, Antoni C Picornell1, Montserrat García-Closas2, Manolis Kogevinas5,6, Alfredo Carrato7,8, Adonina Tardón9, Francisco X Real1,10, Debra Silverman3, Stephen J Chanock3, Núria Malats4

1Spanish National Cancer Research Centre, Madrid, Spain, 2MD Anderson Cancer Center, Houston, TX, USA, 3National Cancer Institute, Bethesda, MD, USA, 4Universitat de Vic, Vic, Spain, 5Centre for Research in Environmental Epidemiology, Barcelona, Spain, 6Institut Municipal d’Investigació Mèdica – Hospital del Mar, Barcelona, Spain, 7Hospital Universitario de Elche, Elche, Spain, 8Hospital Universitario Ramon y Cajal, Madrid, Spain, 9Universidad de Oviedo, Oviedo, Spain, 10Universitat Pompeu Fabra, Barcelona, Spain

This study represents a first attempt, up to our knowledge, to build a predictive model for bladder cancer (BC) risk, combining both genetic and non-genetic data. Bladder cancer is a complex disease that has been associated with several risk factors, being cigarette smoking the most important one. In addition, recent GWAS have identified common variants associated with bladder cancer and reported possible interactions between two of them and tobacco. To eliminate the smoking effect, we focused here on non-smoker individuals to estimate the contribution of the whole genome information to predict risk for BC.

We studied 458 non-smoker subjects from the Spanish Bladder Cancer/EPICURO study, a hospital based case-control study. Individuals were genotyped using HumanHap 1M-probe chip and with TaqMan to assess the presence of GSTM1 gene. After quality control and LD filters, genotypes from 475,289 SNPs and GSTM1 were used. Three Bayesian LASSO threshold models were performed: 1) model including only whole genome information, 2) model including only non-genetic data (age, region and gender) information and 3) model combining both sources of information. The area under the ROC curve was used to evaluate the predictive ability of each model in a 10-fold cross-validation scenario.

Similarly to what has been recently reported for other traits, the model including only environmental covariates outperformed that including only genome information (0.63 vs. 0.56). When adding genome information, the predictive ability did not improve (~0.63). In order to alleviate the potential overfitting, a pre-selection of SNPs was performed, resulting in no improvement of the prediction.

The individual’s unrelatedness, the incomplete penetrance and modest effects of SNPs, as well as the reduced sample size resulting in low statistical power, are among the explanations of the low predictive ability of genetic data for BC risk found in this study.
P-160
PRIORITIZATION OF THE BAYESIAN ALPHABET FOR GENOMIC SELECTION THROUGH ESTIMATING HYPERPARAMETERS

Jicai Jianfeng, Qin Zhang, Jianfeng Liu
1China Agricultural University, College of Animal Science and Technology, Beijing, China, 2Key Laboratory of Animal Genetics, Breeding and Reproduction, Ministry of Agriculture, Beijing, China

In the original Bayesian alphabet for genomic selection, hyperparameters involved in the prior distributions of SNP effects and corresponding variance should be specified ahead of MCMC analyses for estimation of SNP effects. This may be a potential factor influencing accuracy and unbiasedness of genomic estimated breeding values (GEBVs). In the present study, a novel strategy for extension of the original Bayesian alphabet is proposed to address this concern. Specifically, the priors of $\nu$ (the degrees of freedom of the prior distribution of variance of SNP effects) for BayesA, BayesB and BayesC were all assumed to be exponentially distributed, and $\sigma^2$ (scale parameter of prior distribution of variance of SNP effects) was assumed to be proportional to its reciprocal for BayesA and follow a Gamma(1,1) prior for BayesB and BayesC. The probability $\pi$ that a SNP has zero effect in the model was treated as unknown with uniform(0,1) prior for BayesB and BayesC. Accordingly, MCMC was employed to solve the model. The posteriors of $\sigma^2$ and $\pi$ can be standardized as gamma and beta distributions respectively and easily sampled, and $\nu$ can be sampled via an accept-reject algorithm as there is no clear expression for its posterior. Both simulation and real data analyses demonstrated that the extension of BayesA and BayesB generally performed better regarding the accuracy and unbiasedness in contrast to the original Bayesian methods, and the accuracy and unbiasedness of extension to BayesC are similar to those of BayesC. In conclusion, accounting for the accuracy and unbiasedness of GEBV and taking a fundamental approach to $\pi$, our novel strategy for extension of BayesA, BayesB and BayesC has significant advantages over the traditional Bayesian approaches for routine applications in genomic selection.

P-161
GENOME LOCATION SPECIFIC PRIORS IMPROVES GENOMIC PREDICTION: BAYESRS

Rasmus Froberg Brøndum, Guosheng Su, Mogens Sandø Lund, Phillip J. Bowman, Michael E. Goddard, Benjamin J. Hayes
1Department of Molecular Biology and Genetics, Aarhus University, Tjele, Denmark, 2Biosciences Research Division, Department of Primary Industries Victoria, Bundoora, VIC, Australia, 3Faculty of Land and Food Resources, University of Melbourne, Parkville, VIC, Australia

The accuracy of genomic prediction is highly dependent on the size of the reference population. For smaller populations strategies to include information from other populations could improve predictions. BayesRS is a novel approach for genomic predictions with sharing of information across populations. Proportions of a four-distribution mixture for SNP effects are derived in segments of fixed size along the genome in one population and set as location specific prior distributions for proportions of SNP effects in the other population. The hypothesis is that insufficient persistence of LD phase between populations will delude the advantage of a simple pooling of data for genomic prediction, and that information from other populations can be used as a prior instead, in order to improve predictions. The model was tested using 540 Australian Jersey Bulls, 2297 Australian Holstein bulls and 5214 Nordic Holstein bulls. The traits studied were protein-, fat- and milk yield. Genotypic data was Illumina 777K SNPs, real or imputed. Results showed an increase in accuracy of 2% for fat yield in the Jersey population when using BayesRS with a prior derived from Australian Holstein compared to combining reference populations, but decreases were seen for protein and milk yield. An increase of 1-2% for all traits was observed in the Australian Holstein population when using a prior derived from the Nordic Holstein population compared to using no prior information. Results show that for some traits the method would be advantageous compared to pooling of reference data for distantly related populations. For closely
related populations the method does not improve on pooling of the data but it does offer an increased accuracy without an increased computational burden, and provides a general setup for inclusion of location specific priors. For example the approach could be used to include biological information in genomic predictions.

P-162
THE POTENTIAL OF GENOMIC PREDICTION FOR BOAR TAINT REDUCTION IN PIGS

Boris Lukic1,2, Ricardo Pong-Wong2, Suzanne Rowe2, Dirk-Jan de Koning2,3, Ingela Velander4, Chris S. Haley2,5, Alan L. Archibald2, John A. Woolliams1

1Faculty of Agriculture, University of J.J. Strossmayer, Osijek, Croatia, 2The Roslin Institute and R(D)SVS, University of Edinburgh, Roslin, Easter Bush, Midlothian, UK, 3The Swedish University of Agricultural Sciences, Uppsala, Sweden, 4Danske Slagterier, Copenhagen, Denmark, 5MRC Human Genetics Unit, MRC IGMM, University of Edinburgh, Edinburgh, UK

Androstenone and skatole accumulate in the fat of mature non-castrated male pigs and is detected as boar taint - an offensive odour affecting the smell and taste of cooked pork. Genetic variation exists for their concentrations and so selection may provide a more sustainable solution than castration for reducing taint in commercial pigs. Genetic progress is limited by the inability to measure taint prior to selection in young animals. The study objectives were to investigate the potential value of genomic selection methodology for predicting breeding values of androstenone and skatole within a commercial Danish Landrace population. Five hundred male pigs with high skatole (>0.3 μg/g) at slaughter were matched with low skatole male litter mates and measured for androstenone. Phenotypes for both traits were corrected for farm as a fixed effect and age as a covariate prior to genetic analysis. Genotyping was performed using the Illumina SNP60 porcine beadchip. After QC (call rate >0.95, MAF >0.01), 42,916 SNPs and 938 individuals remained. A 5-fold cross-validation study (~751 training and ~187 validation animals per group) was carried out to compare the accuracy of GBLUP, Bayes A, Bayes SSVS and Bayesian Lasso to predict the unobserved phenotypes. Accuracy was calculated as the correlation between the estimated genomic breeding value and phenotype of validation animals. The range of accuracies obtained by different methods was narrow for androstenone, between 0.29 (Bayes SSVS) and 0.30 (Bayes A), but wider for skatole, between 0.21 (GBLUP) and 0.26 (Bayes SSVS). Relative accuracies corrected for h2, were 0.55-0.56 and 0.77-0.96 for androstenone and skatole respectively, although these are biased by the selection of the cases and controls. Increasing the size of the training set should further improve genomic prediction, especially for androstenone. The results demonstrate that genomic selection has great potential for reducing boar taint in commercial pigs.

P-163 – ABSTRACT WITHDRAWN

P-164
POTENTIAL OF GENOMIC SELECTION FOR TRAITS WITH A LIMITED NUMBER OF PHENOTYPES

Ilse van Grevenhof, Piter Bijma, Johan van Arendonk
Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands

Genomic selection is a method to estimated breeding values from genome-wide marker information. Simulations and results of real data suggest that breeding values can be predicted with high accuracy using genetic markers alone. To reach high accuracies, however, large reference populations are needed to estimate marker effects. Often these cannot be realized when traits are difficult or expensive to record, or when population size is small. The value of genomic selection (GS) becomes questionable then. In this study, we compare traditional breeding schemes based on own performance or progeny
information to genomic selection schemes, when the number of phenotypic records is limiting. For this
goal, deterministic simulations were performed using selection index theory. Results showed that genomic
selection schemes suffer more from the Bulmer effect than schemes with selection based on phenotypes,
but that the Bulmer-effect is the same as with selection EBVs estimated with BLUP. To maximize
the accuracy of genomic EBVs when the number of phenotypic records is limiting, the phenotyped
individuals, rather than progeny tested individuals, should be genotyped. When the generation interval
cannot be decreased when implementing GS, large reference populations are required to obtain
equal response as with own performance selection or progeny testing. The accuracy of genomic EBVs,
however, increases non-linearly with the size of the reference population, showing a diminishing-return
relationship. As a consequence, when a GS scheme has a small decrease in generation interval, relatively
small reference population sizes are needed to obtain equal response as with own performance selection
or progeny testing. When the trait of interest cannot be recorded on the selection candidate, GS schemes
are very attractive even when the number of phenotypic records is limited, because traditional breeding
would have to rely on progeny testing schemes with long generation intervals.

P-165
WHICH ANIMALS ARE TO BE DENSELY GENOTYPED IN ORDER TO IMPUTE THE MISSING
GENOTYPES OF SPARSELY GENOTYPED ANIMALS?

Xijiang Yu1, John Woolliams2, Theo Meuwissen1
1Norwegian University of Life Sciences, Ås, Norway, 2Edinburgh University, Edinburgh, UK

Genotyping accounts for a substantial part of genomic selection costs. Using both dense and sparse
SNP chips, together with imputation of missing genotypes, can lower the total genotyping costs. This
study is to rank the candidates for their importance for dense genotyping, when they are used to impute
the missing genotypes of the sparsely genotyped animals. The accuracy of the imputation results were
investigated. We used a real pig pedigree of 13,276 individuals and 12 generations. The missing
genotypes of the animals in the last 3 generations were imputed. These 3,772 animals were called the
target. Candidates to be densely typed were chosen from the whole pedigree. Three lists of animals to
be densely genotyped were compared, each with 200 pigs: (1) sires with most offspring; (2) ancestors
with the largest genetic contributions to the target; (3) candidates that minimise the conditional genetic
variances of the target animals, where the genetic variance matrix is defined proportional to the numerator
relationship matrix and this variance matrix is conditioned on the set of animals that are densely
genotyped.

Ideal populations were simulated with 10,000 generations of random mating. Assuming a mutation rate
of $10^{-8}$/nucleotide/meiosis and effective population size of 200, 400, or 1,000, around 5,000, 11,000, or
33,000 polymorphic SNP loci were generated, respectively. The alleles were then dropped through the
founders into the pig pedigree. Two types of sparse chips were tested, with 100 or 500 randomly chosen
loci. The remaining loci were to be imputed in the target population by the Beagle software. Results
show that minimizing the conditional variance gave equally accurate results as dense genotyping the 200
sires. List (2) gave the highest error rates, probably because these ancestors were rather old and thus did
not share large unrecombined chromosome segments with the target population.
P-166
COMPARISON OF METHODS TO IMPUTE MISSING SNP MARKERS IN NORDIC RED DAIRY CATTLE POPULATION
Peipei Ma1,2, Rasmus Frøberg Brøndum1, Qin Zhang1, Mogens Sandø Lund1, Guosheng Su1
1Department of Molecular Biology and Genetics, Faculty of Science and Technology, Aarhus university, Tjele, Denmark, 2Department of Animal Genetics, College of Animal Science and Technology, China agricultural university, Beijing, China

Genomic selection has been widely used in cattle breeding. Various SNP panels have been developed to meet different requirements of the customers. In many countries, individuals in a cattle population have been genotyped using different panels. To use the SNP data from different panels efficiently for genomic prediction, it is necessary to impute the markers which are absent from a SNP panels but exist in the other panels.

In this study, we compare different imputation methods with regard to imputation accuracy in different scenarios, based on the data of Nordic Red Cattle population. The softwares of Beagle, findhap and Impute2 are applied in two data sets. One is imputed from 3k marker data to 54k marker data, the other is from 54k data to 777k data. Alphalmpute and FImpute are also used in the dataset which is imputed from 3K to 54K. Imputation accuracy is assessed in four groups of validation animals which differ in their relationship with reference animals.

As a whole, imputation accuracy is higher when imputing 54k data to 777k data than imputing 3k data to 54k data. The methods using pedigree information are more sensitive to the relationship between validation animals and reference animals, compared with the methods without using pedigree information. Imputation accuracy is lower for rare alleles than common alleles.

P-167
PERFORMANCE OF DIFFERENT MACHINE LEARNING METHODS AT PREDICTING TRAITS WITH DIFFERENT GENETIC ARCHITECTURE IN THE SPANISH HOLSTEIN POPULATION
Jose Antonio Jiménez-Montero1, Oscar González-Recio2, Rafael Alenda1
1UPM, Madrid, Spain, 2INIA, Madrid, Spain

The objective of this study was to develop and test the first genomic evaluation procedure in the Spanish Holstein population for production and type traits. Five complex traits were included: milk yield (MY), fat yield (FY), protein yield (PY), fat percentage (FP) and udder depth (UD). The progeny adjusted MACE proofs were used as response variables. Bulls were genotyped using the BovineSNP50 Illumina beadchip. After editing a total of 39,714 remained SNPs were included in the model. A two-fold cross-validation scenario was created, with a training set containing 1576 progeny tested bulls born before 2005. Three methods were compared: Bayesian LASSO (BL), L2 Boosting and Bayes A. Their predictive ability was compared in a validation set containing 221 bulls born between 2005 and 2007 that did not have progeny test in 2009. Pearson correlation, bias and MSE were calculated between predicted response and progeny adjusted MACE proofs from December 2011. Genomic predictions showed an increment in accuracy regarding pedigree index of 51%, 58%, 26%, 94% and 24% for MY, FY, PY, FP and UD, respectively. The three methods achieved similar accuracy and MSE. LASSO was the most accurate method for MY (0.59), FY (0.66) and PY (0.58) whereas L2 Boosting showed higher accuracies for FP (0.79) and UD (0.58). BL showed lower MSE for MY, whereas Bayes A was preferable for FY, FY and FP, and Boosting for UD.

Development will continue on the reported traits and will be extended to the rest of traits included in the Spanish genetic evaluations. The collaboration in the EUROGENOMICS consortium, including a reference population of above 22,000 progeny tested bulls is expected to increase notably the accuracy of genomic evaluations as well as reduce the observed bias.
P-168 – ABSTRACT WITHDRAWN

P-169
MIXED MODEL COMPRESSION FACILITATES GENOMIC PREDICTION

Zhengkui Zhou2,3, Zhe Zhang4,5, Meng Li6,1, Gang Guo7, Qin Zhang4, Yachun Wang4, Junya Li3, Peter Bradbury5, Rory Todhunter1, Edward Buckler5, Zhiwu Zhang1
1Cornell University, Ithaca, New York, USA, 2Chinese Academy of Sciences, Beijing, China, 3Chinese Academy of Agricultural Sciences, Beijing, China, 4China Agricultural University, Beijing, China, 5South China Agricultural University, Guangzhou, China, 6Nanjing Agricultural University, Nanjing, China, 7Beijing Sanyuan Breeding Technology Co., Ltd, Beijing, China, 8United States Department of Agriculture (USDA), Ithaca, New York, USA

Dense genetic markers from new genome sequencing technologies make it possible to predict an individual’s genetic merit at an early life stage. This could lead to better health care management and faster breeding cycles for agricultural production. The challenge is that the genetic architectures of complex traits impact the genomic prediction accuracy when using current statistical methods. The conventional kinship method fits well only those traits with a large number of underlying causative mutations. While Bayesian methods fit the opposite situation well when fewer mutations underlie the trait, they are computationally intensive and require a long time to fit models with many genetic markers. We developed a new genomic prediction method based on the compressed mixed linear model that was originally developed for genome wide association studies. The new method is more accurate than both the kinship and Bayesian methods regardless of the underlying genetic architecture. The computing time of the new method remains similar to the kinship method. The method is implemented in an R package (GAPIT) which is free for the public at http://www.maizegenetics.net/gapit.

P-170 – ABSTRACT WITHDRAWN

P-171
A COMPARISON OF STATISTICAL METHODS FOR GENOMIC SELECTION IN A MICE POPULATION

Sandra Aider de Queiroz1, Haroldo Henrique de Rezende Neves2, Roberto Carvalheiro2
1Departamento de Zootecnia, FCAV, Unesp, Jaboticabal, SP, Brazil, 2Gensys Consultores Associados S/C Ltda, Porto Alegre, RS, Brazil

The objective of this research was to compare the predictive performance of ten different statistical methods employed in genomic selection, by using public data from a heterogeneous stock mice population kept by Welcome Trust Centre for Human Genetics. Phenotypes (weight at six weeks, growth slope and body length) and genotypes for 9,917 SNP markers on 1884 animals were analyzed. Data were split in training and testing sets according to two strategies, such that within and across-family predictions were assessed. The following methods were employed to estimate marker effects: ordinary least squares (OLS), ridge regression assuming all markers contribute equally to the additive variance (GBLUP), subset selection (four methods defined by using either Bonferroni correction or false-discovery rate, either treating markers as fixed or random effects), principal components regression (PCA), Random Forest, fast BayesB and LASSO. Predictive ability was assessed as the Pearson’s correlation between predicted and observed phenotypes in the validation set (average of 10 replicates). Within-family predictions were more accurate than across-family predictions. GBLUP had the highest predictive ability in most of the situations investigated and resulted in higher predictive ability than a classical polygenic model, except in the case of within-family predictions for weight. In the case of across-family predictions for body length, the methods Random Forest and PCA had predictive ability about 7% and 11.7% greater than GBLUP. As
Poster Abstracts

a general rule, GBLUP outperformed other methods of estimation of marker effects in terms of predictive performance for the traits investigated and there was evidence that a fraction of this predictive ability derive from genetic relationships among genotyped animals. This superiority of GBLUP may suggest that many loci contribute to genetic variation of weight, body length and growth slope in this mice population.

P-172
SYNBBREED: A FRAMEWORK FOR THE ANALYSIS OF GENOMIC PREDICTION DATA USING R
Valentin Wimmer, Theresa Albrecht, Hans-Jürgen Auinger, Chris-Carolin Schön
Technische Universität München, Freising, Germany

High-throughput genotyping and large scale phenotyping produce massive amounts of data. We present a novel R package named synbreed for the analysis of such data to derive genome-based predictions. It contains a comprehensive collection of functions required to fit and validate genomic prediction models in plant and animal breeding. This covers data processing, visualization and analysis. Thereby, a versatile analysis pipeline is established within one software package. All functions are embedded within the framework of a single, unified data object. The implementation is flexible with respect to a wide range of data formats and models. The package fills an existing gap in the availability of user-friendly software for next-generation genetics research and education. Where necessary, the package provides gateways to other software programs to extend the field of applications. The utility of the package is demonstrated using three large-scale example data sets provided by the synbreedData R package: a simulated data set representing a maize breeding program, a publicly available mice data set and a dairy cattle data set.

P-173
ESTIMATION OF ADDITIVE, EPISTATIC AND DOMINANCE GENETIC VARIANCES AND PREDICTION OF GENETIC MERITS USING GENOME-WIDE MARKERS
Guosheng Su1, Ole F. Christensen1, Tage Ostersen2, Mark Henryon2,3, Mogens S. Lund1
1Aarhus University, Tjele, Denmark, 2Pig Research Centre, Danish Agricultural & Food Council, Copenhagen, Denmark, 3University of Western Australia, Crawley, Australia

Genome-wide markers have been widely used to study genetic architecture and predict breeding values of complex traits. Currently, genetic analysis and genomic prediction based on dense marker data are usually carried out using the statistical models that ignore non-additive effects. However non-additive genetic effects may have an important contribution to total genetic variation of complex traits in human, wild life, model organisms or farm animals. This study extended the genomic relationship matrices to include non-additive genetic relationships, estimated additive, epistatic and dominance genetic variances of daily gain, and assessed the accuracy and unbiasedness of genomic predictions using models with and without non-additive genetic effects in Danish Duroc pigs. Four linear models were used in this study: 1) a simple additive genetic model (MA), 2) a model including both additive and epistatic genetic effects (MAE), 3) a model including both additive and dominance genetic effects (MAD), and 4) a full model including all three genetic components (MAED). Additive, epistatic and dominance genetic relationship matrices were constructed using genome-wide SNP markers. Estimates of narrow-sense heritability were 0.397, 0.373, 0.379 and 0.357 for models MA, MAE, MAD and MAED, respectively. Estimated dominance variance and epistatic variance accounted for 5.6% and 9.5% of the total phenotypic variance, respectively. Based on model MAED, the estimate of broad-sense heritability was 0.506. Reliabilities of genomic predicted breeding values for the animals without performance records were 28.5%, 28.8%, 29.2% and 29.5% for models MA, MAE, MAD and MAED, respectively. In addition, models including non-additive genetic effects reduced bias of genomic predictions.
P-174
COPY NUMBER VARIATIONS ASSOCIATED WITH RESISTANCE/SUSCEPTIBILITY TO PARASITES IN CORRIEDALE SHEEP

Nicolás Grasso¹, Gabriel Ciappesoni², Daniel Castells², Diego Gimeno², Virginia Goldberg¹, Wanda Iriarte¹, Gonzalo Rincón³
¹INIA, Canelones, Uruguay, ²SUL, Montevideo, Uruguay, ³UC-Davis, California, USA

Gastrointestinal parasites (GIPs) are a serious limitation affecting sheep production worldwide. Anthelmintics are rapidly becoming ineffective due to selection pressure on parasite populations that increase resistance to them. An option for stud breeders is the selection of resistant animals to GIPs by the evaluation of the Expected Progeny Differences (EPD) for Faecal Worm Egg Count (FEC). However, infected animals are required and FEC implies a complex procedure, making resistance/susceptibility (R/S) to GIPs a difficult trait to evaluate. Therefore, the availability of genetic markers would simplify assessment of candidates for selection in breeding programs.

Several authors have reported relationships between copy number variations (CNV) and various diseases, and its possible association to complex traits in many different animals.

A total of 54 R and 44 S Corriedale sheep that belong to divergent selection lines for R/S to GIPs were selected by FEC EPD (R<-0.3; S>0.3 LnFEC respectively) and average relatedness coefficient (<4%). The 98 DNA samples were genotyped using the Illumina OvineSNP50 BeadChip. Genotyping intensity files (.idat) were used to carry out the quality control (genotype calls, error rates, P50 GenCall score) and signal intensity files were calculated for each SNP (expressed as Log2 Ratios) with the Genome Studio software (Illumina Inc.). The log2 R intensity files were used to perform a case/control CNV association study with SVS7 Golden Helix software.

The CNV analysis revealed 5 genomic regions showing significant differences between R and S animals over OAR1, OAR12, OAR15, OAR16 and OAR24. The effectiveness of these CNV regions to separate R/S to GIPs is going to be validated using real-time quantitative PCR in 1500 Corriedale sheep.

P-175
PREDICTED GENETIC GAIN FROM USE OF GENOMIC SELECTION AND FOREIGN GENETIC EVALUATIONS IN A SMALL HOLSTEIN POPULATION

Hugo H. Montaldo¹, Adriana García-Ruiz¹, Felipe de Jesús Ruiz-López², Sergio Iván Román-Ponce²
¹Facultad de Medicina Veterinaria y Zootecnia, UNAM, D.F., Mexico, ²Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, SAGARPA, Querétaro, Mexico

The purpose of this study was to predict genetic gain from a breeding program that incorporates genomic selection (GS) and foreign sires’ information for genetic improvement of an economic breeding goal with a heritability of 0.16, in a small Holstein population using young sires. The genetic parameters and population characteristics used were those of the recorded Mexican Holstein population. Deterministic modelling methods based on selection index (SelAction software) were used. Accuracy for sires’ genomic evaluations were obtained as functions of the number of animals in the reference population, assuming an effective population size of 100, a genome size of 30 Morgans and a 50K SNP chip which capture 0.80 of the total genetic variance. As the number of proven sires available was small (500), and their average accuracy was low (0.51), data from 500 to 10000 cows genotyped and with evaluations based on phenotype with an average accuracy of 0.4, were incorporated into the reference population. Average generation intervals for sires and cows were 2.00 and 4.94 years. Incorporation of foreign information (genetic correlation between countries=0.80) from 100 half sibs, or of local phenotypic information into sires genomic selection indices, increased only marginally genetic gains. Genetic gain per year of 0.28 genetic standard deviations (σg) were possible using GS of the best 20 out of 2000 sires with 10000, 0.26 σg with 5000, and 0.24 σg with 2000 cows in the reference population. Genetic gains depended mainly on the number of cows in the reference population and sire selection intensity. Genetic gain per year for a
Poster Abstracts

GS program based on 5000 cows in the reference population would be 202% that obtained from a local young sire selection program based on phenotypic information only, and 130% that obtained from the use of sires from a foreign conventional progeny testing program.

P-176
POWER AND ACCURACY OF POLYGENIC SCORES FOR COMPLEX TRAITS
Frank Dudbridge
London School of Hygiene and Tropical Medicine, London, UK

Polygenic scores have recently been used to infer the presence of genetic effects among an ensemble of markers that do not individually achieve significance in a large scale association study. Markers are selected from an initial discovery study and used to construct a score in an independent replication sample by summing the number of trait increasing alleles for each subject, perhaps weighted by their estimated effect sizes. Association between a trait and this composite score implies that a genetic signal is present. This approach has been used to obtain evidence of a genetic effect when no single markers are significant, to establish a common genetic basis for related disorders, and to construct risk prediction models. In some cases, however, the desired association or discrimination has not been achieved by the polygenic score. In this work, the power of polygenic score analysis is derived from the quantitative genetics model as a function of the sizes of the two samples, bivariate heritability of the trait, selection thresholds for including a marker in the score, and method for weighting effect sizes in the score. Expressions are derived for quantitative and discrete traits, the latter allowing for case/control sampling. It is shown that published studies with significant association of polygenic scores have been well powered, whereas those with negative results can be explained by low sample size. Expressions are also obtained for correlation and area under the ROC curve for genetic predictors estimated from a finite sample. Here it is shown that the theoretical maximum accuracy, known as a function of the heritability, can only be approached when the predictor is estimated from a sample that is orders of magnitude larger than those currently available. Therefore, polygenic scores currently have more utility for association testing than for predicting complex traits.

P-177
THE EFFECT OF LINKAGE DISEQUILIBRIUM, HAPLOTYPES AND FAMILY RELATIONSHIPS ON THE ACCURACY OF DIRECT GENOMIC VALUES
Yvonne Wientjes1,2, Roel Veerkamp1, Mario Calus1
1Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, Lelystad, The Netherlands, 2Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands

Our objective was to investigate effects of linkage disequilibrium (LD), haplotypes and family relationships on accuracy of direct genomic values. Accuracies were predicted using genomic relationships among the 529 cows of the reference population (RP) and their relationships with the selection candidates. Four groups of selection candidates were used, of which genotypes for three groups were simulated using increasing similarity with RP: 1) allele frequency of RP (FREQ); 2) allele frequencies and LD pattern of RP (LD); 3) randomly drawing haploid chromosomes from RP (HAP); 4) animals from RP, thereby being the only group with real family relationships to RP (FAM). Accuracy of FAM was calculated using the remaining 528 animals as RP. The number of effective chromosome segments ($M_e$) was estimated based on genomic and pedigree relationship matrices for each scenario. FREQ used the same assumptions as the Daetwyler-formula, therefore $M_e$ was equal to the number of SNPs used. At a heritability of 0.6, calculated accuracies were 0.093±0.003 (FREQ), 0.168±0.006 (LD), 0.355±0.015 (HAP) and 0.577±0.064 (FAM). At a heritability of 0.1, relative differences across groups were similar. Calculated accuracies were very similar to accuracies predicted with the formula of Daetwyler et al. (2008) using estimated $M_e$ for all scenarios. Variance of accuracy of FAM was much higher compared to other scenarios, due to much
higher variances in relationships with animals in RP. Accuracies of FAM were on average more than fifty percent higher than accuracy of HAP. It is concluded that accuracies can be predicted using empirically estimated $M_e$ and that level of relationship with RP has a much higher effect on the accuracy of direct genomic values compared to linkage disequilibrium per se. Furthermore, increasing length of haplotypes shared with RP animals only improves prediction accuracy when it results in increasing relationship with RP.

P-178
GENOMIC PREDICTION USING HAPLOTYPE BLOCKS BUILT FROM HIGH DENSITY MARKER MAP

Beatriz Cuyabano, Guosheng Su, Mogens Lund
Aarhus University, Tjele, Denmark

Genomic selection using moderate density markers (e.g. 50k) has been widely applied in animal breeding. With the advances and developments of biotechnology, high density (HD) marker chips are now available. When performing genomic prediction using these chips, two major challenges arise: the huge amount of explanatory variables contrasting with few genotyped individuals (usually less than 10k) and the high computational demand required by such datasets. The objective of this study is to reduce explanatory variables without loss of information by dividing the high density markers into haplotype blocks, and evaluate the advantage of using such haplotypes as predictors for genomic predictions. The original HD marker map in this study consists of approximately 500k SNP's (Single Nucleotide Polymorphisms). The haplotype blocks are built based on linkage disequilibrium (LD) between markers, requiring a minimum $D'$ of 0.75 between them. Consequently, circa 118k haplotype blocks are built, generating about 265k haplotype variables in the haplotype block map. This indicates that the number of explanatory variables can be reduced by 45% when using the haplotypes for genomic prediction, compared with using SNP genotypes, thus reducing the number of variables whilst creating predictors in higher LD with causative mutations. Models based in haplotype blocks are expected to improve genomic prediction and such models are compared to single marker prediction models.

P-179
LOCALISING GENOMIC VARIANCE OF PRODUCTION TRAITS IN DAIRY CATTLE USING REGIONAL GENOMIC RELATIONSHIP MAPPING

Yoshitaka Nagamine$^{1,2}$, Ricardo Pong-Wong$^2$, Pau Navarro$^3$, Koichi Hagiya$^1$, Takeshi Yamazaki$^1$, Chris Haley$^{2,3}$
$^1$National Agricultural Research Center, Sapporo, Japan, $^2$The Roslin Institute and R(D)SVS, Midlothian, UK, $^3$MRC Human Genetics Unit, Edinburgh, UK

Large genome-wide association studies have been successful at localising many genomic regions influencing complex traits over the last few years through their association with individual genetic markers. However, for most traits only a small proportion of the variation has been pinpointed in this way. In our study we introduce a new method, Regional Genomic Relationship Mapping, that identifies trait variation associated with short regions of the genome containing a number of genetic markers. We used genomic relationship matrices, all SNPs for whole genomic relationship and consecutive 100 to 10 SNPs for regional genomic relationship matrices, in a model. If there are some significant regions, we may add these regional relationship matrices in a model. We have analysed data from 2590 of Japanese Holstein sires for three production traits – milk, fat and protein yield in 305 days. All three traits showed highly significant region at the beginning of chromosome 14 using 100 SNPs for a window. Therefore, we applied smaller size of window using 20 to 10 SNPs for regional relationship matrix and three traits showed the higher significant values almost the same region again. GWAS approach using single SNP could not find significant SNP in this region for protein yield. Regional breeding values for milk and
fat indicated negative correlation (< -0.9). It suggested that why a highly significant region for both milk and fat have been kept through long selection history. Hence our approach may help identify some of the missing genetic variation and improve understanding of relationship between complex traits.

P-180
POTENTIAL OF GENOMIC SELECTION IN PERENNIAL CROPS: PRELIMINARY RESULTS IN THE CONTEXT OF EUCALYPTUS AND OIL PALM BREEDING

David Cros¹, Marie Denis¹, Benoît Cochard¹, Létizia Camus-Kulandaivelu¹, Tristan Durand-Gasselin², Jean-Marc Bouvet²
¹Cirad Research unit AGAP : Genetic improvement and adaptation of Mediterranean and Tropical Plants, France, ²PalmElit SAS, France

The methodology of selection in plant breeding has markedly evolved with the advent of high throughput molecular technology, the increasingly reasonable cost of genotyping, and the implementation of genomic selection (GS). For perennial crops, the potential of GS is high and gives the opportunity to shorten the breeding cycle by selecting at the juvenile stage using marker information. Here we present preliminary results of GS experiments for two perennials crop, Eucalyptus and oil palm, that play an important economical role in tropical regions.

In the case of Eucalyptus, a simulation study was developed to test the efficiency of GS in the frame of a recurrent selection scheme for clone production over four breeding cycles. Scenarios crossing broad sense heritabilities ($H^2$=0.6 and 0.1), dominance to additive variance ratios (R=0.1; 0.5 and 1) and training population structure were compared using Bayesian LASSO method. Models including dominance effects are all the more relevant when the R ratio and the training population size are high. The genetic gain per unit time with GS was 1.5 to 3 times higher than with phenotypic selection at mature stage for breeding and clone populations.

For oil palm, we implemented a cross-validation approach with 111 individuals of the last generation of a key breeding population, evaluated through progeny tests including 40,000 individuals and genotyped with 140 microsatellites. The accuracy of GS increased when increasing the training population size and reached 0.6-0.7, according to the trait, with a 3:1 ratio for training and validation populations respectively. The small effective population size detected in this breeding population explains the good GS performance even with a limited panel of markers.

Our studies based on two perennials crops presenting different biological patterns and different breeding contexts suggest very promising results of GS for long rotation plant species.

P-181
PERSISTENCE OF LINKAGE PHASE BETWEEN THE CHINESE AND NORDIC HOLSTEINS AND GENOMIC PREDICTION USING THE COMBINED REFERENCE POPULATION

Lei Zhou¹, Xiangdong Ding², Qin Zhang², Yachun Wang², Mogens Lund¹, Guosheng Su¹
¹Aarhus University, Viborg, Denmark, ²China Agriculture University, Beijing, China

This study investigated the persistence of linkage phase between the Chinese and Nordic Holsteins, and the gain in reliabilities of genomic predictions for Chinese Holsteins by including Nordic bulls in reference population. 80 Chinese Holstein bulls, 2,091 Chinese Holstein cows and 4,398 Nordic Holstein bulls were used in the analysis. Totally, 41,838 SNPs with minor allele frequencies higher than 0.01, were used in the analysis. Genomic estimated breeding values (GEBV) of Chinese Holsteins were predicted using a single-trait GBLUP model based on Chinese Holsteins reference alone, and using a two-trait GBLUP model based on a combined reference population including both the Chinese and Nordic Holsteins. The persistence of linkage disequilibrium (LD) phase between two populations, measured as correlation of $r$ (a measure of LD between adjacent markers) was 0.97. Comparing the predictions based
on combined reference population with those based on Chinese Holsteins reference population only, reliabilities of GEBV of 48 Chinese bulls increased from 0.154, 0.224 and 0.110 to 0.473, 0.515 and 0.362 for milk yield, fat yield, and protein yield; and reliabilities of GEBV of Chinese cows increased from 0.118, 0.154 and 0.154 to 0.169, 0.256 and 0.198 for the three traits, respectively. In conclusion, the persistence of linkage disequilibrium phase between Chinese and Nordic Holsteins is very high, combining Nordic Holsteins in reference can substantially improve reliabilities of GEBV for both Chinese bulls and cows.

P-182
PREDICTOR TRAITS IMPROVE ACCURACY OF GENOMIC BREEDING VALUES FOR SCARCELY RECORDED TRAITS
Marcin Pszczola1,2, Roel F. Veerkamp1, Yvette de Haas1, Tomasz Strabel2, Mario P. L. Calus1
1Wageningen UR Livestock Research, Lelystad, The Netherlands, 2Poznan University of Life Sciences, Poznan, Poland

For some traits reference populations may have limited size because of phenotyping costs (e.g. methane emission or dry matter intake (DMI) in dairy cattle). Using predictor traits, a trait of interest can be improved at reduced cost. Impact of using predictor traits in a genomic selection approach has not been studied yet on real data. This study aimed to empirically investigate the effect of using predictor traits on accuracy of direct genomic values (DGV) of DMI based on a small cow reference population. Accuracies were obtained in cross-validation by correlating DGV and phenotypes divided by the square root of heritability. Multivariate genomic BLUP was used to simultaneously evaluate DMI, fat protein corrected milk yield (FPCM) and live weight (LW). DMI was the predicted trait with a small number of observations (869). FPCM had 1,520 and LW 1,309 records and were used as predictors. Four cow reference populations were created by increasing the number of traits with phenotypic records for: 1) only DMI; 2) DMI and FPCM; 3) DMI and LW; or 4) DMI, FPCM, and LW. Evaluated animals (EVA) were cows with phenotypes on: 1) no traits, 2) FPCM, 3) LW, or 4) FPCM and LW. Lowest accuracies were obtained when no observations for EVA were available (from 0.32 to 0.33). With FPCM observations for EVA, accuracy increased to 0.50. Replacing FPCM by LW increased accuracy to 0.57. Highest accuracy was achieved for trivariate analyses (0.63), with information for EVA on both FPCM and LW. Estimates showed most bias were when no observations for EVA were available (slopes of regression of phenotype on DGV ranging from 0.68 to 0.73). Estimates were unbiased with multivariate approach (slopes from 0.99 to 1.12). Using predictor traits in multivariate genomic approach is an inexpensive way to considerably increase accuracy of scarcely recorded traits.

P-183
UTILISING WITHIN FAMILY GENETIC VARIATION WITH GENOMIC SELECTION
Samuel Clark1,2, John Hickey1, Brian Kinghorn1, Julius van der Werf1,2
1School of Environmental and Rural Science, University of New England, Armidale, NSW, Australia, 2Cooperative Research Centre for Sheep Industry Innovation, Armidale, NSW, Australia

Genome wide markers are now widely used for the prediction of breeding values and selection in livestock breeding programs. These genomic estimated breeding values are equal to or more accurate than pedigree based estimated breeding values. In animal breeding programs, an increase in genetic merit due to selection has to be balanced with the need to maintain diversity and manage inbreeding such that long term gains can be achieved. Moreover, selection of animals based on pedigree best linear unbiased prediction information alone may increase inbreeding because of a high estimated breeding value covariance among family members, causing the co-selection of relatives. It has been suggested that genomic information can alleviate this problem as it increases the use of within-family variation in the
prediction of breeding value. The effect of using genomic information in optimal selection was examined based on simulated and real data. Parental coancestry was penalised at different levels to observe changes in genetic merit over varying rates of inbreeding. Genomic breeding values increased genetic gain because they were more accurate than pedigree estimated breeding values. Genomic measures of co-ancestry did not increase genetic gain for varying rates of inbreeding under a half-sib family structure. However, when the population consisted of many large full-sib families, there was an advantage for using genomic relationships to constrain inbreeding. In real data from dairy bulls the additional gain due to selection was because of the increase in accuracy of the breeding value, with no difference between pedigree and genomic constraints on coancestry. Sire and dam information explained 60-70% of the variation in genomic breeding values, therefore 30-40% of variation in genomic breeding values was due to within-family variation.

P-184
ACCURACY OF GENOTYPE IMPUTATION IN SWISS CATTLE BREEDS

Birgit Gredler1, Franz Seefried1, Urs Schuler1, Paul VanRaden2, Mehdi Sargolzaei3,5, John Hickey4
1Qualitas AG, Zug, Switzerland, 2Animal Improvement Programs Laboratory, USDA, Beltsville, USA, 3University of Guelph, Guelph, Canada, 4University of New England, Armidale, Australia, 5Boviteq Alliance, Quebec, Canada

The objective of this study was to evaluate the accuracy of imputation from Illumina Bovine3k Bead Chip (3k) and Illumina BovineLD (6k) to 54k chip information in Swiss dairy cattle breeds. Genotype data comprised of 54k SNP chip data of Original Braunvieh (OB), Brown Swiss (BS), Swiss Fleckvieh (SF), Simmental (SI), and red Holstein (HO). Genotypes of OB and BS (data set BSW) were analysed together as well as SI, SF, and HO (data set MIX). After routine genotype quality checks BSW and MIX included 3,738 animals with 39,841 validated SNP and 4,753 animals with 39,743 validated SNP, respectively. 54k genotypes of animals born between 2008 and 2011 were masked to mimic animals genotyped with the 3k and 6k chips. Methods used for imputation were AlphaImpute, Beagle, FImpute, and Findhap V2. AlphaImpute, FImpute and Findhap use pedigree information whereas Beagle does not directly use pedigree information. The accuracy of imputation was assessed by the squared correlation (R²) between true and imputed genotypes. R² was higher for imputation from 6k than from 3k. R² decreased with lower relationship between the 54k genotyped reference population and the 3k and 6k genotyped imputation candidates. For BSW, average R² was highest (0.96) using AlphaImpute and FImpute when both parents of a 6k genotyped animal were genotyped for the 54k chip. R² was between 0.85 and 0.93 when only the sire was 54k genotyped and lowest (0.79-0.92) when no direct relatives were 54k genotyped. Accuracy of imputation was highly dependent on MAF of the imputed SNP. Using FImpute R² was between 0.67 (MAF < 0.025) and 0.96 (MAF between 0.4 and 0.5). Computation time was lowest using FImpute and Findhap. All programs gave high imputation accuracy where FImpute slightly outperformed the other programs in terms of the R² while AlphaImpute had lowest error rates.

P-185
PREDICTION OF MAIZE TESTCROSS PERFORMANCE ACROSS ENVIRONMENTS

Theresa Albrecht1, Valentin Wimmer1, Hans-Jürgen Auinger1, Christina Lehermeier1, Carsten Knaak2, Milena Ouzunova2, Chris-Carolin Schön1
1Technische Universität München, Freising, Germany, 2KWS SAAT AG, Einbeck, Germany

In maize breeding, genome-based prediction is a promising tool to assess the genetic potential of untested lines without resource consuming field trials. Therefore, the aim of this study was to evaluate the prediction of testcross performance across environments within an advanced cycle breeding population of maize. The analyzed data set comprised 698 doubled haploid lines genotyped with 56,110 single
nucleotide polymorphism markers. Testcross performance was evaluated in four environments for two quantitative traits, grain yield and grain dry matter content. Predictive performance was assessed with genomic best linear unbiased prediction (BLUP) and cross-validation (CV). The data set was divided into 200 different estimation (ES) and test sets (TS) taking into account genotypic and environmental sampling. With genotypic sampling, predictive abilities ranged between 0.39 and 0.76 with an average of 0.62 for grain yield. Accounting for environmental sampling resulted in an average predictive ability of 0.50. Predictive abilities were reduced even more, when the sampling accounted for genotypic and environmental effects simultaneously (0.42). Due to a higher trait heritability, predictive abilities for grain dry matter content were higher than those for grain yield for all CV schemes. These findings give a first insight into the allocation of resources for implementing genomic selection into plant breeding programs.

P-186
GENOMIC SELECTION IN A PUREBRED PIG BREEDING SCHEME: A SIMULATION STUDY
Thierry Tribout1,2, Catherine Larzul1,2, Florence Phocas1,2
1INRA, Jouy-en-Josas, France, 2Agroparistech, Jouy-en-Josas, France

Stochastic simulation was used to compare the efficiency of 3 pig breeding schemes based on traditional or genomic evaluations. The simulated population (1,000 females, 50 males) was selected during 10 years for a synthetic breeding goal including 2 traits with equal heritability (h²=0.4) and economic weights. The reference breeding scheme (BLUP-AM) was based on BLUP-Animal Model evaluations and the phenotyping of 13,770 candidates per year for trait 1 and of 270 relatives per year for trait 2. Under the GE-1TP scenario, selection was based on genomic breeding values (GBV) estimated with one training population (TP) made up of candidate relatives phenotyped for both traits, whose number increased from 1,000 to 3,430 over time. Under the GE-2TP scenario, the GBV for trait 2 were estimated using a TP identical to that of GE-1TP, but the GBV for trait 1 were estimated using a large TP made up of candidates whose number increased from 13,770 to 55,080 over time. Concerning trait 2, GE-1TP generated 125% higher reliabilities, and therefore respectively 116% and 84% higher average annual genetic trends than BLUP-AM. Concerning trait 1, GE-1TP generated 34% lower reliabilities and 37% lower annual genetic trends than BLUP-AM, whereas GE-2TP generated 80% higher reliabilities and 12% higher annual genetic trends than BLUP-AM. Consequently, GE-2TP was 29% more efficient in improving the global breeding goal than BLUP-AM, whereas GE-1TP was only as efficient as the reference scheme. At last, GE-1TP and GE-2TP reduced the inbreeding rate by 22% and 55%, respectively. In conclusion, genomic selection could substantially and durably improve the efficiency of pig breeding schemes, even for traits with currently limited phenotyping capacity resulting in small TP. Conversely, large TP are required to remain as efficient as traditional schemes for traits recorded on numerous candidates.

P-187
HAPLOTYPES BASED ON LOCAL GENEALOGIES FOR GENOMIC PREDICTION
Vahid Edriss1, Rohan L. Fernando0,2, Guosheng Su1, Mogens S. Lund1, Bernt Guldbraudtsen1
1Department of Molecular Biology and Genetics, Aarhus University, Tjele, Denmark, 2Department of Animal Science, Iowa State University, Ames, IOWA, USA

The accuracies of genomic predictions using haplotypes constructed based on local genealogies were investigated. Haplotype based methods were expected to yield higher accuracies when predicting genetic values than when using single marker methods alone. Data contained 4,429 Nordic Holstein bulls which were genotyped with 50k SNP chip and had reliable estimated genetic values. Of these 3,145 were used as the training population. The test population contained the remaining 1,284 animals. Additive genetic values for fertility, milk protein yield and mastitis were predicted. Genomic predictions were done using either haplotypes or single markers as explanatory variables and using four statistical models.
Haplotypes were constructed using the Blossoc software which reconstructs local genealogies without recombination around each SNP. Haplotypes were clustered according to the local genealogy. Additive genetic values were predicted for the test animals using GBLUP, Bayes B, Bayes C and Bayes Cp. Both Bayes C and Bayes Cp assume that a proportion (1-p) of markers belong to the class of markers with larger effects. Bayes C uses a fixed p In Bayes Cp, p is dynamically obtained from the data. Bayes Cp yielded the highest accuracies both using single marker and using haplotype based method. The accuracies for the single marker method were 0.596, 0.650 and 0.629 for fertility, milk protein yield and mastitis. With the haplotype method they were 0.596, 0.658 and 0.629. Using Bayes B or Bayes C with p=0.99, the haplotype approach led to higher accuracy than the single marker approach. On the whole, the haplotype methods showed a small improvement compared to single marker methods. However, the haplotype methods were less sensitive to misspecification of p than the single marker methods.

P-188
COMPARING METHODS FOR ESTIMATING GENE EFFECTS: LEAST SQUARES VS. GENOMIC SELECTION

Jules Hernandez-Sanchez1, Ricardo Pong-Wong2, Gertraude Freyer3, Dimitrios Vagenas4
1IRTA, Lleida, Spain, 2The Roslin Institute and The R(D)SVS, Edinburgh, UK, 3Leibniz Institute for Farm Animal Biology, Dummerstorf, Germany, 4Institute of Health and Biomedical Innovation, Brisbane, Australia

Ascertaining the genetic architecture of complex traits is not trivial because there can be many genes involved, interacting with each other and with the environment. The widely used gene mapping via least squares regression one marker at a time (LSR1) renders biased estimates of effects. The biased estimate at locus $i$ is

$$
q_i = \sum_{j=1}^{M} a_j D_{ij} / p_i q_i
$$

with $j=1...M$ genes, $a_j$ unbiased additive effect, $D_{ij}$ linkage disequilibrium (LD), and $p_i=1-q_i$ allele frequency. In contrast, multi-marker least squares regression (LSRM) estimates are not biased. However, LSRM cannot handle models where the number of parameters (M) is larger than the number of phenotypes. Another potential problem with LSRM is collinearity among loci. Therefore, a variant of LSRM to cope with over-parameterised situations was implemented. Here, fixed sets of markers (5, 10...40, 45) are randomly sampled and AIC used to select the best fitting model (AICLSRM). Genomic selection (GS) methods deal with over-parameterisation assuming random marker effects and prior distributions with a high density of null effects. We simulated data to compare LSR1, LSRM, AICLSRM, Ridge Regression, Bayes-C and Bayesian Lasso in terms of their accuracy (bias) and precision (standard error) of estimated gene effects. Data consisted of 1000 replicates of 1000 normally distributed phenotypes, heritability of 0.5 and 50 genes in LD. The low number of parameters was required for LSRM to work. LSR1 was the most precise but least accurate method. LSRM was the most accurate method. GS methods ranked between LSR1 and LSRM in terms of both accuracy and precision. AICLSRM rendered the most parsimonious model, in which 25% of markers explained most of the genetic variation. More realistic and complex situations are being investigated, for example testing SNPs not genes.

P-189
SENSITIVITY TO PRIOR SPECIFICATION IN BAYESIAN MODELS FOR GENOMIC PREDICTION IN MAIZE

Christina Lehermeier1, Valentin Wimmer1, Theresa Albrecht1, Hans-Jürgen Auinger1, Chris-Carolin Schön1, Volker Schmid2
1Plant Breeding, Technische Universität München, Freising, Germany, 2Department of Statistics, Ludwig-Maximilians-Universität München, Munich, Germany

Many different models have been proposed for genomic prediction of breeding values. The aim of our study was to compare four Bayesian models (Bayesian Ridge, Bayesian Lasso, BayesA and BayesB)
with respect to their sensitivity to prior specification. The models differ in their prior density of marker effects and have different shrinkage properties which are specified by hyperparameters. To evaluate the predictive ability of the models we used an experimental maize data set, which comprises 698 doubled-haploid lines genotyped with 56,110 single nucleotide polymorphism markers and phenotyped for two quantitative traits (grain dry yield, grain dry matter content). The predictive ability of the models with different hyperparameter settings was assessed by five-fold cross-validation. The extent of Bayesian learning was measured as the Hellinger distance between prior and posterior density of marker effects. In BayesA and BayesB the choice of hyperparameters for specification of the respective prior density had a stronger impact on the predictive ability, than in Bayesian Ridge and Bayesian Lasso. The ability of Bayesian learning was reduced in BayesA and BayesB compared to Bayesian Ridge and Bayesian Lasso, which was shown by their significantly smaller Hellinger distances between prior and posterior densities of marker effects. Marker-specific shrinkage did not outperform marker-homogeneous shrinkage in the investigated advanced cycle breeding population of maize with high linkage disequilibrium between markers. We conclude for this data set that BayesA and BayesB are not advantageous over Bayesian Ridge and Bayesian Lasso because they were highly sensitive to prior specification but did not outperform the other models with regard to their predictive ability.

P-190
REALIZED GENOMIC KINSHIP BETWEEN FULL AND HALF SIBS PAIRS IN A COMMERCIAL PIG POPULATION
Marcos Soares Lopes1,2, Fabyano Fonseca Silva2, Barbara Harlizius1, Naomi Duijvesteijn1, Paulo Sávio Lopes2, Simone Facioni Guimaraes2, Egbert Frank Knol1
1Institute for Pig Genetics, Beuningen, The Netherlands, 2Universidade Federal de Viçosa, Viçosa, Brazil

Traditional breeding programs consider a mean pairwise kinship between sibs in order to derive a relationship matrix disregarding variation due to Mendelian sampling. Molecular markers make possible the estimation of the real relationship. The aim of this study was to estimate the realized genomic kinship between full and half sibs in a commercial pig population. A total of 997 boars (502 full sibs and 5,756 half sib pairs) from a Duroc-based line were genotyped with the PorcineSNP60 Beadchip. Genomic kinship was estimated with the R package GenABEL using the full set of markers (n=28,740) after quality control (missing genotype frequency <0.05 and minor allele frequency >0.05). A reduced SNP set with markers in linkage equilibrium (LE markers, n=9,579) was also tested. To select the subset of LE markers, linkage disequilibrium between SNPs was estimated. Tightly linked SNPs (r²>0.5) were excluded using LD-based SNP pruning in PLINK. For comparison, pedigree-based kinship using information from 10 generations was estimated using ENDOG v4.8. The mean pedigree-based kinship for full sibs was 0.27±0.01 (0.26 to 0.34). The mean genomic kinship between full sibs using both SNP sets was 0.24±0.04 ranging from 0.08 to 0.34 (LE markers) and from 0.08 to 0.37 (full set). For half sibs, the mean pedigree-based kinship was 0.15±0.01 (0.14 to 0.24) while the mean genomic coefficient was 0.12±0.03, ranging from 0.02 to 0.28 (LE markers) and from 0.02 to 0.30 (full set). The correlation between genomic and pedigree-based kinship was higher when LE markers were used (0.29 and 0.34 for full sibs and 0.14 and 0.15 for half sibs, using all markers and LE markers, respectively). A genomic relationship matrix estimated using unlinked markers should therefore yield significantly more accurate breeding values for animals without own performance.
P-191
EYE OF THE BEHOLDER: NEW APPROACHES IN QUANTIFYING IRIS COLOUR AND MORPHOLOGY

Andreas Wollstein, Susan Walsh, Stefan Boehringer, Manfred Kayser
1Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands, 2Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands

Several genetic variants are known to be involved in human eye colour variation, mostly from studies considering eye colour as a categorical trait. Eye colour DNA prediction studies have shown that although blue and brown are predictable with high accuracies, eye colours usually classified as intermediate are more difficult. One possible explanation for this is the categorical use of eye colour in almost all gene mapping studies together with the fact that in reality, eye colour represents a continuous trait with blue and brown being on the extremes. Hence, it has been pointed out before that in order to complete genetic knowledge of iris colour variation, particularly regarding the non-blue and non-brown colours, it should be considered as a continuous trait in gene-mapping studies. However, current methods that quantify eye colour are scarce. Here we introduce a newly developed approach for eye colour, which represents the natural distribution of the types of melanin that occur within the iris that has been extracted from high-resolution digital images. We demonstrate that our approach improves the strength of association for DNA variants known to be involved in eye colour. Hence, our new phenotyping approach, which in principle can be applied to all pigmentation types, opens up new possibilities for finding new DNA variants involved in eye colour in future gene-mapping studies. This in turn also has the potential to increase the predictability of non-blue and non-brown eye colours with use in certain disciplines such as forensic genetics.

P-192
ACCURACY OF GENOTYPE IMPUTATION IN CANADIAN YORKSHIRE PIGS USING FIMPUTE SOFTWARE

Alison Kinsman, Mehdi Sargolzaei, Flavio Schenkel, Gordon Vander Voort, Mohsen Jafarikia, Andy Robinson
1University of Guelph, Guelph, ON, Canada, 2L’Alliance Boviteq, Saint-Hyacinthe, QC, Canada, 3Canadian Centre for Swine Improvement, Ottawa, ON, Canada

Using high-density SNP panels for genomic evaluation has become the de facto standard. The current porcine genotyping panel is the Illumina PorcineSNP60 BeadChip, which contains 62,163 SNPs. Large scale application of genomics requires genotyping thousands of animals, which can be cost-prohibitive. A more affordable solution is to use a less costly low-density (LD) panel and then impute the high-density panel genotypes. Genotypes from the 60K SNP panel provided by the Canadian Centre for Swine Improvement (CCSI) on 907 Yorkshire pigs were used to investigate the optimum density of a LD panel for imputation to the 60K panel. A total of 39,526 SNPs were considered after excluding 1,328 SNPs located on the sex chromosomes, 13,985 unmapped SNPs and 7,324 SNPs with minor allele frequency of less than 1% and call rate of less than 85%. Four LD panels of 395, 6K (5,647), 10K (9,882) and 20K (19,763) evenly spaced SNPs were created from the 60K genotypes. FImpute software was used to carry out the imputation analyses. In a scenario with 249 younger animals with LD genotypes (study group) and 658 older ones with 60K genotypes (reference group), imputation accuracy was 63.1%, 91.5%, 94.8%, and 97.7% for 395, 6K, 10K and 20K panels, respectively. The imputation accuracy of the 10K panel decreased by 1.7% when the number of animals in reference group decreased from 658 to 200. For the 10K panel with the availability of 60K genotypes on neither, one or both parents imputation accuracy was of 94.8%, 95.9% and 97.8%, respectively. These preliminary results suggest that a 10K SNP panel might be a practical choice for most animals, particularly if influential parents are genotyped using the denser panel. Further investigation including other major pig breeds, a larger reference population and accuracy of GEBV from imputed genotypes is recommended.
P-193
CORRECTION FOR ERRORS IN GENOTYPE ASSIGNMENT IN CASE-CONTROL STUDIES

Matthew Cooper1, Nicholas de Klerk1, Kathryn Greenop1, Sarra Jamieson1, Denise Anderson1, Frank van Bockxmeer1, Bruce Armstrong2, Elizabeth Milne1
1University of Western Australia, Perth, Australia, 2University of Sydney, Sydney, Australia

Objectives: Genotyping has become more cost-effective and less invasive with the use of buccal cell sampling. However, low or fragmented DNA yields from buccal cells sometimes requires additional whole genome amplification to produce sufficient DNA for typing. In a case-control study of childhood leukaemia, discordance was found between genotypes derived from blood and whole genome amplified buccal DNA samples. We aimed to develop a user-friendly method to correct for this genotype misclassification, as existing methods were not suitable for our purposes.

Methods: Discordance between the results of blood and buccal-derived DNA was able to be assessed, but only in disease cases, some of whom had both blood and buccal samples. Using the misclassification matrices as probability distributions, we sampled likely values for corrected genotypes for controls with only buccal samples, creating multiple datasets for analysis. Each dataset was analysed separately, adjusting for multiple covariates using logistic regression. Regression coefficients were then combined using standard methods for multiple imputed datasets.

Results: Application to synthetic datasets was effective in producing correct odds ratios (ORs) from data with known misclassification. When applied to actual findings from each of six bi-allelic loci, correction altered the ORs in the expected way given the type of misclassification Shown. Increasing the size of the validation data set increased the precision of the effect estimates. Conclusions: Bias arising from differential genotype misclassification can be reduced by correcting results using this method whenever data on concordance of genotyping results with those from a different and probably better DNA source are available.

P-194
GENOMIC SELECTION TO INCREASE GENETIC GAIN IN WHEAT BREEDING

Yongle Li1, David Bonnett1, Jose Crossa1, Jeffrey Endelman2, Juan Burgueno1, Susanne Dreisigacker1, Jesse Poland3, Mark Sorrells1, Jean-Luc Jannink2, Julie Dawson2
1CIMMYT, Texcoco, Mexico city, Mexico, 2Cornell University, Ithaca, New York, USA, 3USDA-ARS, Kansas, USA

Genomic selection is a new tool with potential to increase breeding efficiency. The objectives of this study are 1) Design, implement and evaluate wheat breeding strategies involving genomic selection. 2) Develop high yielding spring wheat lines suitable for international evaluation and use as parents by CIMMYT and other breeding programs. To build a model for genomic selection, we used a training population consisting of 1338 elite, high-yielding CIMMYT lines with three years (2009-2011) of yield data. The training population was genotyped via genotyping-by-sequencing (GBS) with 13k SNPs. Various statistical models were used for prediction, pedigree (or infinitesimal model), reproducing kernel Hilbert spaces (RKHS) with and without pedigree, and ridge regression best linear unbiased prediction (RR-BLUP). Prediction assessment using only the training set by means of a 10-fold cross validation show different correlation between predicted and observed observations with model RKHS with and without pedigree giving the best prediction accuracy. Twenty of the highest yielding elite CIMMYT lines from the training set were selected for intercrossing and 40 crosses were generated. From these populations 1925 F2s were genotyped via GBS as candidates for genomic selection. Genomic Estimated Breeding Values (GEBVs) for the F2s were calculated in a one-step analysis via RR-BLUP. Around 500 F2s with the highest GEBVs were selected for intercrossing. A subset of progeny from these crosses will be genotyped and GEBVs calculated to identify individuals for a second cycle of intercrossing. A large number of the F2s from the original crosses were not genotyped and will be progressed through the conventional selection scheme as
a comparison. A further 1200 new advanced lines will be genotyped to increase the size and diversity of the training population. We estimate that genomic selection could increase the rate of gain of CIMMYT’s wheat breeding programs two fold.

**P-195**
**INVESTIGATION OF THE FACTORS AFFECTING GENOME DATA STRUCTURE AND THE ACCURACY OF GENOMIC EVALUATION IN COMPUTER SIMULATION**

*Motohide Nishio, Toshihiro Okamura, Osamu Sasaki, Akiko Nishiura, Masahiro Satoh*
*NARO Institute of Livestock and Grassland Science, Tsukuba, Ikenodai 2, Japan*

The objectives of the present study were to investigate the factors affecting genome data structure using computer simulation and to evaluate the accuracy of predicting breeding values (BVs) using genome information. The simulated genomes consisted of one chromosome of one Morgan. The positions of 1,000 single nucleotide polymorphisms (SNP) markers and 100 quantitative trait loci (QTL) were assumed randomly across the genome with different mutation rates ($MR = 10^{-3}$, $10^{-4}$, $10^{-5}$). The historical populations of two effective population sizes ($Ne = 100, 500$) with initial minor allele frequencies ($IMAF = 0.0, 0.1, 0.3, 0.5$) were generated for various numbers of generations ($NG = 500, 2000, 5000, 20000$). After $NG$, each animal in each population was simulated for more seven generations ($G0-G6$) with marker genotypes and a phenotype of a performance trait. The heritability was assumed to be 0.3. For each scenario, 20 replicates were performed. Changes of distribution of allele frequencies of SNP and linkage disequilibrium coefficient between markers and QTL in generation $G0$ were used to evaluate the stability of genome data structure. BVs of animals (EBV) at $G6$ were estimated using information on pedigree ($G0-G6$) and trait phenotypes ($G3-G5$). In addition, information on marker genotypes ($G4-G6$) was used to estimate BVs (GBV). Correlation coefficients between true BV and EBV ($r_A$) and GBV ($r_H$) were used to evaluate genomic information. Genome data structure was stable after adequate generations for any IMAF. The numbers of adequate generations were depended on $MR$ and $Ne$. The $MR$ and $Ne$ affected the value of $r_H$, but $NG$ did not affect. The values of $r_H$ were 1.5-1.9 times larger then $r_A$ in various conditions.

**P-196**
**ALLELE FREQUENCY CHANGES WITH GENOMIC, BLUP AND PHENOTYPIC SELECTION AT VARYING NUMBER QTL AND HERITABILITY**

*Huiming Liu, Peer Berg, Anders Christian Sørensen*
*Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, Aarhus University, Tjele, Denmark*

Increases in the allele frequency of favorable alleles in populations drive the genetic improvement in livestock. The availability of SNP chips and genomic selection may give a possibility to optimize long-term genetic response. Our objective is to compare the effect of genomic selection with phenotypic selection and BLUP, and to investigate the advantage of genomic selection in terms of allele frequency changes by a stochastic simulation. We compared the realized allele frequency changes under different value of heritability (0.05 to 0.5) and different number of QTL (10 to 500). The simulated genomes consisted of four chromosomes each with a size of 1M. The number of loci is 10,000, and 5000 of them are biallelic. A total of 2000 generations were simulated until linkage disequilibrium and heterozygosity values were stable for $Ne=200$. The population in the 2000th generation was assigned to be the founder population; each selection method was implemented for 7 generations subsequently.

The results show that irrespective of the number of QTL, GBLUP changes allele frequencies faster than BLUP and phenotypic selection (e.g. 0.0612; 0.0477 and 0.0127 respectively in the situation where $h^2$ is 0.05 and the number of QTL is 10). The differences in allele frequency changes between selection methods were smaller when the $h^2$ is higher. Within each selection method, the mean allele frequencies
change faster when there are less QTL. Likewise genomic selection results in faster changes in allele frequencies at linked markers. Keeping the heritability constant, the accuracies of estimated breeding values (EBV) show that GBLUP performs better compared to BLUP and phenotypic selection, but with decreasing accuracies when more QTL were involved. We conclude that genomic selection results in a signature of selection with faster changes in allele frequencies both at QTL and linked markers, resulting in higher response to selection.

P-197
IMPROVEMENT OF THE GENOMIC MODEL FOR DIRECT AND MATERNAL CALVING EASE IN THE UK

Sophie Eaglen1,2, Eileen Wall1, Mike Coffey1, John Woolliams0,2, Raphael Mrode1
1SAC, Edinburgh, UK, 2Roslin Institute, Edinburgh, UK

This study was conducted to compare accuracies of direct genomic breeding values (DGVs) estimated by univariate GBLUPs models and a bivariate GBLUP model with the aim of potentially optimising the current genomic model for calving ease (CE) in the UK (univariate ridge regression).

Data from 11480 bulls with 50K genotypes were available for analyses as a result of cooperation between the Cooperative Dairy DNA Repository (CDDR), UK AI industry and SAC. Reference and validation populations consisted of 4556 and 1333 bulls for CEd (direct) and 4553 and 532 bulls for CEm (maternal) respectively.

DGVs were estimated by two univariate GBLUP models incorporating a genomic relationship matrix (G). Dependent variables were official deregressed proofs (DRP) weighted by effective daughter contribution. Variance components were adapted from conventional variance component estimation. Accuracies of genomic evaluation with the univariate GBLUP model (correlation between DGV and DRP) were 0.55 for both CEd and CEm.

A bivariate GBLUP model was subsequently analyzed with CEd and CEm fitted as correlated traits. Results showed similar accuracies compared to the univariate GBLUP models; 0.56 for both CEd and CEm respectively. Accuracies are likely similar due to the imposed genetic covariance by this model while DRPs follow from the deregression of univariate models. Both direct and maternal DRPs are now estimated for all individuals, regardless of the existence of an official proof. The next step in this study is an estimation of DRPs while accounting for a direct-maternal genetic covariance.

P-198
OPTIMIZING SNP PANELS FOR GENOMIC PREDICTION BY SIMULATED ANNEALING

Joaquim Casellas
Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

Genomic best linear unbiased prediction (gBLUP) try to predict the genetic merit of livestock by summarising the genetic information from thousands of single nucleotide polymorphisms (SNP) into the genomic relationship matrix (G). Note that current gBLUP parametrizations typically use all SNP to compute G although little is known about existing linkage disequilibrium between each SNP and quantitative trait loci (QLT) underlying each phenotypic trait of interest. Within this context, simulated annealing (SA) was used to optimise the panel of SNP for gBLUP in terms of mean square error (MSE) between simulated and predicted data. The analytical process started with all SNP contributing to G.

For each SA iteration, a new SNP was chosen at random and the MSE was calculated after changing its state (i.e., removed or included to the list of SNP for the calculation of G). This change was accepted if MSE reduced from the previous iteration and the SA process stopped after 1,000 iterations without changes in the list of used/discard SNP. This approach was evaluated on 10 simulated data sets with 1,000 preliminary generations (Ne=100) and five more generations (Ne=200) contributing phenotypic and genomic data (h²=0.5). Genomes had a unique 100 cM chromosome with 5,000 SNP and 500
quantitative trait loci with appropriate mutation probabilities \( (i.e., 10^{-3} \text{ and } 10^{-5}, \text{ respectively}) \). On average, the full model with all SNP reached a MSE of \( 1.370 \pm 0.004 \), whereas this parameter reduced until \( 1.293 \pm 0.005 \) when dropping off between 35% and 47% of the SNP. These results involved a \( \sim 6\% \) reduction of MSE when using appropriate SNP for \( G \) and suggested a very appealing way to improve the statistical performance of gBLUP models when analysing massive genomic data.

P-199

**IMPROVEMENT OF GENOMIC SELECTION USING A RIDGE REGRESSION APPROACH WITH SELECTED MARKERS**

Marcio Resende Jr, 1,2, Patricio R Munoz Del Valle, 2,3, Marcos DV Resende, 4, Dorian J Garrick, 5, Rohan Fernando, 1, John M Davis, 3, Gary Peter, 3, Matias Kirst, 3

1Genetics and Genomics Graduate Program, Gainesville, Florida, USA, 2Plant Molecular and Cellular Biology Graduate Program, Gainesville, Florida, USA, 3School of Forest Resources and Conservation, Gainesville, Florida, USA, 4EMBRAPA Forestry, Colombo, Parana, Brazil, 5Department of Animal Science, Iowa State University, Ames, Iowa, USA

The use of marker-enabled predictions in breeding programs can drastically increase the genetic gain per unit of time. However, different genomic selection methodologies may have different predictive abilities depending on the genetic architecture of the trait. In principle, a method might perform better in the prediction of traits that fit the assumptions under which was constructed. When genomic selection is modeled according to a Ridge Regression BLUP (RR-BLUP) approach, the assumptions for the marker variances follows the ones from the infinitesimal model, even though they are applied to a finite set. In this work, we evaluated a small modification of the RR-BLUP method that utilizes a subset of selected markers, and is referred hereafter as RR-BLUP B. The selection of the subset of markers is based in the group that maximizes the predictive ability in a cross-validation scheme. Once this group is defined, a new model is fit, under the same assumptions, but with a reduced number of covariates. To evaluate this approach we report, for an experimental breeding population of the tree loblolly pine \( (Pinus taeda \text{ L.}) \), a comparison of GWS predictive models for 17 traits with different heritabilities and predicted genetic architectures. Genome-wide selection models included RR-BLUP B and four traditional methods: RR-BLUP, Bayes A, Bayes C\( \pi \) and the Bayesian LASSO. We show that, for most traits, there is limited difference among these five methods in their ability to predict genetic merit. Nonetheless, out of the four traditional methods tested, Bayes C\( \pi \) performed 15% better for fusiform rust resistance – a disease-resistance trait known to be controlled in part by major genes. However, for the same trait, the proposed method RR-BLUP B increased the accuracy (from 0.24 of RR-BLUP to 0.37) outperforming all the other methods when a reduced number of 100 markers were selected and re-fit.

P-200

**WHOLE GENOME QUANTITATIVE TRAIT PREDICTION BY FUNCTIONAL IDENTITY GENOTYPING AND DIRECT ESTIMATION OF GENE-LEVEL ADDITIVE/DOMINANCE EFFECTS**

Serge Sverdlov

University of Washington, Seattle, WA, USA

Existing methods for predicting genotypic values and breeding values for polygenic traits from whole genome SNP data, such as ridge regression and replicating kernel Hilbert spaces (RKHS) techniques, assign additive effects to individual SNP’s or SNP interactions. Classical quantitative genetics decomposes the genotypic value before epistasis into additive and dominance effects, at each gene, or functionally polymorphic locus. We seek to reconcile the two approaches by designing a trait prediction model with both additive and dominance effects, as well as a mechanism for mapping from SNP data to common functional variants.
We introduce an evolutionary model of diploid genotypic effects accounting for dominance, using an infinite allele model and an effect covariance structure based on an analogy to identity by descent (IBD) symmetry patterns. We also introduce a method for probabilistic gene-level imputation of common functional variants from dense unphased SNP data using a large reference set of phased full sequences, such as that made available for human genetics by the 1000 Genomes project. Together, these two methods allow the attribution of individual additive and dominance effects to functional variants of annotated genes, and the bottom-up computation of additive and dominance variances from genotypic effects and allele frequencies. We further consider the applications of this methodology to epistasis and sparsity. The model is applied to the construction of a predictive model for human height from dense SNP data for 1,000 individuals, and the estimation of corresponding allele effects, additive, dominance, and environmental variances, and narrow- and broad-sense heritability.

**P-201 – ABSTRACT WITHDRAWN**

**P-202**

ESTIMATING INDIRECT GENETIC EFFECTS WHEN INTERACTIONS DIFFER BETWEEN KIN AND NON-KIN

Setegn Alemu, Peer Berg, Piter Bijma
Aarhus University, Tjele, Denmark

Social interactions among individuals are common both in wild and domestic populations. With social interactions, trait values of individuals may depend on genes carried by other individuals, a phenomenon known as Indirect Genetic Effects (IGE). IGE may have significant effect on the rate and direction of response to selection. Previous research has shown that genetic parameters for IGEs are estimated most accurately by using groups composed of two families. This result, however, relies on one important assumption, being that an individual has the same IGE on each of its group mates, irrespective of whether it belongs to its own family or to the other family. However, due to kin selection and kin recognition, individuals may have different IGEs on kin vs. non-kin, which violates the model assumptions. As a consequence, the estimated IGEs may be incorrect, and artificial selection based on those estimates will be suboptimal. Hence, we investigated whether genetic parameters for IGEs are statistically identifiable when IGEs differ between kin and non-kin, and developed statistical models to estimate such effects. First we show that the full set of genetic parameters is not identifiable when IGEs differ between kin vs. non-kin; direct and indirect effects on kin cannot be separated. Subsequently we developed an alternative model, and show that this model can identify the total breeding value and the non-kin IGE. Hence, total heritable variance and total breeding value are estimable. The Monte Carlo simulation confirmed that our method yields unbiased estimates. This work, therefore, shows how artificial selection can be used to reduce competitive interactions among individuals, also when IGEs differ between kin and non-kin.

**P-203**

SOCIAL INTERACTIONS AMONG FINISHING PIGS DIVERGENTLY SELECTED FOR SOCIAL GENETIC EFFECTS ON GROWTH IN BARREN AND STRAW HOUSING

Irene Camerlink, J. Elizabeth Bolhuis, Piter Bijma
Wageningen University, Wageningen, The Netherlands

Group housed pigs may influence each other’s behaviour and growth through social interactions. The genetic effect of an individual on traits of other individuals can be estimated, and is known as Social
Breeding Value (SBV) or Indirect Genetic Effect. In pigs, SBV for growth during finishing have been estimated, but consequences of selection for SBV on pig behaviour and realized growth have not been evaluated. We, therefore, studied the effect of a single generation of selection for either high or low SBV on behaviour and growth of pigs, in barren and straw-enriched pens. In five batches, 480 finishing pigs, housed in groups of six, were studied in a GxE setting (80 pens). The average estimated SBV contrast between high and low pigs was 3.6 g/day, corresponding to an expected growth difference of (6−1) x3.6 = 18 g/day between high and low pens. Behaviours were recorded at 12, 16 and 21 weeks of age by 2-min instantaneous scan sampling during 6 h daytime. Low SBV pigs spent more time on aggressive biting (P<0.05), while high SBV pigs tended to engage more in fights (P=0.06), which suggests different behavioural strategies. SBV did not affect tail biting, but low SBV pigs spend more time on ear biting (P<0.05) and chewing on a sack of jute or metal chain (P<0.01). This may indicate that low SBV pigs have a stronger tendency to perform oral manipulation. Straw-bedding strongly reduced time spent on ear biting, tail biting and other oral manipulative behaviours and increased pen exploration and comfort behaviour (all P<0.01). Growth during the finishing period (10 to 23 weeks of age) was not significantly affected by SBV or its interaction with environment. In conclusion, both SBV and enriched housing had positive effects on social behaviours but, unexpectedly, SBV did not affect realized growth in this experiment.

P-204
PHENOTYPIC PLASTICITY OF THE DROSOPHILA TRANSCRIPTOME

Shanshan Zhou1,2, Terry Campbell2,3, Eric Stone2,3, Trudy Mackay2,3, Robert Anholt1,2
1Department of Biology, North Carolina State University, Raleigh, USA, 2Department of Genetics, North Carolina State University, Raleigh, USA, 3W.M. Keck Center for Behavioral Biology, North Carolina State University, Raleigh, USA

Phenotypic plasticity is the ability of a single genotype to produce different phenotypes in response to changing environments. We assessed variation in genome-wide gene expression and four fitness-related phenotypes of an outbred Drosophila melanogaster population under 20 different physiological, social, nutritional, chemical and physical environments, and compared the phenotypically plastic transcripts to genetically variable transcripts in a single environment. The environmentally sensitive transcriptome consists of two transcript categories, which comprise ~15% of expressed transcripts. Class I transcripts are genetically variable and associated with detoxification, metabolism, proteolysis, heat shock proteins, and transcriptional regulation. Class II transcripts have low genetic variance, and show sexually dimorphic expression enriched for reproductive functions. Clustering analysis of Class I transcripts reveal a fragmented modular organization, and distinct environmentally-responsive transcriptional signatures for the four fitness-related traits. Our analysis suggests that a restricted environmentally-responsive segment of the transcriptome preserves the balance between phenotypic plasticity and environmental canalization.

P-205
HERITABILITY OF MORPHOLOGICAL, BEHAVIOURAL AND LIFE-HISTORY TRAITS IN A POPULATION OF WILD BATS WITH HIGH MALE REPRODUCTIVE SKEW

Helen Ward1, Roger Ransome1, Gareth Jones2, Alastair Wilson2, James Cotton2, Stephen Rossiter1
1Queen Mary, University of London, London, UK, 2University of Bristol, Bristol, UK, 3University of Edinburgh, Edinburgh, UK, 4Wellcome Trust Sanger Institute, Cambridge, UK

Bats are long-lived, show extended maternal care, and often live in complex societies. Yet virtually nothing is known about the extent to which quantitative traits in bats are determined by genetic inheritance versus non-genetic determinants (environmental factors and/or maternal effects).

Objectives: To assess the heritability of morphological, life-history and behavioural traits in wild greater
horseshoe bats (*Rhinolophus ferrumequinum*).

**Methods:** We used microsatellites to infer the parentage of 924 pups born into a UK colony since 1993. We built a highly resolved pedigree spanning 10 generations, and applied mixed-effects ‘animal models’ to estimate the genetic variance components of a range of traits.

**Results:** Our preliminary analyses suggest that forearm and digit length are both highly heritable and also strongly correlated with each other, indicating that they are controlled by the same gene(s). Conversely, we have not found any evidence that behavioural patterns such as mate fidelity, or life history traits including survival and reproductive success, are heritable within the colony. Inspection of breeding patterns based on the parentage results reveal that a small minority of males contribute massively to the gene pool; just 4% males (n = 6) fathered 25% (n>170) of all colony pups (n >700) born during a 19-year period.

**Conclusions:** We have demonstrated heritable genetic variation in several quantitative traits within a population of wild bats. We have not however found any evidence for heritability of mating patterns or life-history traits which might start to explain the high levels of male reproductive skew within this population. We discuss these and our related findings, and their implications for future work.

**P-206**

**GENETICS OF ASCITES RESISTANCE AND TOLERANCE IN CHICKEN: A RANDOM REGRESSION APPROACH**

*Antti Kause¹, Sacha van Dalen², Henk Bovenhuis²*

¹MTT Agrifood Research Finland, Biometrical Genetics, Jokioinen, Finland, ²Wageningen University and Research Centre, Animal Breeding and Genomics Centre, Wageningen, The Netherlands

Resistance is the host trait that prevents an infection or reduces pathogen performance in a host, both factors reducing pathogen burden within the host. Tolerance to infections is defined as the ability of the host to limit impact of a given pathogen burden on host health, performance, and ultimately on fitness. Simultaneous genetic analysis of both resistance and tolerance is needed for understanding the ways selection can mould defence strategies against parasites, pathogens, and production diseases. We demonstrate the use of multitrait random regression animal model and covariance functions to estimate genetic parameters for both resistance and tolerance, and illustrate the way individual variation in resistance and tolerance induce genotype re-ranking in host performance across infection-free and infected environments, and modify genetic variation in host performance along increasing infection severity. In a random regression model, tolerance is analyzed as a reaction norm in which host performance (e.g. growth, on y-axis) is regressed against an increasing infection severity (on x-axis), fulfilling the definition of tolerance. This approach was applied to data on ascites and body weight records of 7722 domesticated chicken. Tolerance to ascites displayed significant genetic variance, the estimated breeding values of tolerance slope ranging from strongly negative (very sensitive genotype) to weakly negative (less sensitive). Resistance to ascites had heritability of 0.34. Both traits are hence expected to respond to selection. The two complementary defence strategies, tolerance and resistance, were genetically independent. Although ascites induced elevated phenotypic and genetic variances in body weight of affected birds, heritability displayed negligible changes across healthy and affected birds. Ascites induced moderate genotype re-ranking in body weight, the genetic correlation of healthy birds with mildly affected birds being unity but with severely affected birds 0.45. This study demonstrates a novel approach for exploring genetics of defence traits and their impact on genotype-by-environment interactions.

**P-207 – ABSTRACT WITHDRAWN**

**P-208 – ABSTRACT WITHDRAWN**
P-210  
GENETIC ARCHITECTURE OF MICRO-ENVIRONMENTAL SENSITIVITY OF QUANTITATIVE TRAITS – APPLICATION TO BOVINE SOMATIC CELL SCORE  

Han A. Mulder,2, Ron E. Crump, Mario P.L. Calus, Roel F. Veerkamp  
1Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands, 2Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, Lelystad, The Netherlands  

In the last decade a number of studies have shown that not only the phenotypic trait value, but also micro-environmental sensitivity or environmental variance of traits is under genetic control. Micro-environmental sensitivity can be considered as differential sensitivity of individuals to unknown environmental factors observed as differences in micro-environmental variance. The availability of whole-genome SNP data on relatively large numbers of individuals makes it possible to perform genome-wide association studies to unravel the genetic architecture of micro-environmental sensitivity. The aim of this study was therefore to find SNP associated with micro-environmental sensitivity using a genome-wide association study and to quantify the predictive ability of genome-wide breeding values for micro-environmental sensitivity of bovine somatic cell score (SCS). Bovine SCS is a quantitative trait associated with mastitis and it is known that both its mean and micro-environmental sensitivity are partly under genetic control. A Bayesian stochastic search variable selection multi-SNP model was applied to different measures of micro-environmental sensitivity of SCS and mean SCS of 1563 cows located on research farms in four different countries. After quality control, genotypes of 37,590 SNP were used. No significant SNP were found to be associated with micro-environmental sensitivity, whereas 34 significant SNP were found for mean SCS. Based on 10-fold cross-validation, the accuracy of genome-wide breeding values was 0.4-0.5 for both micro-environmental sensitivity and mean SCS. Micro-environmental sensitivity was estimated to be affected by ~2000 - 2500 independent chromosome segments, whereas ~1700 independent chromosome segments were affecting the mean SCS. About 25-28% of 50-SNP windows were required to explain 50% of the total genetic variance of micro-environmental sensitivity. It can be concluded that micro-environmental sensitivity, at least for SCS, is likely to be determined by many genes and therefore the power to detect QTL for micro-environmental sensitivity is likely to be low.

P-211  
INDIRECT GENETIC EFFECTS IN CATTLE SOCIAL DOMINANCE  

Cristina Sartori, Roberto Mantovani  
University of Padua, Legnaro (PD), Italy  

The genetics of social behaviour has been widely disclosed in recent years, due to the effort provided by the detection of indirect genetic effects (IGEs) in quantitative genetics. In hierarchies assessment, a dominance relationship among two conspecifics is usually established by the outcome of a contest, where the winner acquires a dominant status over the loser. “To be a winner” and “not to be a loser” are two sides of the same coin, thus an individual providing the direct information “good at winning” necessarily carries also the information “bad at losing” as indirect effect, and the correlation among these two components gets a level of -1. Moreover, since each win necessarily implies a loss, the average number of wins in a population cannot increase, and the heritability of the dominant status should be close to zero. Moving from these assumptions, the genetics of social dominance was studied in Aosta Chestnut and Aosta Black Pied breeds (Bos taurus), two Alpine rustic cattle famous for traditional tournaments in which pairs of cows assess the respective dominant status in bloodless fights. The outcomes of about 12,000 dyadic interactions performed by 4,000 individuals in six years of tournaments were analysed in hierarchical models implementing a classical quantitative model or more complex models by introducing
indirect effects. Data were analysed via Bayesian approach on threshold models. The assessment of variances revealed a genetic correlation near to \(-1\) among direct and indirect components. The heritability measured on a liability scale for the direct phenotype was almost 13\%, but decreased to levels close to zero when the total heritable variance was considered. This result supports the relevance of IGEs on fights outcome and the assumption that the mean individual social dominance cannot evolve within a population, due to the evolutionary constraints imposed by the “social environment”.

**P-212**

THE GENETICS OF PHEROMONES; A SEARCH FOR INDIRECT GENETIC EFFECTS AND GENOMIC REGIONS FOR BOAR TAINT

*Naomi Duijvesteijn\(^1,2\), Egbert Frank Knol\(^2\), Piter Bijma\(^1\)

\(^1\)Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands, \(^2\)Institute for Pig Genetics B.V., Beuningen, The Netherlands*

Pheromonal communication plays an important role in the evolution of social signals resulting in sexual selection. Androstenone (AND) is a pheromone released from saliva in pigs, and attracts estrous females. As interaction between animals is important for the expression of AND, the AND phenotype of an animal can be affected by the genotype of another (referred to as an indirect genetic effect; IGE). Besides the function of AND in reproduction, AND is also one of the components that causes boar taint in pork; an unpleasant odour to the meat when cooked or heated. To prevent boar taint, and piglets are usually castrated early in life. Castration is undesirable for reasons of animal welfare, and selection against AND may provide an alternative. Selection against AND, however, could affect reproduction of both males and females. In this study, we investigated the genetic background of AND in boars, including IGEs, and searched for genomic regions affecting the direct and indirect genetic effect. Data included 6,245 boars housed in groups of on average 7 individuals, of which 4,455 had an AND observation. The estimated direct genetic variance was 0.35. The estimated indirect genetic variance was small (0.002) but significant, and contributed 12\% of the total heritable variance. The environmental effects of the group and compartment (separate room within a barn in time) were small, but significant.

A previous association study on partly the same population showed clear associations on SSC6 where several candidate genes were identified such as CYP2A19, SULT2A1 and SULT2B1. The present study validates these results in other genetic lines and narrows the confidence interval using more sophisticated models and using the improved build of Sus Scrofa. This association study is currently on going, and includes a search for genomic regions with IGEs on AND.

**P-213 – ABSTRACT WITHDRAWN**

**P-214**

DIRECT AND INDIRECT GENETIC EFFECTS FOR SURVIVAL IN PUREBRED AND CROSSBRED LAYING HENS

*Katrijn Peeters\(^2\), Tineke T. Eppink\(^1\), Esther E. Ellen\(^1\), Jeroen Visscher\(^2\), Piter Bijma\(^1\)*

\(^1\)Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands, \(^2\)Institut de Sélection Animale B.V., Hendrix Genetics, Boxmeer, The Netherlands*

Abstract can be found in the student symposium abstract section on page 62
P-215

NOVEL METHODS FOR CALCULATING GENOME-WIDE RELATEDNESS MATRICES AND EXAMINING GENOTYPE-ENVIRONMENT INTERACTIONS ACROSS CHROMOSOMES IN WILD PEDIGREE POPULATIONS

Matthew Robinson, Jon Slate
The University of Sheffield, Sheffield, UK

The environment experienced by individuals in the wild can influence both selection pressures acting upon traits, and their underlying genetic architecture. While current methods can provide great insights into the genomic regions underlying trait variation in natural populations, QTL mapping is limited by its ability to only reveal genes of large effect, and SNPs identified by genome-wide association studies often only explain a small fraction of the heritability. Here develop three new methods for calculating genome-wide relatedness matrices in natural pedigreed populations. We apply these to estimate the variance explained by all the SNPs of each chromosome, across a series of ecologically important traits, within a wild bird population. We then use factor analytic models to examine whether the variance explained by each chromosome changes as a function of both spatial and temporal environmental variation. Finally, we also examine whether trait-fitness covariance changes across the same environmental gradients. In this way, we are able to examine whether genotype-environment interactions are likely to influence evolutionary dynamics within a natural population.

P-216

GENETIC AND ENVIRONMENTAL FACTORS AFFECTING INCIDENCE OF FLOPPY KIDS IN NATIVE AND CROSSBRED GOATS OF KERALA.

Radhika G., Rani Alex, Jimcy Joseph, Raghavan K.C., Ajithkumar S.
Kerala Veterinary and Animal Sciences University, Pookode, Wayanad, Kerala, India

Floppy Kid Syndrome (FKS) is characterized by acute onset of sudden weakness, associated with metabolic acidosis in goat kids and such a condition existed in University Goat & Sheep Farm, Kerala, India. If untreated, these kids succumbed to death leading to high preweaning kid mortality rate. Reports of FKS from many parts of the state, initiated a study to investigate genetic and environmental factors affecting this syndrome. 687 kids born to Malabari, Attappadi Black purebred goats and Malabari crossbreds during 2008 to 2010 formed the material for study. Kids were classified into those which showed classical symptoms of FKS (affected animals) and those which were apparently healthy (unaffected). Disease affection data obtained was transformed (arcsine of square root of proportion affected) before analysis. Least square analysis revealed that overall affection rate was 50.38%. Genetic group (p<0.01), year of birth (p<0.01) and dam’s weight at kidding (p<0.05) had significant effect on the incidence of floppy kids in the farm, whereas sex, birth weight, type of birth and season of birth had no significant effect. Among genetic groups, Attappady Black goats were the least affected (35.46%), as these indigenous goats, known for its disease resistance capacity, must have resisted the gastrointestinal infection due to Escherichia coli, Clostridium perfringens and/or Cryptosporidium, which is suggested as the primary cause for FKS. This syndrome is reported in flocks with improved nutrition and milk production. Hence higher incidence of floppy kids among dams with comparatively better body weight at kidding is well justified. Results obtained from this study clearly indicate that genetic and environmental factors influence the incidence of floppy kids in goats.
P-217
CRYPTIC GENETIC VARIATION UNDERLYING STRESS RESISTANCE IN EVOLVED LINES OF CAENORHABDITIS REMANEI

Kristin Sikkink, Rose Reynolds, Catherine Ituarte, Patrick Phillips, William Cresko
University of Oregon, Eugene, OR, USA

When organisms become isolated in novel environments they must acclimate to new and often stressful conditions. If the population is to persist then subsequent adaptive evolution must also occur. But what is the relationship between individual acclimation and adaptation to novel environments? Many organisms acclimate to novel environments through phenotypic plasticity, the ability of a genotype to produce different phenotypes that match the environment. Underlying the exposure of novel phenotypes is cryptic genetic variation (CGV), which does not contribute to an organism’s phenotype until it is exposed by a stressful environment. Thus, variation in phenotypic plasticity is heritable, subject to selection, and can evolve. However, an open question is how important the evolution of plasticity is generally for adaptation to novel environments, and the molecular basis of CGV is poorly understood. To address this question, we used experimental evolution of Caenorhabditis remanei in a stressful environment. Lines were consistently evolved under high temperature or a control environment. The lines also were subject to one of three acute selective regimes – heat shock, oxidative shock, or no acute stress. Replicate lines were evolved under all combinations of long-term and acute stress, allowing us to isolate the effects of each factor. After 20-30 generations of selection, we observed adaptation to acute stress in lines selected for either acute stress. Furthermore, the response to selection was greater in the stressful heat environment, consistent with selection on CGV for acute stress resistance in the novel environment. To understand the molecular basis of the uncovering of CGV underlying acute stress resistance, we used RNA-seq to evaluate changes in gene expression resulting from the exposure of CGV in the novel environment, and to identify the key pathways that harbor CGV. These results provide a comprehensive understanding of the molecular basis of adaptation to a stressful environment.

P-218
GENETIC PARAMETER ESTIMATES OF THE ENVIRONMENTAL VARIATION FOR BODY WEIGHT AT HARVEST SIZE IN A PACIFIC WHITE SHRIMP BREEDING POPULATION

Héctor Castillo-Juárez1, Hugo H. Montaldo2, Gabriel Ricardo Campos-Montes1
1Universidad Autónoma Metropolitana, México D.F., Mexico, 2Universidad Nacional Autónoma de México, México D.F., Mexico

Pedigreed data from eight generations (2003-2010) of a Pacific white shrimp Mexican population Penaeus (Litopenaeus) vannamei selected for 130-day body weight (BW), were used in order to estimate the genetic parameters of the environmental variation (Vbw). A total of 123,549 records from 746 sires and 1022 females in a nested structure were analyzed. An animal model containing animal, female family common environmental random effects and fixed effects was fitted to obtain residuals with ASReml. Residuals (R) were transformed to LR2 as ln(R2). BW and LR2 were analyzed using a bivariate animal model with the same mentioned effects to obtain estimates of genetic parameters for BW and Vbw. Heritability estimate for BW was 0.19 ± 0.03 and for Vbw it was 0.04 ± 0.01. Genetic correlation estimate between BW and Vbw was 0.18 ± 0.14. Genetic correlations estimates between males and female records considered as different traits were close to 1 for both BW and Vbw. Linear genetic trends were positive for both BW (0.35 ± 0.001 g/year) and Vbw (0.02 ± 0.0001 units/year). Unadjusted coefficient of variation (phenotypic) linear trend for BW was also positive (0.59 ± 0.004 %). Selection for body weight at harvest increased the average breeding value for both BW and Vbw in this population and probably may have increased the phenotypic variability of BW as well.
P-219
MODELLING COMPETITION AT THE GENETIC AND NON-GENETIC LEVELS, WITH PARTICULAR APPLICATION TO FOREST GENETIC TRIALS

João Costa e Silva1, Richard Kerr2, Brad Potts1
1Centro de Estudos Florestais, Departamento dos Recursos Naturais, Ambiente e Território, Instituto Superior de Agronomia, Universidade Técnica de Lisboa, Lisboa, Portugal, 2PlantPlan Genetics Pty Ltd, Hobart, TAS, Australia, 3School of Plant Science, University of Tasmania, Hobart, TAS, Australia

Available experimental evidence suggests that there are genetic differences in the abilities of trees to compete for resources, in addition to differences due to micro-site variation. Predictions of response to selection will be biased if this extra heritable variance is not accounted for. The use of indirect genetic effects within the framework of linear mixed model methodology has been proposed for estimating genetic parameters and responses to selection in the presence of genetic competition. In this context, an individual's total breeding value reflects the effects of the direct breeding value on its own phenotype and the competitive breeding value on the phenotype of its neighbours. The present study used simulated data to investigate the use of an incorrect mixed model when competition is present at both genetic and non-genetic levels, in terms of the estimation of (co)variance components and selection response. Different experimental designs that resulted in different genetic relatedness levels within a neighbourhood and mortality were other key issues examined. The additive variance remained unbiased in a model that did not fit competition at the genetic and non-genetic levels, but only for designs associated with low relatedness. Models that fitted competition at one level only resulted in biased competitive and residual (co)variance estimates. The ability to detect the correct model was improved when the applied designs promoted some level of intra-family mixing within a neighbourhood. Selection response changed considerably between selecting on breeding value estimates from a model ignoring genetic competition and total breeding estimates using the correct model.

P-220
DETECTION OF EPISTASIS IN GENOME-WIDE ASSOCIATION STUDIES – TO PURSUE OR TO GIVE UP?

Wenhua Wei, Chris Haley
University of Edinburgh, Edinburgh, UK

The role of epistasis in complex traits has been a long lasting debate. It has been speculated epistasis could be a potential source of ‘the missing heritability’ encountered in most genome-wide association (GWA) studies, despite that epistasis remain largely unexplored due to the computational challenge in enumerating billions of pair-wise tests such studies. Recently Lander and colleagues (PNAS.1119675109) suggested that epistasis, likely acting in the form of pathways, could create substantial phantom heritability but be hard to detect statistically in GWA studies.

We have overcome the computational challenge by releasing a free tool BiForce developed for high throughput analysis of epistasis in either disease or quantitative traits of GWA studies on commonly used computer systems. Using BiForce we have conducted multiple studies of epistasis across a number of traits and GWA populations including Psoriasis, serum uric acid, obesity related and metabolic phenotypes (e.g. body mass index, waist circumference and lipids). We can show: a) Power is the major limiting factor to detect significant epistasis via pair-wise genome scans of GWA populations studied due to relatively small sample sizes; b) Significant epistasis signals likely involve loci with significant marginal effects and tend to exhibit local interactions; c) Evaluating sub-significant interaction signals in pathways can produce useful insights into the biology of the traits; d) Replication of epistasis signals is very hard at the SNP level, hard at the gene level but relatively easy at the pathway level; e) Detecting epistasis via pathway interactions could be a good alternative but requires more efforts to robustly construct pathways and interpret results.

We conclude that analysing epistasis in GWA studies is worthwhile but needs to be enhanced with
pathway based analysis to maximise its value. Besides, meta-analysis of epistasis in GWA populations is essential to boost power of detection.

P-221 – ABSTRACT WITHDRAWN

P-222 – ABSTRACT WITHDRAWN

P-223
EVOLUTION OF GROWTH PATTERN OF NELLORE CATTLE

Patricia Tholon1, Flávia Paterno2, Sandra Queiroz2
1Embrapa Pecuária Sudeste, São Carlos, São Paulo, Brazil, 2Universidade Estadual Paulista / Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal, São Paulo, Brazil

The use of growth functions allows the study of animal growth in adulthood, enabling the correlation between weights at different ages and attainment of adults with smaller size (A), a higher rate of weight gain (K) and younger age at the inflection point (K-1). The selection criteria for growth in Nellore cattle has changed in the last years, according to the demands of the consumer market, providing different patterns within the breed. The objective of this study was to determine the evolution of the growth pattern of Nellore cattle registered at the Brazilian Association of Zebu Breeders, born in 1995, 2000 and 2005.

For this purpose it was determined the mean values of A, scale parameter (B), K and the inflection point (M) which indicatives the onset of puberty, using the nonlinear model Richards. The average asymptotic weight for animals born in 1995, 2000 and 2005 were 844.8, 769.9 and 537.9 kg, respectively. The b values were 0.9927, 0.9928 and 0.9498, respectively. Values of K were 0.000533, 0.000645 and 0.00149 with age and the inflection point of about 59.6, 49.2 and 21.3 months. M values were 0.7231, 0.7205 and 0.8983 for the three birth years considered. The correlations between A and K, A and M and K and M were equal to -0.99, -0.95 and 0.97 in the three periods, showing the strong dependence between the parameters. It can be seen that a clear differentiation occurred in the growth pattern of the animals over time. The search for sexual precocity verified by k-1 was effective in reducing by half the age of onset of puberty, besides the production of adult animals best suited to lighter production systems of Nellore cattle.

P-224
INTERACTIONS BETWEEN VDR AND RXRA GENE POLYMORPHISMS AFFECT METABOLIC TRAITS IN THE 1958 BRITISH BIRTH COHORT

Alana Cavadino, K.S. Vimaleswaran, Diane Berry, Elina Hyppönen
Centre for Paediatric Epidemiology and Biostatistics and MRC Centre of Epidemiology for Child Health, UCL Institute of Child Health, London, UK

Vitamin D receptor (VDR) requires retinoid X receptor (RXRA) as a heterodimer partner to bind to vitamin D response elements, activating vitamin D responsive genes involved in various metabolic pathways. Although studies have investigated the main effects of VDR on metabolic traits, none have examined the interaction between the VDR and RXRA genes for these outcomes. We used a tag SNP approach to probe common genetic variations in VDR and RXRA and determine their role in contributing to metabolic phenotypes.

We examined 22 tag SNPs in VDR and 23 tag SNPs in RXRA using data from 5,321 individuals in the 1958 British Birth Cohort Study. Tag SNPs were selected using genotype data from the CEU population of the International HapMap project, phase II. SNP-SNP interactions on HDL, LDL cholesterol, triglyceride levels, glycated haemoglobin, body mass index (BMI), waist circumference and waist-hip ratio adjusted
Poster Abstracts

for BMI (adjusted WHR) were evaluated using linear regression. A joint likelihood ratio test (LRT) of the main SNP effects and the SNP-SNP interaction effects ($H_0: \beta_{S1} = \beta_{S2} = \beta_{S1S2} = 0$) was used to maximise statistical power. Multiple testing was corrected for using the False Discovery Rate (FDR).

After FDR correction, the joint LRT showed significant effects on triglycerides (7 SNP combinations, $p<0.04$), LDL cholesterol (2 SNP combinations, $p<0.02$), and adjusted WHR (29 SNP combinations, $p<0.05$). Twenty-two VDR SNPs and eight RXRA SNPs were implicated in these results. No significant interactions on other traits were observed after FDR correction.

We found potential evidence for interactions between VDR and RXRA on triglycerides, LDL cholesterol and WHR adjusted for BMI. These results suggest a combined effect of VDR and RXRA via SNP-SNP interactions, which might help our understanding of the pathogenesis of metabolic traits. Replication of these results is currently being investigated in the 1966 Northern Finland Birth Cohort Study ($n=5,400$).

**P-225**

**GENE POLYMORPHISM OF GROWTH HORMONE (GH) IN ONE OF INDONESIAN LOCAL CATTLE (MADURA) AND MADURA CROSSED CATTLE**

Mifta Pratiwi Rachman, Widya Asmara, Tety Hartatik
Gadjah Mada University, Sleman, Daerah Istimewa Yogyakarta, Indonesia

The present study was conducted to identify polymorphisms of the growth hormone (GH) gene in one of the Indonesian local cattle breeds (Madura) and Madura crossed (Limousin-Madura) cattle, and also its association with growth traits. Blood samples of 58 cattle were taken from the population of Madura and Madura crossed, then all cattle were located in the same area in Madura island. A 211 bp fragment of the GH gene (spanning from the fourth intron region to fifth exon) was amplified using GHF (forward) primer: 5’GCTGCTCTGTAGGCGCCCTC-3’, and GHR (reverse): 5’CATGACCGCTAGTGATCTCCC-3’.

Polymorphism was identified with Alul restriction enzyme. The result indicated that two genotypes LL and LV were found at the GH gene in the population of Madura crossed cattle. The frequencies of L and V alleles were 0.64 and 0.04, respectively. Madura and Madura crossed cattle had a significantly differences in their body weight estimation, chest girth, withers height, and body length. The appearance of the LV genotype in Limousin-Madura had no associated effect with body weight estimation and several body measurement yet detected.

Keywords: Polymorphism, growth hormone gene, Madura and Limousin Madura cattle.

**P-226**

**USING TRIBOLIUM CASTANEUM AS A MODEL ORGANISM TO INVESTIGATE SOCIAL INTERACTIONS AMONG INDIVIDUALS**

Esther D. Ellen1, Merel Verhoeven1, Katrijn Peeters1, Rieke Gols2, Michael J. Wade3, Marcel Dicke2, Piter Bijma1
1Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands, 2Laboratory of Entomology, Wageningen University, Wageningen, The Netherlands, 3Department of Biology, Indiana University, Bloomington, Indiana, USA

Nearly all organisms show social interactions. Social interactions may create so-called indirect genetic effects (IGEs), where trait values of individuals depend on genes in other individuals. IGEs contribute to the genetic variation in traits, and may have large effects on response to selection. To investigate the effect of IGEs on the genetic architecture of traits, it would be useful to have a small model organism with short generation time. Here we investigate whether flour beetles (Tribolium castaneum) are a suitable model organism. We will present estimated genetic parameters for pupal body mass (BM) and duration of development (DD). DD is defined as the number of days from placing parents together till pupation. In total, 5,940 beetles of two populations (Bhopal and Ring eye) were used. Five full sib larvae of both populations were placed together
in one vial. Each vial was filled with 1.5 gram standard medium and 0.5 gram α-cellulose to stimulate competition. Data were analysed separately for both populations. Mean BM was 2.37±0.01 mg for Bhopal and 2.34±0.01 mg for Ring eye, and mean DD was 27.0±0.05 days for Bhopal and 25.9±0.04 days for Ring eye. In both populations, females were heavier than males (P<0.0001). Results of a model fitting direct genetic effects yielded a moderate heritability of BM for both Bhopal and Ring eye (0.38±0.07 for Bhopal and 0.39±0.06 for Ring eye). Heritability of DD was low for Ring eye (0.10±0.03) and moderate for Bhopal (0.46±0.08). For both populations, the phenotypic correlation between BM and DD was 0.35±0.02, whereas the genetic correlation was 0.16±0.15 for Bhopal and 0.46±0.14 for Ring eye. These results demonstrate genetic variation in both traits. The next step will be to estimate IGEs for both traits.

P-227
EXPOSURE TO A NOVEL ENVIRONMENT IN CONJUNCTION WITH SERIAL BOTTLENECKS INCREASES PHENOTYPIC AND ADDITIVE VARIATION OF A QUANTITATIVE TRAIT

Taylor Callicrate1, Robert Fleischer2, Frank Siewerdt1
1University of Maryland, College Park, Maryland, USA, 2Smithsonian Institution, Washington, DC, USA

Demographic bottlenecks were replicated under laboratory conditions using Tribolium castaneum in order to determine how additive genetic variance, and therefore adaptive potential, may be affected by a novel environment. A two-level bottleneck was imposed. The first level was a 100-individual bottleneck (five replicates). After three generations at population size 100, the second bottleneck consisted of a single mating pair (15 sub-replicates stemming from each level one replicate). Six migrant individuals from the associated level one population were added to the sub-replicates at generation three. This design emulates an endangered species experiencing significant population decline, followed by a founder population being taken into captivity or translocated. Two growth media (environments), standard (flour) or novel (bran), were used in the sub-replicates to simulate translocation to an environment with a different foraging substrate (bran) and to determine how such a change would affect phenotypic or additive variances. Sub-replicates were managed for slow population growth during six generations (to population size 50) and were maintained at this size for nine additional generations. Pupa weights were measured in the progeny of single pair matings in generation 16 to estimate additive and phenotypic genetic variances. For both growth media, variance in pupa weight was significantly affected by drift from both levels of the bottleneck (flour: N=382, bran: N=470 pupae, both P<0.0001 at each bottleneck level). Total phenotypic variance was determined from the progeny of the single pair matings and was significantly larger for bottlenecks occurring in bran than in flour at the P<0.01 level (F469,381 = 1.29). Additive variance was also larger in bran than flour bottlenecks(F469,381 = 1.38). These results indicate that bottlenecks occurring in a novel environment could potentially increase the effectiveness of selection by bolstering additive variance.

P-228
GENOME-WIDE ASSOCIATION STUDY FOR GENETIC HETEROGENEITY FOR MILK YIELD AND SOMATIC CELL SCORE

Freddy Fikse1, Lars Rönnegård1,2, Han Mulder3, Erling Strandberg1
1Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics, Uppsala, Sweden, 2Dalarna University, Statistics Unit, School of Business and Technology, Borlänge, Sweden, 3Wageningen University, Animal Breeding and Genomics Centre, Wageningen, The Netherlands

Recently, genetic variation in residual variance was reported for both Swedish Holstein and Swedish Red. The aim of this study was to perform a genome-wide association study for this genetic heterogeneity. Breeding values for residual variance for milk yield and somatic cell score were available for 842 Swedish Red bulls. In addition, marker data were available for 701 bulls using the Illumina Bovine SNP50
BeadChip, which includes 54,001 single nucleotide polymorphisms (SNP) markers. After edits on minor allele frequency, call rates and GenCall scores more than 48,000 markers were available to be included in the analyses. A multi-locus Bayesian stochastic search variable selection model was used for the analysis. Here, allele effects follow a priori a mixture distribution, where a small fraction of the markers (prior probability of 5%) has a large effect and the remaining markers have virtually no effect. For milk yield in the Swedish Red breed, two regions with a Bayes factor larger than 10 were identified and a further five regions had a Bayes factor larger than three. For somatic cell score, the corresponding numbers were one and four, respectively. In conclusion, we found a few SNPs associated with residual variance of milk yield and somatic cell score in the Swedish Red breed.

P-229
THE BENEFIT OF USING MOLECULAR COANCESTRY IN THE MANAGEMENT OF POPULATIONS UNDER CONSERVATION AND ITS DEPENDENCY ON EFFECTIVE POPULATION SIZE AND MARKER DENSITY

Fernando Gómez-Romano, Beatriz Villanueva, Ángeles De Cara, Jesús Fernández
Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Madrid, Spain

The consensus method to maintain genetic diversity in populations under conservation programmes is to optimise, for each potential parent, the number of offspring left to the next generation by minimising the global coancestry. Coancestry is usually calculated from genealogical data but molecular markers could be used to replace genealogical coancestry with molecular coancestry. Recent studies have shown that optimising contributions based on coancestry calculated from a large number of SNP markers could maintain higher levels of diversity than optimising contributions based on genealogical data. In this study we investigate how SNP density and effective population size (Ne) impact the effectiveness of using molecular versus genealogical coancestry in maintaining diversity. A base population in mutation-drift equilibrium was obtained after 5000 generations of random mating, and then the population was managed for ten discrete generations using genealogical or SNP information to compute coancestry. The genome was composed of 20 chromosomes of 1 Morgan each. Two types of biallelic loci were simulated: markers (used for management) and non-markers (used for measuring diversity through expected and observed heterozygosity). The number of markers varied across scenarios, and the number of non-markers was always 1000 per Morgan. Levels of genetic diversity maintained when using genealogical coancestry were higher than those maintained when using molecular coancestry at low densities of SNP. However, with high marker densities there was a benefit of using molecular coancestry. The number of markers required for molecular coancestry to outperform genealogical results depended on Ne, and ranged from about 100 markers/Morgan for Ne = 20 to about 500 markers/Morgan for Ne = 160. The diversity maintained using 1000 markers was close to the potential maximum obtained when using the non-markers' genotypes to compute coancestry. Therefore, the expected benefit from increasing the number of markers over 1000 is small.

P-230 – ABSTRACT WITHDRAWN
P-231
SHOULD ONE AIM FOR GENETIC IMPROVEMENT OF HOST RESISTANCE OR TOLERANCE TO INFECTIOUS DISEASE?
Andrea Doeschl-Wilson¹, Ilias Kyriazakis², Beatriz Villanueva³
¹The Roslin Institute, Easter Bush, Midlothian, UK, ²Newcastle University, Newcastle upon Tyne, UK, ³INIA, Madrid, Spain

Recent advances in genomics provide new opportunities for dissecting host genetic response to infectious pathogens and to accelerate genetic improvement. There are two alternative host defence mechanisms to infectious pathogens that could be targeted for genetic improvement - host resistance vs host tolerance. Resistance refers to mechanisms that restrict the within host pathogen reproduction rate, whilst tolerance mechanisms aim at minimising pathogen inflicted damage on the host. Both strategies can have contrasting effects on population performance and epidemiology. Improving host resistance may result in successful disease eradication, whereas this may be difficult for tolerant hosts who harbour pathogens without showing symptoms. However, evidence suggests that increasing host resistance may drive pathogen evolution towards higher virulence. Further, whereas resistance mechanisms are often pathogen specific, tolerance mechanisms are associated with the host, and thus more likely to be generic to a range of pathogens. Thus, improving tolerance may be beneficial if individuals are exposed to a variety of pathogens, and where disease eradication has proven difficult. The aim of this study was to develop a systematic and comprehensive framework for determining under what conditions improving host genetic disease resistance would be favourable over improving tolerance, and vice versa. Influencing factors and evidence for their effects from literature considered in this framework include (i) the breeding goal (e.g. disease eradication or sustainable production), (ii) (knowledge of) the host genetic (co-)variance and architecture underlying resistance or tolerance mechanisms (e.g. single gene or polygenic, pleiotropy), (iii) epidemiological characteristics (e.g. pathogen virulence, transmission patterns, presence of asymptomatic carriers), and (iv) undesirable side-effects. Theoretical concepts will be illustrated with examples from various livestock diseases and required information for anticipating the outcome of selection will be outlined.

P-232
USING PEAK SHAPE TO IMPROVE EFFICIENT AND EFFECTIVE DETECTION OF MULTIPLE QTL IN KNOWN CROSSINGS
Carl Nettelblad, Behrang Mahjani
Uppsala University, Uppsala, Sweden

Objectives: With full genome resequencing or efficient screening of millions of SNPs, existing methods for scanning for QTL in ordered crosses need to be revisited in order to exploit these advances. We analyse linear regression models for assessing QTL effects based on genotype and phenotype data (e.g. the Cockerham model), in order to better understand the behaviour of the function over short- to mid-range distances. This understanding allows better accuracy in QTL detection.

Methods: Linear regression models are frequently used in order to detect QTL, based on metrics derived from the explained variance in the model. We propose one specific such metric, which renders a well-defined expected peak shape around a QTL, assuming full genotype data. Such theoretic endeavours have previously been of limited use, as historically sparse genotype data limited their practical use. We also investigate the variations in peak shape due to finite size population sampling using mixture models.

Results: Based on the knowledge of the expected peak shape and its variation, we discuss various issues, including how many points need to be evaluated in a simultaneous multi-QTL search to with high confidence be sure to detect any pair or triplet with a certain value for the explainable variance. The possible speedup of several orders of magnitude compared to exhaustive search from the use of algorithms such as DIRECT is given a theoretical explanation, quantifying how to control detection power.
Conclusions: The knowledge of the expected ideal peak shape based on our proposed metric can be used to better understand possible detection power. Furthermore, such knowledge can be used to detect errors. Based on simulations and experimental data, we discuss typical behaviour as well as “corner cases” such as identification of extremely closely linked QTL of identical and opposite effect.

P-233
SCAFFOLDING THE HELICONIUS MELPOMENE GENOME WITH RAD SEQUENCING

John W. Davey1, Simon W. Baxter2, Luana S. Maroja4, Durrell K. Kapan3, Chris D. Jiggins2, Mark L. Blaxter1,5
1Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK, 2Department of Zoology, University of Cambridge, Cambridge, UK, 3Department of Entomology & Center for Comparative Genomics, California Academy of Sciences, San Francisco, California, USA, 4Department of Biology, Williams College, Williamstown, Massachusetts, USA, 5The GenePool Genomics Facility, University of Edinburgh, Edinburgh, UK

It is becoming increasingly easy to sequence and assemble draft genomes of reasonable quality with next generation sequencing technologies. However, a substantial bottleneck is assigning draft genome scaffolds to chromosomes in order to produce a chromosomal map of the genome. We have recently completed a map of the Heliconius melpomene genome using a linkage map constructed using RAD Sequencing. 43 intercross progeny were RAD sequenced using the 6 bp restriction enzyme PstI in one HiSeq lane. RAD reads were aligned to the 273 Mb H. melpomene genome. 31 Kb of high quality candidate marker positions were identified, on scaffolds comprising 246 Mb of the genome. A linkage map containing 1,040 unique markers of length 1,455 cM was constructed from the candidate marker data. This map was used to assign 231 Mb (86%) of the genome to 21 chromosomes and was also used to correct 141 misassembled scaffolds. This novel approach is cheap and accessible to any researcher able to produce a mapping cross for a species of interest and we hope it will prove valuable for the assembly of many non-model genomes.

P-234
SEARCHING CANDIDATE MUTATIONS FOR CHICKEN GROWTH USING GENETIC DIVERGENCE ANALYSIS WITHIN QTL REGIONS

Xidan Li1, Muhammad Ahsan1, Andreas E. Lundberg1, Marcin Kierczak1, Paul B. Siegel2, Örjan Carlborg1, Stefan Marklund1
1Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics, Uppsala, Sweden, 2Department of Animal and Poultry Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

QTL mapping is a first step to identify chromosomal regions harbouring genetic polymorphisms that regulate complex traits. Searching the causative mutations for the observed effects is a sometimes daunting task as even after fine mapping of the QTL, millions of basepairs including many genes will typically need to be explored. There is thus a great need for efficient bioinformatics strategies to trace the causative mutation(s). Here, we searched for mutations regulating body weight at 56 days traits in the Virginia chicken lines – an experimental population comprising two lines that have been divergently selected for 8 week body weight for more than 50 generations. Several QTL regions have been mapped in an F2 intercross between the lines, and the regions have subsequently been replicated and fine mapped using an Advanced Intercross Line. Candidate mutations were here sought in the parts of the QTL regions where the highest genetic divergence between the high weight selected (HWS) and low weight selected (LWS) lines was observed. Such regions, 46 Mb or 34% of the QTL regions in total, were identified by comparing the allele frequencies in the genomes of the HWS and LWS lines using both individual 60K SNP chip genotyping of birds and analysis of read proportions with 12X ABI SOLID genome resequencing.
of DNA pools. Gene transcripts in the target segments, obtained using the Ensembl genome browser 65, were used as input in a pathway analysis using the NCBI Entrez gene, Gene Ontology and KEGG tools. Six growth-related pathways were revealed, represented by a total of 19 genes, four of which contained line-specific non-synonymous mutations. These will be prime candidates for further investigations but the target regions will also be further searched for polymorphisms at locations of presumed functional importance using the resequencing data and a Python script pipeline.

P-235 – ABSTRACT WITHDRAWN

P-236
POWER IN GWAS: ADVANCES IN PHENOTYPING

Sophie van der Sluis1, Danielle Posthuma1, Michel G. Nivard2, Matthijs Verhage1, Conor V. Dolan3
1VU University Amsterdam & VU University medical center, Amsterdam, The Netherlands, 2VU University Amsterdam, Amsterdam, The Netherlands, 3University of Amsterdam, Amsterdam, The Netherlands

Despite the considerable heritability of psychiatric traits, identified variants explain little variation, and insight into the role of genes in psychiatric disorders remains poor. Low statistical power is seen as the main reason, resulting in a call for larger samples. We show, however, that better use of (available or to-be-collected) phenotypic information can also increase power considerably.

The case-control phenotype used in genetic association studies is usually obtained by applying clinical cut-off criteria to continuous yet skewed data collected with clinical questionnaires or interviews. Assuming that the common-trait common-variant hypothesis holds (i.e., that complex psychiatric disorders result from the concerted effect of many variants of small effect), this dichotomization of test scores removes all variation present in the test scores among controls, that could be informative for the detection of causal variants. Using simulation, we show that the power to detect causal variants is much higher when the original skewed test scores are used, compared to various forms of categorizations.

Furthermore, clinical instruments include mostly items referring to extreme behavior, and thus mainly distinguish between cases and controls: they provide little information about phenotypic differences among controls, who rarely report such extreme behavior. Our simulations shows that use of phenotypic instruments that also resolve phenotypic differences within the control population could improve the statistical power to detect causal variants even more.

In summary, our studies show that the use of clinical instruments and clinical criteria is very disadvantageous for the power of genetic association studies. If the common-trait common-variants hypothesis holds, new phenotypic instruments should be developed that resolve phenotypic differences across the entire trait range. This would result in increased phenotypic variation, and thus in increased statistical power to detect common causal variants or small effect.

P-237
GENETIC ANALYSIS OF SOW LONGEVITY IN DANISH PIG

Rafael Pimentel Maia1, Rodrigo Labouriau1, Per Madsen1, Bjarne Nielsen2
1Aarhus University, Aarhus, Denmark, 2Danish Pig Production, Copenhagen, Denmark

The main purpose is to present a statistical model that characterizes the longevity of sows and to describe possible genetic trends. The data is composed of 100,027 Yorkshire and 200,084 Landrace pure sows from Danish multiplier herds that had the first parity between 1999 and 2010. Two traits were proposed to measure sows longevity: A) the time from the day of the first parity to the death/culling day and B) the total number of parities. Studying in detail the trait A we demonstrated that the velocity of mortality (hazard function) follows essentially the same pattern over time during the different parities, but with increasing
levels from one lactation to the other. Using a piece-wise constant hazard model for A and a discrete hazard model for B, we shown that the models constructed with each of those traits present similar estimative of fixed effects (age at the first parity, season-year, herd and litter size) and similar random components representing genetic and environment variation. However, the models based on A presented a residual variance 3.5 and 4 times larger than the models based on B. We conclude that the model based on B seems to be a better choice to describe the longevity of the sows.

P-238
GENOME-WIDE ASSOCIATION ANALYSIS FOR OLFACTORY BEHAVIOR IN THE DROSOPHILA MELANOCASTER GENETIC REFERENCE PANEL

Robert Anholt, Gunjan Arya, Shilpa Swarup, Trudy Mackay
North Carolina State University, Raleigh, North Carolina, USA

Olfactory behavior is a quantitative trait determined by multiple segregating genes that are sensitive to the environment. The recently developed Drosophila melanogaster Genetic Reference Panel (DGRP) enables us to capitalize on natural variation to investigate the genetic basis of variation in olfactory behavior. The DGRP consists of 192 wild-derived inbred lines with sequenced genomes and identified sequence variants. Genetic variation among individuals within each line is minimal, whereas genetic variation among the lines is maintained and reflects the variation of the population from which they were derived. The lines are genetically variable for olfactory behavior, and phenotypic variation within the DGRP far exceeds that observed among Drosophila laboratory strains. We have measured olfactory responses of the DGRP lines to 14 odorants and identified 571 genes that harbor SNPs associated with phenotypic variation in responses to one or more odorants. Polymorphisms associated with variation in responses to different odorants do not necessarily reside in olfactory receptors, but cover a range of gene ontology categories, with overrepresentation of genes associated with functions of the nervous system, and 71 genes comprising a network centered on axonal path finding and neural connectivity. This unexpected finding suggests that, unlike physiological measurements of antennal sensilla, variation in behavioral responses to odorants depends on the perceptual integrated neural representations of chemosensory sensations, and variation in responses to different odorants accompanies variation in different molecular neural substrates. Analyses of segregating alleles in advanced intercross DGRP-derived populations and RNAi targeted gene knockdown can be used for functional validation of associated SNPs.

P-239 – ABSTRACT WITHDRAWN

P-240 – ABSTRACT WITHDRAWN

P-241
FRAMEWORK FOR DISSECTING COMPLEX CYTONUCLEAR EPISTASIS BY TWO-DIMENSIONAL GENOME SCAN

Zaixiang Tang1,2, Yonghong Zhang1, Chenwu Xu2
1School of Public Health, Medical College of Soochow University, Suzhou, China, 2Jiangsu Provincial Key Laboratory of Crop Genetics and Physiology; Key Laboratory of Plant Functional Genomics of Ministry of Education, Yangzhou University, Yangzhou, China

Objectives: Epistasis among cytoplasm and nuclear is the primary genetic component of complex quantitative traits. Dissecting biologically interesting cytonuclear epistasis is fundamentally important for understanding the genetic architecture of complex traits. The objective of the current research is
to develop statistical models and methods for mapping and evaluating the QTL which interact with cytoplasm background in the presence of an interaction among nuclear genes.

**Methods:** In this study, to introduce the quantitative variation of cytoplasmic effect and their interactions with nuclear genes, a reciprocal mating design was suggested to create the segregation population that contains different combinations of nuclear genotypes and cytoplasms. A two-dimensional genome scan strategy which treated cytoplasmic effects as a covariate was proposed to evaluate the contribution of cytoplasm, quantitative trait loci (QTL), QTL×QTL interactions, QTL×cytoplasm interactions and QTL×QTL×cytoplasm interactions to the phenotypic variation. The proposed statistical model incorporated important interactions within nuclear and interactions among single nuclear loci and cytoplasm background.

**Results:** By multiple regression analysis, the parameter values and corresponding $p$ value were calculated for each genetic effect. Then, a stepwise approach was suggested to build confidence in candidate QTL on the basis of $q$-value estimation, false discovery rate calculation and Bonferroni adjustment. A fine-scale grid scan strategy was also proposed for further fine grid mapping in interesting chromosome regions. An example, plant height in maize, illustrates the efficiency of the two-dimensional genome scan strategy.

**Conclusions:** Cytoplasmic effect acts as an important and significant contributor and should be taken account of in complex trait mapping. Its inclusion will reduce residual variation and so can enhance our ability to detect QTL. In this article, a framework for dissecting complex cytonuclear epistasis based on a reciprocal mating design was proposed, which provides a new tool to dissect and understand the genetic basis of complex traits.

**P-242**

**GENOME-WIDE ASSOCIATION STUDY OF SLEEP IN DROSOPHILA MELANOGASTER**

*Susan Harbison¹, Lenovia McCoy², Trudy Mackay²*

¹National Heart Lung and Blood Institute, Bethesda, MD, USA, ²North Carolina State University, Raleigh, NC, USA

Sleep is a highly conserved behaviour, found in every animal species in which it has been studied. Paradoxically, the duration and pattern of sleep varies considerably both among and within species. The source of this variation remains unknown. To identify polymorphisms contributing to variation in sleep, we performed a genome-wide association of ~2.5 million single nucleotide polymorphisms (SNPs) with sleep in *D. melanogaster* using the *Drosophila* Genetic Reference Panel (DGRP). The DGRP population consists of wild-derived inbred lines, which facilitated the discovery of genetic variants that impact both mean sleep and the sensitivity of sleep to random environmental fluctuations. Many SNPs had sex-specific or sex-biased effects, and were generally located in non-coding regions. Most SNPs (72.4%) affecting sleep occurred at low frequency and had moderately large effects, suggesting that rare variants have an important role in maintaining genetic variation in sleep. Many novel loci were associated with natural variation in sleep, including genes that interact with the *Epidermal growth factor receptor* pathway. We found SNPs in *Drosophila* homologues that were previously implicated in human sleep, suggesting that genes affecting variation in sleep may be conserved across species. We confirmed the role of novel loci such as *Vmat*, *Ultrabithorax*, *thickveins*, and *frizzled* on sleep. Our discovery of genetic variants that influence environmental sensitivity to sleep may have a wider application to GWA studies, because individuals with genotypes sensitive to small perturbations in the environment will not exhibit consistent phenotypes.
P-243
GENOME-WIDE ASSOCIATION ANALYSIS AND GENETIC ARCHITECTURE OF EGG WEIGHT AND EGG UNIFORMITY IN LAYER CHICKENS

Anna Wolc1,2, Jesus Arango3, Petek Settar3, Janet Fulton3, Neil O’Sullivan3, David Habier2, Rohan Fernando2, Dorian Garrick2, William Hill4, Jack Dekkers2
1Poznan University of Life Sciences, Poznan, Poland, 2Iowa State University, Ames, USA, 3Hy-Line International, Dallas Center, USA, 4University of Edinburgh, Edinburgh, UK

Understanding the factors that determine the phenotype and its variation is of great importance to animal breeders. In this study, genomic prediction model Bayes-B was used to identify genomic regions associated with the mean and standard deviation of egg weight at three ages in a commercial brown egg layer line. A total of 24,425 segregating SNPs were evaluated simultaneously using over 2,900 genotyped individuals or families. Corresponding phenotypic records were represented as individual measurements or family means from full-sib progeny. The posterior distribution of window variances from the MCMC samples was used to describe the trait genetic architecture and for making statistical inferences about regions with the largest effects. A region on chromosome 4 was found to explain a large proportion of genetic variance for the mean (30%) and uniformity (up to 16%). Additional regions with smaller effects on chromosomes 2, 5, 6, 8, 20, 23, 28 and Z were suggestively associated with egg weight mean and a region on chromosome 13 with the age specific standard deviation. The genetic architecture of the analyzed traits was characterized by a limited number of genes or genomic regions with large effects and many regions with small polygenic effects. The region on chromosome 4 can be used to improve both the mean and uniformity of egg weight by marker-assisted selection.

Funded by USDA NIFA AFRI Award 2010-65205-20341 and the EW group.

P-244
GENOME-WIDE ASSOCIATION ANALYSIS OF NATURAL VARIATION IN TERGITE MELANIZATION IN DROSOPHILA MELANOGASTER

Lauren Dembeck1,2, Michael Magwire1,2, Faye Lawrence1, Richard Lyman1, Trudy Mackay1,2
1North Carolina State University, Raleigh, NC, USA, 2W.M. Keck Center for Behavioral Biology, Raleigh, NC, USA

Pigmentation varies within and between species and is often an adaptive trait crucial for fitness. D. melanogaster females generally have light to medium stripes of melanization on the posterior end of each tergite. We measured natural variation in female abdominal pigmentation in 158 sequenced inbred lines of the Drosophila Genetic Reference Panel, derived from the Raleigh, NC population. We visually scored females for the proportion of melanization on tergites 4-6 on a scale from 0 (no melanization) to 5 (total melanization). We found significant genetic variation in melanization for each tergite, with broad sense heritabilities ranging from 0.239 – 0.876. We performed genome-wide association analyses for each tergite using 2.5 million single nucleotide polymorphisms (SNPs). We identified 30 SNPs associated with the proportion of melanization on tergite 6, the most significant of which was in bric-a-brac 1, a gene known to affect the proportion of melanization in female D. melanogaster. We also identified a SNP in the cis-regulatory element of tan, which was previously shown to affect interspecific differences in pigmentation. After accounting for linkage disequilibrium and imputing missing genotypes, we conducted a forward regression to estimate the fraction of variance explained. Five SNPs constituting 22 haplotypes account for 52% of the variance among the lines. We are currently conducting studies to further confirm the effects of these SNPs. This study will provide insight into the genetic architecture of melanization and also shed light on the question of whether genes causing variation within a species are the same as those involved in trait divergence between species.
P-245
A TEST OF THE INFINITESIMAL MODEL

Hossein Jorjani
Interbull Centre, Department of Animal Breeding & Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden

The use of the infinitesimal model (TIM) in quantitative genetics has been motivated mainly by the mathematical traceability of TIM. However, it is not clear whether TIM is far from or rather close to the biological reality. One point of debate is whether TIM is equally valid for low heritability traits as for high heritability traits or not. Evolutionary considerations suggest that the majority of loci affecting fitness related traits are probably homozygote for the favourable allele. Non-fitness traits, however, are affected by loci segregating for different alleles. I postulate that the evolutionary history of these two types of traits causes the distribution of SNP effects to show different patterns depending on the heritability of the trait under study. The aim of this study was to test the conformity of a large number of quantitative traits to TIM. Number of SNPs with large effect (> 2 SD), and location and dispersion parameters of 42826 SNP effects estimated for a large number of traits were correlated to the heritability of these traits. The SNP effects were estimated on training populations of size 2108-5163 dairy bulls genotyped with an array of > 54000 SNPs. For estimation of SNP effects an iterative non-linear model with heavy tailed prior with a curvature parameter of 1.05 for SNP effects (analogous to Bayes A) was used. The bulls originated from 6 Brown Swiss populations. Up to 33 traits from each population, summing up to 140 traits, were analysed. Results indicated that SNP effects have a lower average for traits with high heritability. Further, regression of number of SNPs with large effect on the trait heritability had a highly significant (p < 0.001) negative slope. It may be concluded that the high heritability traits conform better to the infinitesimal model than the low heritability traits.

P-246
UNRAVELLING COMPLEX VARIATION IN MEASURES OF GENERAL INTELLIGENCE AND COGNITIVE DECLINE

Suzanne Rowe1, Gail Davies1, Sarah Harris1, Geraldine McNeill1, John Starr2,4, Dave Liewald1, Ian Deary∗2,5, Albert Tenesa∗1,6
1The Roslin Institute, Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Roslin, Midlothian, UK, 2Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK, 3Institute of Applied Health Sciences, University of Aberdeen, Forres[erhill, Aberdeen, UK, 4Geriatric Medicine Unit, University of Edinburgh, Edinburgh, UK, 5Department of Psychology, University of Edinburgh,, Edinburgh, UK, 6Department of Psychology, University of Edinburgh,, Edinburgh, UK

Heritable genetic variation exists for human intelligence and less certainly for age associated cognitive decline. Rates of cognitive change vary widely in healthy individuals affecting health status, longevity and quality of life. Phenotypic measures of lifetime cognitive changes in healthy individuals are extremely rare. Measures of intelligence were available for 1806 individuals, in childhood (age 11) and at adult age (65-79) enabling construction of a phenotypic measure for lifetime cognitive change. As with many complex traits GWAS have failed to explain more than a few percent of genetic variation in IQ. Objectives were to investigate genomic variation of intelligence using a regional heritability approach. Extending single SNP analyses to the estimation of heritability of larger regions captures additive variation which might elude the stringent significance thresholds necessary for testing each SNP individually. Rare variants not in complete LD with common SNP markers may be captured by estimating the genetic variation from an entire “region” or set of SNPs. After QC, we divided 547 750 SNPs ranked by position into overlapping windows of 101 SNPs. A linear model was constructed to estimate the genetic variance explained by each of the resulting 10 908 genomic regions for fluid and crystallised intelligence at adult age and for lifetime cognitive change.
Related individuals were omitted (>0.025) and PCA fitted as covariates to control population stratification. A likelihood ratio test for a linear mixed model containing a genomic region vs. a null model was used. The ten regions of greatest significance within each trait (non-overlapping) were fitted in a model with a polygenic component comprised of the remaining SNPs in the genome. These 10 regions explained 60, 86 and 42% of the genetic variation for cognitive decline, fluid and crystallised intelligence respectively. Furthermore, these regions contain genes associated with autism, schizophrenia, and cognitive processing.

P-247 – ABSTRACT WITHDRAWN

P-248

HOW IS ENERGY BALANCE OF DAIRY COWS GENETICALLY DETERMINED DURING THE COURSE OF LACTATION?

Nina Buttchereit1, Eckhard Stamer2, Wolfgang Junge1, Jens Tetens1, Georg Thaller1

1Institute of Animal Breeding and Husbandry, Christian-Albrechts-University, Kiel, Germany, 2TiDa Tier und Daten GmbH, Westensee/Brux, Germany

Postpartum energy status of dairy cows is critically important to health and fertility, and continued selection for high milk production will intensify the energy deficit in early lactation unless energy balance (EB) is directly or indirectly included in breeding programs. It remains a major task to develop a suitable breeding strategy to overcome this problem. Therefore, heritability estimates and genetic correlations between daily EB estimates across lactation were evaluated using random regression methodology on data collected from 682 Holstein Friesian primiparous cows recorded between lactation day 11 to 180. Heritability was found to be highest in early lactation. The results indicate a negative association between EB observations in early and mid lactation.

In addition, curves for daily breeding values of EB were generated. The curves show that cows with positive EB breeding values in early lactation are likely to have a comparatively more balanced energy status. Cows with negative EB breeding values at the beginning, on the other hand, swing from one extreme to the other during lactation.

EB might be influenced by different genetic mechanisms within the course of lactation. On average, primiparous cows exhibited a negative EB during the first 42 lactation days. Thus, the first six weeks in milk are of special interest. Whole genome association studies were done using breeding values for day 11, 20, 30 and 42 in milk from 586 cows as phenotypes. Applying a genome-wise significance threshold level of \( p \leq 0.05 \), three SNPs significantly associated with EB were detected on chromosomes 11 and 13. The two SNPs on chromosome 11 show a significant association for all analyzed days, whereas the SNP on chromosome 13 only passed the significance level for day 30. Further studies are on the way to determine whether QTL affecting EB are active only at certain lactation stages.

P-249

SYSTEMS GENETICS OF PAIN- AND RELIEF-LEARNING

Mirjam Appel1, Claus-Jürgen Scholz2, Sinead Savage1, Samet Kocabey1, Sami Gödekdağ1, Hiromu Tanimoto1, Ayse Yarali1

1Max Planck Institute for Neurobiology, Martinsried, Germany, 2Institute of Clinical Biochemistry and Pathobiochemistry, University of Würzburg, Würzburg, Germany

All animals learn the causal relationships between events to adapt their behaviour. Drosophila is a suitable model for revealing the genetic architecture behind such ‘associative learning’, as it combines well established behavioural paradigms with accessibility to genetic intervention. Indeed, over 100 genes have been discovered to affect the flies’ ability to associate an odour with an electric shock. An account
of how these act as a network is yet lacking. Moreover, a whole ‘other side’ of learning has so far been ignored: Flies also associate an odour with the cessation of shock, consequently approaching it as a signal for ‘relief’\(^1\). Relief-learning is conserved across phyla, up to humans\(^2\) and arguably balances out the effects of pain memories in health and psychiatric disease. The genetic bases for relief-learning are entirely unknown. We aim to fill this gap with a systems genetics approach.

We have characterized 38 wild-derived inbred Drosophila strains\(^3\) in pain- and relief-learning, uncovering significant across-strain variation. We then brought these behavioural data together with the available genome-wide transcript abundance data\(^3\) to reveal associations. We can now hypothesize which genes uniquely affect pain- or relief-learning, which ones affect both, particularly, which act on the pain- vs. relief-learning balance. We are now testing a number of these hypotheses using reverse genetics. In parallel, we are looking for associations between genome-wide sequence polymorphisms\(^4\) (SNPs) and learning. At the end, we will integrate the transcript-learning and the SNP-learning associations into gene networks, which will provide superior hypotheses on how pain- and relief-learning work in balance with each other, also paving the road to translational studies.

\(^4\)http://www.hgsc.bcm.tmc.edu/project-species-i-Drosophila_genRefPanel.hgsc

P-250
GENETIC RELATIONSHIPS BETWEEN MASTITIS, OTHER DISEASES, AND PREDICTORS OF MASTITIS IN CANADIAN DAIRY CATTLE

Astrid Koeck\(^1\), Filippo Miglior\(^2\), David Kelton\(^1\), Flavio Schenkel\(^1\)

\(^1\)Centre for Genetic Improvement of Livestock, Department of Animal & Poultry Science, University of Guelph, Guelph, Ontario, Canada, \(^2\)Guelph Food Research Centre, Agriculture and Agri-Food Canada, Guelph, Ontario, Canada, \(^3\)Canadian Dairy Network, Guelph, Ontario, Canada, \(^4\)Department of Population Medicine, Ontario Veterinary College, University of Guelph, Guelph, Ontario, Canada

This study is part of a larger project whose overall objective is to develop genetic evaluations for resistance to mastitis and other diseases in Canada. Health data recorded by producers were available from the national dairy cattle health recording system. Eight diseases are recorded by producers on a voluntary basis: mastitis, displaced abomasum, ketosis, milk fever, retained placenta, metritis, cystic ovaries and lameness. The overall objective of this study was to investigate genetic relationships between mastitis, other diseases, and predictors of mastitis. In the first part of this study, genetic parameters of mastitis and other diseases were estimated. Heritabilities of health traits were low and ranged from 0.01 for lameness to 0.06 for displaced abomasum. Genetic correlations between diseases were mostly positive. The strongest genetic correlations were found between displaced abomasum and ketosis (0.64) and between retained placenta and metritis (0.62). Pearson correlations between breeding values for health traits and other routinely evaluated traits were computed, which revealed noticeable favorable relationships to direct herd life and fertility. In the second part of this study, genetic relationships between mastitis and its predictors were examined. The analyzed predictors of mastitis included mean somatic cell score, alternative somatic cell count traits, udder conformation traits and body condition score. The alternative somatic cell count traits had lower heritability, but stronger genetic correlations with mastitis than mean somatic cell score. Udder depth and fore udder attachment were the most highly correlated udder conformation traits with mastitis. A moderate favorable genetic correlation was found between body condition score and mastitis. Overall, the results of this study showed that producer-recorded health data can be used for genetic evaluations for disease resistance in Canada.
Mouse mandible is a classical model system in quantitative genetics. Its shape is determined by a complex sequence of developmental processes and later remodeling and thus its genetic architecture may integrate contributions from many genomic networks. However, experiments suffered until recently from both low genetic resolution and diversity to fully described the complexity of this genetic architecture.

Geometric morphometrics of 2,000 mice from an 8-way heterogeneous stock was used in combination with a 12K SNP map for high-resolution mapping of QTLs that affect mandible shape. Using model resample averaging, 52 additive QTLs were mapped in an interval of 2.5 Mb in average. Modal effect size is lower than 1% of the total phenotypic variance. Nonetheless, all together these additive QTLs account for most of the additive variation estimated from the pedigree which represents 55% of the total variance and all dimensions of the shape space show additive variation. Moreover, computer simulations suggest that this number of discovered QTLs underestimates the true number of QTLs, perhaps by a large extent. Simulations suggest that although estimations of effect sizes were consistently small, they likely overestimate the true QTL effects.

Altogether, these results underline the highly polygenic nature of the genetic architecture of mandible shape which exceeds by far the number of QTLs found for biochemical, physiological and behavioral traits in the same animals and with directly comparable statistical methods.

MENDELIAN RANDOMIZATION STUDIES TO UNDERSTAND THE ROLE FOR REDUCED ADIPONECTIN LEVELS IN INSULIN RESISTANCE OR TYPE 2 DIABETES

Hanieh Yaghootkar1, Liling Warren2, Debbie Lawlor1, Robert Scott4, Zari Dastani5, Marie-France Hivert6, Peter Henneman1, Claudia Langenberg4, Christopher Gillson2, Aurelian Bidulescu9, Ying Wu1, Leo-Pekka Lyttikäinen10, Christie Ballantyne11, Luudmyra Kedenko12, Olli Raitakari13, Ruth Loos4, Matthew Allison14, Alan Sinaiko9, Bernhard Paulweber12, Karen Mohlke9, Cornelia van Duijn16, Terho Lehtimäki10, Sarah Buxbaum8, John Richards17, James Pankow18, James Meigs19, Colin Palmer20, Andrew Morris21, Nick Wareham4, Markku Laakso12, Mark Walker21, Dawn Waterworth24, Tim Frayling1

1IBCS, Peninsula College of Medicine and Dentistry, Exeter, Devon, UK, 2Quantitative Sciences, GlaxoSmithKline, Research Triangle Park, NC, USA, 3Department of Social Medicine, University of Bristol, Bristol, UK, 4MRC Epidemiology Unit, Institute of Metabolic Science, Cambridge, UK, 5Department of Epidemiology, Biostatistics and Occupational Health, Lady Davis Institute, Jewish General Hospital, McGill University, Montreal, Quebec, Canada, 6Department of Medicine, Université de Sherbrooke, Sherbrooke, Québec, Canada, 7Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands, 8Morehouse School of Medicine Cardiovascular Research Institute, Department of Community Health and Preventive Medicine Atlanta, Georgia, USA, 9Department of Genetics, University of North Carolina, Chapel Hill, NC, USA, 10Department of Clinical Chemistry, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland, 11Baylor College of Medicine and Methodist DeBakey Heart and Vascular Center, Houston, TX, USA, 12First Department of Internal Medicine, St. Johann Spital, Paracelsus Private Medical University Salzburg, Salzburg, Australia, 13Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku and the Department of Clinical Physiology, Turku University Hospital, Turku, Finland, 14Department of Family and Preventive Medicine, University of California San Diego, La Jolla, USA, 15Department of Pediatrics, University of Minnesota, Minneapolis, Minnesota, USA, 16Genetic Epidemiology Unit, Department of Epidemiology, Erasmus Medical Center, Rotterdam, USA, 17Departments of Medicine, Human Genetics, Epidemiology
Adiponectin is strongly inversely associated with insulin resistance. Despite this association, the causal role of adiponectin in insulin resistance and Type 2 Diabetes remains controversial. In an attempt to understand the role of adiponectin we formed a consortium of 18 studies and >40,000 individuals with measures of both adiponectin and insulin resistance and/or Type 2 Diabetes.

First, we used genetic variants, including a low frequency large effect variant, at the ADIPOQ gene in a multivariable regression analysis to find a set of SNPs that explain the most variance in adiponectin levels. Second, we used the genetic variants as an instrument to calculate a regression slope between adiponectin levels and fasting insulin and Type 2 Diabetes.

In preliminary analyses of 2933 individuals from 4 studies we identified 4 SNPs at the ADIPOQ gene which together explain 4% of the variance in adiponectin levels. The four SNPs are correlated with each other with varying degrees of linkage disequilibrium. In multivariable models each individual SNP is associated with adiponectin levels (p<5x10^-8). We observed no association between reduced adiponectin levels and increased risk of Type 2 Diabetes (odds ratio 1.06 (0.70,1.62), p=0.77; 1928 cases vs. 1226 controls) in instrumental variable analyses. There was some suggestion of an inverse association with fasting insulin amongst 2933 non-diabetic individuals from 3 studies, but the estimate was imprecise and did not reach conventional levels of statistical significance: mean difference in fasting insulin per 1SD greater adiponectin -0.14 SD (-0.26,0.01, p=0.08) using instrumental variable analyses.

Our preliminary analyses suggest a causal inverse association of adiponectin with fasting insulin, but highlight the requirement for increased sample size. Results from 40,000+ participants will provide a comprehensive assessment of the causal role of adiponectin in insulin resistance and Type 2 Diabetes.

P-253 – ABSTRACT WITHDRAWN

P-254
COMPOSITION OF GENETIC COVARIANCE BETWEEN TRAITS USING HIGH DENSITY SNP INFORMATION, A QUANTITATIVE PERSPECTIVE

Roel F. Veerkamp, Mario P. L. Calus
Animal Breeding and Genomics Centre, Wageningen Livestock Research, Wageningen, The Netherlands

Genetic association between traits, and correlated responses resulting from those associations, are generally predicted using genetic correlations. However, depending on the composition and magnitude of the genetic covariance, asymmetry of correlated responses might occur and predictions may last for one generation only (Bohren et al. 1966). Genome-wide association studies with denser marker genotypes might be useful to investigate the makeup of the genetic covariance between traits. Therefore, the objective was to investigate the makeup of the genetic covariances between quantitative traits in more detail. Phenotypic records of 1737 heifers of farms in four different countries were used after homogenizing and adjusting for management effects. All cows had a genotype for 37,590 SNPs. Initially, only the milk yield traits were considered, with a univariate Bayesian Stochastic Search Variable Selection (BSSVS) model including a separate polygenic effect. SNP estimated heritability and covariances differed from pedigree based estimates for some of the traits and the SNPs without a significant association explained most of the genetic variances and covariances of the traits. Ten regions were found with an
association with multiple traits, in one of these regions the DGAT1 gene was previously reported with an association with multiple traits. DGAT1 explained up to 41% of the variances of four traits and explained a major part of the correlation between fat yield and fat% and contributes to asymmetry in correlated response between fat yield and fat%. Some of the prior assumptions of the model (few QTL assumed and fitting a polygenic effect), and using a univariate model might have favoured the infinitesimal model like description of the covariances. Therefore, subsequently, a multi-trait BSSVS model was used and prior model assumptions were varied to investigate the effect on the estimated composition of the underlying the covariances.

P-255
LINKING GENETICS OF BRAIN CHANGES TO ALZHEIMER’S DISEASE: A WHOLE GENOME ASSOCIATION SCAN OF HIGH RESOLUTION REGIONAL MRI VOLUME AND THICKNESS MEASURES

Mizanur Khondoker, Richard Dobson, Andy Simmons
Institute of Psychiatry, King’s College London, London, UK

Motivation: Genome-wide association studies (GWAS) have been an effective tool for understanding the genetic basis of many complex diseases. Up until recently there have been few replicated genetic markers for the disease and the genetic markers identified so far are likely to account for only a tiny fraction of the heritability of AD and many more genetic risk alleles are thought to be undiscovered.

Methods: Imaging quantitative trait loci (imaging QTLs), where associations between quantitative brain imaging measures and genetic variants are investigated, has recently emerged as an interesting research area for linking genetics of brain changes to AD. In this study, we extend our previous work by considering a genome-wide association scan of 109 brain-wide regional imaging phenotypes to identify genetic susceptibility loci for AD from a combined set of 1045 subjects from the US based Alzheimer’s Disease Neuroimaging Initiative (ADNI) and the European AddNeuroMed studies. Using multi-SNP Hyperlasso analysis of all SNPs as well as single-SNP imaging QTLs we identified several novel markers showing significant association with AD related changes in regional brain volumes and cortical thickness measures.

Results: Multi-SNP Hyperlasso analysis of the entire set of SNPs identified several novel markers associated with AD at the genome-wide significance level (1 × 10^-8), namely HOMER2 (rs1256429; intronic) EOMES (rs2724509; flanking), JAM2 (rs2829841; intronic) and WEE1 (rs10770042; coding). The analysis also identified the gene ARPP-21 to have possible causal link with AD, as reported in our previous work. The SNP rs1256429 was also the top hit (p = 8.7 × 10^-10) in the single-SNP analysis showing an association with the volume measure of Right Thalamus and AD.

Conclusions: We believe that the markers identified in this study are novel additions to the existing list of genetic variants associated with AD which can be validated in future replicated studies. We have also been able to replicate some earlier findings and have been able to suggest the genetic basis for some genes that have been previously identified as being linked to AD and cognitive performance such as the genes ARPP-21, JAM2 and DISC1.

P-256
PROPERTIES AND POWER OF THE DROSOPHILA SYNTHETIC POPULATION RESOURCE FOR THE ROUTINE DISSECTION OF COMPLEX TRAITS

Elizabeth G. King1, Stuart J. Macdonald2, Anthony D. Long3
1University of California, Irvine, Irvine, CA, USA, 2University of Kansas, Lawrence, KS, USA

The vast majority of genetically based traits have a complex basis with many causative genetic variants that often interact with one another and with environmental factors. This complexity makes identifying the causative genetic variants for any trait a significant challenge and requires the development of new and creative experimental designs. The Drosophila Synthetic Population Resource (DSPR) is a newly
developed multifounder advanced intercross panel consisting of over 1600 recombinant inbred lines (RILs) designed for the genetic dissection of complex traits. This design results in a set of RILs whose genomes are a fine-scale mosaic of segments from the founder lines. Here, we describe the inference of the underlying mosaic founder structure for the full set of RILs from a dense set of semi-codominant RAD markers and use this inference to describe the genetic properties of the resource. In addition, we assess the power of the DSPR by assigning genotypes at a hidden QTL to each RIL based on the inferred founder state and simulating phenotypes for different experimental designs, different genetic architectures, different sample sizes, and QTL of varying effect sizes. We found the DSPR has both high power (e.g. 84% power to detect a 5% QTL) and high mapping resolution (e.g. ~1.5 cM for a 5% QTL). The inferred founder genotypes combined with the full genome sequences of the founders also allows us to infer the full genome sequence for every RIL in the DSPR and compare our approach to a genome-wide association study approach.

P-257

EPISTASIS ASSOCIATION MAPPING FOR ULTRASOUND CARCASS TRAITS IN TROPICAL BEEF CATTLE

Abdirahman Ali¹, Peter Thomson², Mehar Khatkar¹, Herman Raadsma¹, Haja Kadarmideen²
¹Faculty of Veterinary Science, The University of Sydney, Sydney, Australia, ²Division of Genetics and Bioinformatics, Department of Basic Animal and Veterinary Sciences, Faculty of Life Science, University of Copenhagen, Copenhagen, Denmark

Objectives: A genome-wide association study (GWAS) was conducted to identify SNP (single nucleotide polymorphisms) associated with ultrasound scanned carcass traits in Australian tropical beef cattle.

Methods: A total of 565 Brahman heifers were ultrasound-scanned for eye muscle area (SEMA), fat depth at P8 (SP8) and 12/13th rib (SRIB) sites. These measurements were taken at two occasions: i) at the end of first wet season, and ii) at the end of second dry season, when the average age were 18 months and 24 months, respectively. Animals were also genotyped for 9,075 SNPs using ParaAllele/Affymetrix 10k SNP chips. The phenotypes and genotypes were fitted in a linear mixed model, the model included at each SNP position, age, property of origin, and contemporary group (heifers with same experiment and year of birth) effects as fixed and animal and residual effects as random.

Results and Discussions: Direct polygenic heritability estimates were ranged from 0.45 ± 0.17 to 0.65 ± 0.17 for all ultrasound scanned traits. Using a false discovery rate (FDR) cutoff of 10%, 10 SNPs were significantly (P ≤ 1x10⁻⁵) associated with at least one of the phenotype considered. Five out of the 10 SNPs were concentrated on chromosome 14 at 24 -34 Mb related to SP8 and SRIB. SNP Annotation for all significant SNPs was conducted using dbSNP information in Bos Taurus genome with a limit of ± 300kb and revealed protein coding genes of SNTG1, RAPGEF2, TMEM132D and NNT as candidate genes affecting carcass traits considered. We further conducted genome-wide epistasis association-analyses among all significant SNPs and identified a set of epistatic SNP pairs significantly associated with carcass traits. This study provides evidence of additive and epistatic SNP loci linked to variation in carcass traits in beef cattle.

P-258

GWAS OF LOW HERITABLE TRAITS: THE CASE OF SENSORY ATTRIBUTES OF DRY-CURED HAMS

Maria-Dolors Guardia¹, Raquel Quintanilla¹, Josep Reixach³, Arianna Manunza², Anna-Maria Mercader², Marcel Amills², Romi Pena³, Jules Hernandez-Sanchez²
¹IRTA, Catalonia, Spain, ²UAB, Catalonia, Spain, ³Selección Batallé, Catalonia, Spain, ⁴UdL, Catalonia, Spain

Sensory attributes assessed by trained panellists in dry-cured ham have low heritabilities (h²). The main reason is more likely a large error variance (Vₑ) rather than a lack of additive genetic variance (Vₐ). Sensory characteristics are decisive for consumers therefore there is a strong economic interest in
improving them. We evaluated 16 sensory attributes assessed by trained panellists in biceps femoris (BF) and semimembranosus (SM) muscles of 300 dry-cured Spanish hams, including overall liking, 6 texture attributes and 9 flavour attributes. An animal model that included farm, batch and carcass weight as fixed effects was used to obtain V_A and V_E estimates. None of the h^2 were significantly greater than zero. Nevertheless, sensory attributes are strongly associated to several genetically-determined meat traits, such as muscle fibre type (h^2~0.2-0.6) and intramuscular fat content (h^2~0.1-0.7) and composition (h^2~0.3-0.6). Despite the low h^2 obtained, a GWAS of the same sensory attributes rendered many significant hits (p<0.00005), i.e. 37 in BF and 84 in SM. Several of those hits fell on the same haplotype blocks so the effective number of hits is lower. Some hits point at plausible genetic mechanisms underlying sensory traits. For example, the hollow defect is a musty off flavour associated to growth of aerobic bacteria, moulds and mites in hollow cavities located in the adductor muscle and around the coxofemoral joint. The most significant SNP mapped close to RXFP2, a gene from the relaxin family involved in connective tissue relaxation. On the practical side, identification of genes controlling sensory attributes may lead to the development of biosensors that could be used by breeders (improving carcass quality) and abattoirs (paying carcass quality) alike.

P-259 – ABSTRACT WITHDRAWN

P-260
EXOME SEQUENCING IN EARLY-ONSET TRIPLE POSITIVE, TRIPLE NEGATIVE AND FAMILIAL BREAST CANCER

Rosanna Upstill-Goddard, Jane Gibson, William Tapper, Andrew Collins, Diana Eccles
University of Southampton, Southampton, UK

Background: Exome sequencing determines the genomic sequence of most coding exons and has previously demonstrated success in identifying causal genes of Mendelian disorders. Exome sequencing was applied to eight patients with early-onset breast cancer chosen from the POSH cohort. We selected cases with age at diagnosis ≤25 years with specific pathological characteristics that fitted with the observed HER2 overexpressing subtype of breast cancer associated with germline TP53 gene mutations but where no mutation in TP53, BRCA1 or BRCA2 had been found. We also tested exomes from triple-negative breast cancer cases and two cases with a strong family history. We hypothesised that rare mutations in novel breast cancer genes within the TP53 pathway might be causal of cancer in these patients.

Methods: Exome sequencing was carried out on DNA samples from our eight patients. Raw data were aligned against the hg19 human genome reference sequence to identify variants that differed from the reference sequence. Variant annotation was undertaken using the ANNOVAR program. A thorough filtering programme retained only rare exonic variants classified as nonsynonymous or frameshift. All variants were cross-referenced against a comprehensive list of 327 genes of the TP53 pathway. Analysis focussed on variants classified as potential compound heterozygotes, potentially recessive, or potentially deleterious by SIFT score.

Results: Analysis identified 41 rare polymorphisms across 20 genes, of which only two were found in multiple individuals. 23 were potential compound heterozygotes and the remaining 18 were classified as potentially damaging by SIFT; none of the variants retained after filtering was potentially recessive.

Conclusions: No genes were implicated consistently by disease sub-type reflecting the genetic complexity of causative or predisposing factors in each individual case of breast cancer. We have demonstrated the potential of exome sequencing technology as a means to uncover rare, potentially disease-causing, variants in a targeted panel of genes.
P-261
GENETIC ARCHITECTURE OF QUANTITATIVE MORPHOLOGICAL VARIATION IN A WILD POPULATION OF THE GREAT TIT, P. MAJOR

Isabelle De Cauwer¹, Anna W. Santure¹, Matt R. Robinson¹, Ben. C. Sheldon², Jon Slate¹
¹University of Sheffield, SHEFFIELD, UK, ²University of Oxford, OXFORD, UK

The recent developments of genomic resources and improved genotyping technologies provide the means to explore the genetic architecture of quantitative variation in pedigreed wild populations, giving new insight into the mechanisms of adaptation in complex and variable environments. To date, most mapping studies in the wild have found tentative evidence that a substantial part of genetic variation is explained by a restricted number of loci of large effect. By combining ecological data generated from a classical long-term study of the great tit, Parus major, with a newly developed 8K SNP chip, we describe the genetic architecture of a range of morphological traits using (i) a population-based linkage analysis approach and (ii) a genomic relatedness-based variance component partitioning method. We demonstrate that many morphological traits are highly polygenic with little evidence for genes of major effect. This study represents an advance in the dissection of the genetic architecture of complex traits in the wild.

P-262
FINE MAPPING OF TWENTY-SIX REGIONS ASSOCIATED WITH LUNG FUNCTION

María Soler Artigas¹, Louise Wain¹, Nick Shrine¹, Ian Sayers², Ian Half², Martin Tobin¹
¹University of Leicester, Leicester, UK, ²University Hospital of Nottingham, Nottingham, UK

Chronic obstructive pulmonary disease (COPD) is a major health problem worldwide and its diagnosis is based on lung function measures. The analysis of binary disease status has proven to be challenging due to the large sample sizes required in order to detect modest genetic effects. A promising approach to study the genetics of COPD is to analyse the quantitative lung function measures instead. We undertook a meta-analysis of lung function GWAS¹ on over 94,000 individuals and discovered sixteen new loci associated with lung function, increasing the total number of genetic variants associated with lung function to twenty-six¹-³. To date, seven of these twenty-six variants have been associated with COPD⁴-⁷, and as sample sizes increase in genetic association studies of COPD, associations with more of these variants are likely to be reported.

Identifying rare genetic variants in regions of genome-wide association with lung function could provide insights into the molecular pathways of lung function and COPD, and could lead to new targets for treatments. Hypothesizing that the twenty-six variants associated with lung function are also associated with COPD through an effect on lung function, we have designed a study to fine map these regions. We present a cost-effective design to enrich and resequence these twenty-six regions in 300 COPD cases and 300 controls, grouped in 24 pools of 25 individuals each. We describe the enrichment and resequencing approaches, the approaches to alignment and variant calling, and quality control steps, taking into account the additional challenges of the analysis of pooled data. Identification of variants associated with COPD will include analysis of common variants and specific approaches to study rare variants, including collapsing or allele matching methods. Additional genotyping will be undertaken in order to validate the findings.

2. E. Repapi et al., Nature genetics 42, 36 (Jan, 2010).
P-263
USING SINGLE NUCLEOTIDE POLYMORPHISMS TO SEARCH FOR GENOMIC PATTERNS RELATED TO SELECTION IN HEREFORD BEEF CATTLE

Yijian Huang1, Joseph Cassady1, Michael MacNeil1, Christian Maltecca1
1North Carolina State University, Raleigh, NC 27695, USA, 2USDA Agricultural Research Service, Miles City, MT 59301, USA

The Line 1 (L1) is a closed (since 1934) Hereford population that has been subjected to specific selection on postweaning as compared to the general American Hereford population (AHA) where multiple selection objectives have been applied. Different levels of selection and random drift affected these two populations albeit continual migration from L1 to AHA happened since the 1950’s. Genotypes at 50,562 single nucleotide polymorphisms (SNP) from L1 (N = 240) and AHA (N = 311) revealed different linkage disequilibrium (LD) patterns in L1 compared to AHA. Differences in LD between L1 and AHA decreased as SNP spacing increased, indicating the recent divergence of L1 and AHA. Window heterozygosity (Hw) based on non-overlapping sliding windows of six SNPs, was calculated for L1, AHA and a combination of both and revealed genomic regions that approached fixation or conversely had a high proportion of heterozygotes. Putative directional selection signatures were detected in low Hw regions found in all three groups of samples. Low Hw regions co-located with postweaning gain quantitative trait loci (QTL) previously identified. Low Hw regions peculiar of L1 were found which might in turn reflect fixation driven by random drift. One high Hw region overlapped with GSK3B, a gene that has been implicated in insulin-like growth factor regulation and associated with embryonic survival, suggesting possible stabilizing selection. SNP genotypes were further regressed on cumulative selection differentials (CSD) for postweaning gain in L1. Three Hw regions including known QTL for growth-related traits overlapped with significant SNPs for CSD regression. In this study patterns of genomic variation were employed to detect population divergence for two Hereford populations and to identify genomic regions related to selection.

P-264
PEDIGREE-FREE ESTIMATES OF HERITABILITY IN THE WILD: PROMISING PROSPECTS FOR SELFING POPULASSING

Laurene Gay1, Mathieu Siol1,2, Joelle Ronfort1
1Diversity and Adaptation of Mediterranean Species, Montpellier, France, 2Genetics and Ecophysiology of Legume Species, Dijon, F

Understanding adaptation and phenotypic evolution requires an accurate estimation of the level of heritable variation available within populations. This is particularly crucial for selfing populations, where decreased genetic variation is expected. It can be achieved using quantitative genetics, by correlating the phenotypic resemblance with the actual proportion of the genome that two relatives share identical by descent (realized relationship). The later can be predicted with pedigrees (θA: expected value) or molecular markers, by measuring the average number of alleles shared over many loci, arguably a more precise estimation of the ‘realized relationship’. Nevertheless, evolutionary biologists, unlike animal breeders or human biologists, remain cautious about marker-based relationship coefficients and their use to study complex phenotypic traits in populations lacking pedigree information. In this paper, we review published results comparing different pedigree-free methods and use simulations to test animal models using marker-based relationship coefficients, with a special focus on the influence of mating systems. We found that if Ritland’s pairwise regression model is often unreliable, animal models including marker-based relationship matrices seem promising. Our simulations show that in populations with large variance in relationship (small or selfing populations), θA is well predicted by molecular markers and marker-based animal models perform well. Selfing further improves the accuracy because it reinforces the linkage between observed SNPs and causal loci. In large outcrossing populations, however, heritability estimates are highly biased. Then, family focused
or multi-generational sampling might be required to accurately estimate quantitative genetic parameters from marker-based relationship coefficients.

P-265
EVIDENCE OF SHARED POLYGENIC RISK AMONG SMOKING BEHAVIORS AND BODY COMPOSITION

Roseann Peterson, Xiangning (Sam) Chen, Jingchun Chen, Bradley (Todd) Webb, Hermine Maes
Virginia Commonwealth University, Richmond, VA, USA

Abstract can be found in the student symposium abstract section on page 62

P-266
EVIDENCE OF INBREEDING DEPRESSION ON HUMAN HEIGHT

James Wilson1, Niina Eklund2,3, Nicola Pirastu4, Maris Kuningas5, Brian McEvoy6, Tõnu Esko7,8, Tanguy Corre9, Gail Davies1, Ruth McQuillan1
1University of Edinburgh, Edinburgh, UK, 2Institute for Molecular Medicine Finland, Helsinki, Finland, 3National Institute for Health and Welfare, Helsinki, Finland, 4University of Trieste, Trieste, Italy, 5Erasmus University Medical Center, Rotterdam, The Netherlands, 6Queensland Institute of Medical Research, Brisbane, Australia, 7Estonian Genome Center, Tartu, Estonia, 8University of Tartu, Tartu, Estonia, 9San Raffaele Research Institute, Milano, Italy

Stature is a classical and highly heritable complex trait, with 80-90% of variation explained by genetic factors. In recent years, genome-wide association studies (GWAS) have successfully identified many common additive variants influencing human height; however, little attention has been given to the potential role of recessive genetic effects. Here, we investigated genome-wide recessive effects by an analysis of inbreeding depression on adult height in over 35,000 people from 21 different population samples. We found a highly significant inverse association between height and genome-wide homozygosity, equivalent to a height reduction of up to 3 cm in the offspring of first cousins compared with the offspring of unrelated individuals, an effect which remained after controlling for the effects of socio-economic status, an important confounder. There was, however, a high degree of heterogeneity among populations: whereas the direction of the effect was consistent across most population samples, the effect size differed significantly among populations. It is likely that this reflects true biological heterogeneity: whether or not an effect can be observed will depend on both the variance in homozygosity in the population and the chance inheritance of individual recessive genotypes. These results predict that multiple, rare, recessive variants influence human height. Although this exploratory work focuses on height alone, the methodology developed is generally applicable to heritable quantitative traits (QT), paving the way for an investigation into inbreeding effects, and therefore genetic architecture, on a range of QT of biomedical importance.

P-267
EPIDEMIOLOGIC ARCHITECTURE FOR GENES LINKED TO ENVIRONMENT: SMOKING MODIFIES HDL-C, LDL-C, AND TRIGLYCERIDE GWAS-IDENTIFIED ASSOCIATIONS IN THE NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEYS

Logan Dumitrescu1, Robert Goodloe1, Kristin Brown-Gentry1, Dana C. Crawford1,2
1Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA, 2Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN, USA

Despite many successes, lipid-associated variants discovered through genome-wide association studies
GWAS do not account for the majority of heritability estimated for these traits. Epidemiological studies have long indicated that certain environmental factors are capable of shaping lipid distributions in the population. However, environmental modifiers of known genotype-phenotype associations are just recently emerging in the literature. We, as part of the Population Architecture using Genomics and Epidemiology (PAGE), have genotyped GWAS-identified variants in samples collected by the Centers for Disease Control and Prevention for the National Health and Nutrition Examination Surveys (NHANES). NHANES is a cross-sectional survey of Americans representing three major groups: non-Hispanic whites (n=3,866), non-Hispanic blacks (n=1,883), and Mexican Americans (n=2,339). Along with lipid levels, NHANES contains an abundance of environmental variables, including serum cotinine, a widely used marker for tobacco smoke exposure, which can lead to unfavorable alterations of the lipid profile. Using linear regression assuming an additive genetic model, gene-environment interactions were modeled with a multiplicative interaction action term between smoking status (non-smokers: cotinine<15 ng/ml; smokers: cotinine≥15 ng/ml) and 23 GWAS-identified lipid-associated variants for HDL-C, LDL-C, and ln(TG) levels. In models adjusted for age, sex, and marginal effects, eight SNP×smoking interactions associated with one of three lipid traits at p<0.05. The most significant interaction was PLTP rs7679×smoking (p=0.003) for HDL-C levels among B. Smokers with at least one copy of the minor allele (C) had, on average, higher mean HDL-C compared to smokers homozygous for the major allele. Interestingly, this interaction term was not associated in the other two populations (p=0.529 and p=0.532 in W and MA, respectively). PLTP is known to play a major role in HDL-C particle size and composition and studies have shown that PLTP activity is affected by smoking. Our results suggest that smoking may impact GWAS-identified associations for lipid traits and demonstrate the need for further gene-environment interaction studies in diverse populations.

P-269 – ABSTRACT WITHDRAWN

P-269
GENETIC AND ENVIRONMENTAL MODIFIERS OF FIBRINOGEN LEVELS IN DIVERSE POPULATIONS FROM THE THIRD NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY

Janina Jeff1, Kristin Brown-Gentry1, Dana Crawford1,2
1Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA, 2Department of Molecular Physiology & Biophysics, Vanderbilt University, Nashville, TN, USA

Fibrinogen levels are significantly associated with cardiovascular disease (CVD). Up to 50% of the trait variability of fibrinogen levels are attributable to genetic factors. Genes within the fibrinogen gene cluster (FGA, FGB, and FGG) regulate the production of fibrinogen. However, collectively these loci do not account for the expected genetic component for fibrinogen levels. Like many traits associated with CVD, the role of the environment, several loci in the genome or the interaction of amongst genes and/or with environment can account for the trait variability observed in fibrinogen levels. Here we examine the effect of genetic modifiers (gene-gene and gene-environment) with plasma fibrinogen levels in three populations (non-Hispanic whites, non-Hispanic blacks, and Mexican Americans) from the Third National Health and Nutrition Survey (NHANES III). Using linear regression, cross product terms (all pair-wise gene-gene (G×G) or gene-environment (G×E) terms for 25 fibrinogen cluster variants) were added to the regression models for plasma fibrinogen. Tests for G×E interactions between fibrinogen cluster variants and sex, age, smoking status, and BMI separately, were performed with plasma fibrinogen levels. All SNPs were coded additively, stratified by race/ethnicity, and adjusted for the main effect of the SNP (G×G and G×E models) and environmental factor (G×E models). At a liberal significance threshold of p<0.05, we identified a total of 25 and 16 G×G and G×E interactions from 828 and 297 tests, respectively. Overall, greater than of 60% of the significant G×G interactions include intronic rs2227395, located in the FGB gene, the rate-limiting gene in fibrinogen production. We observed several G×E interactions with fibrinogen levels, the most
significant GxE was rs2227434*age in non-Hispanic blacks $\beta=0.004$, $p=3.8E-7$. Here we successfully identify genetic and environmental modifiers in diverse populations for fibrinogen levels. Identifying these modifiers provides important information about intermediate phenotypes with the ultimate goal of predicting CVD outcome.

P-270
WHERE ARE THE VARIANTS WITH INTERMEDIATE EFFECT SIZE FOR HUMAN HEIGHT?

Kathryn Kemper$^1$, Peter Visscher$^{2,3}$, Mike Goddard$^{1,4}$
$^1$University of Melbourne, Melbourne, Australia, $^2$University of Queensland Diamantina Institute, Brisbane, Australia, $^3$The Queensland Brain Institute, Brisbane, Australia, $^4$Victorian Department of Primary Industries, Melbourne, Australia

Genetic architecture describes the number, size of effect, mode of action and allele frequencies for mutations underlying genetic variation in a trait. This study reviewed the known variants underlying height in humans, an archetypal quantitative trait.

We find that described mutations affecting height form two almost disjointed sets. The first set contains over 241 loci with large (>2 s.d.) effects. Examples include $FBN1$ mutations causing overgrowth in Marfan syndrome. Mutations in this set show high allelic heterogeneity, are relatively rare or de novo events and are usually associated with a syndrome of pleiotropic effects. Collectively, these loci were estimated to account for <3% of the phenotypic variance in height.

The second set of loci was identified through genome-wide association studies (GWAS) and had relatively small effects (0.02-0.13 s.d., 1-9mm). Although the causative mutations were seldom identified, loci were often implicated in skeletal development and replicated in independent studies. One study identified 180 loci explaining about 10% of the phenotypic variance in height.

In conclusion, we make two interesting observations. First, there were no reported variants with intermediate effects on height (e.g. effects of 1-2cm). We suggest that these variants are missed because family studies investigate extremely rare but obvious phenotypes while GWAS detect variants with moderate allele frequencies. Second, we suggest that standing genetic variation in height is mostly due to small effect mutations. This is because natural selection acts against mutations with large effect, through pleiotropy, keeping the mutations rare and therefore contributing little to standing genetic variation. For mutations with smaller effects, only a fraction of the total has been detected by GWAS. The remaining loci are missed because of incomplete linkage disequilibrium between mutations and genotyped SNP. We predict that sequence data with large samples are likely to detect low frequency variants with intermediate effect size.

P-271
ESTIMATING THE GERMLINE GENETIC CONTRIBUTION TO COMMON CANCERS

Stuart Macgregor, Yi Lu, Q-MEGA and AMFS investigators, Nicholas Martin, Grant Montgomery, Nicholas Hayward
Queensland Institute of Medical Research, Brisbane, Queensland, Australia

Common cancers are frequently demarcated into ‘hereditary’ and ‘sporadic’ types. This distinction arose from work identifying rare highly penetrant germline mutations causing ‘hereditary’ cancer. These rare mutations are important in particular families but do not explain many cases in the general population. This led to the notion that the numerically dominant set of ‘sporadic’ cancers had a negligible germline genetic component and this was reinforced by large twin studies which showed that for several major cancers there is no significant heritability. The novel cancer susceptibility loci identified by GWAS have a small effect size and the perception persists in some quarters that germline genetic variation is not important in ‘sporadic’ cancers. Here we apply a recently developed method to several cancer GWAS
data sets. This method links case-control status to relatedness derived from dense genetic marker data, as implemented in the program GCTA (Lee et al, AJHG 2011). We demonstrate that for many cancers there is a significant ‘polygenic’ component - this is only estimated from common polymorphisms and hence provides a lower bound on heritability. Specific estimates of variance explained are: Australian melanoma (N=2000 cases, N=2000 controls) variance explained (V_G)=0.29 (s. e. 0.10); USA melanoma (N=2000, 1000) V_G=0.17 (0.10); USA prostate cancer (N=1200, 1200) V_G=0.67 (0.20); USA breast cancer (N=1200, 1200) V_G=0.17 (0.22); USA pancreatic cancer (N=2500, 2500) V_G=0.15 (0.06). These estimates are consistent with results from twin studies showing significant heritability for prostate cancer but for melanoma and pancreatic cancer much stronger statements can now be made. We demonstrate the robustness of the approach by comparing controls from different studies.

P-272 – ABSTRACT WITHDRAWN

P-273

GENOMIC ANALYSIS OF TRAIT VARIATION IN SOAY SHEEP

Camillo Berenos, Phil Ellis, Josephine Pemberton
Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

The genetic architecture of highly heritable morphological traits such as body weight and body size is insufficiently understood. Efforts have been made to identify specific genomic regions that underlie such traits in humans and animal breeding studies using high density SNP chips, and such studies often detect hundreds of significant gene variants. Yet, significant quantitative trait loci, (QTL), often only explain a small fraction of the phenotypic variance, suggesting a truly polygenic architecture, with many rare alleles influencing trait variation. Even less is known about the molecular genetic basis of morphological traits in natural populations. As body size is often tightly linked to fitness, and thus potentially exposed to natural selection, this knowledge is essential for our understanding of the evolutionary potential and trajectories in such traits. We have screened a longitudinal dataset from the well-phenotyped feral population of Soay sheep on St. Kilda on a 50K SNP chip to unravel the genomic architecture of morphological traits. We will here present some preliminary results, using a step-wise approach, first by partitioning the phenotypic variance explained by all SNPs, second by partitioning variance between chromosomes, and third, by identifying specific genomic locations that contribute to trait variation.

P-274

MISSED RATHER THAN MISSING HERITABILITY? TRANSFER AND TRAPPING OF A SMALL-EFFECT MOUSE OBESITY LOCUS INTO THREE CONGENIC STRAINS

Danielle Reed, Cailu Lin, Anna Lysenko, Brad Fesi, Maria Theodorides, Alexander Bachmanov
Monell Chemical Senses Center, Philadelphia, PA, USA

Body composition is the epitome of a complex trait, governed by many genes of small effects which are interactive and sometimes imprinted, and, in addition, it differs by age, sex, and diet. With these issues in mind, we mapped the quantitative trait locus Adip5 (adiposity 5) on mouse chromosome 9 in two F2 intercrosses, and found its effects to be replicable. To isolate Adip5, a chromosome substitution strain was constructed from the parent strains with marker-assisted selection, and it served as the starting point for cycles of backcrossing to create congenic strains. To monitor whether the Adip5 effects remained detectable, incipient congenic mice were phenotyped at each step of backcrossing. Age, sex, diet, imprinting and maternal effects were controlled to reduce sources of variation other than the genetic effect of Adip5. Three resulting congenic strains with partially overlapping donor regions retained the same phenotype detected in the F2 intercross. Comparison of these congenic strains narrowed the region
of Adip5 to 13.5 Mb. This study demonstrates that small-effect obesity loci can be isolated and suggest that ‘missing heritability’ in human studies may be due to the inability to control the environment, subject characteristics and genetic background.

P-275
A LATENT CURVE APPROACH USING PENALIZED REGRESSION TO MODEL PHENOTYPIC PLASTICITY AND GENOTYPE BY ENVIRONMENT INTERACTION

Fred van Eeuwijk², Sabine Schnabel², Martin Boer², Marcos Malosetti², Paul Eilers²
²Biometris, Wageningen, The Netherlands, ²Erasmus MC, Rotterdam, The Netherlands

In plant breeding there exists a strong tradition to model phenotypic plasticity and genotype by environment interaction by linear-bilinear models. Typically a two-way table of genotype by environment (GxE) means is approximated by models containing either or both of a genotypic and environmental main effect combined with one or more bilinear terms. Bilinear terms are products of genotypic and environmental scores that both need to be estimated from the data. Some well-known linear-bilinear models for GxE data are the Finlay Wilkinson model, the AMMI model and the GGE model. A popular interpretation of linear-bilinear models is that by optimizing a least squares loss function, latent environmental variables are estimated to which genotypes differ in sensitivity. A limitation of these models is that they require the genotypic responses to be linear.

The objective of our research is to find a generalization of the linear-bilinear framework for modelling GxE in which we allow latent genotypic responses to be smooth curves of underlying latent environmental characterizations.

Our method consists in using P-spline representations for latent environmental variables in a linear-bilinear framework. An iterative algorithm alternates between fitting the curves for the set of genotypes conditional on the last P-spline representation of the latent environmental variable and updates for the latter.

We will illustrate results of our approach as applied to various examples focussing on goodness of fit and interpretability and compare with traditional linear-bilinear models. Our approach to GxE can be embedded in models for QTL by environment interaction.

As a conclusion we state that our latent curve approach to modelling GxE interaction provides a flexible and powerful extension to the existing GxE tool kit.

P-276 – ABSTRACT WITHDRAWN

P-277
EFFECTS OF DGAT1 ON FAT AND PROTEIN CONTENTS IN MILK OF DAIRY COWS ARE NOT CONSTANT ACROSS LACTATION

Joanna Szyda², Jolanta Komisarek², Ireneusz Antkowiak²
²Department of Animal Genetics, Wroclaw University of Environmental and Life Sciences, Wroclaw, Poland, ²Department of Cattle Breeding and Milk Production, Poznan University of Life Sciences, Poznan, Poland

Objectives: Our study is focused on the investigation of potential changes of additive effects of DGAT1 gene on milk production traits across time. Such changes can be expected by investigating additive polygenic effects on milk production traits estimated in other studies, which clearly indicate changes in the genetic determination of milk production throughout 305-days of lactation.

Methods: The data sets comprise 192 and 475 cows representing Jersey and Holstein-Friesian breeds, respectively. Phenotypes comprise: milk yield, protein and fat contents in milk, recorded for each cow throughout its first lactation. Genotypes of at K232A mutation located within a DGAT1 gene were also
available. The effects of DGAT1 across lactation were estimated using a fixed regression model fitted to test day yields comprising fixed effects of a test day, age of a cow at calving, a season of calving and regression coefficients for the additive effects of DGAT1, as well as a random additive polygenic effect of a cow and a random permanent environmental effect.

**Results:** Using two independent data sets from Holstein-Friesian and Jersey breeds, we show that the additive effects of DGAT1-K232A on fat and protein contents increase with progressing lactation. The increase was more pronounced in Jersey than in Holstein-Friesian. E.g. for protein content the effect of allele GC was close to zero during the first days of lactations and increased up to 0.310% per day in Jersey and 0.104% per day in Holstein-Friesian.

**Conclusions:** Longitudinal modelling of SNP effects allows for a more precise description of the genetic background underlying variation of complex traits. Our result regarding variation of DGAT1 effect over time has direct implications for dairy cattle selection, but on a wider scale it may have a more extensive impact on gene detection in humans and provokes a question on the regulatory mechanisms underlying this phenomenon.

---

**P-278**

**A GENOME-WIDE ASSOCIATION STUDY OF GENERAL COGNITIVE ABILITY: THE CHARGE CONSORTIUM**

Gail Davies1, Stéphanie Debette2, Albert Smith3, Jan Bressler4, Carla Ibrahim-Verbaas5, Josh Bis6, Katja Petrovic7, Lina Zgaga1, Mirna Kirin1, Lei Yu, Qiong Yang8, Jarhi Lahti9, Christopher Oldmeadow11, Caroline Hayward1, Srikanth Velandai12, David Bennett8, Lenore Launer13, Cornelia van Duijn5, Annette Fitzpatrick9, Sudha Seshadri9, Thomas Mosley14, Peter Visscher15, Ian Deary1, for the CHARGE cognitive working group.

1University of Edinburgh, Edinburgh, UK, 2Institut Pasteur de Lille, Lille, France, 3Icelandic Heart Association, Kopavogur, Iceland, 4University of Texas, Texas, USA, 5Erasmus MC University Medical Center, Rotterdam, The Netherlands, 6University of Washington, Seattle, USA, 7Medical University Graz, Graz, Austria, 8Rush Alzheimer’s Disease Center, Chicago, USA, 9Boston University, Boston, USA, 10University of Helsinki, Helsinki, Finland, 11University of Newcastle, Newcastle, Australia, 12Monash University, Victoria, Australia, 13National Institute on Aging, Washington DC, USA, 14The University of Mississippi Medical Center, Jackson, USA, 15University of Queensland, Brisbane, Australia

General cognitive ability is substantially heritable through the human life-course from adolescence to old age. To date, there are almost no replicated genetic associations with general cognitive ability. An exception is the small contribution of *APOE* at older ages. Limitations of previous studies include inadequate sample sizes, population stratification, and differences in cognitive measures. To explore genetic contributions to variation in general cognitive ability in middle-aged and older adults, we conducted genome-wide association analyses in the CHARGE consortium. GWA analyses were performed on 17 cohorts with cognitive phenotypes that allowed the extraction of a general cognitive ability score; Atherosclerosis Risk in Communities, Austrian Stroke Prevention Study, Cardiovascular Health Study, Framingham Heart Study, Lothian Birth Cohorts, Helsinki Birth Cohort, Tasmanian Study of Cognition and Gait, Croatian Isolate Cohorts, Orkney Complex Disease Study, Rotterdam Study, Erasmus Rucphen Family Study, Rush Memory and Ageing Project and Religious Orders Study (N = 24,826). This is the largest sample, to date, for a GWAS of general cognitive ability. Inclusion criteria were being over 45 years of age, absence of clinical stroke or dementia, and Caucasian origin. A general cognitive ability score was created in each cohort by principal components analysis, using at least three cognitive tests that assessed different cognitive domains. The GWA results were meta-analysed using an inverse variance method. No genome-wide significant results were found. Preliminary results replicate findings that general cognitive ability is under polygenic control. Additional analyses are in progress, including gene-based tests and polygenic analyses. The GCTA method will be applied to the largest cohort ARIC (N ~9000) to estimate the proportion of variation accounted for by all genotyped common SNPs. Suggestive
SNP-based and gene-based findings will be followed up in replication samples of N ~12,000. The polygenic nature of this phenotype will be investigated using genetic prediction analyses.

**P-279**

**COMPARISON OF GENOME-WIDE ASSOCIATION AND REGIONAL HERITABILITY MAPPING APPROACHES TO IDENTIFY LOCI UNDERLYING FAECAL EGG COUNT VARIATION IN BLACKFACE LAMBS**

Valentina Riggio, Oswald Matika, Ricardo Pong-Wong, Stephen C. Bishop  
The Roslin Institute and R(D)SVS, Edinburgh, Midlothian, UK

The genetic architecture underlying nematode resistance in Blackface lambs was evaluated comparing genome-wide association (GWAS) and regional heritability mapping (RHM) approaches. RHM estimates heritabilities attributable to small regions of adjacent SNPs. The traits analysed were faecal egg counts (FEC) collected at ca. 16, 20 and 24 weeks of age, and their weighted average, for *Nematodirus* and *Strongyles* nematodes, on 752 lambs. All lambs were genotyped using the ovine 50K SNP chip. After quality control, 44,388 SNPs were available for GWAS and 42,841 for RHM, which utilises only mapped SNPs. The same fixed effects were used in both analyses. For GWAS, association was tested using the mmscore function in GenABEL. In RHM, chromosomes were divided into 100 SNP windows, shifted every 50 SNPs, with the total genomic additive effect estimated from all SNPs and the regional genomic additive effect from SNPs within each window. In total, 868 windows were tested. The null hypothesis assumed no regional variance; variance in each window was tested using a likelihood ratio test. Bonferroni-corrected significance thresholds were p<1.13x10^-6 and p<2.25x10^-5 for genome-wide and suggestive levels for GWAS, and p<2.30x10^-4 and p<2.30x10^-3 for genome-wide and suggestive levels for RHM. For both methods, the strongest evidence for association was found on OAR14 and OAR6 for *Nematodirus* and *Strongyles* FEC traits, respectively. Additionally, results from GWAS indicated suggestive association on OAR2 and OAR4 for *Nematodirus*, whereas RHM found suggestive evidence on OAR2 (a different region from the GWAS result), OAR4, OAR9 for *Nematodirus* and OAR3 and OAR21 for *Strongyles*. This suggests that RHM may capture variation not detected by GWAS. This agrees with reported human results, showing that RHM performs better than GWAS, especially when associated SNPs do not have large enough effects, or sufficiently strong linkage disequilibrium with causative mutations, to be declared significant at the genome-wide level.

**P-280**

**CONSEQUENCES OF GENOMIC IMPRINTING IN LIVESTOCK GENETIC EVALUATION**

Luis Varona1, Joaquim Casellas2, Carlos Moreno1, Juan Altarriba1  
1Universidad de Zaragoza, Zaragoza, Spain, 2Universitat Autonoma de Barcelona, Barcelona, Spain

Genetic evaluation by mixed models becomes the basis for its genetic improvement in livestock. The most widespread methodology accounts for direct polygenic additive genetic effects plus some systematic effects and a residual source of variation. However, the genetic determinism of many phenotypic traits must be also linked with maternal and paternal genetic effects, such as paternal and maternal genomic imprinting. Genetic evaluation models must capture all potential sources of genetic variability. A complete model must include direct additive genetic effects and parental, maybe imprinted, genetic effects. However, when partial imprinting is present, it originates covariance between direct and parental genetics effects that cannot be inferred from the phenotypic and pedigree data. Given that the complete model cannot be properly addressed, the aim of this study was to analyse the consequences of ignoring some of the genetic effects (i.e., direct, maternal and paternal genetic effects) in terms of variance component estimation and breeding value prediction by using beef cattle data sets and simulated populations under several scenarios. Simulated and real data were analysed using a Bayesian approach with several models
that include a) Direct additive effect, b) Direct and paternal effects, c) Direct and maternal effects, d) Paternal and maternal effects, e) Direct, paternal and maternal effects with null correlations between them. The conclusion of this study indicates that paternal or maternal effects may play a relevant role in some relevant traits in livestock production. Further, its absence in the genetic evaluation models may lead to biased parameter estimation and to erroneous ranking of the candidates of selection, with important consequences in the phenotypic performance of future generations. In particular, the cause of negative correlations between direct and maternal effects may be caused by the presence of not considered paternal effects.

P-281
GENETIC BASIS OF EYE SIZE DIFFERENCES IN TWO SISTER SPECIES OF DROSOPHILA

Saad Arif1, Corinna Hopfen1,2, Daniela Nunes1, Maarten Hilbrant1, Alistair McGregor1
1Oxford Brookes University, Oxford, UK, 2Veterinärmedizinische Universität Wien, Vienna, Austria

The evolution of morphological characters remains one of the main foci of contemporary evolutionary biology. Exploring the evolution of morphological characters requires an understanding of the genetic architecture of those characters. Here we try to dissect and understand the genetic architecture of differences in compound eye size between Drosophila mauritiana and D. simulans. Differences in compound eyes such as those of Drosophila can arise due to either more ommatidia (the individual facets of compound eyes) or larger ommatidia. Additionally, negative correlations between face width and eye size have been reported for many insect species, highlighting the shared developmental basis of eye and face tissue. We have found that D. mauritiana tend to have significantly larger eyes compared to D. simulans and this is caused by differences in both ommatidia number and diameter. Our data also show that eye and face width display a negative correlation suggesting a trade-off in the development of these structures. Interestingly Quantitative Trait Loci (QTL) mapping, between D. mauritiana and D. simulans, has revealed a distinct chromosomal basis for differences in eye size and face width. Subsequent fine-scale mapping focusing solely on the eye size differences between these species has allowed us to identify a region of 2 Mb on the X chromosome that is responsible for most of the difference in this trait. This region contains two promising candidate genes based on prior mutant phenotypes and developmental expression, which may be responsible for differences in ommatidia size. We are currently focusing on further fine-scale mapping of other eye and head traits that differ between these species, and functionally analysing candidate gene(s) for ommatidia size variation. The identification of the developmental genetic basis for differences in eye form will broaden out understanding of how complex morphological traits evolve.

P-282
DETECTION AND ANALYSIS OF HETEROGENEITY OF CIS EQTLS ACROSS ETHNIC GROUPS

Christopher Grace1, John Whittaker2, Julie Huxley Jones2, Andrew Morris1
1Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK, 2GlaxoSmithKline, Medicines Research Centre, Stevenage, UK

Trans-ethnic meta-analysis of populations with different ancestry can increase power and facilitate fine-mapping. Association analysis and fixed-effect meta-analysis were performed to locate cis eQTLs in the three Phase II HapMap ethnic groups: 30 CEPH trios (CEU), 30 Yoruba trios (YRI), and 90 Chinese / Japanese individuals (CHB/JPT). 4177 cis eQTLs were identified (FDR: 5%, p < 1x10^-4). Evidence of heterogeneity (Cochran’s Q p < 0.05) was detected at 537 eQTLs, greater than by chance (binomial test p < 2.96x10^-40). 65 cis eQTLs with genome-wide significance (p <5x10^-8) and significant heterogeneity (Cochran’s Q p < 10^-3) were selected for analysis. Two distinct groups were identified: (i) association signals present (p < 0.05) in all samples; (ii)
association signals that are specific to one / two samples. Across 65 eQTLs, 3 were specific to one ethnic group, and 24 specific to two ethnic groups. Specific examples were selected for further analysis. Probe GI_33457307-S (HGNC: TMED4) has peak signal at SNP rs2072183 (p = 7.22x10\(^{-17}\), Cochran’s Q p = 3.4x10\(^{-4}\)) with significant signals in CHB/JPT and CEU (p = 1.45x10\(^{-10}\) and 3.83 x 10\(^{-7}\)), but not in YRI (p = 0.65). Allele frequencies of the samples are of similar size (ASN: 0.36, CEU: 0.28, YRI: 0.22). There is evidence of differences of LD patterns between CHB/JPT, CEU and YRI.

Probe GI_21070998-S (HGNC: STIM2) has a peak signal at SNP rs7688662 (p = 3.96x10\(^{-17}\), Cochran’s Q p = 4.93x10\(^{-4}\)) with a significant signal in CHB/JPT (p = 3.18x10\(^{-14}\)), but not in CEU and YRI (p = 0.07 and 0.35). Allele frequencies in CHB/JPT and CEU are similar (0.39 and 0.31), but differ in YRI (0.7). Further inspection of the region suggests that patterns of LD differ between CHB/JPT and CEU.

These examples indicate that the heterogeneity may be partially explained by differences in LD between the samples.

P-283
GENETIC INTERACTIONS IN THE HUMAN LIVER

Darren J. Fitzpatrick, Derek Greene, Denis Shields
University College Dublin, Dublin, Ireland

Abstract can be found in the student symposium abstract section on page 61

P-284
GENETICS OF SEED AND SEEDLING GROWTH COMBINING GROWTH MODELLING AND QTL MAPPING IN BRASSICA RAPA

Ram Kumar Basnet\(^1\), Dev Tiwari\(^1\), Johan Bucher\(^1\), Steven Groot\(^2\), Maliepaard Chris\(^\prime\), Bonnema Guusje\(^\prime\)

\(^1\)Plant Breeding Department, Wageningen University, Wageningen, The Netherlands, \(^2\)Plant Research International, Wageningen, The Netherlands

Brassica rapa (2n=20, AA) is an important crop grown for vegetable and oil production. Good quality seed resulting in vigorous seedlings is essential for crop establishment under diverse environmental conditions for higher crop productivity. Since seed and seedling vigor are complex traits, it is necessary to measure the longitudinal physiological development over time to unravel the genetics of seed and seedling vigor. In this study, we used a segregating double haploid (DH) population of a cross between Pak Choi (black-seeded vegetable type) and Yellow Sarson (yellow-seeded oil-seed type). For phenotyping, germination was first recorded over time periods at graded levels of salt, drought stress and non-stress conditions, and then the same germinated seeds were used to monitor seedling growth by measuring root length, shoot length and number of lateral roots for the first 10 days of seedling growth under respective conditions. Several QTL for seed germination, root and shoot length over at different growing days were identified across the genome. We will assess linear as well as non-linear growth models, for instance, Gompertz and expo-linear to define the germination behavior, as well as root and shoot growth characteristics of B rapa seedlings. In addition, response to estimated growth parameters on diverse environmental conditions will be explored. These growth parameters will be correlated with genomic regions affecting germination and seedling growth behavior. Genetics of seed and seedling vigor of B. rapa will be explored using two-step approach of growth modelling on longitudinal data and QTL mapping for genomic selection.
P-285
SELECTING THE OPTIMAL SET OF INDIVIDUALS WITHIN A FAMILY-BASED COHORT TO MAXIMISE POWER FOR A LINKAGE TEST

David Porteous, Harry Campbell, Andrew MacLeod
University of Edinburgh, Edinburgh, UK

Objectives: Genome-wide association studies (GWAS) have identified a number of significant genetic variants influencing quantitative traits, but a substantial proportion of the estimated heritability of these traits remains unexplained. Population-based GWAS studies are designed to detect common variation, so rare variants are likely to be overlooked. Using family data is likely to increase the chance of detecting linkage to such variants. Here, theoretical calculations on the optimal set of individuals to maximise power to detect rare variants are applied to a family-based cohort, the Generation Scotland Scottish Family Health Study (GS:SFHS).

Methods: The contribution made to the power of a linkage analysis for all families in GS:SFHS was calculated by approximating the contribution that each family would make the non-centrality parameter (NCP) of the test statistic under a range of population and trait parameters. Several models of selective genotyping were applied to calculate the most efficient use of the available resource in terms of average power contributed per individual across the whole cohort.

Results and Conclusions: For larger families, we find that genotyping selected family members will maximise the average contribution made per individual within a family to the overall power, when compared to genotyping the whole family. We apply this approach to GS:SFHS and calculate that the study would be well powered to replicate previous results and detect linkage to rare variants. When resources are limited in a family-based cohort analysis - e.g. by the total number of individuals available to genotype in a family based cohort - then selectively genotyping family members can result in a more efficient study, and allows for the selection of sub-cohorts for genotyping. We present a simple criterion to select subsets of families to genotype to ensure that a specified QTL effect can be detected with a fixed total sample size.

P-286
OMIC PARTITIONING OF OBESITY TRAITS IN MICE

Alireza Ehsani1, Luc Janss1, David Edwards1, Daniel Pomp2, Mark F. Allan2, Peter Sorensen1
1Aarhus University, tjele, Denmark, 2UNC, chapel Hill, USA

Genome-wide expression profiling, using microarray or sequence based technologies, can be used to identify genes and genetic pathways whose expression patterns influence a complex trait. In many cases in agriculture and biomedicine, multiple phenotypes are measured, and diseases can be characterized by changes in a set of phenotypes. We take here the case of obesity in mice, and present a multivariate approach for genome-wide expression analyses of complex traits is presented. The approach is based on a Bayesian mixture model that simultaneously accounts for genetic relatedness among samples, relatedness among transcripts, observations from multiple correlated traits, and we also consider the use of gene sets. The use of this novel multivariate gene set approach is demonstrated in a genome-wide expression study of growth traits recorded in a mouse F2 population. The experimental data obtained from 440 mice included phenotypic recordings of three growth related traits (body weight, feed intake and feed efficiency), genotypes of 1806 single nucleotide polymorphisms (SNPs) and genome expression profiles of liver tissue samples. These data gives us a unique opportunity to investigate how much of the phenotypic variance in the observed growth traits is explained by variations in the genome itself and by variations in the transcriptome. It also gives an opportunity to illustrate the power of our approach to explore the underlying genetic architecture of growth traits such as the contribution of individual chromosomes and genomic regions to the trait variance and how these regions influences the expression levels of different biological pathways.
P-287
GENETIC DISSECTION OF GROWTH TRAITS IN A CHINESE INDIGENOUS × COMMERCIAL BROILER CROSS

Zheya Sheng1,2, Mats E. Pettersson2, Xiaoxiang Hu1, Örjan Carlborg2, Ning Li1
1State Key Laboratories of AgroBiotechnology, China Agricultural University, Beijing, China, 2Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden

Objectives: The main objective is to explore the genetic basis of growth-related traits we examined in this study to identify alleles from the commercial stock HQLA that could potentially be used for improving the performance of the indigenous breed HB.

Methods: We established an F2 intercross between the slow-growing Huiyang Beard Chicken (HB) and the fast-growing commercial broiler breed “High Quality chicken Line A” (HQLA). With the genotypes from Illumina Chicken 60K SNP chip and phenotypes of 7 classes, all of which are growth-related, we performed a genome-wide linkage analysis based on Haley&Knott regression in both one- and two-dimensional scan.

Results: 7 distinct QTL were detected in this population. Most of them were pleiotropic QTL that influencing correlated traits or acting through several growth phase of the same trait. The phenotypic variances explained by the identified QTL varied from 2% to 25%. In the two-dimensional scan, 6 pairs of loci were found with significant additive-by-additive effects for 5 traits. They explained a range of 4%-6% of phenotypic variances. Furthermore, we found an interesting feature while searching for candidate genes within the epistatic regions. The finding was that 8 of 10 unique interaction regions harbored genes in the same pathway — Ubiquitin Mediated Proteolysis.

Conclusions: The results from Single-QTL scan were in concordance with the correlations between all the pairs of phenotypes. In addition, all the QTL mapped in this study overlap with QTL reported in earlier studies. On the other hand, the QTL pairs we found in two-dimensional scan were not the same regions as in the one-dimensional scan, suggesting the main additive and epistatic components underlying growth-related traits in our population were different. Additionally, the candidate genes we found provide new insights to the genetics of growth related phenotypes.

P-288
THE GENETIC BASIS OF HETEROSIS: MULTIPARENTAL QUANTITATIVE TRAIT LOCI MAPPING REVEALS CONTRASTED LEVELS OF APPARENT OVERDOMINANCE AMONG TRAITS OF AGRONOMICAL INTEREST IN MAIZE (ZEA MAYS L.)

Amandine Lariepe1,2, Laurence Moreau2, Julie Fivet3, Brigitte Mangin4, Sylvain Jasson4, André Gallais3, Alain Charcosset1
1INRA, Gif-Sur-Yvette, France, 2Biogemma, Chappes, France, 3AgroParisTech, Gif-Sur-Yvette, France, 4INRA, Castanet-Tolosan, France

Understanding the genetic bases underlying heterosis is a major issue in maize (Zea mays L.). We extended the North Carolina design III (NCIII) by using three populations of recombinant inbred lines derived from three parental lines belonging to different heterotic pools, crossed with each parental line to obtain nine families of hybrids. A total of 1253 hybrids were evaluated for grain moisture, silking date, plant height, and grain yield. Quantitative trait loci (QTL) mapping was carried out on the six families obtained from crosses to parental lines following the “classical” NCIII method and with a multiparental connected model on the global design, adding the three families obtained from crosses to the nonparental line. Results of the QTL detection highlighted that most of the QTL detected for grain yield displayed apparent overdominance effects and limited differences between heterozygous genotypes, whereas for grain moisture predominance of additive effects was observed. For plant height and silking date results were intermediate. Except for grain yield, most of the QTL identified showed significant additive-by-additive epistatic interactions. High correlation observed between heterosis and the heterozygosity
of hybrids at markers confirms the complex genetic basis and the role of dominance in heterosis. An important proportion of QTL detected were located close to the centromeres. We hypothesized that the lower recombination in these regions favors the detection of linked QTL in repulsion phase leading to apparent overdominance for heterotic traits and linked QTL in coupling phase, reinforcing apparent additive effects of linked QTL for the others.

P-289
GENETIC DISSECTION OF GENOMEWIDE EXPRESSION VARIATION IN DROSOPHILA HEADS

Stuart Macdonald1, Elizabeth King2, Brian Sanderson1, Casey McNeil1, Tony Long2
1University of Kansas, Lawrence, KS, USA, 2University of California, Irvine, Irvine, CA, USA

Drosophila melanogaster is widely employed as a model genetic system to understand fundamental aspects of the control of complex trait variation in populations. In addition, the system is increasingly recognized as an important translational model for the study of human neurodegenerative disease and the action of drugs of abuse. As with all complex, polygenic traits identifying the molecular pathways and causative genes responsible for variation in these phenotypes is challenging. Given the community interest in genetically dissecting neurobehavioral traits in flies, we took advantage of a novel resource for genetic analysis to characterize quantitative variation in transcript abundance in Drosophila heads. The Drosophila Synthetic Population Resource (DPSR) is composed of over 1,600 genotyped Recombinant Inbred Lines (RILs) derived from a pair of highly-recombinant synthetic laboratory populations. These two populations were each initially founded by a different set of eight founder strains, ensuring high functional allelic diversity in the DSPR. We generated 600 heterozygous genotypes - the progeny of intercrosses between DSPR lines from the different populations - isolated RNA from mated adult female heads, and subjected each sample to microarray analysis. These data allow us to construct gene networks and capture the full biological complexity of the pathways involved in gene regulation in the Drosophila head and brain. Genomewide expression QTL (eQTL) analysis also provides a high-resolution picture of the location, effect, and population frequency of loci that influence expression variation in the head, and allows us to examine the relative abundance of cis- and trans-regulatory loci. In addition, researchers using the DSPR to genetically dissect neuronal or behavioral phenotypes will be able to exploit our eQTL data for a systems-level analysis of trait variation, and quickly home in on likely candidate genes.

P-290
DETECTION OF TWO GENOMIC REGIONS ASSOCIATED WITH RESISTANCE TO BOVINE TUBERCULOSIS

Mairead L. Bermingham1, Stephen C. Bishop1, John A. Woolliams1, Ricardo Pong-Wong1, Robin A. Skuce2,3, Stanley W.J. McDowell2, Adrian R. Allen2, Stewart H. McBride2, Jon J. Ryder4, David M. Wright3, Elizabeth J. Glass1
1Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, UK, 2Agri-Food and Biosciences Institute Stormont, Stoney Road, Belfast, Belfast, UK, 3The Queen’s University of Belfast, Department of Veterinary Science, Stormont, Belfast, UK, 4Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool, UK

Bovine tuberculosis (bTB), caused by Mycobacterium bovis, remains a serious problem in the UK. Control involves routine tuberculin skin testing and the slaughter of test-positive animals. The aim of this study was to identify loci associated with bTB resistance. DNA samples were collected from 464 Holstein-Friesian dairy herds in Northern Ireland during 2008-2009. Cases were animals with both a positive skin test reaction and confirmed bTB lesions, and were sampled at slaughter. Age-matched controls (animals negative to the skin test) were sampled from a subset of high-prevalence case herds. After QC edits, 1,151 cattle (592 cases and 559 controls) and genotype data from 617,610 SNPs remained. Genome-
wide association (GWA) using linear and logistic mixed models and regression and regional heritability mapping (RHM) were used to test for associations between SNPs and bTB resistance and also to estimate the variance explained by significant regions. Using GWA mapping, chromosome-wide significant associations for bTB were detected at eight loci (chromosome-wide adjusted P < 0.05), one in an intronic of *Myosin IIIB* (**MYO3B**) (chromosome 2; LD with adjacent SNPs was low, r²<0.05) and seven in intronic regions of protein tyrosine phosphatase, receptor type, T (**PTPRT**) (chromosome 13), respectively. RHM also identified the region comprising the cluster of significant SNPs (chromosome-wide adjusted P < 0.05), but did not confirm the region with the single significant SNP. The estimated genomic heritability for resistance to bTB was 0.21 (SE 0.06). In this sample, which is a selected population of animals, when fitted simultaneously the SNP in **MYO3B** explained 15.2% and regional SNP loci in the **PTPRT** gene collectively explained 5.5% of the additive genetic variance. Human polymorphisms in **MYO3B** and **PTPRT** genes have been linked to autism and autoimmunity respectively. We aim to independently replicate these findings and understand the genetic architecture of the significant regions

**P-291**
**BIOFUEL FROM BARLEY STRAW- THE QUEST TO FIND LOCI UNDERLYING THE SACCHARIFICATION TRAIT**

*Helena Oakey¹, Jordi Comadran², Reza Shafiei¹, Leonardo Gomez³, Caragh Whitehead¹, Simon McQueen-Mason¹, Robbie Waugh², Claire Halpin¹*

¹University of Dundee, Dundee, UK, ²James Hutton Institute, Dundee, UK, ³University of York, York, UK

Improving the sugar yield of barley straw for second generation biofuel production without having any detrimental effects on crop yield or quality is an important step in making biofuel production feasible and competitive.

A diverse genetic pool of 648 elite spring barley varieties with 5 replicates were grown in a polythene tunnel in 2010 and were screened for sugar yields in a simple saccharification assay using a high throughput automated platform. This panel was also genotyped with approximately 9000 SNPs in known barley genes.

The aim of the experiment was twofold - to identify the best elite varieties for biofuel production and to identify loci underpinning the saccharification trait.

Here we discuss a one-stage approach to the association analysis of the large raw data set which allows for environmental (during plant growth) and laboratory (during saccharification assay) variation. Results suggest that accounting for this extraneous variation is just as important as accounting for structure and relatedness between varieties in the quest to ensure that spurious positive associations can be removed.

**P-292**
**THE USE OF NEXT GENERATION SEQUENCING TO IDENTIFY PUTATIVE CAUSATIVE MUTATIONS IN A DEFINED (E)QTL REGION**

*Xidan Li², Xiaodong Liu¹², Javad Nada³⁴, F Turner³, I.C Dunn³, E Le Bihan-Duval³, Dirk-Jan De Koning¹⁴*

¹Swedish University of Agricultural Sciences, Animal Breeding and Genetics, Uppsala, Sweden, ²International Master Program in Bioinformatics, Uppsala University, Uppsala, Sweden, ³INRA, UR83, Recherches Avicoles, Nouzilly, France, ⁴Roslin Institute and R(D)SVS, University of Edinburgh, Genetics and Genomics, Midlothian, UK

Genetical genomics has been shown to be a promising approach for dissecting complex traits. Using targeted genetical genomics, a QTL affecting the initial PH of chicken meat on chromosome 1 has been narrowed down to 16 differentially expressed genes spanning a region of < 1 Mb. Here, we show the application of next generation sequencing (NGS) to identify putative causal mutations in the QTL region. In this study, ten birds that were homozygous for either the high (QQ) or low (qq) QTL allele were
selected for targeted sequencing. DNA was enriched for 1 Mb around the QTL using Agilent Sure Select Target Enrichment. Using paired-end libraries with an average insert size of 300 kb, we obtained about 200X coverage for each of the 10 homozygote genotypes. The candidate SNPs were identified from key regions, such as the 5 base pairs flanking a splicing site, CpG islands, promoter sequences, transcription binding sites, utr regions, and exons. We identified SNPs in 5 splicing sites, 4 utr regions, 3 exons and 1 promoter sequence that are all strong candidates for the QTN. Moreover, the non-synonymous SNPs were ranked based on their postulated effect on protein function, where the 3 mutations in two separate genes were defined to be most likely to explain underlying the genetic mechanism between the two QTL genotypes.

The bioinformatics analysis provides a ranking among the positional candidate genes and the SNPs within these genes. The results suggest that candidate genes and SNPs in highly conserved regions have the most potential to affect protein function. These SNPs are valuable candidates for further study to elucidate the mechanism underlying this QTL and identifying the QTN among many SNPs.

**P-293**

**QTL ANALYSIS OF MULTI-PARENT POPULATIONS**

Maria-João Paulo, Martin Boer, Hans Jansen, Marcos Malosetti, Marco Bink, Fred van Eeuwijk
Biometris, WUR, Wageningen, The Netherlands

In the dissection of complex traits into underlying QTL and corresponding allelic effects, multi-parent populations can be more powerful than bi-parental crosses or association panels. Bi-parental populations are restricted by the small number of alleles present in the two parents while association panels lack power to estimate allelic effects accurately and suffer from hidden population structure.

Multi-parent populations may arise as a subset from a breeding population or these may be created by designed crosses between several founders. To increase recombination rates and thereby mapping resolution, the designed crosses may comprise one or multiple rounds of inter-mating before large offspring populations are produced.

Objective: Development of statistical methodology that is efficient to calculate conditional QTL genotype probabilities given marker information, so-called genetic predictors, in multi-parent populations and efficient in the subsequent QTL analyses including epistatic and environmental interactions.

Methods: Mixed models allowing tests for QTL main effects, QTL by background interactions, and QTL by QTL interactions; Hidden Markov Models to calculate genetic predictors from marker data.

Results: We present the results of three multi-parent populations, namely a complex cross of recombinant inbred lines in Arabidopsis (AMPRIL), an F2 diallel design in Tomato and a Nested Association Mapping panel in Maize (NAM).

Conclusions: In general, more QTLs are likely to segregate in multi-parent populations than in a bi-parental cross.

Our method is powerful for QTL detection in multi-parent populations. Mixed models allow detection of QTLs even when they segregate in only one sub-population. Genetic predictors can potentially contain more information on the parental origin of alleles than markers, because of their ability to combine all the information available in groups of markers.
P-295
MAPPING VARIANCE-CONTROLLING GENES: HOW TO CORRECT FOR POPULATION STRATIFICATION

Xia Shen¹, Lars Rönnegård²,³, Örjan Carlberg²
¹Uppsala University, Uppsala, Sweden, ²Swedish University of Agricultural Sciences, Uppsala, Sweden, ³Dalarna University, Borlänge, Sweden

A number of recent studies have shown that it is possible to detect genes that display a genetic variance heterogeneity in GWAS data. The contributions of these genes to the phenotypic variance in a population will be missed when using standard quantitative genetic models and when appropriately accounted for, it has been shown that they might make important contributions to genetically determined robustness of populations to genetic and environmental perturbations. The genetic mechanisms underlying variance control is still unknown, but it is expected that genes showing variance heterogeneity are involved in gene-by-gene or gene-by-environment interactions, making their detection an important step when aiming to identify complex genetic pathways. Statistical tests for variance heterogeneity might be affected by population stratification in the same way as tests for the mean, which usually results in genomic inflation of false discoveries. Methods such as linear mixed models have been used in standard GWAS analyses, but extension of such techniques to the mapping of variance-controlling genes is not straightforward.

Here, we propose a two-stage correction method for population stratification in variance-controlling gene mapping. The method is based on double hierarchical generalized linear models, and we compare several variants of measuring genotype-specific variance of the phenotype. A simulation study based on a highly stratified real dense marker dataset shows that the method is capable of remove genomic inflation caused by stratification. When used to analyse a real phenotype from the same experimental dataset, we show that the method allows detection of a candidate gene with a strong variance-controlling effect with genome-wide significance after proper control for stratification.

P-296
EVIDENCE FOR BODY MASS INDEX AS A MODIFIER OF GENETIC ASSOCIATIONS FOR THYROID STIMULATING HORMONE LEVELS IN EUROPEAN AMERICANS FROM THE EMERGE NETWORK

Jennifer R. Malinowski⁵, Joshua C. Denny¹,², Suzette J. Bielinski⁵, Melissa A. Basford⁶, Yuki Bradford⁵, High Seng Chai⁵, Peggy L. Peissig¹⁰, David Carrell¹², Jyotishman Pathak³, Russell A. Wilke², Luke Rasmussen¹⁰, Xiaoming Wang⁶, Jennifer A. Pacheco¹³, Abel Kho¹³, Noah Weston¹², Martha Matsumoto⁶, Rongling Li¹⁴, Iftikhar J. Kullo⁸, Christopher G. Chute⁹, Rex L. Chisholm¹³, Eric B. Larson¹², Catherine A. McCarty¹¹,¹⁶, Daniel R. Masys¹, Dan M. Roden²,⁴, Mariza de Andrade⁸, Marylyn D. Ritchie²,¹⁵, Dana C. Crawford, on behalf of the eMERGE Network³
¹Department of Biomedical Informatics, Vanderbilt University, Nashville, TN, USA, ²Department of Medicine, Vanderbilt University, Nashville, TN, USA, ³Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN, USA, ⁴Department of Pharmacology, Vanderbilt University, Nashville, TN, USA, ⁵Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA, ⁶Office of Research, Vanderbilt University, Nashville, TN, USA, ⁷Division of Epidemiology, Mayo Clinic, Rochester, MN, USA, ⁸Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA, ⁹Division of Cardiovascular Diseases, Dept. of Health Sciences Research, Mayo Clinic, Rochester, MN, USA, ¹⁰Biomedical Informatics Research Center, Marshfield Clinic Research Foundation, Marshfield, WI, USA, ¹¹Center for Human Genetics, Marshfield Clinic Research Foundation, Marshfield, WI, USA, ¹²Group Health Research Institute, Seattle, WA, USA, ¹³Department of Medicine and Center for Genetic Medicine, Northwestern University, Chicago, IL, USA, ¹⁴Office of Population Genomics, National Human Genome
Thyroid stimulating hormone (TSH) hormone levels are normally tightly regulated within an individual; relatively small individual variations may indicate thyroid disease. Genome-wide studies (GWAS) have repeatedly suggested that variants in PDE8B and FOXE1 are associated with TSH levels. Variants in FOXE1 have been previously identified to impart susceptibility to thyroid cancer and primary hypothyroidism. PDE8B may influence TSH levels through regulation of cAMP signaling. The Electronic Medical Records and Genomics (eMERGE) Network is a collaboration across institutions with biobanks linked to electronic medical records (EMRs); Phase I eMERGE members were Group Health Cooperative/University of Washington, Marshfield Clinic, Mayo Clinic, Northwestern University, and Vanderbilt University. The eMERGE Network uses EMR-derived phenotypes to perform genomic association studies in diverse populations. At each site, GWAS was conducted for complex diseases and medically relevant quantitative traits; overall, approximately 17,000 subjects were genotyped in Phase I. In this report, we performed a “no additional genotyping” GWAS for a quantitative trait—serum TSH level—in ~5000 European Americans (EAs). Tests of association were performed using linear regression and adjusted for age, sex, body mass index (BMI), and principal components, assuming an additive genetic model. Our results replicate the known association of PDE8B in EAs (rs1382879 p=7.16x10^-18, β =0.09). FOXE1 variants were suggestively significant (rs10759944: p=1.08x10^-6, β =-0.05). Tests of interaction between BMI and 118 genetic variants associated with TSH levels at p<10^-5 were also performed. Significant interaction terms were identified involving variants in NRG1 (rs2466067, p=0.04; rs4298457, p=0.05) and NFIA (rs10489909; p=6.21x10^-3). NRG1 (neuregulin) is expressed in papillary thyroid carcinomas and may regulate thyroid cell proliferation; NFIA encodes a nuclear transcription factor. These results support the previously reported association between PDE8B and serum TSH levels in EAs and suggest that BMI may modify specific genetic associations for serum TSH levels in this population.

P-297
ANALYSIS OF BLOOD LIPID TRAITS USING REGIONAL HERITABILITY MAPPING

Masoud Shirali1,2, Ricardo Pong-Wong2, Pau Navarro3, Caroline Hayward3, Veronique Vitart4, Igor Rudan4,5, Harry Campbell6, Nicholas Hasnig3, Alan Wright3, Chris Haley2,3

1Department of Animal Science, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran, 2Division of Genetics and Genomics, The Roslin Institute and R(D)SVS, University of Edinburgh, Easter Bush, Midlothian, UK, 3MRC Human Genetics Unit, MRC IGMM, University of Edinburgh, Edinburgh, UK, 4Croatian Centre for Global Health, Faculty of Medicine, University of Split, Split, Croatia, 5Centre for Population Health sciences, University of Edinburgh, Edinburgh, UK

Although Genome-wide association studies (GWAS) have identified a substantial number of loci associated with diseases and traits, a considerable amount of heritability has not been explained. A newly developed method called Regional Heritability Mapping (RHM) that integrates variance caused by several rare or common alleles at a locus into a single estimate has been shown to be capable of explaining part of this missing heritability. The objective of this study was to compare the performance of RHM and the standard GWAS single-marker analysis for four blood lipid traits, (Low-density lipoprotein (LDL), High-density lipoprotein (HDL), Cholesterol and Triglyceride) in human populations. A total of 2322 adult volunteers from 3 Southern European populations were genotyped using 300k SNP chips; samples were obtained from an epidemiological field survey. The results from the GWAS analysis failed to identify a single genome-wide significant SNP associated with any of the four traits. On the other hand, the analysis performed with RHM identified two genome-wide significant regions associated with the traits. One region on chromosome 16 was associated with HDL explaining 3% of the heritability and the other one located on chromosome 19 associated with LDL explains 6% of the heritability. These two regions are consistent with the results reported in a previously published large meta-analysis including ~100,000
individuals. The first region on chromosome 16 harbours the most significant SNP (rs3764261) associated with HDL reported in the meta-analysis. The most significant SNP (rs4420638) associated with LDL reported in the meta-analysis was not available in our genotype data set, but it is located in our significant region on chromosome 19. This study shows that RHM has enhanced power to detect trait-associated loci, potentially due to the fact that RHM effectively captured the aggregate variance of all the alleles segregating in the region.

P-298
ASSOCIATION BETWEEN RETINAL ARTERIOLAR TORTUOSITY AND A VARIANT IN COL4A2, ENCODING A COLLAGEN SUBUNIT
Mirna Kirin, Tom MacGillivray, Sarah Wild, Igor Rudan, Caroline Hayward, Alan Wright, Harry Campbell, Veronique Vitart, James Wilson
University of Edinburgh, Edinburgh, UK
Cardiovascular disease is a classic complex disease, influenced by numerous risk factors and their interactions. Much research has therefore focussed on the risk factors or intermediate phenotypes such as blood pressure. Here we analyse a rarely studied endophenotype, retinal vessel tortuosity or degree of curvature, which has been associated with hypertension. The retinal vasculature is thought to share the same physiological and anatomical characteristics as the cerebral and coronary systems, however unlike these can be visualised non-invasively. Retinal arteriolar and venular tortuosities were measured in non-mydriatic fundus images from 808 subjects in the Orkney Complex Disease Study (ORCADES) using the semi-automated retinal vasculature measurement programme SIVA. Vessels were measured up to two disc diameters away from the optic disc. As a quality check, left-right eye correlations were calculated (0.65 for venules, 0.91 for arterioles). Inter and intra-grader coefficients varied from 0.88-0.97. Using pairwise estimates of kinship based on genomic sharing, heritabilities were calculated as 31% for venules and 54% for arterioles. A genome wide association analysis of SNPs imputed to Hapmap revealed four suggestive associations with arteriolar tortuosity, the most significant of which is in COL4A2 at 13q34 (P~10^-6). COL4A2 encodes one of the six subunits of type IV collagen – the major structural component of the basement membrane, which is important in angiogenesis. COL4A2 shares a promoter with the neighbouring paralogue, COL4A1, Mendelian variants in which underlie syndromes which include tortuous retinal vessels. Further evidence that this retinal tortuosity association is real comes from the fact that SNPs in this region show genome-wide significant associations with vessel traits, including coronary artery calcification, coronary heart disease and arterial stiffness. Mouse mutants report small vessel disease, intracerebral haemorrhage and retinal changes. Replication is being sought in two further populations.

P-299
GENOME-WIDE ASSOCIATION STUDY OF HEART TRANSPLANT PATIENTS AND NEPHROTOXICITY FROM IMMUNOSUPPRESSION THERAPY IN A BIOREPOSITORY LINKED TO ELECTRONIC MEDICAL RECORDS
Matthew Oetjens, Erica Bowton, Kelly Birdwell, Joshua Denny, Dan Roden, Dana Crawford
Vanderbilt University, Nashville, TN, USA
Immunosuppression therapy after a heart transplant reduces a patient’s inflammatory immune response toward the allograft, increasing its longevity in the patient. However, many patients with long-term exposure to immunosuppressants are at increased risk for kidney failure. To identify possible genetic predictors of nephrotoxicity, we performed a genome-wide association in 136 heart transplant patients identified in Vanderbilt’s biorepository (BioVU) linked to electronic medical records (EMRs). Our cohort algorithm identified heart transplant recipients in BioVU using the following criteria: three or
more ICD9 codes for heart replaced by transplant or one CPT code for heart transplant, one or more mention of an immunosuppressant, and two or more heart biopsy pathology reports or if the subject died within 12 months of the transplant. Nephrotoxicity was assessed using pre and post-transplant serum creatinine laboratory values. The clinical characteristics of the population are as follows: the mean age was approximately 44 years old, 82% were observer-reported European descent, 21% required dialysis, and 21% died while undergoing treatment. Median pre-transplant creatinine levels were 0.9mg/dl and median post-transplant creatinine levels measured between 6 and 8 months post-transplant were 1.4 mg/dl. This study population has been genotyped on the Illumina’s Human OMNI 1 Quad and the pharmacogenetic-targeted ADME Core Panel. Using linear regression assuming an additive genetic model, we performed single SNP tests of association in this multi-ethnic population for post-transplant creatinine levels adjusted for age and sex. No associations at genome-wide significance were identified. However, log transformed post-transplant creatinine levels were most associated with intragenic SNPs rs17634863 (p = 1.28 x 10-6, β = 0.52), rs13400738 (p = 3.92 x 10-6, β = 0.49), and rs9574774 (p = 4.48 x 10-6, β = -0.57). We have identified potential loci associated with post-transplantation kidney function, but larger sample sizes will be needed to confirm and extend these findings.

P-300 – ABSTRACT WITHDRAWN

P-301
THE IMPACT OF HETEROGENEITY ON THE IDENTIFICATION OF CAUSATIVE GENES IN COMPLEX DISEASES: A PROBABILISTIC ANALYSIS

Yael Berstein, Shane E. McCarthy, W. Richard McCombie
Stanley Institute for Cognitive Genomics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA

Targeted sequencing of genomic coding regions, also known as exome sequencing, have shown outstanding potential in identifying causative genes linked to Mendelian disorders. However, the efficient application of this method in the study of complex genetic disorders still remains a major challenge. Despite the large amount of sequencing data accumulated from complex genetic disorder studies, such as schizophrenia and bipolar disorder, causative genes related to these diseases remain to be discovered. Preliminary studies suggested that some of the most significant barriers in the detection of genes related to complex genetic disorders are the large heterogeneity of the data, both genetic and phenotypic, and the lack of theoretic boundaries to predict errors. Here we introduce probabilistic models to estimate the impact of heterogeneity on the identification of variants related to complex genomic disorders. Several key parameters of case control studies were used as variables in the formulation of the probabilistic models e.g. sample size, the expected amount of causative mutations and the minimal recurrence level required to determine that a variant is a causative mutation. The impact of genetic heterogeneity was measured by the probability of observing a mutation given different combinations of these parameters mimicking the statistical properties of typical case-control studies. Our models show that even with moderate sample sizes, increases in mutation recurrence level and genetic heterogeneity significantly impact the signal of causative mutations over random genetic variation. For example, modelling psychiatric diseases such as Bipolar Disorder where the genetic heterogeneity is assumed upwards of 500 causative mutations, with a minimal recurrence level of 17, over 10,000 individuals will require screening in order to achieve 80% probability of identifying specific causative mutations. Even though these models are theoretic, they are likely to contribute to the understanding and dissection of the main challenges facing future studies of disease-related genes.
P-302
THE GENETIC ARCHITECTURE OF HUMAN OCULAR BIOMETRY – THE EXAMPLE OF CENTRAL CORNEA THICKNESS

Yi LU1, on behalf of CCT Consortium2
QIMR, Brisbane, Australia

The human ocular biometric parameters comprise a set of highly heritable and often correlated quantitative traits. One notable example is central cornea thickness (CCT). It has an estimated heritability up to 95%. Previous Genome-wide association studies (GWAS) have identified 11 CCT-associated loci which account for ~4% of additive variance. Based on a set of 1853 individuals we estimated that ~15.5% (se=18%) of variance can be explained by all common SNPs on GWAS arrays. Therefore, we conducted CCT meta-analysis on an overall sample size of >20,000 samples including European and Asian descendants, with and without eye conditions. The differences in the sample attributes enabled various comparisons between the subsets of meta-analysis results. We brought the number of CCT-associated loci to 35, among which two loci harbor multiple association signals. The novel loci explained another ~4% of heritability. We observed similar CCT distributions in European and Asian populations, and found that underlying genetics were largely shared between populations. Despite a slightly shifted CCT distribution in the group of glaucoma patients, there was significant overlap in CCT loci from samples with and without glaucoma, suggesting similar pathway controlling CCT variation regardless of disease status. Our results also showed that genes harboring rare variants causing Mendelian disorders with clinical feature of extreme cornea thinning (e.g. Brittle Cornea Syndrome, Ehlers-Danlos Syndrome and Osteogenesis Imperfecta) also harbor common variants that affect CCT in the general population. Furthermore, some CCT-associated loci affect susceptibility of common eye disease with mildly reduced CCT, e.g. two loci were nominally associated with glaucoma risk. Finally, using a novel pathway analysis tool we demonstrated that CCT loci converge to collagen and extracellular pathway, which allowed rationalisation of the existing findings. Taken together, this meta-analysis revealed a more complete genetic architecture for CCT, with implications for related ocular biometry and conditions.

P-303
APPLYING A PHENOME-WIDE ASSOCIATION APPROACH TO INVESTIGATE THE EFFECT OF MITOCHONDRIAL DNA LEVELS ON DIVERSE PHENOTYPES IN THE NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEYS

Sabrina Mitchell, Kristin Brown-Gentry, Laura Hunt, Dana Crawford, Deborah Murdock
Vanderbilt University Medical Center, Nashville, TN, USA

Mitochondria are the primary energy producers of the cell and are essential for cellular metabolism. Mitochondrial dysfunction has been implicated in cancer, aging, and neurodegenerative diseases including Alzheimer’s and Parkinson’s. Mitochondria are unique organelles in that they have their own DNA separate from the nuclear DNA. Maintenance and expression of mitochondrial DNA (mtDNA) are essential for mitochondrial function. Indeed, a reduction in levels of mtDNA has been shown to compromise mitochondrial function. Mutations in a number of genes result in decreased mtDNA copy number manifesting in a range of phenotypes, dependent upon the mechanism of mtDNA depletion. Collectively, these are referred to as mtDNA depletion disorders and some of the phenotypes observed in patients with these disorders include encephalopathy, respiratory chain deficiency, and neurodegenerative symptoms. MtDNA levels have also been shown to decrease with increasing age, as has mitochondrial function. The National Health and Nutrition Examination Surveys (NHANES) are large, population-based surveys conducted in the United States by the National Center for Health Statistics at the Centers for Disease Control and Prevention (CDC). Each NHANES collects demographic, lifestyle, and health data from participants using surveys, laboratory measures, and medical examinations. Additionally, DNA samples were collected for participants consenting to Genetic NHANES. Using quantitative real-time PCR
we determined the mitochondrial genome equivalents (MGEs) for a subset of Genetic NHANES samples (n=6692). Preliminary analysis revealed that MGEs ranged from 3-1019 with a mean MGE value of 250. Leveraging the rich phenotypic data contained in this survey we are following up this preliminary analysis using a phenome-wide association approach to determine factors that affect MGEs and determine if MGEs are associated with phenotypes in the NHANES dataset. Results of this study may elucidate factors that influence mtDNA maintenance as well as reveal novel phenotype associations with mtDNA levels, both of which are likely critical to uncovering the role mitochondrial dysfunction plays in the etiology of disease phenotypes.

P-304 – ABSTRACT WITHDRAWN

P-305 – ABSTRACT WITHDRAWN

P-306
CHANGES IN GENOME DIVERSITY AFTER GENOMIC AND PEDIGREE BLUP SELECTION IN LAYER CHICKEN

Marzieh Heidaritabar1, John Bastiaansen1, Addie Vereijken2, Martien Groenen1
1Wageningen University, Animal Breeding and Genomics Centre, Wageningen, Gelderland, The Netherlands, 2Hendrix Genetics, Research & Technology Centre, Wageningen, Gelderland, The Netherlands

Genome-wide changes in allele frequencies were measured in three generations of layer chickens that have been under genomic selection (GS) and pedigree BLUP selection. Genotypic data (60K SNPs) of three layer lines (A, B, and C) were used, consisting of approximately 2,200 animals per line. Within each line, three generations of candidates for GS as well as the first and third generation of BLUP selected animals were genotyped. Changes in allele frequencies were compared between GS and BLUP selection. Allele frequencies for the first and third generations were calculated. The change in allele frequencies were calculated for SNPs with MAF > 0. The drift variance of allele frequencies in the 3rd generation was calculated to identify genomic regions where the allele frequency difference exceeded the changes expected due to drift:

\[ SD_t = \sqrt{\frac{p(1-p)}{N_e t}} \]

Where \( SD_t \) is the standard deviation of the allele frequency after \( t \) generations, \( p \) and \( q \) are the initial allele frequencies of the SNP, and \( N_e \) is the effective population size. Values used for \( N_e \) and \( t \) were 60 and 2, respectively. The normalized change in allele frequencies was calculated for each SNP as:

\[ d_n = \frac{d}{SD_p} \]

Patterns of \( d_n \) were very different between GS and BLUP selection. In GS, in line A, the largest \( d_n \) was observed on chromosome 8. In line B, the largest \( d_n \) was observed on chromosomes 1 and 4 and in line C, the largest \( d_n \) was in chromosome 1. Results from BLUP selection in all lines show no region with an average \( d_n \) above \( 3SD_p \). We conclude that regional changes in allele frequencies observed under GS are large in comparison to changes under BLUP selection.
P-307
RECOMMENDATIONS FOR GENOME-WIDE SEARCH FOR EPISTATIC LOCI
Tatsiana Vaitsiakhovich1, Christine Herold2, Dmitriy Drichel1, André Lacour2, Vitalia Schüller1, Tim Becker1,2
1Institute for Medical Biometry, Informatics and Epidemiology, University of Bonn, Bonn, Germany,
2German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany

Various strategies to identify interacting SNPs in GWAS studies are evaluated in our work. Series of “realistic” disease models are defined on a 2-SNP-genotype table by specification of allele frequencies, penetrances and a minimal distance between available and causal SNPs. We compare the performance of standard single-marker analysis with multi-marker analysis. Then we investigate the relative performance of tests for interaction and tests including both marginal and interaction effects, so-called “tests allowing for interaction”. We compare case-only with case-control tests for interaction. Finally, we contrast allelic and genotypic models. A subtle problem is that tests including marginal effects may become significant because of the marginal effect of just one SNP from a pair. Since our goal is to detect both SNPs of an interaction pair, we embed tests allowing for interaction in a two-step strategy: analysis of all pairs with a test with marginal effects, followed up by an interaction test on pairs that remains significant after multiple testing adjustment. For about 5% of settings, the most efficient strategy is single-marker analysis, typically when allele frequencies are high or causal variant tagging is poor. For another class of models (5%) the most powerful strategy is testing for interaction without inclusion of marginal effects, provided that a case-only test is used. Typically, a 4 d.f. genotypic case-only test is more powerful than a 1 d.f. allelic case-only test. In the remaining majority of scenarios, a hybrid strategy is the most suitable: genome-wide interaction analysis with a combined case-only-interaction/marginal-effects test; follow-up analysis of the significant pairs with a test for interaction, excluding the strongest marginal effect but allowing for marginal effect of the second potential SNP. Allelic and genotypic tests each lead to optimal power in about 45% of scenarios. Therefore, both models are recommended to be used.

P-308
QUANTITATIVE GENETICS OF AUTOPHAGY REGULATION IN DROSOPHILA LARVAE
Izzet Mehmet Akcay1,2, Sibylle Vonesch1,2, Trudy F.C. Mackay3, Ernst Hafen1, Katja Koehler1
1Institute of Molecular Systems Biology, ETHZ, Zurich, Switzerland, 2Molecular Life Sciences Program, Life Sciences Zurich Graduate School, Zurich, Switzerland, 3Department of Genetics, North Carolina State University, Raleigh, USA

Autophagy is a stress response mechanism triggered upon cellular stresses, such as starvation, and plays a role in tissue homeostasis, cell differentiation and animal development. Although the process is highly conserved in all eukaryotes, some aspects of autophagy regulation and function significantly differ between yeast and other organisms, suggesting that there must exist several unidentified genes required for autophagy in higher eukaryotes.

In this study, we aim to find novel regulators of autophagy using a systems genetics approach in Drosophila. We will be screening 180 wild-derived fly lines for their starvation response at larval stages. Lines showing increased sensitivity or resistance will then be tested for changes in autophagic levels. By association mapping and transcript abundance analyses we hope to identify candidate genes and transcriptional networks responsible for the variation in this trait. Candidate genes can then be tested for their role in autophagy using the corresponding mutants or RNAi lines.

A prerequisite to test starvation responses in larval stages is the knowledge on the developmental timing of the different lines. For this purpose we will employ a second screen to determine the time point when the larvae reach the critical weight required to initiate pupariation. This screen will also provide data suited for association mapping of this trait and may reveal some yet unknown signalling pathways regulating Drosophila development.

Given the pleiotropic function of autophagy and the fact that some autophagic genes have subtle, but
distinct effects on animal development, association studies may prove to be an attractive approach to gain further insights into the network regulating autophagy.

P-309
DETAILS OF IMMUNE FUNCTION OF LANDRACE PIGS SELECTED FOR DECREASED MYCOPLASMA PNEUMONIA MORBID LESION

Keiichi Suzuki1, Tomoyuki Shimazu1, Jyunana Riku1, Yuuki Katayama1, Mika Ri1, Takumi Sato1, Haruki Kitazawa1, Hisashi Aso1, Kazuo Katoh1, Yoshihito Suda2, Akiko Sakuma3, Mitsuru Nakajyo2, Hiroshi Uchida2
1Graduate School of Agricultural Science, Tohoku University, Sendai city/Miyagi, Japan, 2Department of Food, Agriculture and Environment, Miyagi University, Sendai city/Miyagi, Japan, 3Miyagi Prefecture Animal Industry Experiment Station, Oosaki city/Miyagi, Japan

Objectives: Breeding for resistance to chronic diseases is one of the most important aspects of animal breeding. We have performed selection experiments for decreased morbid lesions of mycoplasmal pneumonia in Landrace pigs over the past 5 generations. To investigate the immune responses in detail, we compared the temporal changes in various immunity factors and cytokines after vaccination between the Landrace line selected for MPS resistance, and a control line.

Methods: Twelve castrated males of the selected line and 12 control Landrace pigs were used. Blood samples were collected on days -14, -7, 0, 2, 7, and 14. The MPS vaccine was injected twice intramuscularly on days -7 and 0. Blood samples were analyzed for white blood cell count, and a phagocytosis assay and cell population analysis (B, T, and myeloid cells) were performed. In addition, the mitogenic activity of blood cells was analyzed before and after the first sensitization, and cytokines (IL-10, IL-13, IL-17, IFNg, and TNFalpha) concentrations were assessed by ELISA or RIA. Cytokine mRNA expressions were analyses by real-time PCR. Statistical analyses were performed using the SAS MIXED procedure.

Results and conclusions: The MPS-selected line showed a reduction in B cell count and an increase in the myeloid cell (especially neutrophil cell) proportion after MPS vaccination, while, the control line showed contrasting results. Furthermore, the percentage of CD4+ T cells was significantly increased in the control line, even though the total T-cell ratio remained unchanged. Moreover, lymphocyte proliferation was significantly lower in the MPS-resistant line. However, IL-10 and IFNg concentrations were increased after MPS vaccine inoculation and were significantly higher in the MPS-selected line, as reflected by the mRNA levels. These results suggest that MPS-resistance is associated with immunological changes and that the severity of MPS is influenced by the host immunophenotype.

P-310
MAPPING THE GENETIC ARCHITECTURE OF MAIZE HEIGHT AND CORRELATED COMPLEX TRAITS

Jason A. Peiffer1, M. Cinta Romay2, Michael A. Gore3, Sherry Flint-Garcia3,4, Zhiwu Zhang1,2, Mark Millard1,5, Candice Gardner1,5, Michael D. McMullen3,4, James B. Holland1,6, Peter J. Bradbury1,3, Edward S. Buckler1,3
1Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY, USA, 2Institute for Genomic Diversity, Cornell University, Ithaca, NY, USA, 3U.S. Department of Agriculture (USDA)-Agricultural Research Service (USDA-ARS), IA, MO, NC, NY, USA, 4U.S. Department of Agriculture (USDA)-Agricultural Research Service (USDA-ARS), IA, MO, NC, NY, USA, 5Institute for Genomic Diversity, Cornell University, Ithaca, NY, USA, 6U.S. Department of Agriculture (USDA)-Agricultural Research Service (USDA-ARS), IA, MO, NC, NY, USA

Height is one of the most heritable and complex traits of maize. Analogous to findings in human height studies, geneticists have been baffled by this apparent incongruity. How is a trait so heritable, explained so well by parent-offspring regression, nearly irreducibly complex in natural populations? By phenotyping a nested association mapping panel (NAM) of 5,000 recombinant inbred lines and a diverse panel of 2,711 inbreds (AMES) amassed from germplasm resources around the world, we harvested information on
plant height, node counts, and flowering time measurements from the inbreds across 3-12 environments. These populations were genotyped for over 25 million and 680,000 SNPs, respectively. Phenotypic variation was partitioned into components of genetic, environmental, and environmentally conditional genetic variance. Model averaging and ridge regression were used to estimate the distribution of additive genetic variation across the genome and determine the proportion of heritability captured by the SNPs. Over 70% of the heritable variation in the height of these populations was captured by the additive models. Results from the genotype-phenotype maps suggested substantial modularity in the pleiotropy of maize height and flowering genetic architectures as evidenced by nearly nominal correlation of allelic effect estimates. The impact of this modularity and cost of complexity in selection on these traits are discussed. Given rapid developments in high throughput genotyping and phenotyping techniques, a finely resolved empirical understanding of the geometry of genetic covariation is becoming increasingly feasible on a genomic scale. This knowledge will facilitate prediction of the evolvability of these complex genetic systems and foster the wisdom necessary to further optimize crop improvement.

P-311
POPULATION ISOLATES FROM GREECE OFFER POTENTIAL FOR POWERFUL DISEASE GENE MAPPING: THE HELIC-POMAK AND MANOLIS STUDIES

Lorraine Southam1, Aliki-Eleni Farmaki2, Ioanna Tachmazidou1, Kimmo Palin1, William Rayner3,4, Erica Daoutidou2, Ioanna Ntalla2, Kalliope Panoutsopoulou2, George Dedoussis2, Eleftheria Zeggini1
1Wellcome Trust Sanger Institute, Cambridge, UK, 2Harokopio University, Athens, Greece, 3Wellcome Trust Centre for Human Genetics, Oxford, UK, 4Oxford Centre for Diabetes, Endocrinology and Metabolism, Oxford, UK

The study of low-frequency and rare variants can be empowered by focusing on isolated populations, in which rare variants may have increased in frequency and linkage disequilibrium tends to be extended. Sequencing is efficient in isolates, because variants are shared in extended haplotype contexts, supporting accurate imputation. Here we assess sample sets collected from two Greek populations: the Pomak villages are a set of religiously-isolated mountainous villages in the North (population size 11,000); Anogia is a mountainous village on Crete, with high levels of longevity (population size 4,000). 747 and 1118 individuals respectively were typed on the Illumina OmniExpress platform. We calculated genome-wide IBS statistics to assess the degree of relatedness and compared it with the general Greek population (707 samples with OmniExpress data, TEENAGE study). We additionally calculated the proportion of individuals with at least one “surrogate parent” as a means for accurate long-range haplotype phasing and imputation, as proposed by Kong et al, Nature Genetics 2008. We find 1-1.4% of individual pairs with pi-hat>0.05, and ~0.4% with pi-hat>0.1 in the isolates compared to 0% in the general Greek population. We also find that ~80%-82% of subjects have at least one surrogate parent in the isolates, compared to ~1% in the outbred Greek population. We have established the HELIC-Pomak and MANOLIS cohorts as genetic isolates and are currently whole-genome sequencing 250 individuals to enable imputation and subsequent association testing. This approach has the potential to identify novel robust associations with disease-related complex traits.

P-312
LARGE-SCALE STUDY OF THE GENETIC ARCHITECTURE OF COMPLEX TRAITS IN THE RAT

Amelie Baud, Jonathan Flint, Richard Mott
University of Oxford, Oxford, UK

The rat is an important model for many human diseases, particularly where its physiology more closely resembles that of humans than does the mouse (e.g. breast cancer, learning and memory, cardiac function and hypertension). As part of the EURATRANS Consortium, we mapped the quantitative trait loci
contributing to variation in more than 200 measures relevant to a variety of diseases (including anxiety, type 2 diabetes, and multiple sclerosis) and basal physiology, in 2000 heterogeneous stock rats. The rat heterogeneous stock is descended from 8 inbred strains through more than 50 generations of circular breeding. The limited extent of linkage disequilibrium in the population allows high-resolution mapping (95% confidence intervals around 3Mb on average). Because the progenitor strains are known and have been sequenced, we were able to impute the genotypes at all the segregating variants, and test whether each variant could account for the QTL. Whenever that was the case, our analysis significantly reduced the number of candidate genes. We identified 1292 significant QTLs in total, and show that they explain on average 40% of the heritable phenotypic variation. We attribute this relative success compared to human studies to limited allelic variation, and the contribution of common alleles only to phenotypic variation. Finally, we compare the extent to which the same genes and/or the same biological pathways are responsible for phenotypic variation in the rat heterogeneous stock and in a mouse heterogeneous stock in which similar phenotypes have been collected and previously mapped.

P-313
UNTESTED HYBRID PREDICTION USING BOTH INBRED AND HYBRID PERFORMANCE RECORDS

Fabiano Pita, Gregory Gingera, Van Ripley, Carol Powers
Dow AgroSciences, Indianapolis, Indiana, USA

Hybrid systems have been an important tool to explore heterosis in crops through optimal combination of inbred lines. Performance per se of parental inbred lines has been traditionally evaluated with the main goal to minimize hybrid's cost of production. The objective of this work was to evaluate if performance per se of parental inbred lines can be combined with tested hybrid performance to increase accuracy of untested hybrid prediction. A total of 43,446 canola field yield records from inbreds and hybrids across years and locations were used. Yield from inbreds and from tested hybrids was modeled using a mixture mixed model. A covariance structure with a numerator relationship matrix was used to model additive genetic effect for male and female effects (when modeling hybrid records) and inbred effect (when modeling inbred records). A covariance structure among these effects was fitted through a direct product. Dominance relationship matrix was used to model dominance effect for hybrid records. Fixed effect included in the model was a combination of year and location of planting. Heritability estimates were 8.96% for male/female parent and 12.0% for inbred per se effect. Genetic correlation between these effects was 0.37. Dominance variation accounted for 5.8% of total variation. Records from 10% most extensively tested hybrids (165 hybrids) were removed from training dataset and predictions were performed using only inbred per se and hybrid records from remaining hybrids. Correlation between corrected mean from removed hybrids and their predictions was 0.56 when including or not inbred per se information in the prediction model. Low genetic correlation between inbred per se and hybrid performance suggests no benefit on using inbred per se records to predict untested hybrid performance. Also, distinct genes may be responsible to control inbred per se performance and parent’s contribution to hybrid performance.
P-314
SIGNATURE OF FORTY YEARS OF ARTIFICIAL SELECTION IN U.S. HOLSTEIN CATTLE IDENTIFIED BY LONG-RANGE FREQUENCY ANALYSES

Tad Sonstegard1, Shengwen Wang2, Curtis Van Tassell1,2, Li Ma2, George Wiggans3, Brian Crooker2, John Cole3, Brian Mariani4, Eui-Soo Kim1, John Garbe2, Scott Fahrenkrug2, George Liu1, Yang Da2

1Bovine Functional Genomics Laboratory, USDA/ARS, Beltsville, MD 20705, USA, 2Department of Animal Science, University of Minnesota, Saint Paul, MN 55108, USA, 3Animal Improvement Programs Laboratory, USDA/ARS, Beltsville, MD 20705, USA, 4Genetics and In Vitro Fertilization Institute, Fairfax, VA 22031, USA

Three groups of U.S. Holstein cattle were analyzed for selection signature of artificial selection since 1964 using long-range frequency measures. The three groups included Holsteins unselected since 1964, contemporary Holsteins, and an elite line of contemporary Holsteins. Long-range frequencies included allele frequency and heterozygosity differences in 1-8Mb sliding windows of SNP markers of the bovine 50K SNP chip between the unselected group and the two contemporary groups subjected to 40 years of artificial selection. The contemporary Holsteins had significantly lower long-range heterozygosity than the unselected group near a large cluster of 90 RNA genes on Chr21. Some of the microRNA genes in that RNA gene cluster had strong interactions (MFE < -30) with the mRNA of several genes in the telomere region of Chr21 that affected protein yield. The most significant RNA interaction was between microRNA bta-mir-370 and the mRNA of TRM61 (MFE = -43.2). Both the centromere and telomere regions of Chr21 had large long-range allele frequency differences between the selected and unselected groups. Allele frequency differences between the selected and unselected groups identified selection signature near PLGF on Chr10 and in the PRKCA-CACNG5-HELZ region of Chr19. Allele frequency and heterozygosity differences between the elite group, the unselected group and the contemporary Holsteins excluding the elite group identified signature of the elite Holsteins on Chr18 and Chr13. The Chr18 region with elite signature coincided with a 15 Mb region that had a large number of SNP effects on dairy functional traits. The elite signature on Chr13 involved two regions. The first region had elevated allele frequency differences between the elite and unselected groups and was previously reported to have highly significant SNP effect for milk yield. The second region was the 25-30 Mb region with striking decreased long-range heterozygosity in the elite group compared to the unselected group.

P-315
IMPROVED MAPPING OF A BODY WEIGHT QTL ON MOUSE CHROMOSOME 2 DISCOVERED FROM WILD MUS MUSCULUS CASTANEUS

Akira Ishikawa, Sin-ichiro Okuno
Nagoya University, Nagoya, Aichi, Japan

We use body weight and its related traits in mice as model quantitative traits in humans and livestock to study their genetic architecture. We previously revealed many QTLs affecting postnatal body weight and growth by genome-wide QTL analysis in an intersubspecific backcross population between C57BL/6J (B6) and wild Mus musculus castaneus mice capture in the Philippines. Subsequently we developed the B6.Cg-Pbwg1 congenic strain on the B6 genetic background that carries Pbwg1, a prominent growth QTL on a proximal region of mouse chromosome 2, and we developed approximately 20 subcongenic strains derived from B6.Cg-Pbwg1. Using some of the congenic and subcongenic strains developed, several closely linked QTLs affecting body weight and body composition traits were revealed within the 44.1-Mb Pbwg1 congenic region. Among the linked QTLs, Pbwg1.10 has an overdominant effect on body weight causing heterosis. Within a region overlapping with Pbwg1.10, a unique QTL, Pbwg1.12, is located. Its wild-derived allele surprisingly increases body weight despite wild mice having approximately 60% of body weight of B6. In this study, we generated F2 segregating populations between B6 and each of three subcongenic strains in order to map these two QTLs more finely. Phenotypic analyses of the F2
segregating populations suggested that at least one body weight QTL may be located within a maximum region of 5.9 Mb. Exome sequencing revealed many synonymous and nonsynonymous substitutions for 23 genes residing on this region.

P-316
A SYSTEMS GENETICS APPROACH TO UNDERSTANDING INTRA-SPECIES VARIATION IN DROSOPHILA WING SIZE

Sibylle Vonesch1, David Lamparter2, Vasco Medici3, Francesco Crivelli4, Sven Bergmann2, Ernst Hafen1
1Institute of Molecular Systems Biology, ETHZ, Zurich, Switzerland, 2Département de génétique médicale, UNIL, Lausanne, Switzerland, 3Institute of Neuroinformatics, ETHZ and UZH, Zurich, Switzerland, 4Sensory Motor Systems Lab, ETHZ, Zurich, Switzerland

The genetic networks underlying quantitative traits such as height, weight or many common complex diseases, and the mechanisms generating natural variation in these traits are still incompletely understood. The advent of fast and affordable whole genome sequencing combined with the increasing number of genome wide association studies (GWAS) in various organisms is a promising approach to unravel the architecture of such molecular networks.

Environmental effects are the major confounding factor when dealing with quantitative traits. We have developed a standardized protocol for efficient reduction of environmental fluctuations in Drosophila, thus making it a highly suitable organism for studies trying to link genetic and phenotypic variation. We are currently characterizing 192 naturally-derived, inbred Drosophila lines, the DGRP, for wing and body size traits for association mapping in order to identify novel genes affecting organ size and shape and unveil the underlying networks. First results of the association will be presented. This will help us to better understand Drosophila wing development at the systems level.

As a complementary approach, we run an artificial selection experiment to generate extreme wing phenotypes. An automated measurement and selection tool, the FlyCatwalk, is designed for this purpose. This experiment will reveal different strategies of modulating networks to generate a certain phenotype. This combined knowledge will give us valuable insights into the endophenotype, the genetic, transcriptional and metabolic networks underlying genetic variation in natural populations.

P-317 – ABSTRACT WITHDRAWN

P-318
GENETIC ANALYSIS OF CLCuV DISEASE, EARLINESS, YIELD AND FIBER TRAITS IN UPLAND COTTON

Naqib Ullah Khan
Khyber Pakhtunkhwa Agricultural University, Peshawar, KPK, Pakistan

Genetic effects of Cotton Leaf Curl Virus (CLCuV), earliness and yield and fiber traits were studied in a 6 x 6 F1 and F2 diallel hybrids in upland cotton. An additive-dominance model was adequate for most of the traits except bolls per plant and seed cotton yield, where the model was partially adequate. Genetic parameters were estimated following Hayman’s and Mather’s model (D and H genetic components of variance). Additive effects controlled CLCuV resistant and diseased plant %, lint index and lint % in both generations, and earliness in F2 generation. Non-additive inheritance with over-dominance controlled bolls per plant and seed cotton yield in both generations and earliness in F1. Most of the traits presented an unequal proportion of positive (U) and negative (V) alleles in the loci (H2/H1 < 0.25 and F different to zero). The H2/H1 ratios were lower than maximum value (0.25) for most of the traits, which arises when U = V = 0.5 over all loci. Dominance effects (h2) for most of the traits suggested that substantial contribution of dominance was not due to
heterogeneity of loci in these parameters. Broad and narrow sense heritabilities were high for most of the traits. Correlation coefficient between Wr + Vr and mid parental (y) indicated that dominant genes were responsible for increased CLCuV resistant plants %, lint index and lint %, while recessive genes increased the earliness and CLCuV diseased plant % in both generations. However, recessive and dominant genes governed bolls and seed cotton yield per plant in F1 and F2 hybrids, respectively. Genetic gain was encouraging for most traits. Dominant genes contained by cultivar CIM-1100 were reliable for increased CLCuV resistance (due to monogenic dominant nature of viral disease) and was identified as promising parental cultivar to cross combinations.

P-319
BAYESIAN ADAPTIVE MARKOV CHAIN MONTE CARLO ESTIMATION OF GENETIC PARAMETERS

Boby Mathew1, Andrea Bauer1, Petri Koistinen2, Tobias Reetz1, Mikko Sillanpää3, Jens Léon1
1University of Bonn, Bonn, Germany, 2University of Helsinki, Helsinki, Finland, 3University of Oulu, Oulu, Finland

Accurate and fast estimation of genetic parameters underlying quantitative traits using mixed linear models with additive and dominance effects is of great importance in both natural and breeding populations. Here we propose a new fast adaptive Markov Chain Monte Carlo (MCMC) sampling algorithm for the estimation of genetic parameters in the linear mixed model with several random effects. In the learning phase of our algorithm, we use the hybrid Gibbs sampler to learn the covariance structure of the variance components. In the second phase of the algorithm, we use this covariance structure to formulate an effective proposal distribution for a Metropolis–Hastings algorithm, which uses a likelihood function where the random effects have been integrated out. Compared to the hybrid Gibbs sampler, the new algorithm showed better mixing properties and was about twice as fast to run. Our new algorithm was able to detect different modes in the posterior distribution. Also the posterior mode estimates from the adaptive MCMC method were near to the REML (residual maximum likelihood) estimates. Moreover, our exponential prior for inverse variance components was vague and enabled the estimated mode of the posterior variance to be practically zero in agreement with the support from the likelihood (in case of no dominance). The performance of the method is illustrated with simulated data sets and field data in barley.

P-320
CONSTRUCTING GENE NETWORKS UNDERLYING FAT METABOLISM IN PIGS

Nitdia Aznárez1, Jules Hernández1, Ángela Cánovas1, Ramona Natacha Pena1, Marcel Amills2, Raquel Quintanilla1, Anna Mercadé2, Arianna Manunza2
1Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Lleida, Spain, 2Universitat Autònoma de Barcelona (UAB), Bellaterra, Spain

Lipid metabolism in pigs represents a complex system gathering traits related to animal health, carcass performance, and meat quality. In this study, phenotype and gene networks underlying fat metabolism were inferred from global liver expression (GeneChip Porcine Genome arrays, Affymetrix) and high-density SNP (Illumina PorcineSNP60 BeadChip) data of 104 and 350 Duroc pigs, respectively. Ten fat-related traits were measured: serum lipid levels (cholesterol, LDL, HDL and triglycerides), fatness (backfat thickness and lean percentage), intramuscular fat content, and fatty acid composition (SFA, MUFA and PUFA). Two phenotype networks were constructed on the basis of these phenotypes’ associations with either transcriptomic or genomic data. Phenotype network inferred from liver expression was notably denser and showed higher correlation values between traits than phenotype network built from SNP effects. Associating phenotypes with expression levels in the liver, a very lipogenic tissue, could provide a more thoughtful description of physiological relationships between these traits. Besides, only probes and SNPs showing significant associations with two or more traits were retained for the construction
of gene networks: 884 probes, corresponding to 661 genes after annotation, and 867 SNPs that were annotated as the closest gene. Scarce concordances between genes involved in the SNP and the liver expression gene networks were observed. Aiming at disentangling regulation mechanisms underlying lipid metabolism, we performed a functional annotation of genes involved in the liver transcriptomic gene network. Gene Ontology (GO) and subsequent Cluster Analyses allowed identifying a set of 33 genes that enriched a specific cluster grouping several lipid metabolism GO categories including lipid and fatty acid biosynthetic and metabolic processes. Results from co-expression and eQTL studies, currently in progress, will allow going far beyond by connecting genomic and transcriptomic gene networks, as well as by detecting master genes modulating the gene network topology.

P-321
GWAS META-ANALYSIS OF SERUM URATE CONCENTRATIONS AND GOUT IN DIVERSE POPULATIONS

Anna Kottgen1,10, Eva Albrecht2, Alexander Teumer3, Veronique Vitart4, Jan Krumsiek2, Conall O’Seaghdha5, Marina Ciullo6, Caroline Fox7, Mark Caulfield8, Murielle Bochud9, Christian Gieger1

1Freiburg University Hospital, Freiburg, Germany, 2Helmholz Zentrum, Munich, Germany, 3University of Greifswald, Greifswald, Germany, 4MRC HGU,ICMM,University of Edinburgh, Edinburgh, UK, 5Massachusetts General Hospital, Boston, USA, 6CNR IGB, Napoli, Italy, 7Havard Medical School, Boston, USA, 8Queen Mary University of London, London, UK, 9University of Lausane, Lausane, Switzerland, 10John Hopkins Bloomberg School of Public Health, Baltimore, USA

Uric acid is the, biochemically active, final breakdown product of purine oxidation in humans. Hyperuricemia, elevated levels of serum urate, can cause gout, the most prevalent inflammatory arthritis in developed countries. Furthermore, increased levels of serum urate are associated with components of the metabolic syndrome, such as obesity, high blood pressure and insulin resistance, type 2 diabetes and with cardiovascular disease. The heritability of serum urate concentrations is estimated at 50%–70%.

Eleven loci identified by previous genome-wide association studies (GWAS) explain 5-6% of serum urate variance, suggesting that additional loci remain to be identified. We performed a meta-analysis of GWAS on serum urate concentrations among 48 studies with more than 140,000 participants of European ancestry, and on gout among >70,000 individuals (2,125 cases) in the Global Urate Genetics Consortium (GUGC). Secondary analyses included, among others, stratification by sex, pathway analyses, and look-ups of the associated loci in individuals of other ancestries as well as with urate-correlated traits. Replication was performed in 32,813 independent samples.

Altogether, we identified and replicated 28 genome-wide significant SNPs associated with serum urate concentrations, including 18 new loci. Nominal association with gout and fractional excretion of urate was found for 17 and 10 of the urate GWAS SNPs respectively, with consistent directions of effect. Effect sizes on serum urate levels were similar among individuals of Indian ancestry, African Americans and Japanese individuals. The genes implicated highlight new biological pathways offering novel avenues into the treatment and prevention of gout.

P-322
IDENTIFYING SUSCEPTIBILITY LOCI ASSOCIATED WITH HODGKIN’S LYMPHOMA

Miguel Inacio da Silva Filho1, Hauke Thomsen1, Andreas Engert3, Michael Fuchs3, Sabine Ponader3, Elke Pogge von Strandmann1, Asta Försti1, Kari Hemminki1,2

1German Cancer Research Center (DKFZ), Heidelberg, Germany, 2Lund University, Malmö, Sweden, 3University of Cologne, Cologne, Germany

Hodgkin’s lymphoma, a lymphoid carcinoma, is one of the most common tumors in young adults. A genome-wide association study was conducted to identify genetic variants associated with this disease.
The study comprised a total of 2,227 German individuals, 1,001 individuals with Hodgkin’s lymphoma and 1,226 controls. Individuals were genotyped using an Illumina Human Omni Express 12v1 chip (733,202 markers). Data have undergone a stringent quality control including checks for sex, minor allele frequency, individual and marker call rates, heterozygosity, Hardy-Weinberg equilibrium, differential missing rates between cases and controls, relatedness and population substructure. Association tests were carried out by logistic regression analysis assuming an additive model and using the two first principal components as covariates. Two previously identified single nucleotide polymorphisms (SNPs) reached a suggestive significance level in the present study (rs6903608, OR = 1.38, P = 1.24 x 10^{-06} and rs2395185, OR = 0.71, P = 1.05 x 10^{-05}). Both of these SNPs are located at the major histocompatibility complex (MHC) region. In fact, most of the identified susceptibility loci for Hodgkin’s lymphoma to date are in this region, which highlights the importance of the MHC in the disease etiology. Evidence for association of further SNPs in the MHC region and in other regions of the genome were found in the present study. These SNPs are currently being replicated in an independent study population to confirm the findings. The data add important information to the comprehension of the susceptibility of Hodgkin’s lymphoma.

P-323
GENETIC VARIABILITY OF CARCASS AND MEAT QUALITY OF THE TEXEL BREED UNDER GRAZING CONDITIONS
G. Ciappeson1, E.A. Navajas1, R San Julián2, G. Brito2, D. Gimeno3, V. Goldberg1
1INIA, Las Brujas, Canelones, Uruguay, 2INIA, Tacuarembó, Tacuarembó, Uruguay, 3SUL, Montevideo, Montevideo, Uruguay

Due to the relevance of carcass and meat quality traits as breeding objectives of Texel as sire terminal breed, a progeny testing was implemented. In general, these difficult-to-measure traits have received less attention despite their economic relevance, particularly in more extensive environments in which sheep production is mainly located. The aim of this study is to obtain the first preliminary estimations of genetic parameters for carcass quality (CQ) traits of this breed raised on grazing conditions.

Data were recorded on 276 female and male lambs slaughtered in 2009-10-11 with an average of 40kg liveweight and 3.5 of body condition. Hot carcass (HCW), French Rack, Shoulder and Leg weights and GR (indicator of carcass fatness) were measured. Intramuscular fat (IMF%) was assessed at the longissimus dorsi muscle by chemical extraction. Heritabilities (h^2) were estimated by Gibb sampling. The animal model included year-flock, birth type, sex, dam age and age at slaughter (covariate) as fixed effects. Pedigree data comprised 670 animals including 13 sires and 221 dams.

Estimates were of moderate to high magnitude (Table 1), in agreement with other studies. Although results confirm that there is scope for genetic improvement, these preliminary values should be interpreted with caution because of the low number of records and shallow pedigree available. This database is part of the sheep reference populations for genomic studies for complex traits in Uruguay.

Table 1. Estimated statistics of marginal posterior distributions of h^2 estimates

<table>
<thead>
<tr>
<th>Trait</th>
<th>Mean</th>
<th>Median</th>
<th>PSD</th>
<th>95%HPD_L</th>
<th>95%HPD_U</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCW</td>
<td>0.510</td>
<td>0.497</td>
<td>0.205</td>
<td>0.150</td>
<td>0.931</td>
</tr>
<tr>
<td>GR</td>
<td>0.293</td>
<td>0.268</td>
<td>0.173</td>
<td>0.007</td>
<td>0.624</td>
</tr>
<tr>
<td>Rack</td>
<td>0.834</td>
<td>0.866</td>
<td>0.134</td>
<td>0.568</td>
<td>0.999</td>
</tr>
<tr>
<td>Shoulder</td>
<td>0.392</td>
<td>0.364</td>
<td>0.214</td>
<td>0.008</td>
<td>0.800</td>
</tr>
<tr>
<td>Leg</td>
<td>0.334</td>
<td>0.307</td>
<td>0.201</td>
<td>0.000</td>
<td>0.710</td>
</tr>
<tr>
<td>IMF%</td>
<td>0.274</td>
<td>0.240</td>
<td>0.170</td>
<td>0.007</td>
<td>0.606</td>
</tr>
</tbody>
</table>

PSD: posterior standard deviation; 95%HPD: 95% highest posterior density interval Lower (L) -Upper (U) bound
P-324
GENETIC PARAMETERS OF ORGANS TRAITS AND CARCASS WEIGHT OF A PATERNAL BROILER LINE

Guilherme Costa Venturini, Salvador Boccaletti Ramos, Valdecy Aparecida Rocha da Cruz, Thiago Bruno Ribeiro da Silva, Mônica Correa Leduc, Jane de Oliveira Peixoto, Danísio Prado Munari

1FCAV/UNESP, Jaboticabal, SP, Brazil, 2EMBRAPA Suinos e aves, Concórdia, SC, Brazil

Carcass weight (CW) is a component of body weight at 42 days of age (BW). This part-whole relationship explains the genetic association between these traits and suggests that indirect selection for BW can favor CW. Genetic correlations between CW and organs weights can contribute to the understanding of the incidence of metabolic disorders by indirect selection. The aim of this study was to estimate genetic parameters of CW, heart weight (HW), liver weight (LW) and gizzard weight (GW). Data of 1,462 chickens from a paternal broiler line developed by Embrapa Swine and Poultry, Brazil, were used in the analysis. The Restricted Maximum Likelihood Method in a multi-trait animal model that included fixed effects of sex and hatch and the additive genetic and residual random effects was used for the genetic parameters estimate. The heritability estimates varied from 0.21 ± 0.05 (HW) to 0.42 ± 0.07 (GW). All traits could respond to selection. Genetic correlations were 0.26 ± 0.15, 0.37 ± 0.16 and 0.53 ± 0.12 between CW x GW, CW x HW and CW x LW, respectively. The genetics correlations estimates of 0.42 ± 0.14, 0.46 ± 0.15 and 0.55 ± 0.14 were between LW x GW, HW x GW and HW x LW, respectively. The genetic correlations were moderate and positive. The selection process aiming the increase in carcass weight may also increase organs weight, but not in the same proportion. This may contribute to the incidence of metabolic disorders, but there are other causes besides the weight organs increase.

P-325
ENDOPHENOTYPE AND POLYGENIC APPROACH STUDY REVEALS ASSOCIATION BETWEEN NEUROCOGNITIVE GENE SETS AND PSYCHIATRIC DISORDERS

Carla P D Fernandes, Andrea Christoforou, Thomas Espeseth, Sudheer Giddaluru, Thomas Werge, Sven Cichon, Ole A. Andreassen, Vidar M. Steen, Stephanie Le Hellard

1Dr Einar Martens’ Research Group for Biological Psychiatry, Department of Clinical Medicine, University of Bergen, Bergen, Norway, 2Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Helse Bergen HF, Bergen, Norway, 3Department of Psychology, Center for the Study of Human Cognition, University of Oslo, Oslo, Norway, 4Centre for Advanced Study, Oslo, Norway, 5Research Institute of Biological Psychiatry, Copenhagen University Hospital, Roskilde, Denmark, 6Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany, 7Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, Germany, 8Institute of Clinical Medicine, University of Oslo, Oslo, Norway, 9Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway

Psychiatric disorders, such schizophrenia (SCZ) and bipolar disorder (BD), are on the leading edge of the main causes of disability worldwide. Since such patients present great cognitive impairments, it has been suggested that cognitive dysfunction could serve as endophenotype. It has been shown that the manifestation of complex genetic disorders, such as SCZ and BD, may be explained by the small effects of multiple genetic variants, including not only the GWAS top hits but also the weaker association signals that do not reach significance (Valdar, et al. 2006; Lee et al., 2012). In order to validate cognitive dysfunctions as genetic endophenotypes for psychiatric disorders we generated gene sets associated to different neurocognitive domains (general cognition, learning, memory, speed of processing and attention) from the GWAS of a healthy Norwegian sample phenotyped for cognitive abilities. The different gene sets were built after assigning SNPs to genes using the LDsnpR package (Christoforou et al., 2012). The candidate gene sets were then tested against 3 SCZ and 3 BD independent GWASs for gene set enrichment analyses (GSEA). This method, developed for gene expression studies, assesses the enrichment of signal of a priori given gene set in a ranked list of genes (Mootha et al., 2003; Subramanian et al.,
2005). We tested our candidate gene sets against the 6 gene binned GWAS for psychiatric disorders which we ordered by their minimal P-value score (modified with an adjusted Sidak correction). We observed significant enrichment for several neurocognitive gene sets and most interestingly one set replicated in the 3 SCZ datasets and another one in the 3 BD samples.

Our work suggests that the combination of endophenotype and polygenic approach might be important in overcoming the difficulties in psychiatric genetics and may be useful when exploring GWAS outcomes.

P-326

DISENTANGLING THE COMPLEXITY OF INFECTIOUS DISEASE: INSIGHTS FROM QUANTITATIVE GENETIC AND QTL MAPPING STUDIES ON THE OUTCOME OF BACTERIAL INFECTION IN DAPHNIA MAGNA

Matthew Hall, Jarkko Routtu, Dieter Ebert
University of Basel, Basel, Switzerland

Individuals naturally vary in the severity of infectious disease when exposed to a parasite. Dissecting this variation into genetic and environmental components can reveal whether or not this variation depends on the host genotype, parasite genotype or a range of environmental conditions. Complicating this task, however, is that the symptoms of disease result from the combined effect of a series of events, from the initial encounter between a host and parasite, through to the activation of the host immune system and the exploitation of host resources. Here, I report on a series of experiments involving the crustacean Daphnia magna and its parasite Pasteuria ramosa, which attempt to dissect the environmental and genetic basis to a host's ability to fight against infectious disease. First, utilising classical quantitative genetic approaches, I show how parasite resistance consists of two components which vary in environmental sensitivity: one defined by host-parasite compatibility, the other involving the ability of the host the clear the parasite once infection has occurred. Second, based on fine-scale QTL mapping, I build on previous work which indicates that host-parasite compatibility is influenced by a single genomic region and instead reveal how the subsequent ability of the host to clear infection or minimise the severity of disease has a far more complex genetic basis. This work highlights how parasite resistance based on host-parasite genetic compatibility versus the within-host defence cascade have a vastly different genetic basis, as well as how disentangling genetic and environmental factors at different stages of infection improves our understanding of the processes influencing the prevalence and severity of infectious disease.

P-327

FACTOR ANALYSIS OF BREEDING VALUES OF FORTNIGHTLY MILK PRODUCTIONS FOR FIRST AND SECOND LACTATIONS OF HOLSTEIN COWS

Rodrigo Savegnago1,3, Guilherme Nascimento1,4, Sabrina Caetano1,5, Salvador Ramos1,4, Beatriz Nunes1,4, Lenira El Faro2,6, Danisio Munari1,6
1FCAV/UNESP, Jaboticabal, São Paulo, Brazil, 2APTA, Ribeirão Preto, São Paulo, Brazil, 3FAPESP fellowship, São Paulo, Brazil, 4CAPES fellowship, Jaboticabal, São Paulo, Brazil, 5UNESP fellowship, São Paulo, Brazil, 6CNPq fellowship, Brasília, DF, Brazil

The objective of this study was to evaluate the relationship between breeding values of fortnightly milk productions for first and second lactations of Holstein cows. The breeding values were obtained by the Restricted Maximum Likelihood method by single-trait random regression models. The model included the fixed effect of contemporary group (P < 0.01), the covariate age at calving cow (linear and quadratic effect), the lactation’s day (population average curve) fitted using Legendre polynomials of fourth order, the random effect of additive genetic (third order), permanent environment (fifth order) and residual (four classes). Multivariate factor analysis was used to explore the genetic association between the breeding values. It was used 1,709 animals under second lactation because they presented both the records of
fortnightly milk productions in the first and second lactations. Five factors were chosen according to Kaiser criterion (eigenvalues greater than or equal to one) and explained 98.93% of the total additive genetic variation. The factor analysis simplified the complex structure contained in the breeding values database. The linear correlations between the breeding values with each factor revealed that Factors 1, 2, 3, 4 and 5 were more related to breeding values of milk production of two-thirds of the first lactation, half to the end of second lactation, the first half of the second lactation, the last third of the first lactation and the first month of the first lactation, respectively. The results indicated that there was no genetic association between the controls of the first with the second lactation. The breeding values for milk production, related with Factors 1, 4 and 5 were independent of the breeding values for milk production related with Factors 2 and 3, because the factors are orthogonal, i.e. independent.

Financial support: FAPESP, CAPES, UNESP and CNPq

P-328

CHROMOSOME-SPECIFIC ANALYSES OF SNP CHIP DATA FROM A SOUTH AFRICAN MERINO FLOCK DIVERGENTLY SELECTED FOR REPRODUCTIVE POTENTIAL

Lise Sandenbergh1, Schalk Cloete2-3, Rouvay Roodt-Wilding1, Aletta Van der Merwe1

1Department of Genetics, Stellenbosch University, Stellenbosch, South Africa, 2Department of Animal Science, Stellenbosch University, Stellenbosch, South Africa, 3Institute for Animal Production: Elsenburg, Western Cape, South Africa

A South African Merino sheep flock has been divergently selected for their ability to rear multiple offspring. Selection has been applied for more than 8 generations and has resulted in a High line and a Low line that differ markedly in their reproductive output. The causative mutation/s or quantitative trait loci responsible for the difference in reproductive traits between these two lines have not been determined. Several studies have reported mutations in a single ovine gene or closely linked group of genes that result in highly proliferative lines in other breeds. Although lamb rearing ability is assumed to be a complex trait, the chromosomal regions of previously identified mutations could serve as initial candidates. In individuals from divergently selected lines, genes under selection would demonstrate the most genetic differentiation between the lines and would also exhibit a higher degree of homozygosity within lines than expected. In this study, individuals from the divergently selected flock have been genotyped using the OvineSNP50 BeadChip. Sampled individuals were representative of the current genetic composition of the lines and had fairly accurate estimated breeding values for number of lambs weaned and total weight weaned per parity. SNP loci that did not meet quality control thresholds were excluded from the data set. Several loci deviated significantly from Hardy-Weinberg equilibrium, as would be expected in a selected population. Chromosome-specific data subsets were generated and analysed for regions indicated in literature to contain genes involved in sheep reproduction. Loci that differ between the two lines could be identified, indicating that the effect of selection on reproductive traits can be detected on an individual chromosome level. Further investigation into the SNP loci exhibiting the highest degree of divergence between the two lines could lead to the discovery of genes or regions involved in the ability to rear multiple offspring.
P-329
GENETIC PARAMETERS FOR BREECH STRIKE, BODY STRIKE AND BREECH STRIKE INDICATOR TRAITS IN SHEEP

Ansie Scholtz¹, Schalk Cloete¹,², Japie van Wyk³
¹Institute of Animal Production, Elsenburg, Elsenburg, South Africa, ²Department of Animal Sciences, Matieland, South Africa, ³Department of Animal, Wildlife and Grassland Sciences, Bloemfontein, South Africa

The sheep blowfly is an economically important ectoparasite of sheep, and impact on animal health and welfare under pastoral conditions. The absence of flystrike in the breech (ABrS) and on the body (ABoS) was recorded in 2 198 Merino hoggets on the Tygerhoek Research Farm. These traits were analysed together with dag score (DS; n = 1 623) and breech wrinkle score (BWS; n = 2 162) in a four-trait linear-threshold model, using Bayesian inference and Gibbs sampling. The latter traits were scored on respectively a 5-point scale (DS) and a 6-point scale (BWS). Both these traits are considered as indicator traits for ABrS during selection. The incidence of ABoS was very high at 0.986, while the incidence of ABrS was somewhat lower at 0.893. Estimates of direct additive heritability (SE in brackets) amounted to 0.67 (0.40) for ABoS, 0.29 (0.11) for ABrS, 0.27 (0.08) for DS and 0.30 (0.04) for BWS. Direct heritability estimates were significant, except for ABoS, where the very high frequency probably hampered the estimation of genetic variation on the underlying liability scale. Genetic correlations of ABoS and ABrS with DS were negative, amounting to respectively -0.78 (0.32) and -0.19 (0.28). Corresponding correlations with BWS amounted to respectively 0.06 (0.20) and -0.76 (0.19). The genetic correlation between ABoS and ABrS was relatively low and not significant at 0.22 (0.34). DS and BWS were unrelated genetically, at 0.06 (0.15). While it needs to be conceded that the high frequency of ABoS probably compromised the analysis, it is evident that selection for a reduction in BWS will benefit the ABrS.

Further research on the incorporation of direct and indirect selection for ABrS as part of integrated blowfly management is needed in the South African sheep industry, to ensure that the problem is dealt with in a sustainable manner.

P-330
GENETIC ANALYSIS OF CHIOS SHEEP: IDENTIFICATION OF GENETIC VARIANTS AND INVESTIGATION OF THEIR ASSOCIATION WITH MILK TRAITS

Georgia Hadjipavlou¹, Michael Orford², Ouranios Tzamaloukas², Demetrios Chatziplis³, Alkis Koumas¹, Andreas Mavrogenis², Christos Papachristofooru², Despoina Miltiadou³
¹Agricultural Research Institute, Lefkosia, Cyprus, ²Cyprus University of Technology, Lemesos, Cyprus, ³Alexander Technological Educational Institute of Thessaloniki, Thessaloniki, Greece

Dissecting the genetic basis of milk production traits would be of great benefit to the small ruminant dairy industry. The objective of this work was to identify genetic variants at loci that have been mainly mapped in dairy cattle, and investigate their effects on milk traits of Cyprus Chios sheep. For this purpose, 320 Chios ewes from a nucleus flock in Cyprus were genotyped for polymorphisms at the β-lactoglobulin (β-LG), prolactin (PRL), Growth Hormone Receptor (GHR), DGAT1, and Acetyl-Coenzyme A acyltransferase-2 (ACAA2) loci. The GHR F279Y and DGAT1 K232A variants were absent from Chios sheep. The common β-LG variants, A and B, and PRL variants were detected at intermediate frequencies but exhibited no association with milk traits. The entire coding region of ACAA2 was sequenced and a novel SNP (HM537015:g.2982T > C) was identified in the 3’ UTR, with both alleles segregating at similar frequencies. All other exons were monomorphic. Mixed model association analysis, using SNP data from 318 animals from 104 paternal half-sib families and first lactation phenotype and pedigree information on 2405 ewes revealed that this SNP was significantly associated with milk yield. The T allele was associated with increased milk yield, and exhibited partial dominant action. Animals with the g.2982TT or g.2982CT genotype had significantly higher milk yield than those with the g.2982CC genotype, with the
g.2982T allele having an additive effect of 13.4(±4.7) kg and a dominance effect of 7.9(±6.1) kg. Based on estimated allelic effects and sample allele frequencies, the g.2982T>C SNP explained 10% of the additive genetic variance for milk yield. To the best of our knowledge, this is the first study reporting a significant association of an ACAA2 polymorphism with milk yield in ruminants. Whether this (or a closely linked) SNP is a quantitative trait nucleotide for milk yield remains to be determined.

P-331
HAPLOTYPE BASED ASSOCIATION STUDY FOR PIGMENTATION TRAITS IN SHEEP REVEALS SIGNIFICANT INTERACTION BETWEEN LOCI

Elisabeth Jonas1, Peter C. Thomson1, Malcolm Fleet1, Herman W. Raadsma1
1Reprogen - Animal Bioscience, Faculty of Veterinary Science, University of Sydney, Camden, NSW, Australia, 2Swedish University of Agricultural Sciences – Department of Animal Breeding and Genetics, Uppsala, Sweden, 3South Australian Research and Development Institute, Livestock and Farming Systems, Turrettfield Research Centre, Rosedale, SA, Australia

Pigmentation plays a major role for the sheep wool industry and especially Merino wool should be free from pigmented fibres. A genome-wide scan with 200 microsatellite markers was conducted in 164 animals of an ovine Awassi-Merino backcross population identifying QTL regions on chromosomes 2, 6, 8, 19 and 26. SNP markers covering three functional positional candidate genes on chromosomes 2, 6 and 19 were selected from the dataset. Non-informative markers (homozygous in family, inherited grandpaternal allele not clearly assigned) were removed and a total of 112 informative markers were kept to construct grandparental (Merino and Awassi) haplotypes. Association of paternal and maternal haplotypes with pigmentation traits was tested using analysis of variance in SAS (version 9.2) where the paternal haplotype contrast was fitted as a primary fixed effect. All three major haplotype effects and first order interactions were fitted across four colour traits (fibres face, fibres ears, fibre horn, fibres legs) and the combined score (derived principal component).

Haplotypes across two candidate genes on chromosomes 2 and 19 showed significant association (P < 0.05) with pigmentation; animals inheriting the grandparental Awassi allele had more pigmented fibres. Haplotypes on chromosomes 2 and 19 accounted for 25 to 47% and 2 to 16% of the variation in pigmentation, respectively. Significant interaction (P < 0.05) was identified between the haplotypes on OAR 2 and 19, suggesting epistatic effects between both genes, combined models explained 41 to 57% of the variation of fibre pigmentation. The third region on chromosome 6 was significant associated with three traits only, effects sizes were low and effects cryptic with higher pigmentation scores for animals inheriting the Merino allele.

Our study verified previously identified loci associated with fibre pigmentation in sheep and furthermore facilitates proof of epistatic effects as shown by interaction between loci. This study offers scope for marker assisted selection for these traits using a limited range of QTL.

P-332
QTL MAPPING OF BEAK MORPHOLOGY IN THE ZEBRA FINCH

Ulrich Knief1, Holger Schielzeth2,3, Bart Kempenaers1, Hans Ellegren2, Wolfgang Forstmeier1
1Max Planck Institute for Ornithology, Seewiesen (Starnberg), Germany, 2Uppsala University, Uppsala, Sweden, 3Bielefeld University, Bielefeld, Germany

The intra- and interspecific diversity of avian beak morphologies is one of the most compelling examples for the power of natural selection acting on a morphological trait. The development and diversification of the beak has also become a textbook example for evolutionary developmental biology, and variation in expression levels of several genes is known to causally affect beak shape. However, until now no genomic polymorphisms have been identified that are related to beak morphology in birds.
We estimated heritability and genetic correlations for beak length, depth and width and performed a QTL linkage analysis for these traits based on 1,404 informative single nucleotide polymorphisms genotyped in a four generation pedigree of 992 captive zebra finches (Taeniopygia guttata).

Beak size, relative to body size, was found to be sexually dimorphic (larger in males). Heritability estimates ranged from 0.47 for beak length to 0.74 for beak width and genetic correlations were high (ranging from 0.46 between length and width to 0.65 between depth and width). QTL mapping revealed four to five regions of significant or suggestive genome-wide linkage for each of the three beak dimensions (nine different regions in total). All beak dimensions had at least one exclusive QTL linkage peak which demonstrates that there is additive genetic variance for each of the dimensions independent from the other dimensions.

Eight out of eleven genes known to influence beak morphology are located in the nine peak regions. Five QTL do not cover known candidates, thus demonstrating that yet unknown genes or regulatory elements may influence beak morphology in the zebra finch.

P-333
ASSOCIATION OF CIRCADIAN GENE POLYMORPHISMS WITH BIPOLAR DISORDER

Yin-Chieh Lai¹, Ming-Kun Lu², Ru-Band Lu³, Po-Hsiu Kuo²
¹Institute of Epidemiology and Preventive Medicine, NTU, Taipei, Taiwan, ²Department of Health, JiaNan Mental Hospital, Tainan, Taiwan, ³Department of Psychiatry, NCKU, Tainan, Taiwan

Objectives: Circadian rhythm is an endogenously driven around 24-hour cycle in biochemical, physiological, and behavioural processes, which controlled by an autonomous circadian clock. Genetic variants in circadian genes were previously reported to be associated with bipolar disorder (BPD). However, results are inconsistent across different samples. The current study aimed to examine the associations between circadian genes and BPD in the Han Chinese population.

Methods: A case-control association study was conducted for patients with BPD and healthy controls. Patients' were ascertained from three hospitals in 2008 to 2010. Healthy controls were recruited in the community and were screened for mood disturbances and other major psychiatric disorders. We selected 9 circadian genes from prior association studies for BPD, including BMAL1, BHLHB2, CLOCK, CRY1, CRY2, NR1D1, RORA, RORB, and TIMELESS. Thirty three SNPs were genotyped in 280 BPD patients and 200 controls using GoldenGate assays. Single marker and multi-loci association analyses were performed using PLINK. Pair-wise interactions among SNPs were tested using the operation based screening and testing.

Results: All SNPs were in Hardy-Weinberg equilibrium. The call rate was high, ranged from 95.4% to 99.8% for the 33 SNPs. Among 9 circadian genes, there were two SNPs associated with BPD. Associations were observed for rs4774388 in RORA (OR=1.54, P = 0.02), and rs2291738 in TIMELESS (OR = 0.47, P = 0.04) gene. Suggestive interactions were found between RORA-RORB genes, and RORA-NR1D1 genes.

Conclusions: The circadian system includes three activation and feedback loops, in which the ROR regulation showed association with BPD. Our results suggest that the circadian pathway can be important candidates for BPD in the Han Chinese population. Further replication studies and basic research are needed to investigate the functional property of associated genes to contribute on understanding the pathogenesis of bipolar illness.
P-334
GENETIC ARCHITECTURE OF CLINICAL MASTITIS TRAITS IN DAIRY CATTLE
Goutam Sahana, Bernt Guldbrandtsen, Mogens Sandø Lund
Aarhus University, Department of Molecular Biology and Genetics, DK-8830, Tjele, Denmark

A dense SNP panel was used to predict the genetic merit of an individual for selection in livestock. The accuracy of genomic predictions depends in part on the genetic architecture of the trait, in particular the number of loci affecting the trait and distribution of their effects. Here we investigate the genetic architecture of clinical mastitis and somatic cell score traits in dairy cattle using a high density (HD) SNP panel. Mastitis, an inflammation of the mammary gland most commonly caused by bacterial infection, is a frequent disease in dairy cattle. Clinical mastitis and somatic cell score from first three lactations were studied for association with SNP markers in 4,200 progeny-tested Nordic Holstein bulls. Single trait breeding values were used as phenotypes. All the individuals were genotyped with BovineSNP50 Beadchip. Part of this population was also genotyped with the BovineHD BeadChip. A total of 648,219 SNPs passed the quality control criteria for genotypes from the high density SNP panel. All the 4,200 individuals’ genotypes were imputed to the high density SNP panel using the software Beagle. The associations between the phenotypes and SNPs were estimated by a linear mixed model analysis. After Bonferroni correction 12,372 SNP exhibited genome-wide significant associations with mastitis related traits. A total 61 QTL regions on 22 chromosomes associated with mastitis related traits were identified. The SNP with highest effect explained 5.6% of the variance of the predicted breeding values for the first lactation clinical mastitis.

P-335
IDENTIFICATION OF POLYMORPHIC MODIFIERS OF RAS AND WNT SIGNALING IN C. ELEGANS
Tobias Schmid1, Fabienne Largey1, Miriam Rodriguez Sanchez2, Basten Snoek2, Jan Kammenga2, Alex Hajnal1
1Institute of Molecular Life Sciences, University of Zürich, Zürich, Switzerland, 2Laboratory of Nematology, Wageningen University, Wageningen, The Netherlands

The genetic background has a strong influence on the pathogenesis of many complex, polygenic human diseases. Therefore, detailed genetic studies are needed to decipher the underlying genetic risk factors in human diseases.

We use the Nematode C. elegans as a model to study the interplay between cancer-related signaling pathways and genetic background. We focus on the RAS, Notch and WNT pathway genes, whose human orthologues have been shown to be involved in several complex diseases. We use the development of the hermaphrodite vulva as readout for pathway activity, where stronger RAS, NOTCH or Wnt signaling activity leads to more than one vulva (multivulva) and lower activity to a vulvaless animal.

We are addressing the question of how the genetic background influences the output of these signaling networks. For this purpose, we crossed mutants in the RAS, Notch and Wnt pathways originally isolated in the N2 Bristol background into the CB4856 strain isolated in Hawaii and established recombinant inbred lines (RILs). We then measured pathway activity of RAS, Notch or Wnt of each RIL by scoring the number of differentiated vulval cells.

Two RIL sets containing a mutation in either the Ras homologue let-60(gf) or the β-catenin homologue bar-1(ga80) in a mixed Hawaii/Bristol background show strong variation in the penetrance and expressivity of the phenotype when compared to the mutants let-60(n1046) or bar-1(ga80) in the “pure” Bristol background. This indicates the presence of polymorphic modifiers affecting RAS and Wnt signaling. Subsequently, we have identified by QTL mapping several genomic regions that contain those modifiers.

We are now investigating polymorphic candidate genes in these regions for their role in RAS, Notch and Wnt signaling. To identify the casual genes, we combine RNAi knock down, eQTL analysis and transgenic studies.

This work was supported by the EU-FP7 HEALTH project PANACEA, contract nr. 222936.
P-336
MINING FOR USEFUL VARIATION IN BREAD WHEAT COLLECTIONS USING QUANTITATIVE TRAIT MAPPING AND ASSOCIATION MAPPING

Luzie U. Wingen, Simon Orford, Richard Gorum, Sue Freeman, Jun Ma, Simon Griffiths
John Innes Centre, Norwich, Norfolk, UK

Food security is becoming a critical issue both in the UK and worldwide due to rapid population expansion, dietary changes and declining stocks of fossil fuels. Total wheat grain production over the next 50 years must exceed that previously produced. Improving the wheat grain production necessarily means the search for new and useful allelic variation on a larger scale.

In this context, we aim to identify useful traits for modern breeding in two wheat collections:
1) The AE Watkins Collection, which consists of about 800 wheat landraces purchased in the 1930s from local markets in 32 countries across the world.
2) The GEDIFLUX collection, which includes nearly 500 modern European wheat varieties which were sown in major acreages in the years 1945 to 2000.

Both collections have been genotyped with nearly 50 SSR markers and the genotyping is currently extended with SNP markers (KASPar technology).

Association analyses for the traits height, heading date and grain yield have revealed possible loci of interest in the AE Watkins collection, and some quantitative trait locations have been identified in bi-parental genetic material. Near isogenic lines for these loci are currently developed. For the GEDIFLUX collection extensive phenotyping has been conducted on yield and yield component traits, and association mapping results will be presented.

All promising loci will be evaluated for their usefulness before they may enter a breeding programme. This work is funded by the UK DEFRA Wheat Genetic Improvement Network (WGIn) and BBSRC Enhancing Diversity in UK Wheat Through a Public Sector Prebreeding Programme.

P-337
HERITABILITY OF RESISTANCE TO PANCREAS DISEASE IN ATLANTIC SALMON FRY

Serap Gonen1, Ashie Norris1,2, Petter Arnesen1,2, Stephen C. Bishop1, Ross D. Houston1
1The Roslin Institute, University of Edinburgh, Midlothian, EH25 9RG, Scotland, UK, 2Marine Harvest, Sandviksboder 78a, 5035 Bergen, Norway

Pancreas disease (PD) is a viral disease caused by a salmonid alphavirus which can cause morbidity and mortality in farmed Atlantic salmon, with mortality rates varying between 10-50% in post-smolts. A genetic component to host resistance to PD has previously been demonstrated, with heritability estimated at 0.21 in post-smolts (seawater lifecycle stage). To further investigate the genetic basis of salmon resistance to PD, we analysed data from 6540 PD-challenged fry (freshwater lifecycle stage) with known pedigree. The overall mortality in the experiment was 67.9%. Using ASReml we fitted a series of (sire + dam) models, both on the observed scale and also using logit or probit link functions. The dam variance was significantly higher than the sire variance, suggesting possible maternal effects associated with PD mortality. Heritabilities were therefore calculated using the sire variance only. Heritability for mortality was estimated to be 0.31 (± 0.09) on the observed scale, which corresponds to 0.52 when transformed to the underlying scale. Fitting logit link and probit link functions gave heritability estimates of 0.46 (± 0.13) and 0.49 (± 0.14) respectively. Possible future steps include looking at the genetic correlation between PD mortality and other performance traits such as body weight, fillet percent, fat content and body length. This may also help tease out the cause of the maternal effects on PD survival.
P-338
MAPPING OF MILK TRAITS IN THE COMBINED NORDIC RED DAIRY CATTLE POPULATION BY GENOME-WIDE ASSOCIATION

Nina Schulman1, Goutam Sahana2, Miika Tapio1, Mogens Lund2, Johanna Vilkki1
1MTT, Agrifood Research Finland, Biotechnology and Food Research Finland, Jokioinen, Finland, 2Aarhus University, Faculty of Science and Technology, Molecular biology and Genetics, Foulum, Denmark

We have previously detected QTL areas associated with female fertility traits in Nordic red dairy cattle. In order to get better understanding of these QTL areas we carried out single marker association mapping on three milk traits: milk yield, protein yield and fat yield to see if they coincide with the fertility QTL. The analyzed data included 4115 Nordic red AI bulls. Genotypes were obtained with the Illumina BovineSNP50 panel and a total of 38388 informative, high-quality SNP markers were used. The data was analyzed using single marker association with a mixed model, where the marker effect is fixed and animal effect is random. The Nordic red population includes admixed animals that have proportions from Finnish Ayrshire, Swedish SRB and Danish RDM cattle. Therefore the population stratification effect was accounted for by adding the sub population proportions estimated by the software STRUCTURE as fixed effect. The Bonferroni 5% significance threshold based on the number of markers was used to declare significance.

We detected a total of 26 QTL areas on 12 different chromosomes. Three of the areas (BTA4, 12, 14) were detected in more than one of the milk traits. A highly significant QTL peak close to the DGAT1 gene was detected. On BTA12 the most significant SNP associated with all three milk traits was identical with the most significant SNP associated with female fertility. No other QTL areas were overlapping with female fertility traits.

P-339
FIELD STUDY OF SOW PERFORMANCE DURING A PORCINE REPRODUCTION AND RESPIRATORY SYNDROME (PRRS) OUTBREAK

Nader Deeb1, Nicholas Boddicker3, Jerome Geiger2, Saad Sebbana1
1Genus plc, Hendersonville, TN, USA, 2PIC, Hendersonville, TN, USA, 3Iowa State University, Ames, IA, USA

Porcine reproductive and respiratory syndrome (PRRS) reduces growth and increases morbidity in growing pigs. Poor reproductive performances (RP) in breeding sows are also directly associated with PRRS virus (PRRSv) infection. Recent research has focused on identifying genomic regions associated with the host response to PRRSv infection in growing pigs, but little is invested on sow RP, mainly due to the long-term nature of data collection and associated costs. This study represents an opportunistic approach to collect phenotypes from sow herds during an outbreak of PRRSv that will pave the way for a comprehensive genome-wide association study to validate known genomic regions and identify regions specific for RP. Numerous sow herds become infected with PRRSv, despite biosecurity regulations. Several potential herds with excellent recording practices and detailed databases were identified. Once a PRRS outbreak was confirmed within the sow herd, tissue samples were collected from all sows. Historic and current RP data was extracted from the database. Records are limited to numbers of born alive, mummies, still births and abortions. Here we share preliminary results from a genome scan using selected markers and discuss some of the practical limitations of using such data. Gestation stage during outbreak needs to be considered when selecting animals to be genotyped. Historic information is important, as sows that have been previously affected by PRRSv may have less dramatic response to new infection. Many markers discovered using a growing-pig model were found to also affect RP during an outbreak, and several were unique to the sow RP model. This research is the first step in commercial validation of marker’s effects and expanding existing findings to build a broad genomic tool to aid in the selection for PRRS resilience in growing pigs and sow RP.
P-340
A GENOME-WIDE ASSOCIATION STUDY OF VISUOSPATIAL ATTENTION IN HEALTHY INDIVIDUALS

Sudheer Giddaluru1,2, Thomas Espeseth3,4, Pamela Greenwood1,5, Andrea Christoforou1,2, Astrid Lundervold6, Vidar M. Steen1,3, Ivar Reinvang1,7, Raja Parasuraman2, Stephanie Le Hellard2,3
1Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway, 2Dr. Einar Martens Research Group for Biological Psychiatry, Department of Clinical Medicine, University of Bergen, Bergen, Norway, 3Centre for Advanced Study, Oslo, Norway, 4Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway, 5Arch Lab, Psychology Department, George Mason University, Fairfax, VA, USA, 6Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway, 7Center for the Study of Human Cognition, Department of Psychology, University of Oslo, Oslo, Norway

There is accumulating evidence suggesting a substantial genetic influence on individual differences in general and specific cognitive abilities. Understanding the genetic factors affecting individual differences in cognition may help in elucidating the mechanisms underlying brain functions and their dysfunctions in disorders such as neurodegenerative or psychiatric disorders. In spite of their many pitfalls due to sample size, power and multiple testing, genome-wide association studies (GWASs) offer the possibility of identifying candidate genes without a priori biological hypotheses. Here we describe a GWAS for a visuospatial attention task in a sample of healthy adults, the Norwegian Cognitive NeuroGenetics (NCNG, n=643) sample, followed by replication in a second sample with the exact same phenotype (George Mason University sample, GMU, n=450). Visuospatial attention was manipulated in cued discrimination task (CDT), measuring response times (RT) to categorize the target as consonant or vowel after presentation of valid, invalid or neutral location cues. The RT was then tested for association to ~554,000 SNPs, using linear regression. No single SNP reached genome-wide significance (best $P=6.43 \times 10^{-6}$ in invalid RT). SNPs were then assigned to genes; using linkage disequilibrium based binning approach, as implemented in the LDsnpR tool, and scored. The most significant SNPs and genes were selected for replication in the GMU sample (2490 SNPs). Replication was observed at both the SNP and gene levels ($P<0.05$). The top SNPs showing replication after meta analysis include rs2248321, rs3741434, rs3779470 ($P=1.75 \times 10^{-6}$; $2.19 \times 10^{-6}$; $2.80 \times 10^{-6}$) and mapped to the intersectin 1 (ITSN1), the retinoic acid receptor, gamma (RARG), and the tachykinin, precursor 1 (TAC1) genes. At the gene level, replication was observed for ITSN1, TAC1, SH3 domain containing ring finger 2 (SH3RF2), Rho GTPase activating protein 19 (ARHGAP19), and growth factor receptor-bound protein 10 (GRB10). The findings provide new insight into the genetics underlying visuospatial attention emphasizing the contribution of multiple variants of modest effect.

P-341
GENETIC PARAMETERS FOR GESTATION LENGTH, BODY WEIGHT AT BIRTH AND AT 210 DAYS OF AGE AND ACCUMULATED PRODUCTIVITY IN BRAZILIAN NELORE BEEF CATTLE

Tatiane Chud1,3, Sabrina Caetano1,7, Marcos Buzanskas1,6, Daniela Grossi4, Diego Guidolin1,3, Camila Souza1,3, Rayssildo Lôbo2,6, Danisio Munari1,6
1FCAV/UNESP, Jaboticabal, São Paulo, Brazil, 2ANCP, Ribeirão Preto, São Paulo, Brazil, 3CAPES fellowship, Jaboticabal, São Paulo, Brazil, 4University of Guelph, Guelph, Ontario, Canada, 5FAPESP fellowship, São Paulo, Brazil, 6CNPq fellowship, Brasília, DF, Brazil, 7UNESP fellowship, São Paulo, Brazil

In this work, the aim was to estimate genetic parameters for gestation length (GL), body weight at birth (WB), body weight at 210 days of age (W210) and accumulated productivity (ACP) in Nelore beef cattle. The trait GL has been recently included as criterion in the selection index used in some important Brazilian Nelore cattle breeding programs. ACP is an index that evaluates female productivity, considering calf body weight at weaning and number of offspring produced. For GL, WB and W210 the model included the fixed effects of contemporary group (CG) and age of calving and, as random effects,
direct and maternal genetic effect, permanent environmental of the dam and residual. For ACP, the model included the fixed effect of CG and the random effects of additive genetic and residual. Variance components were estimated by Restricted Maximum Likelihood method by the two-trait animal model. Heritability estimates and standard errors, within parenthesis, of direct effect for GL, WB, W210 and ACP were 0.40 (0.02), 0.32 (0.02), 0.36 (0.02) and 0.11 (0.02), respectively. Maternal heritability estimates for GL, WB, W210 and ACP were 0.06 (0.01), 0.10 (0.01) and 0.13 (0.01). Permanent environmental fraction of variance were 0.04 (0.01) for GL, 0.01 (0.01) for WB and 0.14 (0.01) for W210. Genetic correlations ranged from -0.14 (0.01) between GL with ACP to 0.31 (0.05) between WB with W210. Genetic trend for GL indicated a decrease of 0.02 days/year. For WB and W210 the genetic gains were 73 g/year and 845 g/year, respectively. The selection for reduced the GL may be result in low body weight at birth and selection for W210 may lead to an indirect selection for PN.

Financial support: CAPES, CNPq, FAPESP and UNESP

P-342

LONGEVITY STUDY CONSIDERING THE COW’S AGE AT LAST CALVING IN NELLORE BEEF CATTLE

Sabrina Caetano1,8, Guilherme Rosa2, Rodrigo Savegnago1,7, Natália Grupioni1,5, Luiz Bezerra3, Raysildo Lôbo3,6, Cláudia Paz4,6, Danísio Munari1,6
1FCAV/UNESP, Jaboticabal, São Paulo, Brazil, 2University of Wisconsin, Madison, WI, USA, 3ANCP, Ribeirão Preto, São Paulo, Brazil, 4APTA, Ribeirão Preto, São Paulo, Brazil, 5CAPES fellowship, Jaboticabal, São Paulo, Brazil, 6CNPq fellowship, Brasília, DF, Brazil, 7FAPESP fellowship, São Paulo, Brazil, 8UNESP fellowship, São Paulo, Brazil

The economic cost of beef production depends mainly on cows’ reproductive performance. To produce more kilos of calf per year, the cows have to be long-lived. Then, to assess the longevity, an alternative method considering all the cows in the herd was studied. The variable of the cow’s age at last calving (ALD), i.e. each cow’s last calving recorded in the database, was used because this is a trait of interest to producers and forms part of most databases controls on farms. This was done using a criterion that made it possible to include cows not only at their first but also at their ninth calving. The criterion used was the difference between the date of each cow’s last calving and the date of the last calving on each farm. If this date was greater than 36 months, the cow was considered to have failed and was discarded. If not, this cow was censored, thus indicating that future calving remained possible for this cow. The aim of this study was to estimate heritability and genetic value for longevity among cows in herds of Nellore breed, from ALD, by means of survival analysis methodology, and to estimate the risk function relating to the cow’s continuation in the herd, taking into consideration the covariables involved. The survival model used for the analyses was the proportional hazards model, and the base risk was given by Weibull distribution. The covariables involved were: age at first calving, year of birth, season of birth and farm. The heritability estimate obtained (0.25) indicated that ALD is a relevant trait for assessing cows’ longevity in the herd, as well as being easy to measure.

Financial support: CAPES, CNPq, FAPESP and UNESP.
Osteoarthritis is a leading cause of disability worldwide and is increasing to epidemic proportions as world populations age and as obesity rates rise. Increases in OA prevalence have not been accompanied by the advances in our knowledge of OA pathogenesis that are needed to effectively prevent and treat this complex disease. We measured two serum biomarkers of OA – hyaluronic acid (HA) and cartilage oligomeric matric protein (COMP) – in 935 San Antonio Family Study (SAFS) participants in large extended families and used existing data for 1,000,000 SNPs to conduct a GWAS to test for association between quantitative variation in circulating concentrations of HA and COMP and specific SNP variants. Many of the most strongly associated SNPs (p=4x10^{-7} to 7.8x10^{-6}) showed a unifying theme of relevance to skeletal system and cartilage development. HA GWAS revealed 1) three SNPs implicating FOXN4, a member of the winged helix/forkhead transcription factor family whose members are important to embryonic skeletal system development, 2) rs16907395, located within ETV6 that associates with osteochondrogenic transcription factor RUNX1, and 3) rs6482369, an eQTL for TMEM198 -- a membrane protein required for Wnt signalling which is critical to bone and cartilage formation – and which is located within the KIAA1217 gene, a human homolog of the mouse skt gene which is important in mouse embryonic skeletal patterning. COMP GWAS revealed 1) three SNPs implicating ZNF521, a transcription factor involved in regulating osteoblast commitment and differentiation, 2) three SNPs near CSGALNACK1 which is critical to chondroitin sulfate biosynthesis and aggrecan formation in cartilage. These results implicate a broad spectrum of skeletal development signalling processes in OA.
P-344
MAPPING OF PARENT-OF-ORIGIN EFFECTS USING DENSE SNP MARKERS IN PIG

Aisaku Arakawa¹, Masaaki Taniguchi¹, Motohide Nishio², Naohiko Okumura³, Toshimi Matsumoto¹, Kensuke Hirose¹, Kazuo Fukawa⁴, Tetsuya Ito⁴, Takeshi Hayashi³, Hirohide Uenishi¹, Satoshi Mikawa¹
¹National Institute of Agrobiological Sciences, Tsukuba, Ibaraki, Japan, ²National Institute of livestock and Grassland Science, Tsukuba, Ibaraki, Japan, ³JATAFF Institute, Tsukuba, Ibaraki, Japan, ⁴Central Research Institute for Feed and Livestock ZEN-NOH, Kamishihoro, Hokkaido, Japan, ⁵National Agricultural Research Center, Tsukuba, Ibaraki, Japan

Parent-of-origin effects due to genomic imprinting have been often observed in some traits of livestock animals. However, very little is known about the extent of their contribution to genetic variations in quantitative traits. Recently, genome-wide association studies based on a large number of SNPs across the whole genome enable to identify many loci affecting quantitative traits, such as human height. In this study, we investigate chromosomes and genomic regions showing the parent-of-origin effects using genome-wide SNP markers for average daily gain (ADG) and backfat thickness (BFT) in pig. A total of 1059 animals that have phenotypes and genotypes were obtained from a closed population of Duroc breed. We searched chromosomes with different genetic variances between the parental origins from SNP haplotypes using Bayesian ridge regression. As a result, the paternal genetic variances were larger and less than maternal ones on chromosome 6 for ADG and BFT and chromosome 12 for ADG, respectively. For mapping the loci where paternal and maternal haplotypes were differently expressed, the haplotype effects descended from each parent were reestimated using BayesB with \( \pi = 0.995 \). Consequently, we confirmed the parent-of-origin regions where the paternal haplotypes explained larger genetic variance around 134 Mb on chromosome 6 in both traits and the maternal haplotypes explained larger genetic variance around 15 Mb on chromosome 12 in ADG. Further research is in progress for fine-mapping of these QTLs.

P-345
DECIPHERING WHEAT GRAIN PROTEIN CONTENT: GENETIC ANALYSIS OF TEMPORAL DYNAMIC NITROGEN CONTENT RESORPTION IN FLAG LEAVES

Ingrid Vilmus², Pierre Roumet¹, Jacques David¹, Martin Ecarnot¹
¹INRA, Montpellier, Languedoc-Roussillon, France, ²Syngenta Seeds, Saint Sauveur, Midi-Pyrénées, France

Grain Protein content (GPC) is a decisive feature for durum wheat industrial development. This trait is complex since it results from many nested mechanisms occurring during the plant cycle. Since large amounts of nitrogen fertilizer are required to cover plant nitrogen supply, improvement of plant nitrogen use efficiency appears as an alternative to increase the durum wheat GPC.

To obtain a better understanding of this trait, we propose here to analyse the temporal dynamics of nitrogen resorption from flowering to plant maturity.

We phenotyped 282 Rils coming from a half diallel (4 parental lines) in a greenhouse. The flag leaf was chosen as the organ contributing the most to the GPC to document the nitrogen resorption mechanism. About 30 to 40 measurements were carried out with a portable near infrared spectrometer calibrated to infer nitrogen content and leaf mass area (LMA) from spectra data (\( R^2 \) are 0.93 and 0.94 respectively, Ecarnot et Roumet., submitted ).

Each obtained curve was modelled with a nonlinear mean square method using a Gauss-Newton algorithm to fit a logistic equation. They were summed up in some variables linked to senescence. In addition, flag leaves were measured at flowering and the GPC, the grain number and the grain weight were measured after harvest.

This population was mapped with 529 polymorphic loci. The genetic map spans 2082 cM distributed in 26 linkage groups (Vaissayre et al., 2012).

A mixed model (ASReml-R package) including the pedigree was used to estimate variance components.
Heritability values range from 0.17 to 0.40 for senescence variables and from 0.31 to 0.50 for other variables. Genetic correlations between variables were estimated and qtls investigated using MCQTL software (Mangin et al., 2007). Using those results, the efficiency of a breeding strategy taking into account nitrogen resorption parameters to increase the GPC is discussed.

**P-346**

**TRAJECTORY OF GENETIC VARIANCE DUE TO MUTATION AND FOUNDER POPULATION IN A SELECTION EXPERIMENT FOR BODY WEIGHT GAIN IN MICE**

*Nuzul Widyas, Just Jensen, Vivi Nielsen*

*Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark*

Selection was performed for three to nine weeks body weight gain in 13 generations of experimental mice. In each generation a line consisted of eight full sib families. Parents for the next generation were chosen within each litter. Mating was performed ensuring maximum avoidance of inbreeding. A total of 18 lines were included in the experiment. Nine lines were allotted to each of the two treatment diets containing standard and severely reduced protein level (19.3 and 5.1 % crude protein respectively). Within each diet three lines were selected upwards, three lines were selected downwards and the other three lines were kept as controls. In total the experiment comprised 2440 litters and around 15170 individual mice.

Genetic variance change during the experiment due to unavoidable accumulation of inbreeding, genetic drift and mutations. Potential accumulation of positive mutations maintained in the population is due to the selection conducted in each selection line. Bayesian statistical methods are used to estimate the trajectory of genetic variance over generations for different lines. Genetic transmission model is expanded including mutation into variance-covariance additive matrix following the model introduced by Wray (1990). Further analysis employing the mentioned model aimed to separate the trajectory into effects due to inbreeding and genetic drift from the effects due to de novo mutations accumulated during the experiment.

Mutation is assumed to be responsible for the continuous response to long term selection. Thus, the results have implications for commercial breeding programs in order to better ensure that advantageous mutations are selected.

Key words: selection, mice, genetic variance, mutation, Bayesian.

**P-347**

**PREDICTING QTL ALLELE IDENTITY BETWEEN CHROMOSOMES: IMPACT ON MAPPING ACCURACY**

*Julien Jacquin, Jean-Michel Elsen, Hélène Gilbert*

*INRA, Castanet-Tolosan, France*

Approaches were proposed to locate markers in association with trait variability with a similar principle: chromosomes with local molecular identity exhibit better chance to harbor the same allele for a QTL than chromosomes having local differences. Four approaches to quantify local identity between chromosomes were compared for their ability to correctly predict QTL allelic identity, and for their mapping accuracy. Comparisons were performed on 150 simulated dataset obtained from a set of 235 porcine chromosomes (969 markers) and a 26 generations pedigree. Methods to predict allelic identity between chromosomes were 1) identity by state (IBS) of the neighbor marker of the tested position (IBS_m), 2) IBS of the 6-marker haplotype surrounding the tested position (IBS_hap), 3) similarity score (Li and Jiang, 2005) between haplotype pairs (SCORE), 4) IBD probability between haplotype pairs (Meuwissen and Goddard, 2001) (PIBD). The two first methods gave 1 and 0 values, whereas the others gave 0 to 1 continuous values. These predictions were collected in matrices $M_\text{IBS}$ and compared to the actual QTL identity between the chromosomes ($M_{\text{QTL}}$). First, geometrical approaches (distance, angle and Euclidian norm between matrices)
Poster Abstracts

were examined. Secondly, mean square error (MSE) of the QTL position estimate, obtained from mixed models including a random QTL effect with covariance described by the \( M_{id} \) matrices, were estimated. Finally, the Kullback-Leibler measure (KL) of the difference between the phenotypes distributions knowing \( M_{QTL} \) and knowing \( M_{id} \) at the retained QTL position was computed. The geometrical results suggest that the \( \text{SCORE} \) is the best method to predict the QTL allelic identity between chromosomes, and \( \text{IBS\_hap} \) the worst. However, MSE and KL were reduced for \( \text{IBS\_hap} \) and \( \text{SCORE} \) compared to the other methods, suggesting better mapping accuracy. From these results, a strategy can be proposed to threshold optimally the matrices and maximize mapping accuracy.

**P-348**

**GENETICS OF OZONE INDUCED CELL DEATH IN ARABIDOPSIS**

**Johanna Leppälä**¹, **Mikael Brosché**¹,²

¹Division of Plant Biology, Department of Biosciences, University of Helsinki, Helsinki, Finland, ²Institute of Technology, University of Tartu, Tartu, Estonia

Plants constantly face both biotic (e.g. pathogen) and abiotic (e.g. drought, heat) stresses in nature. To increase their fitness plants have developed ways to survive through stressful periods. One of the ultimate ways is programmed cell death (PCD), which for example prevent the spread of some pathogens by restricting access to nutrients. One signal intermediate in plant defence, including PCD, are reactive oxygen species. The atmospheric pollutant ozone (O₃) can be used to study signalling activated by reactive oxygen species and treatment of plants with short, high concentrations of O₃ activates PCD. The molecular responses to acute O₃ exposure are overlapping with responses to other stresses. We have used acute O₃ exposure to trigger PCD in hundreds of Arabidopsis thaliana natural ecotypes included in the 1001 genomes sequencing project. Earlier studies have shown variation in response to O₃ exposure between a smaller set of ecotypes and that stomata play an important role in plant O₃ responses by varying the speed of stomatal closure between ecotypes. We have conducted an association mapping study to identify the genes involved in natural variation of O₃ induced cell death and stomatal signalling.

**P-349**

**INTERCONTINENTAL COMPARISON OF GWAS RESULTS SUPPORTS THE INFINITESIMAL VERSION OF THE COMMON DISEASE/COMMON VARIANT PARADIGM**

**Urko M. Marigorta**¹, **Arcadi Navarro**¹,²

¹Institut De Biologia Evolutiva - UPF-CSIC-PRBB, Barcelona, Catalonia, Spain, ²National Institute for Bioinformatics - UPF, Barcelona, Spain

Genome-wide association studies (GWAS) constitute a powerful tool to detect genetic variants that increase susceptibility to complex disease. GWAS have allowed the detection of thousands of risk alleles, raising hopes for personalized medicine. However, heated debates are still open on the interpretation and relevance of GWAS results, such as their utility for individual risk prediction; the potential role of rare variants as producers of artefactual GWAS results or the extrapolation of GWAS results from one population to another. Current evidence contributing to these debates comes from either re-sequencing efforts aimed to capture rare variants or multi-ethnic replication efforts for a few risk variants; while meta-analysis of GWAS data have either ignored population heterogeneity or focused in narrow set of traits. To address these questions, we have selected 27 complex diseases for which two or more GWAS performed in Europeans and at least one GWAS performed in East Asians were available (as to February 2012). After careful selection of 198 disease-associated SNPs discovered in European GWAS, we have studied their replication in other GWAS. In total, we have gathered 540 replication attempts. We report three main results. First, independently of the trait under study, SNPs found in Europeans present extensive rates of replication in both European (≈ 90%) and East Asians (≈ 50%) populations.
Second, the risk allele appears to be shared among populations, being the Odds Ratios highly correlated. Finally, genomic regions harbouring SNPs that consistently replicate between continents resemble more in their LD patterns. Our results show that (1) GWAS results are overwhelmingly replicated; (2) that common variants underlie the vast majority of significant results from GWAS; and (3) that the major contributors to genetic risk of most complex diseases, as unveiled by GWAS so far, are shared among continents.

P-350
COMPARISON OF ASSOCIATION MAPPING METHODS IN ADMIXTURE POPULATION

Naveen Kumar Kadri, Bernt Guldbrandtsen, Peter Sørensen, Goutam Sahana
Aarhus University, Department of Molecular Biology and Genetics, Tjele, Denmark

Population stratification occurs when the individuals in a sample differ in their genetic backgrounds. If these differences are not accounted for, they often lead to false positive results in genome wide association studies (GWAS) to detect the quantitative trait loci (QTL). The objective of this study was to compare GWAS methods for their power and control on false positive detections (type 1 errors) in the presence of population stratification. An admixed population of 2,000 animals was simulated by mating individuals from two populations that had been isolated for 30 generations. Simulations included 40,000 SNP markers distributed evenly over 20 chromosomes. Two QTL, jointly explaining 10% of the phenotypic variation, were located on one chromosome. All the other chromosomes had no QTL. One hundred replicates were analyzed using four different models: 1) a linear mixed model including pedigree based relationship (LMMped), 2) a linear mixed model including genomic relationship (LMMgmat), 3) an LMMped including population structure using STRUCTURE (LMMstr) and 4) an LMMped including population structure through principal components (LMMpca). All four methods were able to control the false positive rates. LMMgmat was conservative and suffered a slight loss of power, probably due to the use of the SNPs at QTL regions for estimating the relationships. A QTL was considered detected, when a SNP within 5 Mb region of the simulated QTL showed significance at $\alpha=0.05$ after Bonferroni correction. All the models had a lower power (~5%) to detect small-effect QTL (<5% of $\sigma^2$) and a higher power (60-65%) to detect large-effect QTL (5% to 10% of $\sigma^2$). Overall LMMped, LMMstr and LMMpca showed the same power, detecting 35.5% of the QTL and LMMgmat showed a slightly lower power of 32.5%. Models when included only the population structure without the relationship (pedigree/genomic) were not able to control the false positive rates at the nominal significance level.

P-351
NOVEL INSIGHT INTO THE GENOMIC ARCHITECTURE OF FEED EFFICIENCY IN GROWING PIGS

Mahmoud Shirali1,4, Carol-Anne Duthie1, Andrea Doeschl-Wilson2, Pieter Knap3, Egbert Kanis4, Johan van Arendonk4, Rainer Roehe6
1SAC, Edinburgh, UK, 2The Roslin Institute, R(D)SVS, University of Edinburgh, Easter Bush, Midlothian, UK, 3PIC International Group, Schleswig, Germany, 4Wageningen University, Wageningen, The Netherlands

The aim of this study was to understand the genomic association of residual energy intake (REI) with growth, feed conversion ratio (FCR), average daily feed intake (ADFI) and nitrogen excretion (NE) using quantitative trait loci (QTL) information at different stages and over the entire growth. Data were available on 315 F2 pigs from a three-generation full-sib design, developed by crossing Pietrain sires with a commercial dam line. The REI was estimated using regression by adjusting metabolisable energy intake for systematic effects and pre-adjusted average daily protein and lipid deposition. Genomic association criteria were that QTLs showed similar modes of inheritance within the same marker interval. At 60-90kg, a dominance QTL for REI was identified on SSC14, explaining 6.1% of the phenotypic variance. This QTL was favourably associated with ADFI (-0.18 ±0.05, kg/d) and NE (-4.18 ±1.06, g/d). At 90-120kg, three
QTLs for REI were detected on SSC2, SSC4 and SSC7, explaining 3.5%, 3.7% and 2.9% of the phenotypic variance, respectively. The QTL on SSC2 expressed additive effect showing favourable associations with NE (-130 ±30, g/d) and FCR (-0.18 ±0.05). Whereas, the QTL on SSC4 showed genomic imprinting, the QTL on SSC7 expressed dominance and a favourable association with FCR (-0.22 ±0.08). At 120-140kg, an additive QTL for REI was detected on SSC8, explaining 5.5% of the phenotypic variance and had a favourable influence on NE (-2.56 ±0.77, g/d). This study indicates that different growth stages are influenced by different REI QTL (no QTL detected over the entire growth), suggesting different genomic architecture of feed efficiency at different growth stages. The REI was genetically independent from growth and therefore influenced efficiency of feed utilisation. The detected REI QTLs are of high value for improvement of feed efficiency and reduction of nitrogen excretion, and consequently enhance the environmental sustainability of pig production.

P-352
EPISTASIS VERSUS PLEIOTROPY IN THE DETERMINISM OF GENETIC CORRELATIONS: A QTL ANALYSIS FOR ZINC TOLERANCE TRAITS IN ARABIDOPSIS HALLERI (BRASSICACEAE)

Isabelle Decombeix1, Luis-Miguel Chevin2, Angélique Bourceaux1, Cécile Godé1, Hélène Frérot1
1Université Lille 1, Villeneuve d’Ascq, France, 2Centre National de la Recherche Scientifique, Montpellier, France

Genetic correlations between traits can dramatically constrain adaptive trajectories of natural populations, and alter responses to selection. While the simplest evolutionary predictions assume a constant genetic covariance G matrix, for more accuracy the evolution of G should also be included. This requires understanding the genetic mechanisms underlying trait correlations. One the one hand, pleiotropy at individual loci can result in genetic correlations at the genomic scale. On the other hand, some loci can epistatically modify the genetic correlations contributed by other loci segregating in the genome. To empirically quantify the relative importance of these two mechanisms, we performed a QTL analysis on six root and shoot traits underlying zinc tolerance in the pseudo-metallophyte Arabidopsis halleri (Brassicaceae). The progeny from a backcross between A. halleri (zinc-tolerant) and A. lyrata petraea (non-tolerant) was cultivated on two contrasted zinc treatments, and we searched for QTL with effects on (i) the mean of several traits (individual locus pleiotropy), and (ii) the genetic covariance of traits (epistatic pleiotropy). We detected several correlation QTLs (revealing epistatic pleiotropy), which interestingly acted in a species-and environment-specific way.

P-353
GENOME WIDE ASSOCIATION ANALYSIS IN ITALIAN BROWN SWISS DAIRY CATTLE FOR SOMATIC CELL COUNT

Alessandro Bagnato1, Morris Soller2, Ehud Lipkin2, Antonia Bianca Samore1, Dinesh Velayutham1, Fausta Schiavini1, Attilio Rossoni1, Marlies Alexandra Dolezal1
1Department VSA, University of Milan, Milan, Italy, 2Department of Genetics, The Hebrew University of Jerusalem, Jerusalem, Israel, A.N.A.R.B, Bussolengo (VR), Italy

Mastitis is one of the most costly diseases in dairy cattle and a huge concern to animal welfare. Milk Somatic Cell Count (MSCC) is an indirect measure widely used for years to select individuals to reduce mastitis susceptibility in dairy cattle. The purpose of this study was to identify regions carrying loci for mastitis resistance in the Italian Brown Swiss dairy cattle population. We report on a whole genome association study on a total of 1030 bulls genotyped with Illumina’s Bovine 50k v1 SNP chip that had phenotypes in the form of de-regressed breeding values for MSCC. After data filtering a total of 35,566 SNPs were retained for the analysis with MAF >0.02, call rate >0.90 at SNP and bull level. SNPs were anchored to UMD3.1 autosomes based on the recently published map by Fadista and Bendixen (2012).
Stratification in the population was corrected for using a principal component approach. The success of correction for stratification was empirically assessed based on quantile-quantile plots of expected versus observed P-values. We employed single SNP regression and multiple SNP regression in sliding windows of three to five SNPs. Significance was declared employing a false discovery rate approach. Several QTL regions were found across the genome. The most interesting regions were located on BTA1, BTA2, BTA5, BTA6, BTA7, BTA13, BTA20 and BTA21. Reanalysis of the data after imputation of genotypes to Illumina’s HD chip interrogating 735,238 loci will further explore the identified associated regions in the Brown Swiss dairy cattle breed.

This study was supported by the FP7 framework project QUANTOMICS contract n. 222664-2.

P-354

GLYCOMICS MEETS GENOMICS – ANALYSIS OF PROTEIN GLYCOSYLATION IN POPULATION ISOLATES

Jennifer Huffman1, Abdelkader Essafi1, Veronique Vitart1, Ana Muzinic6, Barbara Adamczyk5, Gaya Thanabalasingh4, Caroline Hayward4, Ulf Gyllensten2, Harry Campbell3, Jim Wilson3, Mark McCarthy4, Katharine Owen4, Nicholas Hastie1, Pauline Ruedd3, Gordan Lauc6, Igor Rudan3,8, Alan Wright1

1MRC Hufman Genetics Unit; IGMM, University of Edinburgh, Edinburgh, Scotland, UK, 2Department of Genetics and Pathology, Rubeck Laboratory, Uppsala University, Uppsala, Sweden, 3Centre for Population Health Sciences, University of Edinburgh, Edinburgh, Scotland, UK, 4Oxford Centre for Diabetes, Endocrinology and Metabolism, and Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK, 5National Institute for Bioprocessing Research and Training, Dublin-Oxford Glycobiology Lab, University College Dublin, Dublin, Ireland, 6Genos Ltd., Glycobiology Division, Zagreb, Croatia, 7Faculty of Pharmacy and Biochemistry, University of Zagreb, Zagreb, Croatia, 8Faculty of Medicine, University of Split, Split, Croatia

The majority of human proteins are post-translationally modified by covalent addition of one or more complex oligosaccharides (glycans). Alterations in glycosylation processing are associated with numerous diseases and glycans are attracting increasing attention both as disease biomarkers and as targets for novel therapeutic approaches. Using a recently developed high-throughput HPLC analysis method we performed a genome-wide association study of 33 directly measured and 13 derived glycosylation features in 3533 individuals from 4 European populations. Polymorphisms at six loci (FUT8, FUT6/FUT3, HNF1a, MCAT5, B3GAT1 and SLC9A9) were found to affect plasma levels of N-glycans. Several of these gene products have well characterised roles in glycosylation, however, SLC9A9 and HNF1a were novel findings. SLC9A9 encodes a proton pump which affects pH in the endosomal compartment and it was recently reported that changes in Golgi pH can impair protein sialylation giving a possible mechanism for the observed association. Using ChIP and RNAi we showed HNF1a to be a master regulator of N-glycan fucosylation by controlling the levels of various fucosyltransferases and enzymes needed to produce GDP-fucose in HepG2 cells.

Mutations within HNF1a cause a mendelian form of type 2 diabetes; Maturity Onset Diabetes of the Young (MODY). Treatment differs between HNF1a-MODY and type 2 diabetes and currently the only definitive diagnostic tool is DNA sequencing. We investigated the role of glycans as biomarkers for prioritisation of potential HNF1a-MODY patients for molecular diagnosis. C-statistic measures of discriminative accuracy were 0.94 for HNF1a-MODY vs type 1 diabetes and 0.90 for HNF1a-MODY vs type 2 diabetes for our best marker which translates into 88% sensitivity and 83% specificity to discriminate HNF1a-MODY from other types of diabetes. A few additional glycan biomarkers had very high C-statistics (>0.95) but were less accurate using serum vs plasma.
P-355
THE USE OF A MULTI-LOCUS MIXED MODEL APPROACH FOR GWAS REVEALS ASSOCIATIONS FOR METABOLIC TRAITS IN THE TOMATO, SOLANUM LYCOPERSICUM

Christopher Sauvage1, Vincent Segura2, Guillaume Bauchet1, Rebecca Stevens1, Thi Phuc Do3, Alisdair Fernie1, Mathilde Causse1
1INRA, Montfavet, France, 2INRA, Orleans, France, 3Max-Planck-Institut fur Molekulare Pflanzenphysiologie, Golm, Germany

Genome-wide association studies have been successful in identifying genes involved in polygenic traits notably in crops, towards their improvement. We propose to apply in a major crop, such as tomato, a recently developed multi-locus mixed model as a general method for mapping complex traits in structured populations (Segura et al., 2012). Among major crops, tomato (Solanum lycopersicum), is an highly valuable vegetable in the world for its health value (Causse et 2010). SNP beadchip (Hamilton et al., 2012) are available and enable GWAS for traits of interest. Recently, a pilote study defined the optimal conditions for GWAS by using cherry tomato accessions (Ranc et al., 2012).

In our study, we examined a core collection of 180 tomato varieties composed of 20 wild accessions (S. pimpinellifolium), 130 admixed accessions (S. cerasiforme) and 30 domesticated accessions (S. lycopersicum). Multi-locus GWAS analysis was conducted using the MLMM package (Segura et al., 2012) with 7700 SNP markers and a set of sugar-related, Vitamin C-related and morphological traits as well as a broad range of metabolites related to central carbon metabolism.

The present study is the first one in tomato reporting associations for a large set of traits at the genome scale. We found significant associations for 89 loci with a total of 19 traits including fresh weight, sucrose, ascorbate dehydrogenase, malate or citrate, notably. Identified loci were also concordant with published quantitative trait loci (i.e. Malate), while new loci were identified (for Tocopherol). Moreover, several related metabolites, such as citrate and malate (both involved in Krebs cycle) displayed two identical associations.

These results (1) provide a list of candidate loci to be functionally validated and (2) provide a powerful analytical approach for finding genetic variants that can be directly used for crop improvement and deciphering the genetic architecture of complex trait.

P-356
THE USE OF INFINITESIMAL MULTIPlicative MODEL TO INVESTIGATE THE DETECTION OF EPISTASIS IN CROSSBRED ANIMAL POPULATIONS

Asko Maki-Tanila
MTT Agrifood Research Finland, Jokioinen, Finland

There is plenty of evidence on animal populations that quantitative genetic variation is based on a very large number of loci with infinitesimally small effects. We can consider epistasis as a synergy or positive interaction between loci and use a multiplicative model. The analysis on the multiplicative infinitesimal model is restricted to the levels of interaction resulting in the maximum coefficient of phenotypic variation of 40%. Then with a very large number of loci only up to 2-5% of the genetic variation is due to epistasis. Such low proportions would be hard to detect. On the other hand, it is not uncommon to find epistatic variation in crossbred populations. The infinitesimal multiplicative model is not giving any support for such observations as the coefficient of variation is not expected to be any higher due to crossing. The QTL investigations with crossbreeding designs involving very diverge selected purebred lines or populations have findings on apparent epistatic effects. In the purebred populations, there initially may have been rare alleles or more recently mutations favoured by selection. The fixation of such alleles would be accelerated when there is positive interaction with the polygenic background. It is most likely that fixation may occur at different loci in the lines. Depending on the size of the effect and the level of interaction, such QTL alleles could be easily detected in an F2 design. With moderate gene and
interaction effects due to the major loci, up to 10-20% of the genetic variation could be due to epistasis. The crosses of highly diverged populations are most suitable for searching epistatic QTL, as it is likely that at several loci there would be segregation of major alleles (with equal frequency) with their impact being augmented by positive interaction with the rest of the genome.

P-357
INFERENGE OF POPULATION STRUCTURE IN A SUGARCANE PANEL USING SNP DATA

Renato Rodrigues Silva¹, Luciano Costa e Silva¹, João Ricardo Bachega Feijó Rosa¹, Thiago Gibbin Marconi², Melissa de Oliveira Santos - Garcia³, Anete Pereira de Souza³, Antonio Augusto Franco Garcia¹
¹University of São Paulo/ ESALQ, Piracicaba, São Paulo, Brazil, ²University of Campinas, Campinas, São Paulo, Brazil

Although there already exist previous studies inferring population structure in sugarcane, none have yet used SNP markers generated from high-throughput platforms. Our goal in this study is to infer population structure in sugarcane using SNP markers. We scored 1,033 SNP markers on a panel of 142 accessions using the Sequenom iPLEX MassARRAY platform. Each biallelic SNP was recorded as a pair of scalar values, each value being the intensity of an allele which is expected to be proportional to allelic dosage. SNP genotype calling was made using software SuperMASSA. The model implemented in this software allows to infer genotypes even whether ploidy is unknown. The population structure was analyzed via principal components analysis (PCA) and UPGMA hierarchical clustering method, based on Roger’s modified distance matrix. The top two principal components revealed some structure. The first component separated out RB72454 offsprings from others and allowed to infer presence of outlier IN84-58 (Saccharum spontaneum). The second component allocated various SP80 cultivars (half or full siblings) apart from other accessions. Results from cluster analysis indicated a group comprised by the same SP80 cultivars identified through PCA. Due to PCA removing noise from data, it was more effective than UPGMA method to identify structure. Our PCA results agree with the genealogy of accessions and served well to find population structure in this study. This analysis showed some evidence of narrow genetic distances between accessions, corroborating the known reduced genetic basis of this panel, mostly due to recurrent crosses between related individuals. We hope these informations could be important for association mapping.

P-358
CROSS-SPECIES EQTL-MAPPING TO IDENTIFY HOST-PATHOGEN INTERACTIONS

Dahlia Nielsen¹, David Bird¹, Valerie Williamson²
¹North Carolina State University, Raleigh, NC, USA, ²University of California, Davis, Davis, CA, USA

Expression QTL (eQTL) mapping, in which QTL mapping is performed using gene expression measurements as the quantitative traits of interest, allows investigators to identify genetic loci involved with gene expression regulation. This approach has been successfully applied in a number of different organisms. We have extended this approach to examining interactions across species: we consider a mapping population of pathogens during active infection of a host, and assay both organisms simultaneously. In this way, we are able to make connections between genotypes of the pathogen and gene expression levels of the host during an active infection. We are using the model plant-pathogen system Medicago truncatula and the root knot nematode Meloidogyne hapla. A mapping population of M. hapla has been developed; we use these lines to inoculate genetically homogeneous plants, with each plant receiving an individual nematode line. Infected root samples are dissected at a fixed time point post-inoculation, and quantification is performed by high-throughput sequencing of cDNA (RNA-seq). This technology allows us to assay transcription levels of both the host and the pathogen, in addition to providing information on sequence variation, such as SNPs, within the pathogen genome. Results
comparing plants infected with the parents of the nematode mapping population will be presented, along with some preliminary results from the mapping population.

P-359
A NOVEL APPROACH FOR STUDYING THE GENETIC BASIS OF INBREEDING DEPRESSION IN NATURAL POPULATIONS

Holger Schielzeth, Jochen B.W. Wolf, Bart Kempenaers, Wolfgang Forstmeier, Hans Ellegren
Uppsala University, Uppsala, Sweden, Max Planck Institute for Ornithology, Seewiesen, Germany

Objectives: Heterozygosity-fitness correlations (HFC) and inbreeding depression are common phenomena that have been documented in many natural animal and plant populations. Their genetic basis, however, has remained elusive. Two possible underlying genetic dominance architectures can give rise to HFC and inbreeding depression: overdominance (‘overdominance hypothesis’) and directional overdominance caused by recessive deleterious mutations (‘dominance hypothesis’).

Methods: We studied HFC in a captive population of almost 1,200 zebra finches that was maintained at low levels of inbreeding. We focused on 11 traits that have been found to show positive HFC, albeit with varying effect sizes. Individuals were genotyped for more than 1,200 informative autosomal SNP markers. We developed a novel two-step analytical approach that begins with marker-by-marker analysis and proceeds by evaluating the overall pattern of small effects across the whole genome.

Results and Conclusions: Local effects were generally small, which suggests a highly polygenic basis of HFC in our population. We find that several morphological traits show evidence of directional dominance in the genome-wide analysis across all markers. Although our analytical is very data-hungry, it holds great promises to better understand the genetic basis of HFC in natural populations that have been genotyped for biallelic markers.

P-360
INVESTIGATING SOURCES OF HERITABILITY IN A YEAST CROSS

Joshua Bloom, Ian Ehrenreich, Leonid Kruglyak
Princeton University, Princeton, NJ, USA, University of Southern California, Los Angeles, CA, USA

The genetic basis of complex traits remains poorly understood. We dissect the genetic architecture of many complex traits using a large panel of S. cerevisiae segregants from a biparental cross. Our experimental design has high statistical power for QTL detection and allows us to address many of the basic questions about the genetics of complex traits, including the number of loci underlying a trait, the distribution of allelic effect sizes, the prevalence of genetic interactions, and the distribution of allele frequencies in a population. To facilitate linkage mapping we develop a high-throughput phenotyping assay to measure growth for thousands of individual yeast strains and a novel method for rapid, very highly multiplexed genotyping of these strains using whole genome sequencing. We estimate the relative contributions of additive and non-additive genetic components of trait variance, providing limits and expectations for QTL models that aim to resolve these components of variance to specific genetic loci.

P-361
GENOMIC DISSECTION OF INBREEDING DEPRESSION: LESSONS FROM THE BULL FERTILITY TRAITS

Ino Curik, Maja Ferencakovic, Johann Soelkner
Department of Animal Science, Faculty of Agriculture, University of Zagreb, Zagreb, Croatia, Department of Sustainable Agricultural Systems, Division of Livestock Sciences

Regression of performance (phenotype) on the individual pedigree inbreeding coefficients has been the
standard procedure used to quantify inbreeding depression in humans and animals. In Austrian Fleckvieh bulls we have estimated inbreeding depression for the sperm quality traits (volume, sperm concentration, percentage of viable spermatozoa and motility) with emphases on the comparison of models based on pedigree inbreeding versus models based on various genomic estimates of inbreeding (BovineSNP50 BeadChip). Furthermore, we also analyzed models with chromosomal inbreeding coefficients (1 to 29) as well as models with combination of several chromosomal inbreeding coefficients. Mixed models were used with a bull as a random effect and all other effects (age of the bull, semen collector, month and year of collection and number of ejaculates per bull per day), including pedigree or genomic inbreeding coefficients (covariable) as fixed effects. Inbreeding depression was found for the volume and percentage of viable spermatozoa, although between two artificial insemination centres some results were heterogeneous. In general, models based on pedigree were comparable to models based on genomic inbreeding (runs of homozygosity) while models based on individual heterozygosity were inferior. In contrast, for all traits analyzed, models with various chromosomal inbreeding coefficients outperformed models based on overall genomic inbreeding as well as models based on pedigree. However, while genomic inbreeding coefficients have been shown to be powerful, still much work has left to complete our understanding on architecture of inbreeding depression.

P-362
DETECTION OF TRACES OF SELECTION WITH NUMEROUS SNP IN SMALL EXPERIMENTAL CHICKEN POPULATIONS UNDERGOING DIRECTIONAL SELECTION
Frédéric Hospital, Bertrand Bed’Hom, Mathieu Gautier, Licia Silveri, Nicolas Bruneau, Jean-Luc Coville, David Gourichon, Marie-Hélène Marie-Hélène
INRA - GABI, Jouy en Josas, France

Three lines of White Leghorn Chickens have been selected for 12 generations for one of three different immune response traits, high antibody response (ND3), cell mediated activity (PHA) and phagocytic activity (CC). Line ND3-L was selected on ND3, line PHA-L was selected for PHA, and line CC-L for CC. A fourth line was a contemporary random bred Control maintained throughout the selection experiment (Minozzi et al. 2008). Each generation, 200 chicks per line were hatched (800 chicks in total) in a single batch. Selection for each trait was done by within-family mass selection based on individual phenotype. Heritabilities estimated for the three selection criteria ND3, PHA and CC were 0.35, 0.13 and 0.15, respectively, and correlations between the traits were not significant.

Individuals from the three selected lines and the control line at generation G9, as well as individuals from the founding population (G0) were sampled (about 20 individuals/line) and genotyped with a 60K SNP chip. We present the use of this dataset to detect traces of selection in the three selected Chicken lines. An original Bayesian method was designed to detect the possible effects of selection by comparing the SNP allele frequencies between generations G0 and G9 for each line. The method was able to pinpoint a dwarfing gene that is known to have undergone strong selection in the experiment – hence serving as a validation control. In addition it highlights numerous SNPs that seem to behave non-neutrally, hence providing candidate regions for future search for selected genes.

While classical approaches generally focus on ‘historical’ (long term) traces of selection, this experiment demonstrates that it is possible to detect short-term selection in experimental population using SNPs. Minozzi, G, et al., BMC GENETICS, 9:5, 2008.
P-363
GENETIC ARCHITECTURE ANALYSIS OF MULTIPLE TRAITS FOR A. THALIANA BASED ON WHOLE GENOME SEQUENCING

Quan Long1, Bjarni Vilhjalmsson1, Dazhe Meng1, Qingrun Zhang1, Oliver Stegle2, Alexander Platzer1, Viktoria Nizhynska1, Envel Kerdaffrec1, Arthur Korte1, Laura Sedman1, Fernando Rabanal1, Nick Fulcher1, Glenda Willems3,4, Ortrun Mittelsten Scheid1, Karel Riha1, Magnus Nordborg1,3
1Gergor Mendel Institute, Vienna, Austria, 2Max Planck institues in Tübingen, Tübingen, Germany, 3University of Southern California, Los Angeles, USA, 4SESVANDERHAVE, Tienen, Belgium

Objectives: Understanding the general genetic architecture (instead of identifying particular variants that are causal) of Arabidopsis thaliana traits based on full genomes.

Methods: We sequenced ~240 Arabidopsis thaliana inbred lines with Illumina at deep coverage (ranging from 20 to 60) and measured multiple phenotypic traits like germination, genome size from flow cytometry, telomere length, as well as salt tolerance. Together with many published phenotypes including flowering time etc., we can compare the genetic architecture of different traits. To analyze the genetic architecture, we developed a linear mixed model in which there are two random terms that are assumed to be Gaussian with covariance matrices as local and global kinship matrices calculated by the variants detected from sequences. In contrast to the standard mixed model, we do not include fixed terms (except for the intercept). The rationale is that, hoping the global kinship term can guard against population structure, the local kinship term provides an estimate of phenotypic variance explained by the focal genomic region. This model is similar to local variance component analysis in Visscher’s lab (Yang et all Nat Genet 2010), however jointly accounting for the population structure. In our analysis, the relative contributions of local and global kinship terms to the heritability will be estimated by maximal likelihood ratio as well as a Bayesian approach.

Results: Different traits perform quite differently in the genetic architecture analysis. Some are controlled by many small-effect regions and therefore can be assumed to be approximately infinitesimal; while some are contributed by a limited number of big-effect regions. Some traits do not show big difference when using SNPs only or indels included to calculate local kinship; while some do have more phenotypic variance explained when indels are considered. Some traits show genetic heterogeneity with multiple different variants contributing to the variance independently. Some traits are likely to be altered by gene-gene interactions.

Conclusions: Sequence based whole genome analysis reveals important insights into genetic architecture of A. thaliana.

P-364
WHOLE-GENOME SEARCHING FOR NON-ADDITIVE GENETIC VARIANCE

Matthew Conomos, Bruce Weir
University of Washington, Seattle, WA, USA

Objective: In a recent series of papers, Visscher and colleagues have used whole-genome SNP arrays to estimate relatedness and inbreeding coefficients and then used these estimates to estimate additive genetic variation for traits such as human height. They have been able to account for much of the “missing heritability” that results from concentrating only on SNPs that reach genome-wide significance thresholds for association with these traits. We have extended this approach to allow for non-additivity, both dominance and additive by additive.

Method: As a preliminary step we have had to estimate the full set of nine coefficients of identity by descent (Jacquard coefficients), using maximum likelihood as described previously by Anderson and Weir. We use human height data collected by the GENEVA consortium to illustrate our approach.

Results: We find evidence for non-zero values of all none identity-by-descent coefficients, and this leads to non-zero components of non-additive genetic variance.
Conclusion: The methodology of Visscher and colleagues can be extended to allow for non-additive genetic variance.

P-365

INVESTIGATION OF POSSIBLE QUANTITATIVE TRAIT MODIFIERS FOR AGE-RELATED MACULAR DEGENERATION IN THE THIRD NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY AS PART OF THE EPIDEMIOLOGIC ARCHITECTURE FOR GENES LINKED TO ENVIRONMENT (EAGLE)

Nicole Restrepo1, Kristin Brown-Gentry1, Kylee Spencer2, Ping Mayo1, Melissa Allen1, Nathalie Schnetz-Boutaud3, Dana Crawford4
1Vanderbilt University, Nashville, TN, USA, 2Heidelberg University, Tiffin, Ohio, USA

Age-related Macular Degeneration (AMD) is the leading cause of blindness in elderly Americans. The most widely replicated loci are \textit{CFH} Y402H and \textit{ARMS2}, yet little data exist for these variants in diverse populations. A major goal of the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) is to describe the genetic architecture of common diseases across diverse populations. Targeted genotyping was performed for AMD-associated SNPs in \textit{CFH} (rs1061170), \textit{C2} (rs547154), and \textit{ARMS2} (rs10490924) in 7,159 participants from the Third National Health and Nutrition Examination Survey (NHANES III), a cross-sectional, population-based survey of Americans that includes non-Hispanic whites (W; 2,631); non-Hispanic blacks (B; 2,108), and Mexican Americans (MA; 2,073). Overall, \textit{CFH} rs1061170 (C) was less frequent in MA (0.19) compared with W (0.37) and B (0.36). Conversely, the \textit{ARMS2} rs10490924 (T) frequency was similar across the three groups: 0.22 (W), 0.24 (B), and 0.25 (MA). \textit{C2} was more frequent in (B) 0.20 compared with (W) 0.11 and (MA) 0.11. AMD cases were defined as participants >60 years presenting with early or late AMD determined by fundus photography. Controls were participants >60 years without AMD. For W, B, and MA, a total of 190, 30, and 47 cases and 664, 209, and 270 controls were identified, respectively. As previously reported, \textit{CFH} rs1061170 (OR=1.32; p=0.03) and \textit{ARMS2} rs10490924 (OR=1.89; p=0.0001) were associated with AMD in W in logistic regression adjusted for age, sex, BMI, and smoking status while \textit{ARMS2} rs10490924 was associated with B (OR=0.43; p=0.04) and MA (OR=1.6; p=0.04). We tested for gene-environment interactions involving quantitative traits previously associated with AMD risk (carotenoids, vitamin A, beta carotene, alpha carotene, and BMI). Interestingly in W, BMI modified AMD risk through \textit{C2} rs547154 (OR\textsubscript{interaction}=0.86; p=0.0004). Although underpowered, the present study suggests quantitative traits known to be associated with AMD risk may modify genetic risk.

P-366

DEVELOPMENT OF MOLECULAR BREEDING FOR WOOD PHYSICAL TRAITS IN CANADIAN SPRUCES: POTENTIAL OF ASSOCIATION GENETICS AND GENOMIC SELECTION

Jean Beaulieu1,2, Trevor Doerksen1,2, John MacKay2, Jean Bousquet2
1Natural Resources Canada, Canadian Wood Fibre Centre, Quebec City, Quebec, Canada, 2Université Laval, Arborea, and Centre for Forest Research, Quebec City, Quebec, Canada

Spruces extend over most of Canada’s boreal biome and are the most important forest resource for the forest industry. There are active spruce breeding programs and emphasis has been put on improvement of growth and adaptive traits. The inclusion of wood quality traits in selection criteria is just beginning. Given the time delays and high costs for evaluating these traits, genomics-assisted selection could contribute positively in increasing gains per unit of time. In order to develop molecular breeding tools for wood traits in white spruce, a discovery population of 1700 white spruces has been assembled from a 30-year-old provenance-progeny trial comprising 215 open-pollinated families. Increment cores were collected and wood physical traits assessed using SilviScan technology. A genotyping assay including SNPs belonging to about 2500 candidate genes was developed to genotype the entire population. A total of about 7000 SNPs

241
were retained for analysis as they met quality and minor allele frequency criteria. Association studies were carried out using a mixed linear model approach on single and multiple SNPs. Results of SNP-by-SNP analyses showed that between 325 and 600 SNPs were significantly associated ($P \leq 0.05$) with single traits related to earlywood, latewood or total wood before correction for false-discovery rate. The variation explained by most significant SNPs for each trait varied between 1.1 and 3.3%. Multilocus Bayesian mixed models analyses were also tested and the results indicated that about 20 to 25% of the phenotypic variation could be explained using 40 to 60 SNPs per trait.

Prediction of genetic effects was also tested with models that included 1) pedigree information, 2) genomic information, and 3) both pedigree and genomic information. The predictive value of the models was estimated by cross-validation with within-family samples. Estimates of accuracy (correlation between individual genetic values and their estimates) were weak to moderate, in agreement with expectations for this population in low linkage disequilibrium. Nevertheless, their amplitude suggests that economically useful gains could be obtained if selection delays are reduced.

**P-367**
THE RE-DISCOVERY OF THE DOMINANCE VARIATION BY USING THE OBSERVED RELATIONSHIP MATRIX AND ITS IMPLICATIONS IN BREEDING

*Patricio Munoz*, *Marcio Resende Jr*, *Marcos Resende*, *Salvador Gezan*, *Matias Kirst*, *Gary Peter*

*1University of Florida, Gainesville, Florida, USA, 2EMBRAPA Forestry, Colombo, PR, Brazil, 3Universidade Federal de Viçosa, Viçosa, MG, Brazil*

Dominance effects have been neglected in animal and plant breeding, mainly because the variance estimates are usually small and non-significant compared to the additive effect. Consequently, many breeding programs rely only on breeding values to make progress. It is well known that the partition of the genetic variance into additive, dominance and epistasis is not perfect and estimations are confounded with each other. The Objective of this study is to test the use of a relationship matrix derived from molecular markers (O) to separate the genetic variance components. For the complex trait tree height we compare different BLUP models under additive and full (additive plus non-additive) assumptions with the use of either the pedigree relationship matrix (A) or the O matrix in a clonal breeding population of *Pinus taeda*. Our results show that the additive-BLUP model with the O matrix (add-O) increased the accuracy and heritability by 10% compared to additive-BLUP with the A matrix (add-A). A full-BLUP model with A matrix (full-A) could not partition the variance components, leading to the conclusion that almost all the genetic variation is additive (80%). However, a full-BLUP with O matrix (full-O) partitioned the genetic components, changing the conclusions about the proportion of additive (40%) and dominance (41%) components. With ten-time cross validation on all models, add-A (0.64), add-O (0.67) and full-A (0.65) have similar BV-predictive ability, but the full-O model (0.84) showed a predictive ability 26% greater than add-O. The increase in prediction capacity is not only due to the use of the O matrix but also to the correction for non-additive effects. We show that the use of the O matrix in a BLUP context (GBLUP) efficiently partitions the genetic variance components, changing from the purely additive conclusions to a scenario where dominance and additive effects are equally important.

**P-368**
MULTIPLE RISK FACTORS FOR HUMAN CARDIOVASCULAR DISEASE AND OBESITY ARE HERITABLE IN A SINGLE, EXTENDED PEDIGREE OF RHESUS MACAQUES

*Amanda Vinson*, *Asia D. Mitchell*, *Jacob Silver*, *David Toffey*

*1Oregon National Primate Research Center, Oregon Health & Science University, Beaverton, OR, USA, 2Department of Molecular and Medical Genetics, Oregon Health & Science University, Portland, OR, USA*

Large, managed populations of rhesus macaques are a powerful resource for dissecting the genetic
architecture of quantitative risk factors for complex human disease. Indeed, macaques have natural advantages that make them ideal for this purpose, including their substantial genetic and physiological similarity to humans, breeding strategies and short generation times that quickly produce large paternal half-sibships and other informative relative classes, and their location within a rigorously-managed environment that will likely enhance genetic signal over noise. Based on these advantages, we have characterized an unascertained pedigree of Indian-origin rhesus macaques for quantitative genetic analysis and gene-mapping of complex phenotypes. This pedigree was developed to optimize several criteria: animals must have parentage confirmed by genotyping, belong to a minimum 3-generation lineage, be available for sampling, not have common ancestors outside the pedigree, and must form a single, extended pedigree structure. The resulting pedigree contains 1,289 macaques spanning 6 generations, with 800 females and 489 males. Using a maximum likelihood-based variance decomposition approach, power analyses indicate >80% power to detect heritabilities as low as 0.10, and QTL effect sizes as low as 0.14 with a LOD score of 3. Using blood samples and multiple measures of morphometry and adiposity collected on >850 of these macaques to date, we assessed heritability for cholesterol levels and two risk factors for obesity in a subset of this pedigree. We found significant heritability for total cholesterol (h²=0.257, P=0.032, N=193), LDL cholesterol (h²=0.252, P=0.030, N=193), and triglyceride levels (h²=0.197, P=0.034, N=193), as well as abdominal circumference (h²=0.286, P=0.002, N=476), and body mass index (BMI; h²=0.292, P=0.00005, N=479). These results indicate that additive genetic effects on these phenotypes can be detected and measured in this pedigree, and support the need for expanded analysis of these and other important risk factors for complex human disease in this population.

P-369
NEXT GENERATION EXOME SEQUENCING OF PAEDIATRIC INFLAMMATORY BOWEL DISEASE PATIENTS IDENTIFIES RARE AND NOVEL VARIANTS IN CANDIDATE GENES

Katja Christodoulou¹, Anthony Wiskin², Jane Gibson¹, William Tapper¹, Claire Willis², Nadeem Afza², John Holloway¹, Michael Simpson³, R Mark Beattie², Andrew Collins¹, Sarah Ennis¹
¹University of Southampton, Southampton, UK, ²University Hospital Southampton NHS Foundation Trust, Southampton, UK, ³King’s College London, London, UK

Objectives: Multiple genes have been implicated by association studies in altering inflammatory bowel disease (IBD) predisposition. Paediatric patients often manifest more extensive disease and a particularly severe disease course. It is likely that genetic predisposition plays a more substantial role in this group. We applied exome sequencing analysis to identify the spectrum of rare and novel variation in known IBD susceptibility genes in eight individual cases of childhood onset severe disease.

Methods: DNA samples from the eight patients underwent targeted exome capture and sequencing. Data were processed through an analytical pipeline to align sequence reads, conduct quality checks, and identify and annotate variants where patient sequence differed from the reference sequence. For each patient, we catalogued the entire complement of rare variation within strongly associated candidate genes.

Results: Across our panel of 169 known IBD susceptibility genes, we found approximately 300 variants in 104 genes. Excluding splicing and HLA-class variants, 59 variants across 40 of these genes were classified as rare, with an alternative allele frequency of <5%, of which 17 were novel.

Conclusions: For each of the eight individuals studied we have identified all non-synonymous, truncating and frameshift mutations across all known IBD genes. Only two patients with early onset Crohn’s disease exhibit rare deleterious variations within NOD2; the previously described R702W variant is the sole NOD2 variant in one patient while the second patient also carries the L1007 frameshift insertion. Both patients harbour other potentially damaging mutations in the GSDMB, ERAP2 and SEC16A genes. The two patients severely affected with ulcerative colitis exhibit a distinct profile whereby both carry detrimental variation in the BACH2 and IL10 genes not seen in other patients. A unique profile of rare and potentially damaging variants is evident for each patient with this complex disease.
P-370
THE ARABIDOPSIS RECOMBINATION LANDSCAPE IS ROBUST AGAINST HERITABLE EPIGENETIC PERTURBATIONS

Maria Colomé-Tatché1, René Wardenaar1, Sandra Cortijo2, Ritsert C. Jansen1, Vincent Colot2, Frank Johannes1
1Groningen Bioinformatics Centre, University of Groningen, Groningen, The Netherlands, 2Institut de Biologie de l’École Normale Supérieure, Paris, France

Recombination is an essential aspect of genome evolution. In Arabidopsis thaliana, meiotic cross-over patterns are tightly controlled along chromosomes and correlate with DNA methylation and sequence features. Here we test the robustness of these patterns in response to an experimental perturbation of global cytosine methylation in a nearly isogenic population of epigenetic recombination inbred lines (epiRILs). Using genome-wide methylation data from 126 epiRILs we derive a core recombination map on the basis of 867 meiotically stable Differentially Methylated Regions (DMRs) covering 91.5% of the total genome. Despite the small number of sequence polymorphisms between the lines, and the segregation of many large hypomethylated sequences, global and local recombination rates in the epiRILs were remarkably similar to those of 17 recently published F2 natural populations. Our data reveals a highly streamlined recombination landscape in this species.

P-371
DATA MIRRORING BAYESIAN INFERENCE FROM NEAR-COMPLETE SEQUENCES

Miguel Pérez-Enciso2,1, Sebas Ramos-Onsins1
1Center for Research in Agrigenomics, Bellaterra, Spain, 2ICREA-Universitat Autonoma Barcelona, Bellaterra, Spain

Thanks to recombination, the complete sequence from even a single individual recapitulates the demographic history of the whole population to which it belongs. Now this topic has become a most pressing challenge given the high throughput of current sequencing technologies. Here we have developed an approximate Bayesian computation (ABC) approach to infer genetic parameters that is suited for both pooled and individual sequences. First we consider the most difficult scenario, that of pooling. Note that, for a given depth x, the actual number of chromosomes j sequenced in a pool can vary between 1 and min(x, nc), nc being the total number of chromosomes in pool. Rarely, j approaches nc, even at high fold x. Moreover, depth varies continuously and stochastically along the genome. To circumvent these difficulties, we propose to simulate data that mirrors exactly the depth and coverage actually found. The method consists of splitting the genome in discrete windows, simulate complete data of window’s length via coalescence or forward algorithms, and simulate observed data following exactly the same pattern as observed. For the statistics, we consider the number of SNPs, number of fixed differences with respect to an outgroup, and other statistics based on allele frequency spectrum. Next we perform ABC for each window. There are several options to combine the different posterior densities, we propose to consider simply the average posterior across windows because of its simplicity and numerical stability. Finally, windows – or annotated regions – that depart from the average posterior can be selected for follow up as potential candidates for selection. We illustrate the method with data from a pool of 9 Iberian pigs sequenced at an average 7x and a single individual (4x approx.), together with Potamocherus porcus sequence as outgroup. Previous data reveals that porcine NPAR is a depauperate place in terms of genetic variability so we wish to clarify whether selection or demographic effects are responsible for this observation.
P-372
RECOVERY OF THE GENETIC BACKGROUND IN INTROGRESSED POPULATIONS

Carmen Amador1, Miguel Angel Toro2, Theo HE Meuwissen3, Jesus Fernandez1
1INIA, Madrid, Spain, 2ETSIA, UPM, Madrid, Spain, 3UMB, Ås, Norway

Maintenance of populations genetically differentiated can be crucial in both wild and domestic species for several reasons (conservational purposes, economic interest, etc.). If a population, intended to be preserved, becomes introgressed by individuals from a foreign population, methods for recovering the native genetic background must be implemented. This study tested, through computer simulations, the power of recovery of a native population under several scenarios where a variable number of exogenous individuals mixed for a few generations (using different sources of information). A first group of simulations considered the whole pedigree as known, and strategies based on founders’ representation were implemented. The second group of simulations employed different sets of molecular information (and molecular measures) in order to choose from the available individuals those contributing to next generation. Both types of information led to a substantial recovery of the genetic background, with limitations. However, the former implied pedigree has to be completely recorded, which is not always possible. The efficiency of the markers was highly dependent on the number of alleles and the difference of their frequencies between the native and the foreign population. The main side effect of the methods was a significant increase in the inbreeding of the resulting population. When genome-wide information is available, an alternative and successful strategy is relying on molecular coancestry to determine contributions, which allows for substantial recovery of native background with a better control of inbreeding increase. Results showed that certain recovery of an introgressed genetic background is possible by applying different methods that use the available information. An increase of inbreeding may become an unavoidable side effect, and the time of acting is crucial to maintain the largest possible amount of native genome.

P-373
ELEVATED CONCENTRATION OF SOLUBLE UROKINASE-TYPE PLASMINOGEN ACTIVATOR RECEPTOR IN THE BLOOD OF PATIENTS WITH PELVIC INFLAMMATORY DISEASE

Hsiu Ting Tsai, Shun-Fa Yang, Po-Hui Wang
School of Nursing, Chung Shan Medical University, Taichung, Taiwan

Objectives: The purpose of this study was to estimate the expressions of urokinase-type plasminogen activator (u-PA), soluble urokinase-type plasminogen activator receptor (su-PAR), plasminogen activator inhibitor-1 (PAI-1) in plasma, and the gene polymorphisms in patients with pelvic inflammatory disease (PID) and healthy controls.

Methods: The enzyme-linked immunosorbent assay (ELISA) and polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) were respectively used to measure the plasma levels and polymorphisms of u-PA, u-PAR, and PAI-1 among seventy healthy controls and in Sixty-four PID patients before and after they received routine treatment protocols.

Results: We found the level of plasma soluble u-PAR was significantly elevated in PID patients compared to that in normal controls and decreased significantly compared to that in same patients after they received treatment. The increased expression of u-PAR was significantly correlated with the cell counts of WBC in blood and the levels of plasma CRP as well as u-PA of PID patients before they received treatment, also, significantly correlation between plasma concentration of u-PA and u-PAI-1 was found among PID patients before they received treatment. There was no association between genetic polymorphisms and their gene expression levels and PID susceptibility.

Conclusions: Elevated plasma u-PAR could be a biological marker for the diagnosis and to be a new strategy for target therapy of pelvic inflammatory disease.
EVALUATION OF A GENETIC DIFFERENTIATION APPROACH TO SELECTION MAPPING

Pam Wiener¹, Mohammad Edriss², John Williams¹, John Woolliams¹, Andy Law¹, Beatriz Gutierrez³, Alan Archibald¹, Chris Haley⁷,¹, Hendrik-Jan Megens⁵, Richard Crooijmans⁵, Martien Groenen⁵, Rob Ogden⁶, Samantha Wilkinson¹

¹The Roslin Institute, Edinburgh, UK, ²Isfahan University of Technology, Isfahan, Iran, ³Parco Tecnologico Padano, Lodi, Italy, ⁴Universidad de León, León, Spain, ⁵Wageningen University, Wageningen, The Netherlands, ⁶Royal Zoological Society of Scotland, Edinburgh, UK, ⁷Western General Hospital, Edinburgh, UK

Scanning the genome with high density SNP markers has become a standard approach for identifying regions of the genome showing substantial between-population genetic differentiation, and thus evidence of diversifying selection. This is potentially a valuable (and cost-effective) approach for identifying loci with large phenotypic effects. However, few studies have attempted to address the power or efficacy of such an approach. In this study, the patterns of allele frequency differences between livestock breeds were evaluated in light of results from linkage mapping studies in experimental populations, in which regions of the genome associated with trait differences between founders have been identified. In a detailed comparison of data from two cattle breeds, the overall consistency between genetic differentiation and QTL signal from a linkage mapping analysis was low. While concordance between the two datasets was seen for chromosomes carrying QTL with strong statistical support, such as those carrying genes associated with coat color, an association was not seen in a broader comparison including chromosomes carrying less significant QTL. Furthermore, markers with substantial between-breed genetic differentiation were not found in some chromosomal regions with QTL explaining a large proportion of the phenotypic variation. Investigation of other livestock data revealed similar patterns. These results suggest that simple genomic diversity scans may be limited in their power to detect regions that are associated with quantitative phenotypic differences between populations and that improved methods will be required.

DOES SEQUENCING AT LOW COVERAGE PRODUCE RELIABLE GENOTYPES IN DAIRY CATTLE?

Hans D. Daetwyler¹, Charlotte L. Anderson¹, David Coote², Amanda J. Chamberlain¹, Ben J. Hayes¹,³

¹Biosciences Research Division, Department of Primary Industries Victoria, Bundoora, Vic, Australia, ²Department of Forest and Ecosystem Science, The University of Melbourne, Richmond, Vic, Australia, ³La Trobe University, Bundoora, Vic, Australia

The use of sequence data to identify variants for subsequent genomic prediction and genome-wide association offers great promise. A key question is the depth of coverage required for reliable results. Sequencing at lower fold coverage has been proposed to reduce costs, but this could also increase the proportion of heterozygotes erroneously called as homozygotes. Wrong calls may be reduced using population level data through phasing algorithms which can accommodate genotype probabilities. We compared four fold (4x) and higher fold (HF) coverage sequence genotype calls in a sample of 57 Holstein and Jersey individuals in a 1.61Mb region on chromosome 18. Sequence reads from Illumina HiSeq were aligned using BWA and animals sequenced at HF coverage were cut to 4x using Picard. The HF mean coverage was 10.1x (range 4.0 – 52.4). Variants (single nucleotide polymorphisms; insertions and deletions) were identified and genotypes were called in the 4x and full coverage data sets using Samtools-0.1.18. Filtered variants were then analysed with the Beagle phasing software to test whether it improved genotype calls. Two comparisons were performed pre and post Beagle. First, the 4x and HF genotypes were compared to Illumina Bovine 800k array genotypes in 36 individuals. Sequence 4x genotype concordance was increased by 8.0% using Beagle (mean 88.4%) and HF also increase by 8.1% (mean 91.3%). Second, 4x genotypes pre and post Beagle were compared to post Beagle HF genotypes in each animal. In the 9583 variants common to both sets, agreement of 4x with HF post Beagle genotypes
was improved by 5.5% (to 91.1%). We intend to repeat the experiment using a whole chromosome before the conference. Our results demonstrate a reduction in the accuracy of genotype calls with 4x fold coverage as expected. This reduction can be partially overcome using a phasing software such as Beagle.

P-376
PATTERNS OF MUTATION AND SELECTION LINKED TO NUCLEOSOME POSITIONING IN THE HUMAN LINEAGE

Colin Semple, James Prendergast
MRC Human Genetics Unit, Edinburgh, UK

Primate rates of sequence divergence are correlated with aspects of chromatin structure, including nucleosome positioning, as has been observed in other eukaryotes. However it has generally been unclear to what extent this phenomenon reflects variation in mutational spectra across the genome, or the action of selection. In a recent study we investigated the strengths and modes of selection associated with nucleosome positioning during recent evolution in the human lineage, through the comparison of interspecies and intraspecies rates of divergence. Although mutational biases appear to be present, we were also able to detect selection on the genomic sequences underlying nucleosomes for the first time. These signatures of selection were not restricted to nucleosomes in close proximity to exons, but appeared to be widespread across the genome. Mutational bias and recent selection linked to nucleosome positioning has important implications for human genome evolution and disease, but we could only speculate on the biology underlying these patterns. What are the likely functional consequences? Also, are there particular genomic regions that account for a disproportionately large amount of the selection detected? Here using additional data we further investigate nucleosome positioning and human variation to shed light on these questions.

P-377
IDENTIFICATION OF DELETIONS IN CHICKEN USING MARKER SEGREGATION IN AN F2 INTERCROSS

Anna Johansson1, Lucy Crooks1,2, Stefan Marklund1, Örjan Carlbort1
1Swedish University of Agricultural Sciences, Uppsala, Sweden, 2Wellcome Trust Sanger Institute, Hinxton, UK

Marker null alleles are alleles of a marker that lack a scorable phenotype in a genotyping assay. This can result from e.g. a mutation at a primer site, but also potentially indicate complete absence of the chromosomal segment containing the marker locus. When markers are scored, the presence of a null allele is indistinguishable from a deletion of the entire locus. Null-alleles are usually discarded in genome-wide genetic studies, but this might potentially lead to a loss of important information about large scale deletions segregating in the studied population. Here, we used data from 10,150 genotyped SNPs in two F2 families from an intercross between White Leghorn and Red Junglefowl chickens to study how common null-alleles are in the population and if any of these are caused by deletions. Several genomic regions with non-Mendelian inheritance of markers and patterns of missing values expected from null alleles were identified. In total 325 (3.2%) of the markers indicated possible null alleles. Approximately half were indicated by different homozygotes in parent and offspring and the other half by the pattern of missing values. We considered cases where there were null alleles in the same individuals at adjacent markers as deletion candidates. Four such regions, all involving several individuals, were found and their sizes ranged from 5.3 to 157.4 kb. To confirm the deletions, 22 individuals were genotyped using a larger SNP chip containing 57 636 SNPs. For these SNPs also intensity data were available. Sufficient genotype information was obtained in three of the four regions and based on this data two of the deletions were confirmed. By analysing data that would normally be discarded in a complex trait genetic study, we were
able to identify several large deletions that contain several genes that could contribute to the phenotypic difference between the studied populations.

**P-378**

**RNA POLYMERASE V - A NOVEL CAPACITOR OF PHENOTYPIC VARIATION IN PLANTS**

Janne Lempe1, Alessandra M.V. Sullivan1, Daniel A. Skelly1, Jennifer A. Lachowiec1, John A. Stamatoyannopoulos1, Josh M. Akey1

1Department of Genome Sciences, University of Washington, Seattle, WA, USA, 2Department of Plant Breeding and Genetics, Max Planck Institute for Plant Breeding Research, Cologne, Germany

Natural selection acts on phenotype rather than genotype. Yet most individuals express a robust wild-type phenotype despite considerable genetic variation among them. To understand how biological systems can be robust but flexible enough to evolve, we aim to characterize molecular mechanisms that increase phenotypic variation by revealing cryptic genetic variation and increasing developmental noise. Previous work demonstrated that perturbation of a single gene – the environmentally responsive chaperone HSP90 – reveals cryptic genetic variation and increases developmental noise in plants, flies, fish, and yeast. HSP90-dependent genetic loci occur frequently in complex seedling and mature plant traits, suggesting that buffered variation plays an important role in shaping development and possibly evolutionary trajectories.

We hypothesized that other similarly central cellular machines, specifically proteins functioning in the small RNA pathway, act like the capacitor HSP90 in revealing cryptic genetic variation and maintaining developmental stability. We identified RNA polymerase V as a novel capacitor. Pol V functions in small RNA-directed DNA methylation and maintenance of heterochromatin. When mutated, development was reproducibly destabilized and variation of quantitative traits increased under several experimental conditions. This increase in phenotypic variation was accompanied by an increase in expression space – the number of expressed base pairs was significantly higher compared to controls. Genetic analysis indicated a role of histone de-acetylation rather than DNA methylation in Pol V mediated phenotypic robustness.

We addressed whether HSP90 and Pol V affect phenotypic robustness in a similar or different manner by comparing revealed genetic loci and RNA seq data for plants reduced in either HSP90 or Pol V function. We found that buffered genetic loci differed between both capacitors whereas significant overlap in expression responses existed. Our data suggest that phenotypic buffering is modular at the whole-organism level but that plants with reduced robustness share common molecular features.

**P-379**

**QUANTITATIVE GENETICS OF GENE EXPRESSION**

Joseph Powell1,2, Anjali Henders3, Allan McRea1,2, Emmanouil Dermitzakis3, Grant Montgomery2, Peter Visscher1,4

1University of Queensland Diamantina Institute, Brisbane, Australia, 2Queensland Institute of Medical Research, Brisbane, Australia, 3University of Geneva Medical School, Geneva, Switzerland, 4The Queensland Brain Institute, Brisbane, Australia

Variation in gene expression has been convincingly shown to be controlled, in part, by genetic variants. Understanding the genetic architecture underlying natural variation in gene expression can provide insights in evolution and medicine. Here we present results from a study to investigate the underlying genetic architecture influencing gene expression variation in humans using data from the Brisbane Systems Genetics Study (BSGS). BSGS comprises of a total of 962 individuals from 314 families, for which we have high-density genotype data, collected using Illumina 610-Quad chips, gene expression, measured in whole blood using Illumina HT12-v4.0 arrays, as well as a wide range of phenotypic traits.
Families consist of combinations of both monozygotic and dizygotic twin pairs, their siblings, and, for 72 families, both parents. A significant advantage of the inclusion of parents is improved power to disentangle environmental, additive genetic, non-additive genetic and parent-of-origin effects of gene expression and measured phenotypes.

For each of the ~18,000 probes that remain after quality control, genetic and environmental components of variance were estimated from the correlations of expression values between relatives. Averaged over probes, the observed resemblance between relatives fits a simple model of a common nuclear family effect (which explains ~5% of phenotypic variance on average) and additive genetic variation (~20% on average).

We also estimated additive genetic variance from SNP data, independent of the estimated from the pedigree. The ability to predict narrow sense heritability ($h^2$) from SNP effects is highly variable across probes, most likely due to sampling variance and real differences in genetic architecture of gene expression. For 200 probes with large cis-acting expression QTL (eQTL), the correlation between the proportion of variance explained by SNPs and $h^2$ estimated from pedigree data was 0.85, consistent additive genetic variation causing the resemblance between relatives for those probes.

P-380

**JOINT MODELLING OF CONFOUNDING FACTORS AND PROMINENT GENETIC REGULATORS PROVIDES INCREASED ACCURACY IN GENETICAL GENOMICS STUDIES**

Nicolo Fusi1,3, Oliver Stegle2, Neil Lawrence1,3

1Sheffield Institute for Translational Neuroscience, University of Sheffield, Sheffield, UK, 2Machine Learning & Computational Biology Research Group, Max Planck Institute for Developmental Biology, Tubingen, Germany, 3Department of Computer Science, University of Sheffield, Sheffield, UK

Expression quantitative trait loci (eQTL) studies are an integral tool to investigate the genetic component of gene expression variation.

A major challenge in the analysis of such studies are hidden confounding factors, such as unobserved covariates or unknown subtle environmental perturbations.

These factors can induce a pronounced artifactual correlation structure in the expression profiles, which may create spurious false associations or mask real genetic association signals. Several approaches to account for these confounding factors have been proposed, greatly increasing the sensitivity in recovering direct genetic (cis) associations between variable genetic loci and the expression levels of individual genes.

Crucially, these existing techniques largely rely on the true association signals being orthogonal to the confounding variation. We found that when studying indirect (trans) genetic effects, for example from master regulators, their association signals can overlap with confounding factors estimated using existing methods. This technical overlap can lead to overcorrection, erroneously explaining away true associations as confounders.

We developed PANAMA (Probabilistic ANAlysis of genoMic dAta), a novel probabilistic model to account for confounding factors within an eQTL analysis.

In contrast to previous methods, PANAMA learns hidden factors jointly with the effect of prominent genetic regulators. The proposed model consistently performs better than alternative methods, and finds in particular substantially more trans regulators. Importantly, our approach not only identifies a greater number of associations, but also yields hits that are biologically more plausible and can be better reproduced between independent studies.

References:

Fusi et al., Joint Modelling of Confounding Factors and Prominent Genetic Regulators Provides Increased Accuracy in Genetical Genomics Studies. PLoS Comput Biol

Stegle et al., Using Probabilistic Estimation of Expression Residuals (PEER) to obtain increased power and interpretability of gene expression analyses. Nature Protocols
P-381
WHOLE GENOME SEQUENCES PROVIDE A NEW MOLECULAR CLOCK FOR THE HUMAN Y CHROMOSOME

Ross Fraser, James Wilson
University of Edinburgh, Edinburgh, UK

The human Y chromosome has the potential to be the definitive genetic marker, with approximately 29 Mb of euchromatic sequence inherited as one block, and an inferred mutation rate of about one SNP per generation. Whole genome sequencing should provide both unprecedented genealogical resolution and a much better molecular clock for the inference of coalescence times. However, only a small proportion of this variation has been determined: for example, about half the population of the British Isles belongs to the Y-chromosome paragroup S145*, with no further SNP-defined subclades available. Moreover, dating has relied hitherto on microsatellite markers, which have a complex mutation process and for which evolutionary mutation rates derived from a chimp-human comparison differ by a factor of three in comparison to those inferred from deep-rooting pedigrees.

In order to release the potential of the Y chromosome, we have analysed a number of publicly available whole genomes, aligned the sequences to the optimally alignable segments of the Y chromosome (~10 Mb) and called the genotypes. As dating relies on the number of SNPs, the quantitative fidelity of these calls is paramount, and we discuss various technical issues with utilising next generation sequence data from different sources, including a comparison of next generation sequence alignment and genotype calling algorithms.

Our technique, calibrated by comparison to both the chimpanzee sequence and a published six generation pedigree, allows the accurate determination of a SNP mutation rate and estimation of the ages of key nodes in the tree, including a minimum age for the R1b group, which includes 100 million European men and has been variously inferred to be Palaeolithic or Neolithic in age. We also report hundreds of new markers which will define new subgroups and allow considerably finer evolutionary inference.

P-382
GENERATION SCOTLAND: A NEW GENOMICS RESOURCE NOW AVAILABLE FOR RESEARCH

Shona Kerr1, David Porteous1, Anna Dominiczak2, Lynne Hocking1, Blair Smith4, Andrew Morris4
1University of Edinburgh, Edinburgh, UK, 2University of Glasgow, Glasgow, UK, 3University of Aberdeen, Aberdeen, UK, 4University of Dundee, Dundee, UK

Generation Scotland (GS, www.generationscotland.org) is a collaboration between the Scottish Universities and the NHS, funded by the CSO. GS has recruited over 30,000 participants to three cohorts. The GS:Donor DNA Databank recruited 5,000 volunteers through the NHS Blood Transfusion Service [Kerr et al, BMC Medical Genetics 2010, 11:166], collecting minimal phenotype, plasma and DNA. GS:21CGH recruited 2,000 individuals whose grandparents were born in Scotland. They attended a clinic, provided health and lifestyle data, underwent a series of clinical measures, and provided DNA, plasma and PBL cells. For the Generation Scotland: Scottish Family Health Study (GS:SFHS), over 24,000 volunteers in ~7,000 family groups attended study clinics and provided detailed health and lifestyle information. The high fidelity phenotyping in GS:SFHS includes cardiovascular, pain, cognitive and mental health measurements and many quantitative traits [Smith et al, BMC Medical Genetics 2006, 7:74].

Whole blood, DNA, urine and serum were systematically collected and stored. By targeting families, the statistical power to measure heritability and discover and validate causal genetic variants is enhanced. The already rich GS dataset can be linked with NHS health records and participants have consented to be re-contacted for further research. GS has successfully implemented an access process with high governance standards, for researchers to submit collaboration proposals. Projects range from epidemiological studies in pain, cognitive function and mental health, genetic replication studies of lung function and genome-
P-383
EXPRESSION OF POLYEMBRYONY IN FLUTED PUMPKIN (TELFAIRIA OCCIDENTALIS HOOK. F.)

Josiah Onovo¹, Vincent Onyia²
¹Nasarawa State University, Keffi, Nasarawa State, Nigeria, ²University of Nigeria, Nsukka, Enugu State, Nigeria

Objectives i) To investigate the forms and frequencies of polyembryony in fluted pumpkin. (ii) To access the range of expressivity of the trait.

Methods The investigation was carried out at the laboratory and research farms of the Department of Crop Science, University of Nigeria, Nsukka. In 2004, a total of 1,836 seeds of fluted pumpkin were pre-sprouted and carefully examined with a hand lens for the number of emerging embryos. In 2005, a total of 1,151 fresh seedlings from the seeds of 2004 harvest were examined.

Results Four embryo types were revealed in fluted pumpkin. They include single and three forms of polyembryony namely twin (biembryony), triple (triembryony) and quadruple (tetraembryony). The range of expressivity of the trait ranges from complete loss of side shoots (i.e. monoembryony) to possession of three side shoots (i.e. tetraembryony). Some of the plants that exhibited the polyembryonic trait had rudimentary side shoots, which subsequently degenerated with time. The frequencies of the embryo types in both years are presented in the Table.

<table>
<thead>
<tr>
<th>Embryo Type</th>
<th>Single</th>
<th>Twin</th>
<th>Triple</th>
<th>Quadruple</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>No. (%)</td>
<td>No. (%)</td>
<td>No. (%)</td>
<td>No. (%)</td>
<td>No. (%)</td>
</tr>
<tr>
<td>2004</td>
<td>319</td>
<td>17.37</td>
<td>479</td>
<td>26.09</td>
<td>1001</td>
</tr>
<tr>
<td>2005</td>
<td>150</td>
<td>13.03</td>
<td>244</td>
<td>21.20</td>
<td>702</td>
</tr>
</tbody>
</table>

Conclusion
Four embryo types exist in fluted pumpkin namely single, twin, triple and quadruple. These traits are heritable and occur at different frequencies. The consistent high penetrance of the polyembryonic trait is indicative of evolutionary fitness of the trait, and thus has a tendency to increase in frequency over generations.
P-384
LD-BASED DETECTION OF CRYPTIC ERROR STRUCTURES IN POPULATION SEQUENCING DATA

Oliver Venn1, Adam Auton1,2, Adi Fledel-Alon3, Susanne Pfeifer1, Laure Segurel1,5, Teresa Street4, Ellen Leifler2, Rory Bowden1,4, Ivy Aneas3, John Broxholme1, Peter Humburg1, Zamin Iqbal1, Gerton Lunter1, Julian Maller1,4, Ryan Hernandez7, Cord Melton3, Aarti Venkat3,5, Marcelo Nobrega3, Ronald Bontrop8, Simon Myers1,4, Peter Donnelly1,4, Ivy Aneas3, John Broxholme1, Peter Humburg1, Zamin Iqbal1, Gerton Lunter1, Julian Maller1,4, Ryan Hernandez7, Cord Melton3, Aarti Venkat3,5, Marcelo Nobrega3, Ronald Bontrop8, Simon Myers1,4, Peter Donnelly1,4, Molly Przeworski3,5, Gil McVean1,4

1Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK, 2Department of Genetics, Albert Einstein College of Medicine, Bronx, USA, 3Department of Human Genetics, University of Chicago, Chicago, USA, 4Department of Statistics, University of Oxford, Oxford, UK, 5Howard Hughes Medical Institute, University of Chicago, Chicago, USA, 6Oxford Biomedical Research Centre, Nuffield Department of Medicine, University of Oxford, Oxford, UK, 7Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, USA, 8Department of Comparative Genetics and Refinement, Biomedical Primate Research Center, Rijswijk, NL, UK, 9Department of Ecology and Evolution, University of Chicago, Chicago, USA

The analysis of genetic variation called from high-throughput population sequencing is an increasingly popular approach to study biological processes. However, a major challenge in interpreting such information is the detection of erroneous, mis-assembled, or incorrectly genotyped genetic variants. We present an automated methodology to identify erroneous variants by integrating linkage disequilibrium information from genetic variation data.

We motivate this approach by describing our experience of constructing the first fine-scale genetic map of Western chimpanzees, *Pan troglodytes verus*, from high throughput sequence data from ten individuals. We show examples of incongruous linkage disequilibrium structure around cryptic error classes that are missed by state-of-the-art variant filters. We identify two new error classes of clustered false positive SNPs, generated by systematic mis-assembly and by underrepresentation of repeat DNA classes in the chimpanzee reference genome. We explain how we overcame these issues, enabling a comparison of recombination rates in human and chimpanzee. We observe rapid evolution of fine-scale rates, resulting in no overlap of hotspots between human and chimpanzee, despite similarly elevated rates around CpG islands and decreased rates around genes. In contrast, broad-rates are largely conserved, with exceptions, particularly in regions of chromosome rearrangement. In addition we present an experiment to validate our LD-based filtering approach by comparing our cleaned chimpanzee data to a new sequencing experiment of an extended two-generation chimpanzee pedigree.

P-385 – ABSTRACT WITHDRAWN

P-386
USING DE NOVO ASSEMBLY TO FIND OUT SV

Qian Mao
BGI, London, UK

The International HapMap project and genome-wide association studies (GWAS) have allowed extensive research based on characterization of single-nucleotide polymorphisms (SNPs). This was followed by work to identify and characterize structural variations (SV), including insertions, deletions, inversions and other DNA sequence rearrangements. A large number of such variations have been discovered in the human genome putatively having equal or greater functional impact than SNPs. Here we use whole-genome de novo assembly of second-generation sequencing reads to map structural variation in an Asian genome and an African genome. Our approach identifies small- and intermediate-size homozygous variants (1–50 kb) including insertions, deletions, inversions and their precise breakpoints, and in contrast to other methods, can resolve complex rearrangements. In total, we identified 277,243 SVs ranging in length from 1–23 kb.
Validation using computational and experimental methods suggests that we achieve overall <6% false-positive rate and <10% false-negative rate in genomic regions that can be assembled, which outperforms other methods. Analysis of the SVs in the genomes of 106 individuals sequenced as part of the 1000 Genomes Project suggests that SVs account for a greater fraction of the diversity between individuals than do SNPs. These findings demonstrate that whole-genome de novo assembly is a feasible approach to deriving more comprehensive maps of genetic variation.

P-387
DERMATOGLYPHIC VARIATIONS IN A NIGERIAN POPULATION

Hannah Etta1, Ibukunle Petu2, Idorenyin Etukudoh3
1cross River University of Technology, Calabar, Cross River State, Nigeria, 2national Open University of Nigeria, Victoria Island, Lagos, Nigeria, 3university of Calabar, Calabar, Cross River State, Nigeria

Dermatoglyphic traits have been used to evaluate population structure and micro differentiation in several populations. Dermatoglyphs (finger-prints), are an example of somatic hereditary traits whose expression is variable and therefore can be used as a basis for establishing and confirming the historical relationship within populations. Unilateral thumbprints were collected from a random population of Cross Riverians from the South-South geo-political region of Nigeria. Digital pattern types were classified into four categories defined as: whorls (W), compound (C), arches (A) and loops (L). A total of 897 thumbprints were collected on white paper using a non-spreading ink-pad. Patterns were identified using glass hand lens. Chi square test showed non significant (P>0.05) differences between sexes (SX) and finger print type (FPT) but a high significant (P<0.05) difference between ethnicity (EN) and finger print type (FPT). Frequency distributions gave whorl (W) as the most predominant FPT in Cross River state. Correlations were not significant (P>0.05) between EN and FPT. All analyses were done with the aid of the SPSS ver. 18 analytical software. Interestingly, pattern-types in males and females showed significant (P<0.05) variations throughout the 18 Local Government Areas studied. This study presents the first ever documented record of dermatoglyphs among the Cross River State local government areas of Nigeria and can serve as a basis for future population and genetic studies.

P-388
IDENTIFICATION OF COPY NUMBER VARIATION IN THE IBERIAN PIG GENOME

Ana I Fernández1, M. Carmen Rodríguez1, Yuliaxis Ramayo-Caldas2, Beatriz Villanueva1, Luis Silió1
1Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Madrid, Spain, 2Facultad de Veterinaria, Universidad Autónoma de Barcelona, Barcelona, Spain

Copy number variants (CNVs) are DNA segments, ranging from 1Kb to several Mb, with variable number of copies among individuals. CNVs are considered an important source of genetic variation, covering even more nucleotide content per genome than SNPs. Different methodologies can be applied to identify these variants including those based on the hybridization signal intensity of whole genome SNP chips. The goal of the present study was to identify CNV regions (CNVRs) in two strains of the Iberian breed, Torbiscopal and Guadyerbas, using the Porcine SNP60 BeadChip. We analyzed a total of 61 Torbiscopal and 99 Guadyerbas related animals. Data analyses were carried out for the SNPs located in autosomes and mapped in the latest version of the genome, Sscrofa10.2. We used the software PennCNV and took into account trio pedigree information (52 trios in Torbiscopal and 69 in Guadyerbas). A total of 83 CNVRs were identified in Torbiscopal and 51 CNVRs in Guadyerbas. Four of them overlapped between both related strains. However, given the limited number of SNPs contained in the porcine chip, probably the number of CNVs identified underestimates the actual number. The size of the CNVRs ranged from 3Kb to 1.5Mb and included a minimum of three SNPs and a maximum of 26. The CNVRs represented around 0.61% and 0.35% of the Torbiscopal and the Guadyerbas genomes, respectively. When compared these results with
those obtained in a previous study (www.ensembl.org/biomart) that included a small set of Guadyerbas animals, we observed that 11 out of CNVRs detected in Torbiscal and 12 out of CNVRs detected in Guadyerbas overlap in both studies. Moreover, several of the annotated genes within the detected CNVRs, such as the OR4K15, OR4F6, PCDH15, ARAP2, TBX1, ACA2, ATP6V1E1, BCL6 and TMEM168 genes have been also identified within human CNVs.

**P-389**

**NEXT GENERATION SEQUENCING OF THE DISC1 LOCUS**

_Pippa Thomson_1,2, Jennifer Parla_3, Allan McRae_4, Melissa Kramer_3, Kamna Ramakrishnan_1, Jianchao Yao_5, Dinesh Soares_1, Shane McCarthy_1, Stewart Morris_3, Laura Cardone_3, Steve Cass_1, Elena Ghian_1, William Hennah_1, Kathryn Evans_1,2, Danea Rebolini_1, Kirsty Millar_1, Maureen Bell_1, Sarah Harris_1,2, John Starr_7, Walter Muir_1, Donald Macintyre_1, Andrew McIntosh_1, James Watson_3, Peter Visscher_4, Ian Deary_2, Douglas Blackwood_6, W Richard McCombie_3, David Porteous_1,2

1Medical Genetics Section, University of Edinburgh, Molecular Medicine Centre, Western General Hospital, Crewe Road, Edinburgh, UK, 2Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh, UK, 3Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA, 4University of Queensland Diamantina Institute, Princess Alexandra Hospital, Ipswich Road, Woolloongabba, Australia, 5Institute for Molecular Medicine Finland FIMM, Biomedicum Helsinki 2U, Tukholmankatu 8, 00290 Helsinki, Finland, 6Department of Psychiatry, Royal Edinburgh Hospital, Morningside Park, Edinburgh, UK, 7Centre for Cognitive Ageing and Cognitive Epidemiology, Geriatric Medicine Unit, University of Edinburgh, Royal Victoria, Craigleith Road, Edinburgh, UK

_Disrupted in Schizophrenia 1, DISC1_, was originally identified at the site of a balanced t(1;11) translocation in a Scottish family affected by schizophrenia, bipolar disorder and recurrent major depression. Co-segregation of the translocation with major mental disorder is highly significant (LOD = 7.1). Both positive and null association findings for common sequence variants and mental illness have since been reported for _DISC1_. Here, we report the detection of 2,718 single nucleotide polymorphisms, of which >60% are novel, by sequencing 528 Kb of the TRAX/DISC1 locus in 653 cases of schizophrenia, bipolar and major depressive disorder, and 889 population controls (the Lothian Birth Cohort 1936, LBC1936), the latter with quantitative trait measures of cognitive ability, anxiety, depression, and neuroticism. There was no evidence for association of common variants with schizophrenia or bipolar disorder, but a gene-wide empirical _p_ = 0.026 (Odds Ratio = 3.48, 95%CI = 1.95-6.23, unadjusted _p_-value = 6.3 x 10^-5) was observed between an intronic variant and major depressive disorder. We report multiple rare coding mutations, association of _DISC1_ with major depressive disorder, and an elevated burden of putative regulatory variants nominally associated with quantitative traits linked to depression and cognitive ability in LBC1936.

**P-390**

**GENOME-WIDE ANALYSIS OF NATURAL REGULATORY VARIATION UNDER CONTRASTING ENVIRONMENTS IN A. THALIANA**

_Francisco Cubillos_1, Oliver Stegle_0,2, Cecile Grondin_1, Matthieu Canut_1, Isabelle Gy_1, Olivier Loudet_1

1INRA, Versailles, France, 2Max Planck Institute for Developmental Biology, Tübingen, Germany

One of the central goals in evolutionary genetics is to understand how the distinct allelic variants can shape phenotypic diversity, either across environments or in an environment-specific manner. Genome-wide association studies (GWAS) and linkage mapping strategies have identified countless numbers of QTLs in plants; from these, however, only a small number of essentially large-effect loci have been molecularly characterised. Moreover, many of them have identified changes that dramatically alter the
encoded protein sequence. Here, we measure the effects of non-coding changes by estimating genome-wide levels of allele-specific expression (ASE) in A. thaliana Col x Cvi F1 hybrid individuals (a robust test for cis-acting regulation). Furthermore, in order to account for genetic (trans-effect) and environmental interactions, hybrids and parental accessions were grown under well-watered and water-deficit conditions, as contrasting treatments. Subsequently, we developed a model to estimate the contribution of the genotype, the environment and their interactions, on the final gene expression levels. In this way, we are now able to isolate the effect of each component and identify genes exhibiting ASE and distinct environmental responses. We revealed that, in this cross and in the non-stress environment, more than 30% of expressed genes show detectable levels of significant cis-regulatory variation, whilst 15% show signatures of being trans-regulated. We found strong correlation between conditions for genes being regulated in cis, demonstrating the robust nature of local variation towards environmental perturbation. On the contrary, a higher number of genes were trans-regulated under water-deficit, likely due to a stress-specific response in pathways associated to drought. Identification of regulatory gene networks and accession-specific responses, represents an important step towards bridging the gap between non-coding changes and natural phenotypic variation.

P-391
SMOKING ALTERS DNA METHYLATION IN MULTIPLE LOCI

Loukia G. Tsaprouni, Tsun-Po Yung, Jordana Bell, Katherine Dick, Stavroula Kanoni, Eshwar Meduri, Chris Nelson, Cardiogenics Consortium, Tim Spector, Nilesh Samani, Panos Deloukas

1Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK, 2King’s College London, Department of Twin Research and Genetic Epidemiology, UK, 3University of Leicester, Department of Cardiovascular Sciences, UK

Smoking is a major risk factor in most common diseases impacting mortality. Genome wide association studies have linked nicotine dependence and smoking behaviour to increased risk of cardiovascular, pulmonary and malignant diseases whilst recent evidence show that global DNA methylation is influenced by smoking behaviour. We assessed methylation patterns in peripheral-blood DNA in 473 individuals (Cardiogenics study) of which 185 are current or ex-smokers with the use of the Illumina Human Methylation 450K BeadChip. A linear model analysis confirmed an effect of smoking on methylation in two previously reported genes: Coagulation factor 2 (F2RL3) and G-protein coupled receptor 15 (GPR15). F2RL3 was significantly associated with methylation in smokers (p=2.17×10−18 and p=10−11) which replicates the association reported earlier this year. GPR15 was also associated with current and previous smoking history (p=2.61×10−14). We obtained 22 signals at an FDR of 0.01% corresponding to 10 loci. Interestingly, one of the top signals is in a region of chromosome 2 (233283397-233285959) located close to the neuronal acetylcholine receptor subunit D (CHRND) gene. We then undertook replication in an independent sample of 437 Twins from the TwinsUK registry which comprises 150 current or ex-smokers. After applying Bonferroni correction, 20 of the 22 signals replicated at p < 2.27×10−3 including F2RL3, GPR15 and the chromosome 2 region. Levels of methylation were reduced in smokers in all seven different probes that cover the chromosome 2 region (p=7.31×10−9 – 8.62×10−33). Finally, we assessed the intersection between methylated QTLs (metQTLs; p < 10−4) versus the top 22 probes from the linear model analysis described above. Both F2RL3 and the chr2 region had a metQTL in blood associated with methylation. However, GPR15 did not have a metQTL at genome wide significance. Meta-analysis of the Cardiogenics and Twins data sets may allow detection of further smaller effects between smoking and DNA methylation.
P-392
GENOME-WIDE LINKAGE DISEQUILIBRIUM MAPS IN THREE DANISH COMMERCIAL PIG BREEDS AND COMPARISON OF LD ACROSS BREEDS

Lei Wang1, Tage Ostersen2, David Edwards1
1Department of Molecular Biology and Genetics, Aarhus University, Tjele, Denmark, 2Pig Research Center, Copenhagen, Denmark

Linkage disequilibrium (LD) measures the degree of association of alleles at different loci in a population. Using single nucleotide polymorphism (SNP) information to characterize genetic variation requires that the SNPs are in strong LD with neighbouring genetic variants. The pattern of LD reflects both the population history and local recombination rates. It is therefore important to characterize LD.

The aim of this study is to compare the pattern and extent of LD in three Danish commercial pig breeds (Duroc, Landrace, and Yorkshire). A large sample of pigs from the three breeds (4249 Duroc, 1979 Landrace and 2123 Yorkshire) has been genotyped using the Illumina PorcineSNP60 BeadChip, which contains approx. 62000 SNPs spanning the porcine genome. After pre-filtering approx. 32000 SNPs are used in the analysis.

For each breed, we calculated average r² between adjacent SNPs across the genome, which were 0.55 for Duroc, and 0.50 for both of Landrace and Yorkshire. All pairwise r² for each autosomal chromosome were estimated, and used to construct LD maps for the three breeds. We examined the relations between physical distance between SNP pairs with r², and used this to display patterns of LD over the chromosomes, enabling comparisons of LD to be made between different chromosomal regions and between breeds. Furthermore, we estimated persistence of phase between breeds, finding that Landrace and Yorkshire are more closely related to each other than to Duroc.

P-393
GENOMIC DISSECTION OF SMALL RNAS IN WILD RICE (O. RUFIPOGON): LESSONS FOR RICE DOMESTICATION

Yu Wang1, Longbiao Guo2, Yun Liu1, Yijie Gui1, Xinghua Wei2, Qianhao Zhu3, Longjiang Fan1
1Department of Agronomy, Zhejiang University, Hangzhou, China, 2State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou, China, 3CSIRO Plant Industry, Canberra, Australia

Objectives: Due to the comprehensive regulatary function of miRNA in plants, whether MIRNA genes were under artificial selection during crop domestication was an interesting question. Several studies on role of miRNAs in rice domestication have been carried out from cultivated rice but not from its wild progenitor.

Methods: In this study, we characterized miRNAs and other small RNAs in the genome of O. rufipogon through de novo genome assembly of a Chinese wild rice (Dongxiang), and its small RNA populations and degradomes etc by high-throughput sequencing and comparative genomic analysis.

Results: Based on the Illumina reads from wild rice genome by this study, at least 28 miRNA loci were not covered by any reads from the cultivated rice. On the contrary, 21 and 23 known rice miRNAs were not found or disabled in the wild genome, respectively. Targets of those unique miRNAs were predicted by both in silico and degradome ways. Moreover, expression of miRNAs between wild and cultivated rice presented unique patterns based on miRNA array and high-throughput data.

Conclusions: Our results illustrated a complex evolutionary process of MIRNA genes during rice domestication.
P-394

IMPROVING THE GENOME ASSEMBLY OF THE MOSQUITO Aedes aegypti

Punita Juneja1, Jewelna Osei-Poku2, Shwen Ho1, Arnab Pain1, Frank Jiggins1
1University of Cambridge, Cambridge, UK, 2King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

Aedes aegypti is an important mosquito model for studying the genetics of vector competence. It is a natural vector of dengue and yellow fever and a laboratory model for studying filarial nematodes such as Brugia malayi. The current assembly of the Ae. aegypti genome is composed of 4,758 supercontigs, many of which have not been assigned to chromosomes. Improving upon this genome assembly will open up future lines of research. We aim to assemble a genetic map by using RAD-seq to develop markers and genotype individuals from a mapping population. This genetic map will be used to join supercontigs and potentially assign them to chromosomes. We are also using RNA-seq to improve the annotation of the genome. This improved genome assembly will eventually be used to map resistance of Ae. aegypti to B. malayi, with an ultimate aim of identifying candidate loci that control vector competence.

P-395

LARGE-SCALE DNA METHYLATION ANALYSIS IN PRIMARY HUMAN ADIPOSE TISSUE AND ITS IMPACT ON OBESITY PHENOTYPES

Johanna K. Sandling1, Asa K. Hedman2, Eshwar Meduri1,3, Elin Grundberg1,3, Sarah Keildson2, James Nisbet1, Magdalena Sekowska1, Alicja Wilk1, Amy Barrett5, Kerrin S. Small1,3, Emmanouil T. Dermitzakis4, Timothy D. Spector1, Mark I. McCarthy2,5, Panos Deloukas1
1Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK, 2Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK, 3Department of Twin Research and Genetic Epidemiology, King’s College London, London, UK, 4Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland, 5Oxford Centre for Diabetes, Endocrinology & Metabolism, University of Oxford, Churchill Hospital, Oxford, UK

Epigenetic mechanisms are increasingly being recognised to play a role in obesity phenotypes, however, most studies to date performed in man have been limited either to candidate regions, small numbers of samples, or by choice of tissue. Furthermore, such studies had limited power to separate environmental, genetic and epigenetic influences. To overcome these limitations we explored the influence of DNA methylation on obesity traits at a genome-wide scale in primary tissue from twins. Bisulphite converted DNA from subcutaneous adipose tissue for 662 well phenotyped female twins from the MuTHER study (http://www.muther.ac.uk) was assayed on HumanMethylation450 BeadChips. We focused our analysis on the top 10% variably methylated sites (N=31,665), and investigated association with obesity related phenotypes using linear mixed effects models.

We found the largest number of associations for the DEXA derived per cent trunk fat mass trait (PTFM), with methylation at almost ten thousand sites associated with this phenotype. A similar number of associations was observed for the correlated trait body mass index. More than three thousand loci were associated with blood levels of insulin, triglycerides, and high-density lipoprotein, respectively. In contrast, virtually no associations were observed for low-density lipoprotein, total cholesterol, and glucose. For PTFM one third of the associated sites were annotated to genes, and over a thousand genes had multiple probes associated with the trait. For the vast majority of these genes the direction of effect was consistent across sites. Pathway analysis revealed that these genes were predominantly associated with lipid metabolism, molecular transport, and small molecule biochemistry. Some of these signals were annotated to obesity genes such as leptin and its receptor.

In conclusion, we find methylation at a large number of sites in the genome to be associated with obesity phenotypes. Assessing whether this effect is mediated by changes in gene expression is under way.
P-396
COPY NUMBER ABERRATION DETECTION WITH NEXT GENERATION SEQUENCING DATA: CHALLENGES AND SOLUTIONS

S.W. Wendy Wong¹, Sergii Ivakhno¹, Thomas Royce², Michael Eberle²
¹Illumina Cambridge Ltd., Little Chesterford, UK, ²Illumina, Inc., San Diego, USA

Over the last few years, next generation sequencing analysis has become more and more popular in the area of cancer analysis. Estimation of copy number aberration (CNA) with next generation sequencing data has advantage over more traditional SNP array methods in that it gives much higher resolution in detecting CNV regions, as well as more accurate identification of higher copy numbers. The challenges in CNA calling include contamination of normal cells, aneuploidy in cancer, etc. Here we discuss the challenges and describe the existing algorithms that are designed for identifying levels of contamination, genome ploidy, copy numbers, genotypes of copy number variants in matched cancer/normal samples. We will also present the comparison of the performance of the publicly available and Illumina internal tools with Illumina sequencing technology data.

P-397
MOLECULAR DETECTION OF FIVE MISMATCH REPAIR GENE DEFECTS IN SINGAPORE COHORT OF LYNCH SYNDROME FAMILIES

Yanqun Liu, Xuewei Goh, Min Hoe Chew, Soo Yong Tan, Tien Tau Loi, Poh Koon Koh, Choong Leong Tang
Singapore General Hospital, Singapore, Singapore

Objective: Germline defects of several mismatch repair gene (MMR) genes underlie many cases of Lynch Syndrome (LS). We aimed to gain comprehensive genetic and epigenetic profiles of LS families in Singapore, which might ultimately facilitate the effective molecular diagnosis of LS in Singapore and the region.

Methods: Sixty one unrelated families were selected as fulfilling Amsterdam I or II criteria, forming the largest cohort of clinically-diagnosed LS in Singapore. IHC was performed on MLH1, MSH2, MSH6 and PMS2. Tumor microsatellite stability was assessed with a panel of markers recommended in the revised Bethesda criteria. Mutations in exons, splice-site junctions and promoters of 5 common MMR genes were scanned by high resolution melting (HRM) assay followed by DNA sequencing, large fragment deletions/duplications and promoter methylation were evaluated by multiplex ligation-dependent probe amplification.

Results: A total of 61 variants were detected (17 in MLH1, 19 in MSH2, 7 in MSH6, 11 in PMS2 and 7 in PMS1) in 41 out of 61 families (67.2%). Twenty five families (41%) harbored disease-causing mutations and other 16 families (26.2%) carried variants of uncertain pathogenic significance. A large deletion covering the promoter and exons 1-6 of MSH2 was detected in one family. The promoter methylation of MLH1 was found in another family. Twenty one novel variants were identified. The mutation types across 5 MMR genes were highly heterogeneous without indication of the presence of hot spots. Concordance between MMR mutations and IHC was 67.7%, while 41.2% of patients with MSS tumors were mutation carriers in which 42.9% of mutations were disease-causing.

Conclusion: MLH1 and MSH2 are the most common genes harboring inherited defects. The recommended marker panel in the revised Bethesda criteria might be inadequate for LS cases in Singapore, while IHC is a much better screening tool prior to thorough searching for MMR mutation.
P-398
GENETIC ASSOCIATION OF DNA METHYLATION IN ADIPOSE TISSUE AND ITS EFFECT ON GENE EXPRESSION IN TWINS

Eshwar Meduri1,3, Elin Grundberg1,3, Johanna K Sandling1, Asa K Hedman2, Sarah Keildson2, James Nisbet1, Magdalena Sekowska1, Alicja Wilk1, Amy Barrett3, Kerrin Small1,3, So-Youn Shin1, Emmanouil T Dermitzakis4, Mark I McCarthy2,5, Timothy D Spector3, Panos Deloukas1
1Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge, UK, 2Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK, 3Department of Twin Research and Genetic Epidemiology, Kings College London, London, UK, 4Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland, 5Oxford Centre for Diabetes, University of Oxford, Oxford, UK

There is a considerable amount of evidence that both environmental and genetic factors play a significant role in modulating DNA methylation, which in turn impacts gene expression and phenotypic variation. Thus far, most DNA methylation QTL studies have mainly interrogated a set of 27,000 CpG sites across human genes linked to cancer. Here we explore DNA methylation in ~480,000 CpG sites (Illumina450K-array), covering not only genes but also several other genomic features, across a well powered sample of 662 subcutaneous adipose tissue from phenotypically well characterized twins (MuTHER resource; www.muther.ac.uk). All samples have been imputed to the 1000 Genomes and profiled for gene expression (Illumina HT-12).

Focusing on the top 10% of the differentially methylated CpGs (N=31,665) and using 5,328,305 SNPs (MAF>5%, info>0.8) we found an abundance of methylation quantitative trait loci (metQTLs) acting in cis (defined as 100KB window flanking the CpG site). There is strong enrichment of top signals in close proximity (mean 25kb) of the probe with the majority (70%) of probes located in gene promoters, CpG Islands, shores or shelves. Under the assumption that methylation affects gene expression, we performed cis-eQTL analysis and took the top signals (N=4539) at FDR 1% to assess how many of them overlap with top metQTL signals. We found 270 (6%) of the top signals overlapping, suggesting a shared mechanism between expression and methylation. To understand the relationship between expression and methylation on a global scale, we then took variable probes in gene promoters (5245 genes) from the Illumina 450K array and compared them to the probes associated with corresponding genes on the Illumina HT-12 expression array. We discovered significant associations between expression and methylation in 370 genes (7%, average beta=-0.03). Future analysis include utilizing the twin design to estimate heritability of DNA methylation as well as integrating disease associated GWAS signals with metQTLs.

P-399
TO EVALUATE AND APPLY NOVEL STATISTICAL TECHNIQUES FOR DIVERSITY BASED GENE-MAPPING

John Stainton1, Chris Haley4,1, Brian Charlesworth2, Andreas Kranis1, Kellie Watson3, Pamela Wiener1
1Roslin Institute, Edinburgh, UK, 2Institute for Evolutionary Biology, University of Edinburgh, Edinburgh, UK, 3Aviagen, Edinburgh, UK, 4Medical Research Council, Human Genetics Unit, Edinburgh, UK

The hitchhiking effect reduces genetic variation at sites linked to those under positive selection. Statistical methods can be used to find significant areas of low genetic variation which can then be investigated to find potential genes of interest. Domesticated animals are particularly well suited to these methods, due to years of strong human driven selection for beneficial traits. A study on genetic variation in domesticated animals may find important regions for further study and allow the evaluation of the power and performance of these statistical methods in a way data from other species does not.

Broiler chickens are selected for a variety of traits, including fast growth, high feed conversion ratio and a high proportion of breast meat. In order to discover possible signatures of selection, genetic differentiation (FST) and heterozygosity tests were carried out on SNP data provided by an industrial partner. This data
includes 9 broiler lines genotyped using a 12k SNP chip and 1 line genotyped with a 42k chip. Several regions where selection may be taking place have been identified. Further research includes identifying candidate selected genes within these regions and analysis of additional datasets.

P-400
A HIGH RESOLUTION MAP OF COPY NUMBER VARIABLE REGIONS IN BROWN SWISS DAIRY CATTLE

Marlies Alexandra Dolezal1, Laura Pellegrino1, Antonia Bianca Samorè1, Morris Soller2, Alexandra Graf1, Enrico Santus2, Markus Jaritz3, Alessandro Bagnato1
1Department VSA, University of Milan, Milan, Italy, 2Department of Genetics, The Hebrew University of Jerusalem, Jerusalem, Israel, 3FH Campus Vienna, University of Applied Sciences, Vienna, Vienna, Austria, 4ANARB, Associazione Nazionale Allevatori Razza Bruna, Bussolelengo, VR, Italy

Recent studies have identified copy number variations (CNVs) that are functionally significant within mammalian genomes. This study presents a high resolution genome map of copy number variable regions in Brown Swiss dairy cattle, a prerequisite to studying associations of CNVs to production and functional traits in this breed. The top 192 bulls based on their marginal genetic contribution, identified via pedigree analysis, were genotyped with Illumina’s bovine HD chip. CNVs were identified with PennCNV software employing a hidden markov model based on total signal intensity (log R ratio) and allelic intensity ratio (B Allele Frequency) reported by Illumina’s Genomestudio software at each of 735238 loci anchored to bovine autosomes on the UMD3.1 assembly. Differences in hybridization efficiency due to the sequence composition flanking each SNP were corrected for to reduce false positive calls. After filtering bulls with poor signal-to-noise measures and CNVs overlapping centromeric and telomeric regions, 8275 CNVs encompassing 2221 homozygous losses, 3405 heterozygous losses, 2636 heterozygous duplications and 13 homozygous duplications, copy number zero, one, three and four respectively are segregating among 164 Brown Swiss bulls. Considering CNVs within 100 kb to represent the same region we summarized these CNVs into 2034 copy number variable regions (CNVRs) of which 949 and 892 constitute gain and loss events and 139 are of complex nature. Mean and median CNVR length is 40.74 kb and 12.53 kb, respectively. CNVRs cover a total of 82.87 Mb of sequence which corresponds to 3.33% of the bovine autosome in Brown Swiss dairy cattle. Copy number variable regions were annotated with the Ensembl v66 bovine gene list.

This study was supported by the FP7 framework project QUANTOMICS contract n. 222664-2.

P-401
RECOMBINATION RATE AND HOTSPOTS ANALYSIS BASED ON POPULATION SEQUENCING OF A THALIANA

Qingrun Zhang1, Matt Horton1,4, Quan Long1, Adam Auton2, Dazhe Meng1,3, Magnus Nordborg1,3
1Gregor Mendel Institute, Vienna, Austria, 2Albert Einstein College of Medicine, New York, USA, 3University of Southern California, Los Angeles, USA, 4University of Chicago, Chicago, USA

Objectives: To analyze the recombination dynamics of A thaliana from full genomes.

Methods: We sequenced ~180 Arabidopsis thaliana inbred lines with Illumina at deep coverage (ranging from 20 to 60). LD pattern based MCMC estimation of recombination rate (LDhat) has been applied. After that we estimate the recombination hotspots using LDhot. Several motif finders have been applied to search causal motif that correlated to recombination hotspots.

Results: We find the genome scale recombination rate distribution is correlated with polymorphism pattern. The whole genome polymorphism pattern can be fit into a background selection model using the estimated recombination hotspots. There are changes of recombination rates/hotspots at the regions that structural variants are present, indicating two key insights: (1) the structural variants change the coordinate
system of the reference genome therefore generates some artifacts in the recombination rates estimation based on LD; (2) in the regions that have real high recombination rates, it is more likely to generate more structural variants.

**Conclusions:** Sequence based whole genome analysis reveals important insight into recombination rate and hotspots of *A. thaliana*.

**P-402**

**IDENTIFICATION OF COPY NUMBER VARIANTS ASSOCIATED WITH BACK FAT THICKNESS IN PIGS USING A SELECTIVE GENOTYPING APPROACH**

Giuseppina Schiavo¹, Marlies Dolezal², Pier Luigi Martelli³, Giuliano Galimberti³, Daniela Giovanna Calò⁴, Emilio Scotti⁵, Luca Buttazzoni⁵, Rita Casadio³, Alessandro Bagnato², Vincenzo Russo¹, Luca Fontanesi²

¹Dept. of Agro-Food Science and Technology, University of Bologna, Bologna, Italy; ²Department VSA, University of Milano, Milano, Italy; ³Biocomputing Group, University of Bologna, Bologna, Italy; ⁴Department of Statistical Sciences “Paolo Fortunati”, University of Bologna, Bologna, Italy; ⁵CRA, Centro di Ricerca per la Produzione delle Carni e il Miglioramento Genetico, Roma, Italy

Back fat thickness (BFT) in pigs affects carcass value and consumers’ acceptance of pork. Following market requests, pig breeding programs for fresh pork consumption are usually designed to reduce BFT and increase lean meat percentage with some exceptions for specialized markets. For these reasons BFT is routinely measured in selected nuclei that could be also useful to model and understand biology aspects of human obesity, considering the physiological closeness between human and pig. We have recently carried out an extensive candidate gene analysis and a genome wide association study to identify genes and chromosome regions affecting BFT in the Italian Large White pig population. Here we evaluated if copy number variants (CNVs) are associated with BFT in Italian Large White pigs. Pigs with extreme and divergent estimated breeding values (EBVs) for BFT were selected among a performance tested population of approx. 12,000 animals and genotyped with the Illumina PorcineSNP60K Beadchip. CNVs were called using pennCNV using stringent criteria. Fifteen copy number variation regions (CNVRs) were identified only in the positive BFT-EBV group whereas 12 CNVRs were reported only in the negative BFT-EBV tail. Other CNVRs were more frequent (P<0.05) in one tail or in the other one. Identified CNVRs included genes involved in fat metabolism, growth regulation, immune system, and neuronal regulation of eating behaviour. These results provide additional insights in mechanisms affecting fat deposition in pigs useful to understand aspects of human obesity.

**P-403 – ABSTRACT WITHDRAWN**
<table>
<thead>
<tr>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbe Maleyki, MJ</td>
<td>P-403</td>
</tr>
<tr>
<td>Abdelayed, Mary</td>
<td>P-94</td>
</tr>
<tr>
<td>Abegaz, Solomon</td>
<td>P-201</td>
</tr>
<tr>
<td>Abera, Habtamu</td>
<td>P-201</td>
</tr>
<tr>
<td>Adamczyk, Barbara</td>
<td>P-354</td>
</tr>
<tr>
<td>Ådnøy, Tormod</td>
<td>P-3</td>
</tr>
<tr>
<td>Adriaensen, Frank</td>
<td>P-134</td>
</tr>
<tr>
<td>Afzal, Nadeem</td>
<td>P-369</td>
</tr>
<tr>
<td>Agartz, Ingrid</td>
<td>P-57</td>
</tr>
<tr>
<td>Ahmad, Ilyas</td>
<td>P-385</td>
</tr>
<tr>
<td>Ahmad Zubaidi, A L</td>
<td>P-403</td>
</tr>
<tr>
<td>Ahola, Virpi</td>
<td>P-100</td>
</tr>
<tr>
<td>Ahsan, Muhammad</td>
<td>P-234</td>
</tr>
<tr>
<td>Aidar de Queiroz, Sandra</td>
<td>P-170, P-171</td>
</tr>
<tr>
<td>Ajithkumar, S</td>
<td>P-216</td>
</tr>
<tr>
<td>Akay, Izzet Mehmet</td>
<td>P-308</td>
</tr>
<tr>
<td>Akey, Josh M</td>
<td>O-18</td>
</tr>
<tr>
<td>Albers, Cornelis</td>
<td>P-63</td>
</tr>
<tr>
<td>Albert, Frank</td>
<td>P-321</td>
</tr>
<tr>
<td>Albrecht, Eva</td>
<td>P-172, P-185, P-189</td>
</tr>
<tr>
<td>Albrecht, Theresa</td>
<td>P-202</td>
</tr>
<tr>
<td>Alemu, Setegn</td>
<td>P-167</td>
</tr>
<tr>
<td>Alenda, Rafael</td>
<td>P-128</td>
</tr>
<tr>
<td>Alexandrov, Nickolai</td>
<td>P-216</td>
</tr>
<tr>
<td>Alex, Rani</td>
<td>P-257</td>
</tr>
<tr>
<td>Ali, Abdirahman</td>
<td>P-120, P-85</td>
</tr>
<tr>
<td>Alía, Ricardo</td>
<td>P-290</td>
</tr>
<tr>
<td>Allen, Adrian R</td>
<td>P-365</td>
</tr>
<tr>
<td>Allen, Melissa</td>
<td>P-122</td>
</tr>
<tr>
<td>Allen, Scott</td>
<td>P-146</td>
</tr>
<tr>
<td>Allison, David B</td>
<td>P-128</td>
</tr>
<tr>
<td>Allison, Gordon</td>
<td>P-252</td>
</tr>
<tr>
<td>Allison, Matthew</td>
<td>P-343</td>
</tr>
<tr>
<td>Almasy, Laura</td>
<td>P-280</td>
</tr>
<tr>
<td>Altarriba, Juan</td>
<td>P-87, P-91</td>
</tr>
<tr>
<td>Álvarez-Castro, José</td>
<td>P-372</td>
</tr>
<tr>
<td>Amador, Carmen</td>
<td>P-258, P-259, P-320</td>
</tr>
<tr>
<td>Amills, Marcel</td>
<td>P-240</td>
</tr>
<tr>
<td>Amiri, Esmaeil</td>
<td>O-54</td>
</tr>
<tr>
<td>Anderson, Charlotte</td>
<td>P-375</td>
</tr>
<tr>
<td>Anderson, Charlotte L</td>
<td>P-109</td>
</tr>
<tr>
<td>Anderson, David J</td>
<td>P-193</td>
</tr>
<tr>
<td>Anderson, Denise</td>
<td>O-32</td>
</tr>
<tr>
<td>Anderson, Susan</td>
<td>O-36, P-135, P-52</td>
</tr>
<tr>
<td>Andersson, Leif</td>
<td>P-110</td>
</tr>
<tr>
<td>Andjelkovic, Marko</td>
<td>P-50, P-57, P-325</td>
</tr>
<tr>
<td>Andreassen, Ole</td>
<td>P-384</td>
</tr>
<tr>
<td>Aneas, Ivy</td>
<td>O-39, P-204, P-238</td>
</tr>
<tr>
<td>Anholt, Robert</td>
<td>P-277</td>
</tr>
<tr>
<td>Antkowiak, Ireneusz</td>
<td>P-249</td>
</tr>
<tr>
<td>Appel, Mirjam</td>
<td>P-344</td>
</tr>
<tr>
<td>Arakawa, Aisaku</td>
<td>P-243</td>
</tr>
<tr>
<td>Arango, Jesus</td>
<td>P-162, P-374</td>
</tr>
<tr>
<td>Archibald, Alan</td>
<td>P-16</td>
</tr>
<tr>
<td>Arief, Vivi</td>
<td>P-281</td>
</tr>
<tr>
<td>Arië, Saad</td>
<td>P-193</td>
</tr>
<tr>
<td>Armstrong, Bruce</td>
<td>P-337</td>
</tr>
<tr>
<td>Arnesen, Petter</td>
<td>P-56</td>
</tr>
<tr>
<td>Arnevik, Espen</td>
<td>P-221, P-222</td>
</tr>
<tr>
<td>Arun, Meyyazhagan</td>
<td>P-221, P-222</td>
</tr>
<tr>
<td>Arun, S</td>
<td>P-238</td>
</tr>
<tr>
<td>Arya, Gunjan</td>
<td>P-136</td>
</tr>
<tr>
<td>Aryan, Golnar</td>
<td>P-136</td>
</tr>
<tr>
<td>Aryan, Hajar</td>
<td>P-309</td>
</tr>
<tr>
<td>Asabere, Akwasi</td>
<td>O-15</td>
</tr>
<tr>
<td>Aschard, Hugoé</td>
<td>O-25</td>
</tr>
<tr>
<td>Aslaveinjead, Ali Asghar</td>
<td>P-28</td>
</tr>
<tr>
<td>Asmara, Widya</td>
<td>P-225</td>
</tr>
<tr>
<td>Aso, Hisashi</td>
<td>P-309</td>
</tr>
<tr>
<td>Atlin, Gary</td>
<td>O-56</td>
</tr>
<tr>
<td>Augusto Franco Garcia, Antonio</td>
<td>P-357</td>
</tr>
<tr>
<td>Auinger, Hans-Jürgen</td>
<td>P-172, P-185, P-189</td>
</tr>
<tr>
<td>Aulchenko, Yuri</td>
<td>P-9</td>
</tr>
<tr>
<td>Auton, Adam</td>
<td>P-384, P-401</td>
</tr>
<tr>
<td>Auvinen, Petri</td>
<td>P-100</td>
</tr>
<tr>
<td>Auvray, Renoit</td>
<td>P-26</td>
</tr>
<tr>
<td>Avila, Victoria</td>
<td>P-113</td>
</tr>
<tr>
<td>Axenovich, Tatiana</td>
<td>P-9</td>
</tr>
<tr>
<td>Aylor, David</td>
<td>P-106</td>
</tr>
<tr>
<td>Ayroles, Julien F</td>
<td>O-55</td>
</tr>
<tr>
<td>Azhar, Aysha</td>
<td>P-385</td>
</tr>
<tr>
<td>Aznárez, Nitidia</td>
<td>P-320</td>
</tr>
<tr>
<td>Bachmanov, Alexander</td>
<td>P-274</td>
</tr>
<tr>
<td>Bagnato, Alessandro</td>
<td>P-353, P-400, P-402</td>
</tr>
<tr>
<td>Bahreini Behzadi, Mohammad Reza</td>
<td>P-27, P-28</td>
</tr>
<tr>
<td>Balachandar, Villingiri</td>
<td>P-221, P-222</td>
</tr>
<tr>
<td>Balamuralikrishnan, Balasubramanian</td>
<td>P-221, P-222</td>
</tr>
<tr>
<td>Balding, David</td>
<td>O-3</td>
</tr>
<tr>
<td>Baldwin, Ian</td>
<td>P-114</td>
</tr>
<tr>
<td>Ballantyne, Christie</td>
<td>P-252</td>
</tr>
<tr>
<td>Barrett, Amy</td>
<td>P-395, P-398</td>
</tr>
<tr>
<td>Basford, Kaye</td>
<td>P-16</td>
</tr>
<tr>
<td>Basford, Melissa A</td>
<td>P-296</td>
</tr>
<tr>
<td>Basnet, Ram Kumar</td>
<td>P-284</td>
</tr>
<tr>
<td>Bastiaansen, John</td>
<td>P-306</td>
</tr>
<tr>
<td>Bastin, Catherine</td>
<td>P-68</td>
</tr>
<tr>
<td>Bauchet, Guillaume</td>
<td>P-355</td>
</tr>
<tr>
<td>Baud, Amelie</td>
<td>P-312</td>
</tr>
<tr>
<td>Bauer, Andrea</td>
<td>P-319</td>
</tr>
<tr>
<td>Baxter, Simon W</td>
<td>P-233</td>
</tr>
<tr>
<td>Beattie, R Mark</td>
<td>P-369</td>
</tr>
<tr>
<td>Beaulieu, Jean</td>
<td>P-366</td>
</tr>
</tbody>
</table>
4th International Conference on Quantitative Genetics

Becker, Tim P-23, P-307
Bed’Hom, Bertrand P-362
Bell, Jordana P-77, P-391
Bell, Maureen P-389
Bell, Timothy P-106
Belonogova, Nadezhda P-9
Benedictus, Lindert P-74
Bennett, David P-278
Bennewitz, Jörn P-143, P-276
Benyamin, Beben O-43
Berenos, Camillo P-273
Bergelson, Joy O-29, P-81, P-130
Bergmann, Sven P-316
Berg, Peer P-196, P-202
Bermingham, Mairead L P-290
Berry, Diane SO-2, P-46, P-224
Bersabé, Diego P-111
Berstein, Yael P-301
Bevers, Roel PJ O-38
Beyene, Joseph O-56
Bezerra, Luiz P-342
Bharadwaj, Dwapayanan P-44
Bhuiyan, Md Atiquur Rahman P-67
Bidulescu, Aurelian P-252
Bielinski, Suzette J P-296
Bijma, Piter O-28, O-51, P-164, P-202, P-203, P-212, P-214, P-226 P-19, P-293
Bink, Marco P-278
Bird, David P-358
Birdwell, Kelly P-299
Bishop, Stephen C O-50, P-21, P-279, P-290, P-337
Bis, Josh P-278
Bix, Mark P-40
Blackwood, Douglas O-32, P-60, P-61, P-389
Blangero, John P-343
Blaxter, Mark L P-233
Blekhman, Ran P-77
Bloom, Joshua P-360
Blows, Mark O-5, P-105
Bochud, Murielle P-321
Bodlicker, Nicholas P-339
Bodensen, Natacha O-29
Boehringer, Stefan P-191
Boer, Martin P-156, P-275, P-293
Boerwinkle, Eric O-42
Bogdanovic-Sakran, Nada P-73
Boichard, Didier O-54
Bolhuis, J Elizabeth P-203
Bolstad, Ingeborg P-57
Bolund, Elisabeth P-92
Bond, Gareth P-30
Bongiorno, Fabricio P-2
Bonnett, David P-194
Bontrop, Ronald P-384
Boomsma, Dorret P-213
Bouquet, Alban P-154
Bourdeaux, Angélique P-352
Bousquet, Jean P-366
Boutin, S P-79
Bouvet, Jean-Marc P-124, P-180
Bouwhuys, Sandra P-82, P-92
Bovenhuis, Henk P-206
Bowden, Rory P-384
Bowman, Phillip J P-161
Bowton, Erica P-299
Bradbury, Peter P-169, P-310
Bradford, Yuki P-296
Braathar, Ariel P-42
Bravo, Luis E P-157
Brawner, Jeremy SO-1, P-43
Brenchley, Rachel C P-348
Bressler, Jan P-323
Brites, Michael P-161, P-166
Brosché, Mikael P-348
Brown, Andrew P-57
Brown-Gentry, Kristin P-267, P-269, P-303, P-365
Brown, Katherine P-125
Broxholme, John P-384
Brunak, Søren P-6
Bruneau, Nicolas P-362
Bruun, Camilla S P-59
Bryan, Glenn P-12
Bucher, Johan P-284
Buckler, Edward Q-2, O-22, O-56, P-169, P-310
Burgueno, Juan O-22, O-56, P-141, P-194
Burke, Terry P-119
Burnham, Nancy P-153
Buttazzoni, Luca P-402
Butchereit, Nina P-248
Buschbaum, Sarah P-252
Byers, Diane L O-6
Caballero, Armando P-96, P-108, P-113
Cabrera, Claudia P-53
Caetano, Sabrina P-327, P-342
Cai, Gengqian P-153
Calle, M Luz P-159
Callicrate, Taylor P-227
<table>
<thead>
<tr>
<th>Authors</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calò, Daniela Giovanna</td>
<td>P-402, O-39</td>
</tr>
<tr>
<td>Calbsbeek, Brittny</td>
<td>P-68, P-152, P-177, P-182, P-210, P-254</td>
</tr>
<tr>
<td>Calus, Mario</td>
<td>P-75, P-203</td>
</tr>
<tr>
<td>Camargo, Anyela V</td>
<td>O-60, P-18, P-21, P-53, P-285, P-297, P-298, P-354</td>
</tr>
<tr>
<td>Camerlink, Irene</td>
<td>P-204</td>
</tr>
<tr>
<td>Campbell, Harry</td>
<td>P-218</td>
</tr>
<tr>
<td>Campbell, Terry</td>
<td>P-180</td>
</tr>
<tr>
<td>Campos-Montes, Gabriel Ricardo</td>
<td>P-320</td>
</tr>
<tr>
<td>Camus-Kulandaivelu, Lètizia</td>
<td>P-2, P-37</td>
</tr>
<tr>
<td>Cánovas, Ángela</td>
<td>P-390</td>
</tr>
<tr>
<td>Cantet, Rodolfo J C</td>
<td>P-148</td>
</tr>
<tr>
<td>Canut, Matthieu</td>
<td>P-389</td>
</tr>
<tr>
<td>Carabano, Maria Jesus</td>
<td>O-8, O-40, P-39, P-49, P-234, P-287, P-295, P-377</td>
</tr>
<tr>
<td>Cardone, Laura</td>
<td>P-343</td>
</tr>
<tr>
<td>Carlberg, Örjan</td>
<td>P-159</td>
</tr>
<tr>
<td>Carless, Melanie</td>
<td>P-296</td>
</tr>
<tr>
<td>Carrato, Alfredo</td>
<td>P-93</td>
</tr>
<tr>
<td>Carrell, David</td>
<td>P-170, P-171</td>
</tr>
<tr>
<td>Carter, Patrick</td>
<td>P-402</td>
</tr>
<tr>
<td>Carvalheiro, Roberto</td>
<td>P-198, P-280</td>
</tr>
<tr>
<td>Casado, Rita</td>
<td>O-60</td>
</tr>
<tr>
<td>Casellas, Joaquim</td>
<td>P-263</td>
</tr>
<tr>
<td>Casey, Graham</td>
<td>P-389</td>
</tr>
<tr>
<td>Cassady, Joseph</td>
<td>P-174</td>
</tr>
<tr>
<td>Cass, Steve</td>
<td>P-218</td>
</tr>
<tr>
<td>Castells, Daniel</td>
<td>P-321</td>
</tr>
<tr>
<td>Castillo-Juárez, Héctor</td>
<td>P-355</td>
</tr>
<tr>
<td>Caulfield, Mark</td>
<td>P-224</td>
</tr>
<tr>
<td>Causse, Mathilde</td>
<td>P-15, P-268</td>
</tr>
<tr>
<td>Cavadino, Alana</td>
<td>P-55</td>
</tr>
<tr>
<td>Cavanagh, Colin</td>
<td>P-305</td>
</tr>
<tr>
<td>Chahota, RK</td>
<td>P-296</td>
</tr>
<tr>
<td>Chai, High Seng</td>
<td>P-375</td>
</tr>
<tr>
<td>Chamberlain, Amanda J</td>
<td>O-35</td>
</tr>
<tr>
<td>Chambers, John</td>
<td>O-27</td>
</tr>
<tr>
<td>Chang, Yu</td>
<td>P-159</td>
</tr>
<tr>
<td>Chanock, Stephen J</td>
<td>P-131</td>
</tr>
<tr>
<td>Chanetpe, Stephane</td>
<td>P-288</td>
</tr>
<tr>
<td>Charcosset, Alain</td>
<td>P-399</td>
</tr>
<tr>
<td>Charlesworth, Brian</td>
<td>P-35</td>
</tr>
<tr>
<td>Charlier, Carole</td>
<td>P-88</td>
</tr>
<tr>
<td>Charmantier, Anne</td>
<td>P-330</td>
</tr>
<tr>
<td>Chatziplis, Demetrios</td>
<td>P-235</td>
</tr>
<tr>
<td>Chaudhary, Garima</td>
<td>O-17</td>
</tr>
<tr>
<td>Chavali, Sreenivas</td>
<td>P-132</td>
</tr>
<tr>
<td>Cheli, Ivo</td>
<td>Chen, Guo-Bo</td>
</tr>
<tr>
<td>Chen, Jingchun</td>
<td>P-265</td>
</tr>
<tr>
<td>Chenoweth, Steve</td>
<td>P-122</td>
</tr>
<tr>
<td>Chen, Ruoyan</td>
<td>O-27</td>
</tr>
<tr>
<td>Chen, Xiangning (Sam)</td>
<td>P-265</td>
</tr>
<tr>
<td>Chesler, Elissa</td>
<td>P-106</td>
</tr>
<tr>
<td>Cheverud, James</td>
<td>P-102</td>
</tr>
<tr>
<td>Chevin, Luis-Miguel</td>
<td>P-352, P-97</td>
</tr>
<tr>
<td>Chew, Min Hoe</td>
<td>P-397</td>
</tr>
<tr>
<td>Chisholm, Rex L</td>
<td>P-296</td>
</tr>
<tr>
<td>Chopin, Nicolas</td>
<td>P-19</td>
</tr>
<tr>
<td>Choudary, Ahsan</td>
<td>P-343</td>
</tr>
<tr>
<td>Chris, Maleipaaard</td>
<td>P-284</td>
</tr>
<tr>
<td>Christensen, Ole F</td>
<td>P-173</td>
</tr>
<tr>
<td>Christodoulou, Katja</td>
<td>P-369</td>
</tr>
<tr>
<td>Christoforou, Andrea</td>
<td>O-14, P-50, P-325, P-340</td>
</tr>
<tr>
<td>Chute, Christopher G</td>
<td>P-296</td>
</tr>
<tr>
<td>Ciappesoni, Gabriel</td>
<td>P-174, P-323</td>
</tr>
<tr>
<td>Cicho'n, Mariusz</td>
<td>P-137</td>
</tr>
<tr>
<td>Cichon, Sven</td>
<td>P-325, P-50</td>
</tr>
<tr>
<td>Cirera, Susanna</td>
<td>P-59</td>
</tr>
<tr>
<td>Ciullo, Marina</td>
<td>P-321</td>
</tr>
<tr>
<td>Clark, Andrew</td>
<td>O-42, P-77</td>
</tr>
<tr>
<td>Clark, Samuel</td>
<td>P-183</td>
</tr>
<tr>
<td>Cleveland, Matthew</td>
<td>O-22, P-147</td>
</tr>
<tr>
<td>Clifton-Brown, John</td>
<td>P-128</td>
</tr>
<tr>
<td>Climent, José</td>
<td>P-85</td>
</tr>
<tr>
<td>Cloete, Schalk</td>
<td>P-47, P-328, P-329</td>
</tr>
<tr>
<td>Coan, Heather</td>
<td>P-343</td>
</tr>
<tr>
<td>Cochard, Benoît</td>
<td>P-180</td>
</tr>
<tr>
<td>Coffey, Mike</td>
<td>P-197</td>
</tr>
<tr>
<td>Coin, Lachlan</td>
<td>O-27, P-65</td>
</tr>
<tr>
<td>Cole, John</td>
<td>P-314</td>
</tr>
<tr>
<td>Colinet, Frédéric</td>
<td>P-68</td>
</tr>
<tr>
<td>Collins, Andrew</td>
<td>P-260, P-369</td>
</tr>
<tr>
<td>Colomé-Tatché, Maria</td>
<td>P-129, P-370</td>
</tr>
<tr>
<td>Colot, Vincent</td>
<td>P-370</td>
</tr>
<tr>
<td>Coltman, DW</td>
<td>P-79</td>
</tr>
<tr>
<td>Comadran, Jordi</td>
<td>P-291</td>
</tr>
<tr>
<td>Concordet, Didier</td>
<td>O-20</td>
</tr>
<tr>
<td>Conomos, Matthew</td>
<td>P-364</td>
</tr>
<tr>
<td>Cooper, Matthew</td>
<td>P-193</td>
</tr>
<tr>
<td>Coote, David</td>
<td>O-54, P-375</td>
</tr>
<tr>
<td>Corre, Tanguy</td>
<td>P-266</td>
</tr>
<tr>
<td>Cortijo, Sandra</td>
<td>P-370</td>
</tr>
<tr>
<td>Cossins, Andrew R</td>
<td>O-38</td>
</tr>
<tr>
<td>Costa e Silva, João</td>
<td>P-219</td>
</tr>
<tr>
<td>Costa e Silva, Luciano</td>
<td>P-357</td>
</tr>
<tr>
<td>Cotton, James</td>
<td>P-205</td>
</tr>
<tr>
<td>Coville, Jean-Luc</td>
<td>P-362</td>
</tr>
</tbody>
</table>
Authors

Crawford, Dana  P-267, P-269, P-299, P-303, P-365
Cresko, William  P-217
Crivelli, Francesco  P-316
Crooijmans, Richard  P-374
Crooker, Brian  P-314
Crooks, Lucy  P-377
Cros, David  P-124, P-180
Cross, Jose  O-22, O-56, P-141, P-194
Crump, Ron E  P-210
Cubillos, Francisco  P-390
Cule, Erika  P-144
Curik, Ivo  P-361
Curran, Joanne  P-343
Cuyabano, Beatriz  P-178
da Cruz, Valdey Aparecida Rocha  P-324
Daetwyler, Hans  O-54, P-375
Dagnachew, Binyam  P-3
Dalh, Kjell  P-385
Dambrova, Maija  P-64
Daoutidou, Erica  P-311
da Silva Filho, Miguel Inacio  P-322
da Silva, Thiago Bruno Ribeiro  P-324
Dastani, Zari  P-252
Davey, John W  P-233
David, Ingrid  O-20
David, Jacques  P-345
Davies, Gail  P-246, P-266, P-278
Davis, John M  P-199
Dawson, Julie  P-194
Da, Yang  P-314
de Andrade, Mariza  P-296
Deary, Ian  P-246, P-278, P-389
De Baets, Bernard  P-33
Debette, Stéphanie  P-278
Debnath, Marlies Alexandra  P-107
De Cara, Ángeles  P-80, P-229
De Cauwer, Isabelle  P-261
Decombes, Isabelle  P-352
De Donato, Marcos  P-163
Dedoussis, George  P-311
Deeb, Nader  P-339
de Haas, Yvette  P-152, P-182
De Iorio, Maria  P-144
Dekkers, Jack  P-243
de Klerk, Nicholas  P-193
de Koning Dirk-Jan  P-76, P-162, P-292, P-213
de Kort, Janneke  P-16
de los Campos, Gustavo  O-22, O-24, O-56, P-141, P-146
Deloukas, Panos  O-18, P-391, P-395, P-398
Dembeck, Lauren  P-244
de Mestral Bezanson, Louise  P-126
Denis, Marie  P-180
Denny, Joshua  P-296, P-299
de Oliveira Santos - Garcia, Melissa  P-357
de Rezende Neves, Haroldo  P-170, P-171
Henrique  P-379, P-395, P-398
Demirtzakis, Emmanouil T  O-34
Diaz, Clara  P-148
Dick, Danielle  P-208
Dicke, Marcel  P-226
Dick, Katherine  P-391
Dieters, Mark  P-16, P-43
Diezel, Celia  P-114
Dillmann, Christine  P-98, P-181
Ding, Xiangdong  P-50, P-57
Djurovic, Srdjan  P-255
Dobson, Richard  P-26
Dodd, Ken  P-366
Dorfls, Trevor  P-213, P-236
Doeschl-Wilson, Andrea  P-353, P-400, P-402
Dominiczak, Anna  P-382
Donato, Celeste  P-73
Dondrup, Michael  P-50
Do, Ngoc Duy  P-115
Donnelly, Peter  P-384
Donnison, Iain  P-128
Doonan, John H  P-75
Do, Thi Phuc  P-355
Douglas, Alex  P-121
Dreisinger, Susanne  P-141, P-194
Drechel, Dmitriy  P-307
Drobnjak, Szymon  P-137
Druet, Tom  P-155, P-35
Duchen, Pablo  P-118
Dudbridge, Frank  P-176
Dugdale, Hannah L  P-119
Duijvestijn, Naomi  P-190, P-212
Dumitrescu, Logan  P-267
Dunlop, Malcolm  O-60
Dunn, IC  P-292
Durand-Gasselin, Tristan  P-180
Durbén, Richard  P-117
Duthie, Carol-Anne  P-351
D. Valente, Bruno  P-55
Dyer, Tom  P-343
Authors Index

Eaglen, Sophie P-197
Eberle, Michael P-396
Ebert, Dieter P-326
Ecarnot, Martin P-345
Eccles, Diana P-260
Edriss, Mohammad P-374
Edriss, Vahid P-187
Edwards, David P-286, P-392
Edwards, Stefan P-396
Eghbalsaied, Shahin SO-3, P-41, P-48
Ehrenreich, Ian P-38
Ehsani, Alireza P-360
Eilers, Paul P-286
Eklund, Niina P-266
Eldik, van, Peter P-74
Eleftherhorinou, Hariklia P-65
El Faro, Lenira P-327
Elken Søderby, Ida P-332, P-359
Eldik, van, Peter P-226
Elliott, Phil P-214
Elsen, Jean-Michel P-273
Emes, Richard O-20, P-347
Endelman, Jeffrey P-125
Engen, Steinar P-194
Engert, Andreas P-95
Ennis, Sarah P-322
Epik, Tineke P-369
Eppink, Tineke T P-214
E Silva, Luciano P-24
Esko, Tönu P-266
Espeseth, Thomas P-14, P-325, P-340
Essafi, Abdelkader P-354
Etta, Hannah P-387
Etukudoh, Idorenyin P-387
Evans, Kathryn P-328
Everett-Hincks, Julie P-322
Fahrenkrug, Scott P-387
Fairbairn, Daphne J P-387
Fairfax, Ben P-328
F. Allan, Mark P-26
Fan, Longjiang P-393
Fan, Zhang O-27
Farlow, Ashley O-14
Farnaki, Aliki-Eleni O-27
Farnir, Frederic O-14
Farquharson, Colin O-27
Farrar, Kerrie P-311
Felleki, Majbritt P-35
Ferencakovic, Maja P-76
Fernandes, Carla P-128
Fernandez, Ana I P-25
Fernandez, Jesus P-361
Fernández Martín, Jesús P-325
Fernando, Rohan P-325
Fernie, Alisdair P-325
Fest, Brad P-325
Fievet, Julie P-199
Fikse, Freddy P-243
Fitzpatrick, Annette P-252
Fitzpatrick, Darren J P-252
Flavell, Richard P-352
Fledel-Alon, Adi P-352
Fleet, Malcolm P-352
Fleischer, Robert P-352
Fleming, Katy P-352
Flinn-Garcia, Sherry P-352
Flint, Jonathan P-352
Fisar, Tina P-352
Foster, Katharina P-352
Fontanesi, Luca P-352
Forai, Petar P-322
Forst, Asta P-352
Forsmeier, Wolfgang P-352
Fortes, Marina P-352
Fox, Caroline P-352
Froster, Hunter P-352
Fraser, Ross P-352
Frayling, Tim P-352
Fredholm, Merete P-352
Freeman, Sue P-352
Frérot, Hélène P-352
Freyer, Gertraude P-352
Fries, Michael P-352
Fries, Ruedi P-352
Fuchs, Michael P-352
Fukawa, Kazuo P-352
Fulcher, Nick P-352
Fulssack, Phillip P-352
Fulton, Janet P-352
Fusi, Nicolo P-352
Galimberti, Giuliano P-352
Gallais, André P-352
Gallinger, Steve P-352
Garbe, John P-352
Garcia, Antonio P-352
Garcia-Closas, Montserrat P-352
García Cortés, Luis Alberto P-352
García-Dorado, Aurora P-352
García-Ruiz, Adriana P-352

Authors
Gardner, Candice P-310
Garrick, Dorian O-57, P-163, P-199, P-243
Gautier, Mathieu P-362
Gauzere, Julie P-103
Gay, Alan P P-75
Gay, Laurene P-103, P-264
Geiger, Jerome P-339
Gelernter, Joel P-56
Gelley, Laura O-32
Gengler, Nicolas P-71
George, Andrew P-17, P-268
Georges, Michel P-35
Gerstein, Mark O-15
Gean, Salvador P-367
Ghian, Elena O-32, P-389
Ghita, Elena P-89
Gianola, Daniel O-55
Gibbs, Richard A P-55
Gibson, Greg O-31
Gibson, Jane P-260, P-369
Giddaluru, Sudheer P-50, P-325, P-340
Gieger, Christian P-347
Gilson, Christopher P-252
Gilmour, Arthur P-25
Gimeno, Diego P-174, P-323
Gingera, Gregory P-313
Glahn, David P-60
Glass, Elizabeth J P-290
Glaubitz, Jeff O-45, P-161, P-247
Goddard, Michael O-54, P-270
Godé, Cécile P-352
Gödeckdak, Sami P-249
Goh, Xuewei P-397
Goldberg, Virginia P-174, P-323
Golden, Darach P-36
Goldstein, Pavel P-207
Gols, Rieta P-226
Gomez, Leonardo P-291
Gómez-Romano, Fernando P-229
Gonen, Serap P-337
González-Martinez, Santiago C P-120
Gonzalez-Recio, Oscar P-66, P-167
Goodloe, Robert P-267
Goodrich, Julia P-77
Goonewardene, Laki P-22
Gore, Michael A P-310
Goring, Harald P-343
Gorrell, JC P-79
Gorum, Richard P-336
Gourichon, David P-362
Grabher, Manfred P-135
Grace, Christopher P-282
Graf, Alexandra P-400
Gras, Mikhail Alexandru P-86, P-89
Grasso, Nicolás P-174
Gray, Alan P-11
Gredler, Birgit P-184
Greene, Derek P-283
Green, Kathry P-193
Greenwood, Pamela P-340
Griffiths, Simon P-336
Grooten, Martien P-306, P-374
Grondin, Cecile P-390
Groot, Steven P-284
Grùs, Horia P-86, P-89
Grubert, Fabian O-15
Grundberg, Elin P-395, P-398
Grupioni, Natália P-342
Guadagnini, Maria-Dolores P-258
Guilhaume, François P-55
Guimaraes, Simone Facioni P-35
Gutierrez, Beatriz P-390
Gy, Isabelle P-284
Gylensten, Ulf P-354
Habegger, Lukas O-15
Habier, David O-57, P-243
Hackett, Christine P-12
Hadfield, Jarrod P-10
Hadjipavlou, Georgianna O-30
Haesaert, Geert P-33
Haesen, Ernst P-33
Hagiy, Koichi P-33
Haja, N Kadarimeen P-33
Hajnal, Alex P-179
Haley, Chris P-115
Hall, Ian P-33
Gold, Michael A P-308, P-316
Hall, Ian P-330
Hall, Jeremy P-60, P-61
Hall, Matthew P-326
Halpin, Claire P-291
Hancock, Angela P-130
Hansen, Thomas F P-87
Hansen, Torben P-44
Hanski, Ilkka P-100
Haque, Abdul P-73
Harbison, Susan P-242
Hardy, John O-13
Hariharan, Manoj O-15
Harlizius, Barbara P-190
Harris, Bevin P-246, P-389
Harris, Sarah P-225
Hartatik, Tety P-225
Harun, Abdul Rahim P-67
Hastie, Nicholas P-21, P-297, P-354
Havill, Lorena P-343
Hawkins, Sarah P-128
Hayashi, Takeshi P-344
Hayes, Ben O-45, O-54, P-155, P-161, P-375
Hayward, Caroline P-18, P-21, P-53, P-278, P-297, P-298, P-354, P-271
Hayward, Nicholas P-395, P-398
Hedman, Asa K O-15
Heffelfinger, Chris P-306
Heidariatabar, Marzieh O-14
Hellmann, Ines O-41
Hemani, Gibran P-322
Hemminki, Kari P-379
Henders, Anjali P-389
Hennah, William P-252
Henneman, Peter P-154, P-173
Henryon, Mark P-126, P-127
Herbinger, Christophe P-56
Herman, Aryeh P-320
Hernández, Jules P-384
Hernandez, Ryan P-188, P-258, P-259
Hernandez-Sanchez, Jules P-157
Hernandez-Suarez, Gustavo P-307
Herold, Christine P-157
Herran, Oscar P-157
Herranz, Jesús P-159
Heuven, Henri P-74
Hickey, John O-22, O-56, P-147, P-183, P-184
Hilibrant, Maarten P-281
Hill, William P-243
Hingorani, Aron P-46
Hinrichs, Dirk P-276
Hirose, Kensuke P-344
Hivert, Marie-France P-252
Hocking, Lynne P-382
Hodgson, Edward P-128
Holand, Anna Marie P-31
Holcman, Antonija P-72
Holland, James B P-310
Holloway, John P-369
Holm, Jens-Christian P-44
Hongo, Mi-Young O-15
Hopfen, Corinna P-281
Hopper, John O-60
Horton, Matthew O-29, P-401
Horvat, Simon P-76
Ho, Shwen P-394
Hospital, Frédéric P-362
Hough, Denise P-47
Houlston, Richard O-60
Houston, Ross D P-337
Hovig, Eivind P-57
Huang, Emma P-15
Huang, Wen O-39
Huang, Yijian P-147, P-263
Huber, Christian O-14
Huber, Wolfgang O-15
Huffman, Jennifer P-18, P-21, P-53, P-354
Hulsegge, Ina O-54
Humbug, Peter P-384
Humphries, MM P-79
Hunt, Laura P-303
Hu, Xiaoxiang P-287
Huxley Jones, Julie P-282
Hypponen, Elina P-46, P-224
Ibrahim-Verbaas, Carla P-278
Iftikhar, Tayyba P-73
Illingworth, Chris P-101
Imumori, Ikhide P-163
Iqbal, Zamin P-384
Irani, Alireza P-136
Iriarte, Wanda P-174
Irwijn, Kristen P-93
Ishikawa, Akira P-272, P-315
Isik, Fikret P-145
Ito, Tetsuya P-344
Itoarte, Catherine P-217
Ivakhno, Sergii P-396
Izanwati, Rohayu P-403
Jackson, Phillip P-230
Jacquin, Julien P-347
Jafarkia, Mohsen P-192
Jahan Bakhshi, Abbas P-300
Authors

James, Marianne P-121
Jamieson, Sarra P-193
Jannink, Jean-Luc P-194
Janousova, Eva P-14
Jansen, Hans P-293
Jansen, Ritsert C P-370
Jans, Luc P-4, P-48, P-286
Jaramillo, Juan Pablo P-120
Jaritz, Markus P-400
Jasson, Sylvain P-288
Jeff, Janina P-269
Jensen, Elaine P-128
Jensen, Henrik P-31
Jensen, Jimmy P-57
Jensen, Just P-346
Jensen, Per O-36, P-52
Jianfeng, Jicai P-160
Jiang, Li P-41, P-48
Jiggins, Chris D P-233
Jiggins, Frank P-394
Jiménez-Montero, Jose Antonio O-16, P-129, P-370
Johannes, Frank P-6
Johansen, Morten O-8, P-377
Johansson, Anna P-142
Johnsson, Martin O-36, P-52
Jonas, Elisabeth P-331
Jonassen, Inge P-50
Jones, Gareth P-205
Jørgensen, Claus B P-59
Jorjani, Hossein P-245
Joseph, Jimcy P-216
Joshi, Peter P-18
Juga, Jarmo P-154
Juliette, de Meaux P-81
Juneja, Punita P-394
Junge, Wolfgang P-248
Kadarmideen, Haja P-44, P-59, P-257
Kadri, Naveen Kumar P-350
Kairudin, Narimah Md P-67
Kammenga, Jan O-38, P-45, P-335
Kämpe, Olle O-36
Kandel, Purna Bhadra P-71
Kanis, Egbert P-351
Kanoni, Stavroula P-391
Kapan, Durrell K P-233
Kaprio, Jaakko P-208
Karczewski, Konrad O-15
Karlskov-Mortensen, Peter P-59
Karoui, Sofiene P-148
Karthick Kumar, Alagamuthu P-221, P-222

Karupaiah, Tilakavathi P-67
Kasowski, Maya Q-15
Katayama, Yuuki P-309
Katoh, Kazuo P-309
Kause, Antti P-206
Kayser, Manfred P-191
Kazdova, Ludmila P-58
Kedenko, Lyudmyla P-252
Keildson, Sarah P-395, P-398
Keinan, Alon O-42
Kelton, David P-250
Kempenaers, Bart P-332, P-359
Kemper, Kathryn P-270
Kenig, Bojan P-110
Kent, Jack P-343
Kerdaffrec, Envel P-363
Kerr, Richard P-219
Kerr, Shona P-382
Khan, Naqib Ullah P-318
Khansefid, Majid P-139
Khatkar, Mehar S P-10, P-257
Kho, Abel P-296
Khondoker, Mizanur P-255
Kierczak, Marcin P-234, P-49
Kim, Eui-Soo P-314
King, Elizabeth G P-256, P-289
Kinghorn, Brian O-22, P-32, P-183
Kinsman, Alison P-192
Kirin, Mirna P-278, P-298
Kirkwood, Carl P-73
Kirst, Matias P-199, P-367
Kitazawa, Haruki P-309
Kizilkaya, Kadir P-163
Klar, Joakim P-385
Klein, Etienne P-103
Klementidis, Yann C O-24, P-146
Klovins, Janis P-64
Knaak, Carsten P-185
Knap, Pieter P-351
Knieß, Ulrich P-332
Knight, Julian P-62
Knol, Egbert Frank P-190, P-212
Knott, Sara O-41
Knowles, David P-117
Kocabey, Samet P-249
Koeck, Astrid P-250
Koehler, Katja P-308
Koets, Ad P-74
Kogelman, Lisette JA P-59
Koegvinas, Manolis P-159
Koh, Poh Koon P-397
Koistinen, Petri
Komdeur, Jan
Komisarek, Jolanta
Konrade, Ilze
Korbel, Jan
Korol, Abraham
Korsten, Peter
Korte, Arthur
Koskinen, Patrik
Kottgen, Anna
Koumas, Alkis
Kovac, Milena
Kozich, Viktor
Kraft, Peter
Kram, Melissa
Krmik, Julius
Kruuk, Loeske EB
Kuittinen, Helmi
Kullo, Ifikhar J
Kumar, Naresh
Kumar, Suresh
Kuningas, Maris
Kuo, Po-Hsiu
Kurtz, Theodore
Kvalnes, Thomas
Kvist, Jouni
Kyriazakis, Illias
Laakso, Markku
Labouriau, Rodrigo
Lachowiec, Jennifer A
Lacour, Andre
Lahti, Jarbi
Lai, Yin-Chieh
Lamichhaney, Sangeet
Lamparter, David
Landa, Vladimir
Langenberg, Claudia
Largey, Fabienne
Lariepe, Amandine
Larson, Eric B
Larzul, Catherine
Lata, Sneh
Lauc, Gordan
Launer, Lenore
Law, Andy
Lawlor, Debbie
Lawrence, Faye
Lawrence, Neil
Lawrie, Stephen
Lazar, Cristina
Le Bihan-Duval, E
Ledur, Mónica Correa
Lee, Dongwhan
Lee, Hong
Lee, Youngjo
Leffler, Ellen
Legarra, Andres
Le Hellard, Stephanie
Lehermeier, Christina
Lehnert, Sigrid
Lehtimäki, Terho
Lehtonen, Rainer
Lempe, Janne
Léon, Jens
Leppäälä, Johanna
Le, Quang Nam
Le Rouzic, Arnaud
Levin, Michael
Ley, Ruth
Lian, Shangli
Liewald, Dave
Li, Huihui
Li, Jiasheng
Li, Junya
Li, Lin
Li, Meng
Lin, Cailu
Lindström, Sara
Li, Ping
Li, Yuxin
Li, Zengting
Li, Xidan
Li, Yingrui
Li, Yongle
Löbo, Rainsido
Loi, Tien Tau

Authors

Authors Index

Koistinen, Petri  P-319
Komdeur, Jan  P-119
Komisarek, Jolanta  P-277
Konrade, Ilze  P-64
Korbel, Jan  O-15
Korol, Abraham  P-207
Korsten, Peter  P-134
Korte, Arthur  P-294, P-363
Koskinen, Patrik  P-100
Kottgen, Anna  P-321
Koumas, Alkis  P-330
Kovac, Milena  P-72
Kozich, Viktor  P-58
Kraft, Peter  O-25
Kram, Melissa  O-32, P-389
Krmik, Julius  O-22, P-399
Kruuk, Loeske EB  P-19
Kuittinen, Helmi  P-133
Kullo, Ifikhar J  P-123
Kumar, Naresh  P-296
Kumar, Suresh  P-305
Kuningas, Maris  P-343
Kuo, Po-Hsiu  P-266
Kurtz, Theodore  P-333
Kvalnes, Thomas  P-58
Kvist, Jouni  P-95
Laakso, Markku  P-100
Kyriazakis, Illias  P-231
Labouriau, Rodrigo  P-252
Lachowiec, Jennifer A  P-237
Lacour, Andre  P-378
Lahti, Jarbi  P-233
Lai, Yin-Chieh  P-287
Lamichhaney, Sangeet  P-316
Landa, Vladimir  P-252
Langenberg, Claudia  P-316
Largey, Fabienne  P-335
Lariepe, Amandine  P-288
Larson, Eric B  P-296
Larzul, Catherine  P-186
Lata, Sneh  P-354
Lauc, Gordan  P-278
Launer, Lenore  P-374
Law, Andy  P-252
Lawlor, Debbie  P-244
Lawrence, Faye  P-380
Lawrie, Stephen  P-60, P-61
Lazar, Cristina  P-86, P-89
Le Bihan-Duval, E  P-292
Ledur, Mónica Correa  P-324
Lee, Dongwhan  P-25
Lee, Hong  P-25
Lee, Youngjo  P-247
Leffler, Ellen  P-25
Legarra, Andres  P-384
Le Hellard, Stephanie  P-148
Lehermeier, Christina  P-14, P-50, P-325, P-340
Lehnert, Sigrid  P-183, P-189
Lehtimäki, Terho  P-149
Lehtonen, Rainer  P-252
Lempe, Janne  P-100
Léon, Jens  P-378
Leppäälä, Johanna  P-319
Le, Quang Nam  P-348
Le Rouzic, Arnaud  P-115
Levin, Michael  P-87, P-91
Ley, Ruth  P-65
Lian, Shangli  P-77
Liewald, Dave  P-40
Li, Huihui  P-246
Li, Jiasheng  P-112
Li, Junya  P-169
Li, Lin  P-112
Li, Meng  P-169
Lin, Cailu  P-274
Lindström, Sara  O-25
Li, Ping-I  P-287
Li, Yi  P-253
Lipkin, Ehud  P-353
Lipschutz-Powell, Debby  O-28
Li, Rongling  P-296
Listman, Jennifer  P-56
Liti, Gianni  P-117
Liu, George  P-314
Liu, Huiming  P-196
Liu, Jianfeng  P-160, P-98
Liu, Xiaodong  P-292
Liu, Yanqing  P-397
Liu, Yun  P-393
Liu, Zengting  P-240
Li, Xidan  P-39, P-234, P-292
Li, Yingrui  O-27
Li, Yongle  P-194
Löbo, Rainsido  P-104, P-342
Loi, Tien Tau  P-397
Authors

4th International Conference on Quantitative Genetics

Long, Anthony D  P-256
Long, Quan  O-14, P-294, P-363, P-401
Long, Tony  P-289
Loos, Ruth  P-252
Lopes, Marcos Soares  P-190
Lopes, Paulo Sávio  P-76
Lopes Pinto, Fernando  P-159
López de Maturana, Evangelina  P-113
López-Fanjul, Carlos  P-390
Loudet, Olivier  P-125
Louis, Ed  P-20
Lukic, Boris  P-162
Lu, Ming-Kun  P-82, P-92
Lummaa, Virpi  P-234
Lundberg, Andreas E  P-14, P-340
Lundervold, Astri J  O-54, P-161, P-166, P-173, P-178, P-181, P-187, P-334, P-338
Lund, Mogens Sandø  P-384
Lunter, Gerton  P-333
Lu, Ru-Band  P-271, P-302
Lu, Yi  P-244
Lyman, Richard  P-274
Lysenko, Anna  P-252
Lyytikäinen, Leo-Pekka  P-127
Macdonald, Stuart J  P-298
MacEachern, Kathryn  P-271
MacGillivray, Tom  P-389
Macgregor, Stuart  P-366
Macintyre, Donald  O-33, O-39, O-55, P-204, P-238, P-242, P-244, P-308
MacKay, John  P-285
Mackay, Trudy  P-155
MacLeod, Andrew  P-263
MacLeod, Iona  P-76
MacNeil, Michael  O-17
MacRae, Vicky  P-237, P-41
Madsen, Per  P-33
Maenhout, Steven  P-265
Maes, Hermine  O-39, P-244
Magwire, Michael  P-232
Mahjani, Behrang  P-385
Mahmood Baig, Shahid  P-336
Ma, Jun  P-201
Maksha, Yosphe  P-62
Makino, Seiko  P-356
Maki-Tanila, Asko  P-159
Malats, Núria  O-42, P-314
Ma, Li  P-296
Malinowski, Jennifer R  P-384
Maller, Julian  P-19, P-156, P-275, P-293
Malosetti, Marcos  P-72
Malovrh, Špela  P-145, P-147, P-263
Maltecca, Christian  P-22
Manafiazar, Ghader  P-140
Manfredi, Eduardo  P-288
Mangin, Brigitte  P-221, P-222
Manikantan, Pappusamy  P-99, P-211
Manovani, Roberto  P-258, P-259, P-320
Manunza, Arianna  P-386
Mao, Qian  P-166
Ma, Peipei  P-24
Margarido, Gabriel  P-314
Mariani, Brian  P-362
Marie-Hélène, Marie-Hélène  P-349
Marigorta, Urko M  P-112
Marilyn, Warburton  P-234, P-377
Marklund, Stefan  P-59
Mark, Thomas  P-233
Maroja, Luana S  P-402
Martelli, Pier Luigi  P-135
Martinez Barrio, Alvaro  P-121
Martinez-Padilla, Jesus  O-43, P-271
Martin, Nicholas  P-31
Martino, Sara  P-296
Masy, Daniel R  P-159
Mata, Douglas  P-319
Mathew, Bobby  P-21, P-279
Matika, Oswald  P-296
Matsumoto, Martha  P-344
Matsumoto, Toshimi  P-50
Mattheisen, Manuel  P-134
Mattyszen, Erik  P-50, P-57
Mattingsdal, Morten  P-330
Mavrogenis, Andreas  O-32
Mavruk, Senem  P-128
Ma, Xue-Feng  P-365
Mayo, Ping  P-98
Ma, Yulong  P-79
McAdam, AG  P-290
McBride, Stewart H  P-153
McCarthy, Linda  O-37, P-354, P-395, P-398
McCarthy, Mark  P-301, P-389
McCarthy, Shane E  P-296
McCarty, Catherine A  O-32, P-301, P-389
McCormie, Richard  P-306
McCoy, Lenovia P-242
McDonald, Bruce P-83
McDowell, Stanley WJ P-290
McEvoy, Brian P-266
McFadden, Thomas P-22
McFarlane, SE SO-7, P-79
McGregor, Alistair P-281
McIntosh, Andrew O-32, P-60, P-61, P-389
McKeand, Steve P-145
McKirdy, James P-61
McLean, Karen P-12
McMullen, Michael D P-310
McNicol, Stephen P-289
McNeill, Geraldine P-246
McQueen-Mason, Simon P-291
McQuillan, Ruth P-266
McRea, Allan P-379, P-389
McVeal, Gil P-379, P-389
Medici, Vasco P-316
Meduri, Eshwar P-391, P-395, P-398
Megens, Hendrik-Jan P-374
Mehrabani Yeganeh, Hassan P-374
Meigs, James P-300
Meinshaussen, Nicolai P-252
Melle, Ingrid P-30
Melton, Cord P-57
Meng, Dazhe O-14, P-363, P-401
Meng, Lei P-1
Mercadé, Anna P-320
Mercader, Anna-Maria P-258
Meuwissen, Theo P-3, P-138, P-165, P-372
Meyer, Karin P-8
Meyers, Jacqueline P-208
Michael, Iloje P-7
Middelberg, Rita O-43
Miglior, Filippo P-250
Mikawa, Satoshi P-344
Millard, Mark P-310
Millar, Kirsty P-389
Miller, Darla P-106
Miller, Rebecca R O-6
Milne, Elizabeth P-193
Miliadou, Despoina P-330
Mitchell, Asia D P-368
Mitchell, Sabrina P-303
Mitchell, Sharon O-22, O-56
Mittelsten Scheid, Ortrun P-363
Mohana Devi, Subramaniya P-221, P-222
Mohlke, Karen P-252
Mollah, Md Bazlur Rahman P-272
Montaldo, Hugo H P-175, P-218
Montana, Giovanni P-14
Montgomery, Grant O-43, P-271, P-379
Moradi Shahrbabak, Mohammed P-300
Moreau, Laurence P-288
Moreno, Carlos P-280
Morris, Andrew P-252, P-282, P-382
Morris, Stewart O-32, P-389
Mosley, Thomas P-278
Mott, Richard P-312
Mrode, Raphael P-197
Muddassir Ali, Muhammad O-32, P-389
Muir, Walter O-53
Mukherjee, Sayan O-21
Mulder, Han P-210, P-228
Muller, Stephanie P-324, P-327, P-342
Munari, Danisio P-199
Munoz Del Valle, Patricio R P-367
Munoz, Patricio P-303
Murdock, Deborah P-121
Murray-Dickson, Gill P-221, P-222
Mustaq Ahamed, Shafi Ahammed P-101
Khan P-354
Muzinic, Ana P-63
Muzzey, Dale P-384
Myers, Simon P-292
Nagamine, Yoshitaka P-21, P-179
Nakajyo, Mitsuru P-309
Nascimento, Guilherme P-327
Navajas, EA P-323
Navarro, Arcadi P-349
Navarro, Nicolas P-251
Navarro, Pau P-18, P-21, P-29, P-53, P-179, P-297
Nejati Javaremi, Ardeshir P-300
Nelson, Chris P-391
Nelson, Ronald P-39
Nettelblad, Carl P-232
Newcombe, Paul P-153
Newman, Sheryl-Anne P-26
Ng, Esther P-62
Nguyen, Trong Binh P-115
Nguyen, Van Ba P-115
Nicolella, Daniel P-343
Nidelet, Thibault O-34
Nielsen, Bjarne P-237
Nielsen, Dahlia P-358
Nielsen, Rasmus O-27
Nielsen, Vivi 
Nisbet, James 
Nishio, Motohide 
Nishiura, Akiko 
Nivard, Michel G 
Nizhynska, Viktoria 
Nobrega, Marcelo 
Nordborg, Magnus 
Norouzian, Mohammad Ali 
Norris, Ashie 
Nötchen, Markus M 
Ntalla, Ioanna 
Nunes, Beatriz 
Nunes, Daniela 
Oakey, Helena 
Ober, Ulrike 
O’Brien, Deborah 
Oddou-Muratorio, Sylvie 
Odegård, Jørgen 
Odet, Fanny 
Oetjens, Matthew 
Ogden, Rob 
Okamura, Toshihiro 
Okine, Erasmus 
Okumura, Naohiko 
Okuno, Sin-ichi 
Oldmeadow, Christopher 
Oliveira, Henrique 
Onovo, Josiah 
Onyia, Vincent 
O’Reilly, Patrick 
Orford, Michael 
Orford, Simon 
Oscier, David 
O’Seaghdha, Conall 
Osei-Poku, Jewelna 
Ostensen, Tage 
O’Sullivan, Neil 
Ouwehand, Willem 
Ouwehand, Willem H 
Ouzunova, Milena 
Owen, Katharine 
Paaby, Annalise 
Pacheco, Jennifer A 
Pain, Arnab 
Palin, Kimmo 
Palmer, Colin 
Pankow, James 
Panoutsopoulou, Kalliopi 
Papachristoforou, Christos 
Papadopoulou, Dimitra 
Papmeyer, Martina 
Parasuraman, Raja 
Pardo-Manuel de Villena, Fernando 
Parker, Anton 
Parla, Jennifer 
Parts, Leopold 
Pass, Chloë 
Paterno, Flavia 
Pathak, Jyotishman 
Paul, Dirk 
Paulin, Lars 
Paulo, Maria João 
Paulweber, Bernhard 
Paz, Claudia 
Peculis, Raitis 
Pedersen, Louise 
Pedersen, Oluf 
Peeters, Katrijn 
Pegolo, Newton 
Peiffer, Jason A 
Peissig, Peggy L 
Peixoto, Jane de Oliveira 
Pellegrino, Laura 
Pelmus, Rodica Stefania 
Pemberton, Josephine 
Peñagaricano, Francisco 
Pena, Ramona Natacha 
Pena, Romi 
Pereira de Souza, Anete 
Pérez-Enciso, Miguel 
Pérez-Figueroa, Andrés 
Perez, Jesus 
Perez, Paulinho 
Perez, Paulino 
Peter, Gary 
Petersen, Thomas 
Peterson, Roseann 
Peters, Sunday 
Petrovich, Katja 
Pettay, Jenni E 
Pettersson, Mats 
Petu, Ibukunle 
Pfeifer, Susanne 
Pham, Doan Lan 
Phillips, Patrick 
Phocas, Florence 
Piano, Fabio 
Picornell, Antoni C 
Pierney, Stuart 
Pillai, Meenu
Pimentel Maia, Rafael P-237
Pinto, Carlos P-36
Pirastu, Nicola P-266
Pistol, Gina Cecilia P-89
Pita, Fabiano P-313
Platt, Alexander P-294
Platzer, Alexander O-14, P-363
Poland, Jesse P-194
Pomp, Daniel O-44, P-286
Ponader, Sabine P-322
Pong-Wong, Ricardo P-21, P-29, P-162, P-179, P-188, P-279, P-290, P-297
Porteous, David O-32, P-285, P-382, P-389
Porth, Ilga P-42
Posthuma, Danielle P-236
Potts, Brad P-219
Powell, Joseph P-379
Powers, Carol P-313
Pravenec, Michal P-58
Prendergast, James P-376
Przeworski, Molly P-384
Pszczola, Marcin P-182
Punternoll, Pál P-50
Q-MEGA and AMFS investigators, Queiroz, Sandra P-271
Quesada, Humberto P-223
Quintanilla, Raquel P-90
Raadsma, Herman W P-258, P-259, P-320
P-10, P-94, P-257, P-331
O-14, P-363
Rabanal, Fernando P-225
Rachman, Mifta Pratiwi P-216
Radhika, G P-135
Rafati, Nima P-32
Rafique, Sehrish O-32
Raghavan, K C P-216
Raina, Anupuma P-235
Raitakari, Olli P-252
Ramakrishnan, Kamna P-389
Ramasesha, Adakalavan O-13
Ramayo-Caldas, Yuliaxis P-388
Ramos-Osins, Sebastián P-371
Ramos, Salvador P-327, P-324
Randi, Nima P-10
Ransome, Roger P-205
Rasmussen, Luke P-296
Ratnam, Wickneswari P-67
Rayner, William P-311
Real, Francisco X P-159
Rebolini, Danea P-389
Reecy, James P-163
Reed, Danielle P-274
Reetz, Tobias P-319
Reid, Jane O-11
Reiner, Benaim, Anat P-207
Reinhardt, Fritz P-240
Reinwang, Ivar P-14, P-340
Reixach, Josep P-258
Remington, David P-123
Remula, Saana P-123
Rendon, Augusto O-18
Repapi, Emmanuelela P-30
Resende Jr., Marcio P-199, P-367
Resende, Marcos DV P-199, P-367
Restrepo, Nicole P-365
Reverter, Antonio P-149
Reynolds, Rose P-217
Ricardo Bachega Feijó Rosa, João P-357
Richards, John P-252
Richardson, David S P-119
Richards, Stephen O-55
Ridhwan, Muhamad P-403
Rietschel, Marcella P-50
Rigio, Valentina P-21, P-279
Riha, Karel P-363
Riku, Jyunana P-309
Ri, Mika P-309
Rincón, Gonzalo P-174
Ripley, Van P-313
Ritchie, Marylyn D P-296
Ritland, Carol P-42
Ritland, Kermit P-42
Rivas, María José P-90
Robatzek, Silke P-81
Robert, Alexandre P-131
Robinson, Andy P-192
Robinson, Matthew P-215
Robinson, Matt R P-261
Robson, Paul P-128
Rocha, Dominique O-54
Rockman, Matthew O-7
Rodon, Dan M P-296, P-299
Rodrigues Silva, Renato P-357
Rodríguez, M Carmen P-388
Rodríguez Perez, Iria P-36
Rodríguez-Quilón, Isabel P-120
Rodríguez Sanchez, Miriam P-335
Roeh, Rainer P-351
Roff, Derek A P-133
Rojas, Patria P-157
Rolán-Álvarez, Emilio P-90
Romaniuk, Liana P-60
<table>
<thead>
<tr>
<th>Authors</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Román-Ponce, Sergio Iván</td>
<td>P-175</td>
</tr>
<tr>
<td>Romay, M Cinta</td>
<td>P-310</td>
</tr>
<tr>
<td>Ronfort, Joelle</td>
<td>P-264</td>
</tr>
<tr>
<td>Rönnegård, Lars</td>
<td>P-5, P-25, P-228, P-295</td>
</tr>
<tr>
<td>Rooodt-Wilding, Rouvay</td>
<td>P-328</td>
</tr>
<tr>
<td>Rosa, Guilherme</td>
<td>P-342</td>
</tr>
<tr>
<td>Rose, Richard</td>
<td>P-208</td>
</tr>
<tr>
<td>Rossiter, Stephen</td>
<td>P-205</td>
</tr>
<tr>
<td>Rossoni, Attilio</td>
<td>P-353</td>
</tr>
<tr>
<td>Rothman, Nathaniel</td>
<td>P-159</td>
</tr>
<tr>
<td>Roumet, Pierre</td>
<td>P-345</td>
</tr>
<tr>
<td>Routtu, Jarkko</td>
<td>P-326</td>
</tr>
<tr>
<td>Rowe, Suzanne</td>
<td>P-162, P-246</td>
</tr>
<tr>
<td>Rowlatt, Amy</td>
<td>P-69</td>
</tr>
<tr>
<td>Royce, Thomas</td>
<td>P-396</td>
</tr>
<tr>
<td>Rozowsky, Joel</td>
<td>O-15</td>
</tr>
<tr>
<td>Rubin, Carl-Johan</td>
<td>P-135</td>
</tr>
<tr>
<td>Rudan, Igor</td>
<td>P-18, P-21, P-53, P-297, P-298, P-354</td>
</tr>
<tr>
<td>Rudd, Pauline</td>
<td>P-354</td>
</tr>
<tr>
<td>Ruiz-López, Felipe de Jesús</td>
<td>O-21</td>
</tr>
<tr>
<td>Runcie, Daniel</td>
<td>P-402</td>
</tr>
<tr>
<td>Russo, Vincenzo</td>
<td>O-32</td>
</tr>
<tr>
<td>Ryan, Niamh</td>
<td>P-290</td>
</tr>
<tr>
<td>Ryder, Jon J</td>
<td>P-135</td>
</tr>
<tr>
<td>Ryman, Nils</td>
<td>O-13</td>
</tr>
<tr>
<td>Ryten, Mina</td>
<td>P-100</td>
</tr>
<tr>
<td>Saastamoinen, Marjo</td>
<td>P-67</td>
</tr>
<tr>
<td>Sabu, Kalluvettankuzhy K</td>
<td>P-95</td>
</tr>
<tr>
<td>Saether, Bernt-Erik</td>
<td>P-334, P-338, P-350</td>
</tr>
<tr>
<td>Sahana, Goutam</td>
<td>P-309</td>
</tr>
<tr>
<td>Sakuma, Akiko</td>
<td>P-391</td>
</tr>
<tr>
<td>Samani, Nilesh</td>
<td>P-353, P-400</td>
</tr>
<tr>
<td>Samore, Antonia Bianca</td>
<td>P-157</td>
</tr>
<tr>
<td>Sanabria, Carolina</td>
<td>P-136</td>
</tr>
<tr>
<td>Sanati, Mohamad Hossein</td>
<td>P-124</td>
</tr>
<tr>
<td>Sanchez, Leopoldo</td>
<td>P-108</td>
</tr>
<tr>
<td>Sánchez-Molano, Enrique</td>
<td>P-157</td>
</tr>
<tr>
<td>Sanchez, Myriam</td>
<td>P-328</td>
</tr>
<tr>
<td>Sandenbergh, Lise</td>
<td>P-289</td>
</tr>
<tr>
<td>Sanderson, Brian</td>
<td>P-395, P-398</td>
</tr>
<tr>
<td>Sandling, Johanna K</td>
<td>P-323</td>
</tr>
<tr>
<td>San Julián, R</td>
<td>P-221, P-222</td>
</tr>
<tr>
<td>Sankar, Kathanan</td>
<td>P-37</td>
</tr>
<tr>
<td>Santos-Cristal de Sivak, Mónica</td>
<td>P-120, P-85</td>
</tr>
<tr>
<td>Santos-del-Blanco, Luis</td>
<td>P-261</td>
</tr>
<tr>
<td>Santure, Anna W</td>
<td>P-400</td>
</tr>
<tr>
<td>Santus, Enrico</td>
<td>P-184, P-192</td>
</tr>
<tr>
<td>Sargolzaei, Mehdi</td>
<td>P-35</td>
</tr>
<tr>
<td>Sartelet, Arnaud</td>
<td>P-211, P-99</td>
</tr>
<tr>
<td>Sartori, Cristina</td>
<td>P-195, P-34</td>
</tr>
<tr>
<td>Sasaki, Osamu</td>
<td>P-221, P-222</td>
</tr>
<tr>
<td>Sasikala, Kesavarao</td>
<td>P-34, P-195</td>
</tr>
<tr>
<td>Satoh, Masahiro</td>
<td>P-309</td>
</tr>
<tr>
<td>Sato, Takumi</td>
<td>P-90</td>
</tr>
<tr>
<td>Saura, María</td>
<td>P-355</td>
</tr>
<tr>
<td>Sauvage, Christopher</td>
<td>P-249</td>
</tr>
<tr>
<td>Savage, Sinead</td>
<td>P-327, P-342</td>
</tr>
<tr>
<td>Savagno, Rodrigo</td>
<td>P-123</td>
</tr>
<tr>
<td>Savolainen, Outi</td>
<td>P-262</td>
</tr>
<tr>
<td>Sayers, Ian</td>
<td>O-46</td>
</tr>
<tr>
<td>Schadt, Eric</td>
<td>P-19</td>
</tr>
<tr>
<td>Schäfer, Christian</td>
<td>P-192, P-250</td>
</tr>
<tr>
<td>Schenkel, Flavio</td>
<td>P-353</td>
</tr>
<tr>
<td>Schiavini, Fausta</td>
<td>P-402</td>
</tr>
<tr>
<td>Schiavo, Giuseppina</td>
<td>P-332, P-359</td>
</tr>
<tr>
<td>Schielzeth, Holger</td>
<td>O-53</td>
</tr>
<tr>
<td>Schinchel, Allan</td>
<td>O-55</td>
</tr>
<tr>
<td>Schlather, Martin</td>
<td>P-335</td>
</tr>
<tr>
<td>Schmid, Tobias</td>
<td>P-189</td>
</tr>
<tr>
<td>Schmit, Volker</td>
<td>P-275</td>
</tr>
<tr>
<td>Schnabel, Sabine</td>
<td>P-365</td>
</tr>
<tr>
<td>Schnetz-Boutaud, Nathalie</td>
<td>O-59</td>
</tr>
<tr>
<td>Schoen, Chris-Carolin</td>
<td>P-329</td>
</tr>
<tr>
<td>Scholtz, Anie</td>
<td>P-249</td>
</tr>
<tr>
<td>Scholz, Claus-Jürgen</td>
<td>P-172, P-185, P-189</td>
</tr>
<tr>
<td>Schön, Chris-Carolin</td>
<td>P-158</td>
</tr>
<tr>
<td>Schopen, Ghyslain</td>
<td>O-26</td>
</tr>
<tr>
<td>Schork, Nicholas</td>
<td>P-158</td>
</tr>
<tr>
<td>Schroten, Chris</td>
<td>P-184</td>
</tr>
<tr>
<td>Schuler, Urs</td>
<td>P-307</td>
</tr>
<tr>
<td>Schüller, Vitalia</td>
<td>P-338</td>
</tr>
<tr>
<td>Schulman, Nina</td>
<td>P-402</td>
</tr>
<tr>
<td>Scotti, Emilio</td>
<td>P-252</td>
</tr>
<tr>
<td>Scott, Robert</td>
<td>P-339</td>
</tr>
<tr>
<td>Sebbana, Saad</td>
<td>P-363</td>
</tr>
<tr>
<td>Sedman, Laura</td>
<td>P-184</td>
</tr>
<tr>
<td>Seefried, Franz</td>
<td>P-294, P-355</td>
</tr>
<tr>
<td>Segura, Vincent</td>
<td>P-384</td>
</tr>
<tr>
<td>Segurel, Laure</td>
<td>P-395, P-398</td>
</tr>
<tr>
<td>Sekowska, Magdalena</td>
<td>O-56</td>
</tr>
<tr>
<td>Semagn, Kassa</td>
<td>P-376</td>
</tr>
<tr>
<td>Semple, Colin</td>
<td>P-157</td>
</tr>
<tr>
<td>Serrano, Martha</td>
<td>P-57</td>
</tr>
<tr>
<td>Server, Andres</td>
<td>P-278</td>
</tr>
<tr>
<td>Seshadi, Sudha</td>
<td>P-243</td>
</tr>
<tr>
<td>Settar, Petek</td>
<td>P-291</td>
</tr>
<tr>
<td>Shafiei, Reza</td>
<td>P-17</td>
</tr>
<tr>
<td>Shah, Rohan</td>
<td>P-305</td>
</tr>
<tr>
<td>Sharma, TR</td>
<td>P-261</td>
</tr>
<tr>
<td>Sheldon, Ben C</td>
<td>P-287</td>
</tr>
</tbody>
</table>
Shen, Xia  P-39, P-295
Shields, Denis  P-283
Shimazu, Tomoyuki  P-309
Shi, Minyi  O-15
Shin, So-Youn  P-398
Shirali, Mahmoud  P-351
Shirali, Masoud  P-21, P-297
Shojaei Arani, Mohammad  P-129
Shrine, Nick  P-262
Siegel, Paul  O-8, P-234
Siewerdt, Frank  P-227
Sigvartsen, Niels  P-57
Sikkink, Kristin  P-217
Silio, Luis  P-388
Sillanpää, Mikko  P-319
Silva, Fabiano Fonseca  P-190
Silveri, Licia  P-362
Silver, Jacob  P-368
Silverman, Debra  P-159
Simianer, Henner  O-55
Simmons, Andy  P-255
Simonsen, Anna  P-105
Simpson, Michael  P-369
Sinaiko, Alan  P-252
Sing, Charles  O-42
Siol, Mathieu  P-264
Skapare, Elina  P-64
Skarman, Axel  P-48
Skelly, Daniel A  P-378
Skuce, Robin A  P-290
Slate, Jon  P-215, P-261
Slavov, Gancheo  P-128
Small, Kerrin  P-398
Small, Kerrin S  P-395
Smith, Albert  P-278
Smith, Blair  P-382
Smith, Colin  O-13
Snoek, L Basten  O-38, P-335
Snyder, Mike  O-15
Soares, Dinesh  O-32, P-389
Soellner, Johann  P-361
Soler Artigas, María  P-262
Soller, Morris  P-353, P-400
Sonstegard, Tad  P-314
Sood, BC  P-305
Soranzo, Nicole  O-18, O-35
Sørensen, Anders Christian  P-154, P-196
Sørensen, Daniel  O-24, P-4
Sørensen, Peter  P-41, P-48, P-286, P-350
Sorrells, Mark  P-194
Southam, Lorraine  P-311
Soyeurt, Hélène  P-68, P-71
Spector, Tim  P-77, P-391, P-395, P-398
Speed, Doug  O-3
Spencer, Kylee  P-365
Sperone, Felice  P-130
Sprooten, Emma  P-60, P-61
Spyropoulou, Zoi  P-75
Stainton, John  P-399
Stamatoyannopoulos, John A  P-378
Stamenkovic-Radak, Marina  P-110
Stamer, Eckhard  P-248
Starr, John  P-246, P-389
Stein, Vidar M  P-14, P-50, P-325, P-340
Stefansson, Tryggvi  P-83
Stegle, Oliver  P-363, P-380, P-390
Steinsland, Ingelin  P-31
Stephan, Wolfgang  P-118
Stephens, Jonathan  O-18
Sterken, Mark G  O-38
Stevens, Rebecca  P-355
Stewart, Isobel  O-60, P-11
Stinchcombe, John  P-105
Stinnett, Sandy  P-153
Stone, Eric  O-39, O-55, P-204
Stothard, Paul  O-54
Strabel, Tomasz  P-182
Strandberg, Erling  P-228
Street, Teresa  P-384
Strefford, Jonathan  P-30
Stricker, Christian  O-55
Suda, Yoshihito  P-309
Su, Guosheng  P-161, P-166, P-173, P-178, P-181, P-187
Sullivan, Alessandra MV  P-378
Sundström, Görel  P-135
Sunduimijid, Bolormaa  P-149
Suresh Kumar, Shanmugam  P-221, P-222
Sussman, Jess  P-61
Sussmann, Jessika  P-60
Suzuki, Keiichi  P-309
Sverdlov, Serge  P-200
Svishcheva, Gulnara  P-9
Swaller, Timothy  P-128
Swart, Pieter  P-47
Swarup, Shilpa  P-238
Szyda, Joanna  P-277
Taha, Suketoshi  P-112
Tachmazidou, Ioanna  P-311
Tamini, Rulla  O-25
<table>
<thead>
<tr>
<th>Authors</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanaskovic, Marija</td>
<td>P-110</td>
</tr>
<tr>
<td>Tandon, Nikhil</td>
<td>P-44</td>
</tr>
<tr>
<td>Tang, Choong Leong</td>
<td>P-397</td>
</tr>
<tr>
<td>Tang, Jimmy</td>
<td>P-36</td>
</tr>
<tr>
<td>Tang, Zaixiang</td>
<td>P-241</td>
</tr>
<tr>
<td>Taniguchi, Masaaki</td>
<td>P-344</td>
</tr>
<tr>
<td>Tanimoto, Hiromu</td>
<td>P-249</td>
</tr>
<tr>
<td>Tan, Soo Yong</td>
<td>P-397</td>
</tr>
<tr>
<td>Tapio, Miika</td>
<td>P-338</td>
</tr>
<tr>
<td>Tapper, William</td>
<td>P-260, P-369</td>
</tr>
<tr>
<td>Tararu, Ionelia</td>
<td>P-89</td>
</tr>
<tr>
<td>Tardón, Adonina</td>
<td>P-159</td>
</tr>
<tr>
<td>Tariq, Muhammad</td>
<td>P-385</td>
</tr>
<tr>
<td>Tarlinton, Rachael</td>
<td>P-125</td>
</tr>
<tr>
<td>Tenesa, Albert</td>
<td>O-60, P-11, P-157, P-69, P-246</td>
</tr>
<tr>
<td>Teótonio, Henrique</td>
<td>P-132</td>
</tr>
<tr>
<td>Teplitsky, Celine</td>
<td>P-131</td>
</tr>
<tr>
<td>Terčič, Dušan</td>
<td>P-72</td>
</tr>
<tr>
<td>Ter Braak, Cajo</td>
<td>P-19</td>
</tr>
<tr>
<td>Tetens, Jens</td>
<td>P-240, P-248</td>
</tr>
<tr>
<td>Teumer, Alexander</td>
<td>P-321</td>
</tr>
<tr>
<td>Thuiller, Georg</td>
<td>P-240, P-248, P-276</td>
</tr>
<tr>
<td>Thanabalasingham, Gaya</td>
<td>P-354</td>
</tr>
<tr>
<td>Theodorides, Maria</td>
<td>P-274</td>
</tr>
<tr>
<td>Tholon, Patricia</td>
<td>P-223</td>
</tr>
<tr>
<td>Thomas-Jones, Sian</td>
<td>P-128</td>
</tr>
<tr>
<td>Thomas, Milton</td>
<td>P-163</td>
</tr>
<tr>
<td>Thomsen, Bo</td>
<td>P-48</td>
</tr>
<tr>
<td>Thomsen, Hauke</td>
<td>P-322</td>
</tr>
<tr>
<td>Thomson, Peter C</td>
<td>P-10, P-94, P-257, P-331</td>
</tr>
<tr>
<td>Thomson, Pippa</td>
<td>O-32, P-61, P-389</td>
</tr>
<tr>
<td>Tier, Bruce</td>
<td>P-151</td>
</tr>
<tr>
<td>Tiwari, Dev</td>
<td>P-284</td>
</tr>
<tr>
<td>Tobin, Martin</td>
<td>P-262</td>
</tr>
<tr>
<td>Todhunter, Rory</td>
<td>P-169</td>
</tr>
<tr>
<td>Toffey, David</td>
<td>P-368</td>
</tr>
<tr>
<td>Toivainen, Tuomas</td>
<td>P-123</td>
</tr>
<tr>
<td>Tomlinson, Ian</td>
<td>O-60</td>
</tr>
<tr>
<td>Tompkins, Emily</td>
<td>SO-8, P-109</td>
</tr>
<tr>
<td>Toro, Miguel Angel</td>
<td>P-80, P-372</td>
</tr>
<tr>
<td>Torrance, Helen</td>
<td>O-32</td>
</tr>
<tr>
<td>Trabzuni, Daniel</td>
<td>O-13</td>
</tr>
<tr>
<td>Tran, Thi Thu Thuy</td>
<td>P-115</td>
</tr>
<tr>
<td>Tran, Xuan Hoan</td>
<td>P-115</td>
</tr>
<tr>
<td>Tribout, Thierry</td>
<td>P-186</td>
</tr>
<tr>
<td>Trivedi, Trupti</td>
<td>P-153</td>
</tr>
<tr>
<td>Troukhan, Maxim</td>
<td>P-128</td>
</tr>
<tr>
<td>Tsai, Hsiu Ting</td>
<td>P-373</td>
</tr>
<tr>
<td>Tsaprouni, Loukia G</td>
<td>P-391</td>
</tr>
<tr>
<td>Turner, F</td>
<td>P-292</td>
</tr>
<tr>
<td>Tzamaloukas, Ouranios</td>
<td>P-330</td>
</tr>
<tr>
<td>Uchida, Hiroshi</td>
<td>P-309</td>
</tr>
<tr>
<td>Uduak, Ogundu</td>
<td>P-7</td>
</tr>
<tr>
<td>Uemoto, Yoshinobu</td>
<td>P-21, P-29</td>
</tr>
<tr>
<td>Uenishi, Hirohide</td>
<td>P-344</td>
</tr>
<tr>
<td>Ugarte, Eva</td>
<td>P-66</td>
</tr>
<tr>
<td>Umana, Adriana</td>
<td>P-157</td>
</tr>
<tr>
<td>Upstill-Goddard, Rosanna</td>
<td>P-260</td>
</tr>
<tr>
<td>Urban, Alexander</td>
<td>O-15</td>
</tr>
<tr>
<td>Vagenas, Dimitrios</td>
<td>P-188</td>
</tr>
<tr>
<td>Vaitsiakhovich, Tatsiana</td>
<td>P-307</td>
</tr>
<tr>
<td>Valentine, John</td>
<td>P-128</td>
</tr>
<tr>
<td>van 't Hof, Arjen E</td>
<td>O-38</td>
</tr>
<tr>
<td>van Arendonk, Johan</td>
<td>P-164, P-351</td>
</tr>
<tr>
<td>van Bockxmeer, Frank</td>
<td>P-193</td>
</tr>
<tr>
<td>van Dalen, Sacha</td>
<td>P-206</td>
</tr>
<tr>
<td>van der Harst, Pim</td>
<td>O-35</td>
</tr>
<tr>
<td>Vanderick, Sylvie</td>
<td>P-26</td>
</tr>
<tr>
<td>Van der Merwe, Aletta</td>
<td>P-328</td>
</tr>
<tr>
<td>van der Sluis, Sophie</td>
<td>P-236</td>
</tr>
<tr>
<td>Vander Voort, Gordon</td>
<td>P-192</td>
</tr>
<tr>
<td>van der Werf, Julius</td>
<td>O-22, P-183</td>
</tr>
<tr>
<td>van Duijn, Cornelia</td>
<td>P-9, P-252, P-278, P-19</td>
</tr>
<tr>
<td>Van Eeuwijk, Fred</td>
<td>P-156, P-275, P-293</td>
</tr>
<tr>
<td>van Grevenhof, Ilse</td>
<td>P-164</td>
</tr>
<tr>
<td>van Overveld, Thijs</td>
<td>P-134</td>
</tr>
<tr>
<td>VanRaden, Paul</td>
<td>P-184</td>
</tr>
<tr>
<td>Van Tassell, Curt</td>
<td>O-54, P-314</td>
</tr>
<tr>
<td>van Wyk, Japie</td>
<td>P-329</td>
</tr>
<tr>
<td>Varona, Luis</td>
<td>P-280</td>
</tr>
<tr>
<td>Vaughan, Joseph</td>
<td>P-317</td>
</tr>
<tr>
<td>Vazquez, Ana Ines</td>
<td>O-24, P-146</td>
</tr>
<tr>
<td>Veerkamp, Roel</td>
<td>O-54, P-152, P-177, P-182, P-210, P-254</td>
</tr>
<tr>
<td>Velandai, Srikanth</td>
<td>P-278</td>
</tr>
<tr>
<td>Velander, Ingela</td>
<td>P-162</td>
</tr>
<tr>
<td>Velayutham, Dinesh</td>
<td>P-353</td>
</tr>
<tr>
<td>Venkat, Aarti</td>
<td>P-384</td>
</tr>
<tr>
<td>Venn, Oliver</td>
<td>P-384</td>
</tr>
<tr>
<td>Vensko, Steven</td>
<td>O-39</td>
</tr>
<tr>
<td>Ventrurini, Guilherme Costa</td>
<td>P-324</td>
</tr>
<tr>
<td>Verblya, Arunas</td>
<td>P-150</td>
</tr>
<tr>
<td>Verblya, Klara</td>
<td>P-150</td>
</tr>
<tr>
<td>Vereijken, Addie</td>
<td>P-306</td>
</tr>
<tr>
<td>Vergara, Pablo</td>
<td>P-121</td>
</tr>
<tr>
<td>Verhage, Matthijs</td>
<td>P-236</td>
</tr>
<tr>
<td>Verhoeven, Merel</td>
<td>P-226</td>
</tr>
<tr>
<td>Verne, Sebastian</td>
<td>P-42</td>
</tr>
<tr>
<td>Verweij, Karin</td>
<td>P-116</td>
</tr>
</tbody>
</table>
Authors

Authors Index

Vesimäki, Timo  P-123
Vetter, Madlen  P-81
Vicován, Gabriel  P-86
Víctor Mela Obinna, Okoro  P-7
Villholmsson, Bjarni  O-14, P-294, P-363
Vilki, Johanna  P-338
Villanueva, Beatriz  P-66, P-80 P-229, P-231, P-388, P-345
Vilmus, Ingrid  P-46, P-224
Vimalawar, Karani P-368
Vinson, Amanda  P-214
Visscher, Jeroen  O-1, O-43, P-247, P-270, P-278, P-379, P-389
Vittart, Veronique  P-18, P-21, P-53, P-297, P-298, P-321, P-354, P-75
Vlachonasios, Konstantinos  P-118
Voigt, Sussane  P-308, P-316
Vonesch, Sibylle  O-18
von Strandmann, Elke Pogge  P-322
Voss, Katrin  P-14
Vounou, Maria  P-115
Vu, Chi Cuong  P-226
Wade, Michael J  P-262
Wain, Louise  P-317
Waites, Richard  P-56
Walderhaug, Espen  P-252
Walker, Mark  O-13
Walker, Robert  P-197
Wall, Eileen  P-191
Walsh, Susan  P-1
Wang, Jiankang  P-392
Wang, Lei  P-373
Wang, Po-Hui  P-314
Wang, Shengwen  P-296
Wang, Xiaoming  P-169, P-181
Wang, Yachun  O-27
Wang, Yong  P-393
Wang, Yu  P-22
Wang, Zhiquan  P-370
Wardenaar, René  P-205
Ward, Helen  P-252
Wareham, Nick  P-252
Warren, Liling  O-15
Waszak, Sebastian  P-252
Waterworth, Dawn  O-32, P-389
Watson, James  P-399
Watson, Kellie  P-291
Waugh, Robbie  O-13
Weale, Michael  P-123
Webb, Bradley (Todd)  P-125
Webster, Lucy  P-121
Wei, Julong  P-98
Weir, Bruce  P-364
Weissman, Jonathan  P-63
Weissman, Sherman  O-15
Wei, Wenhua  P-220
Wei, Xianning  P-230
Wei, Xinghua  P-393
Wellmann, Robin  P-143, P-276
Wen, Weiwei  P-112
Wenceslau, Raphael  P-55
Werge, Thomas  P-325
Westbye, Lars T  P-14
Weston, Noah  P-296
Whalley, Heather  P-60, P-61
Whetten, Ross  P-145
White, Amelia  O-7
Whithead, Caragh  P-291
Whitfield, John  O-43, P-46, P-153, P-282
Widyas, Nuzul  P-346
Wiener, Pam  P-374
Wiener, Pamela  P-399
Wentjes, Yvonne  P-172
Wiggans, George  P-314
Wilches, Ricardo  P-118
Wild, Sarah  P-298
Wilk, Alicja  P-395, P-398
Wilke, Russell A  P-296
Wilkinson, Samantha  P-374
Willems, Glenda  P-363
Williams, John  P-374
Williamson, Valerie  P-358
Willis, Claire  P-369
Willi, Yvonne  P-83, P-84
Wilson, Alastair J  P-133, P-205
Wilson, James F  P-18, P-53, P-266, P-298, P-381
Wilson, Jim  P-354
Wimmer, Valentin  P-172, P-185, P-189
Wingen, Luzie U  P-336
Winkelmann, Anne  P-142
Winn, John  P-369
Wiskin, Anthony  P-133
Wolak, Matthew E  P-243
Wolc, Anna  P-359
Wolf, Jochen BW  P-191
Wollstein, Andreas  P-396
Wong, SW Wendy  O-28, P-162, P-290, P-165, P-197, P-374
Wootliams, John A

278
<table>
<thead>
<tr>
<th>Authors</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wray, Naomi</td>
<td>P-247</td>
</tr>
<tr>
<td>Wright, Alan</td>
<td>P-21, P-18, P-53, P-297, P-298, P-354</td>
</tr>
<tr>
<td>Wright, David M</td>
<td>P-290</td>
</tr>
<tr>
<td>Wright, Dominic</td>
<td>O-36, P-52</td>
</tr>
<tr>
<td>Wright, Victoria</td>
<td>P-65</td>
</tr>
<tr>
<td>Wurzelmann, John</td>
<td>P-153</td>
</tr>
<tr>
<td>Wu, Weiren</td>
<td>P-209</td>
</tr>
<tr>
<td>Wu, Ying</td>
<td>P-252</td>
</tr>
<tr>
<td>Xu, Chenwu</td>
<td>P-241</td>
</tr>
<tr>
<td>Xu, Chun-Fang</td>
<td>P-153</td>
</tr>
<tr>
<td>Xu, Shuqing</td>
<td>P-114</td>
</tr>
<tr>
<td>Yaghootkar, Hanieh</td>
<td>P-252</td>
</tr>
<tr>
<td>Yamazaki, Takeshi</td>
<td>P-179</td>
</tr>
<tr>
<td>Yang, Jian</td>
<td>P-209, P-247</td>
</tr>
<tr>
<td>Yang, Qiong</td>
<td>P-278</td>
</tr>
<tr>
<td>Yang, Rong-Cai</td>
<td>O-49, P-91</td>
</tr>
<tr>
<td>Yang, Shun-Fa</td>
<td>P-373</td>
</tr>
<tr>
<td>Yang, Xiaohong</td>
<td>P-112</td>
</tr>
<tr>
<td>Yan, Jianbing</td>
<td>P-112</td>
</tr>
<tr>
<td>Yao, Jianchao</td>
<td>O-32, P-389</td>
</tr>
<tr>
<td>Yarali, Ayse</td>
<td>P-249</td>
</tr>
<tr>
<td>Yeh, Hesther</td>
<td>P-42</td>
</tr>
<tr>
<td>Ye, Kai</td>
<td>O-14</td>
</tr>
<tr>
<td>Yu, Lei</td>
<td>P-278</td>
</tr>
<tr>
<td>Yung, Tsun-Po</td>
<td>P-391</td>
</tr>
<tr>
<td>Yu, Xijiang</td>
<td>P-165</td>
</tr>
<tr>
<td>Zabaleta, Jovanny</td>
<td>P-157</td>
</tr>
<tr>
<td>Zahri, M K</td>
<td>P-403</td>
</tr>
<tr>
<td>Zain, Abdullah Md</td>
<td>P-67</td>
</tr>
<tr>
<td>Zaitlen, Noah</td>
<td>O-25</td>
</tr>
<tr>
<td>Zapata-Valenzuela, Jaime</td>
<td>P-145</td>
</tr>
<tr>
<td>Zeggini, Eleftheria</td>
<td>P-311</td>
</tr>
<tr>
<td>Zeron-Medina, Jorge</td>
<td>P-30</td>
</tr>
<tr>
<td>Zgaga, Lina</td>
<td>P-278</td>
</tr>
<tr>
<td>Zhang, Liling</td>
<td>P-239</td>
</tr>
<tr>
<td>Zhang, Liping</td>
<td>P-128</td>
</tr>
<tr>
<td>Zhang, Luyan</td>
<td>P-1</td>
</tr>
<tr>
<td>Zhang, Qin</td>
<td>P-160, P-166, P-169, P-181, P-98</td>
</tr>
<tr>
<td>Zhang, Qingrun</td>
<td>O-14, P-363, P-401</td>
</tr>
<tr>
<td>Zhang, Yonghong</td>
<td>P-241</td>
</tr>
<tr>
<td>Zhang, Yuandan</td>
<td>P-149</td>
</tr>
<tr>
<td>Zhang, Zhe</td>
<td>P-169</td>
</tr>
<tr>
<td>Zhang, Zhiwu</td>
<td>P-169, P-310</td>
</tr>
<tr>
<td>Zhang, Zhiyan</td>
<td>P-35</td>
</tr>
<tr>
<td>Zhao, Jing Hua</td>
<td>P-20</td>
</tr>
<tr>
<td>Zheng, Yan</td>
<td>P-209</td>
</tr>
<tr>
<td>Zhou, Lei</td>
<td>P-181</td>
</tr>
<tr>
<td>Zhou, Shanshan</td>
<td>P-204</td>
</tr>
<tr>
<td>Zhou, Zhengkui</td>
<td>P-169</td>
</tr>
<tr>
<td>Zhu, Dianhui</td>
<td>O-55</td>
</tr>
<tr>
<td>Zhu, Jun</td>
<td>O-48, P-209</td>
</tr>
<tr>
<td>Zhu, Qianhao</td>
<td>P-393</td>
</tr>
<tr>
<td>Zidek, Vaclav</td>
<td>P-58</td>
</tr>
<tr>
<td>Zietsch, Brendan</td>
<td>P-116</td>
</tr>
</tbody>
</table>
Aqua Gen manages and further develops the world first family-based breeding programs of Atlantic salmon and rainbow trout. Today’s eggs from the Aqua Gen strain are selected for excellent performances on growth, disease resistance, robustness and filet quality by use of state-of-the art methods within quantitative and molecular genetics. Aqua Gen was nominated as “Norway’s most innovative company” in 2010 by the Research Council of Norway.
Delivering success now and for the future

Commitment to research and development is one of the foundations of Aviagen’s success:

- Breeding programme focussed on delivering balanced products
- Management of genetic resources for balancing genetic response and long-term fitness of elite lines
- Unrivalled product portfolio
- Balanced products meet diverse global requirements
- Working with universities, research institutes and external experts to enhance knowledge base
- Incorporating cutting-edge science to routine breeding programmes
Gene polymorphism of growth hormone (GH) in one of Indonesian cattle (Madura) and Madura crossed cattle

Mits Prastiwi Rachman1, Widy Aasmari1, and Tety Maruti2

1Laboratory of Animal Breeding, Faculty of Animal Science Universitas Gadjah Mada, Yogyakarta, Indonesia.
2Laboratory of Microbiology, Faculty of Veterinary Medicine Universitas Gadjah Mada, Yogyakarta, Indonesia.

Corresponding author: mits_rachman@yahoo.com

INTRODUCTION

Madura cattle was one of Indonesian local beef cattle which has good reproductive performance also adaptable to Indonesian environment and its feeding management (Huitema, 1982). Madura cattle has been crossed with Limousine cattle by AI (Artificial Insemination) program, and the offspring was known as Limousine-Madura (Madura crossed). Mostly Indonesian traditional farmer indicated that both of cattle have differences on their body growth. Growth hormone helps in body growth and development through protein synthesis, protein deposition in tissues and organs (Gluckman et al., 2007). It has been assigned to 19q26-ter position of bovine chromosome (Hediger et al., 1990), with five exons and four introns. Several study was reported that there were polymorphism of growth hormone gene in local cattle also its crossed and the present investigation was carried out to find polymorphism of growth hormone gene among Madura cattle and Madura crossed cattle based on the differences of their body growth in Indonesian traditional farm.

OBJECTIVES

The objective of the present study was to identify polymorphism of growth hormone (GH) gene in one of Indonesian local cattle (Madura) and Madura crossed (Limousine-Madura) cattle, also its association with growth traits.

METHODS

Whole genome isolation from blood
(Genomic DNA Isolation Kit, Genet Bio)

Amplification of 211 bp growth hormone gene fragment

Primer: GH (forward) 5’GATGCGTCTGAGGGGCCTGC-3’
(reverse) 5’TATATATCTCAGTACGGCTGCTCG-3’

(Enzyme) Taq polymerase

PCR program:

Pre-denaturation: 95°C for 5 min
Denaturation: 95°C for 50 sec
Annealing: 65°C for 30 sec
Extention: 72°C for 50 sec
Post Extension: 72°C for 5 min

Digestion PCR product using restriction enzyme Alul (5’-AG-3’)

RESULTS

The PCR amplification generated a 211 bp segment of growth hormone gene in Madura, Madura crossed and are shown in Figure 1. Two different restriction patterns were obtained corresponding to two different genotypes, LL and LV. Two fragments of 159 bp, 52 bp, and 211 bp were found in individual with LV genotype whereas none restriction fragment with 211 bp were observed in LL genotype. In Madura crossed cattle, the genotype frequency of LL homozygotes was found to be 91.67%, whereas frequency of LV heterozygotes was 8.33%. Allelic frequency for V allele was 0.04, whereas that of L allele was 0.96.

CONCLUSIONS

In conclusion, it may be noted that growth hormone gene is polymorphic in crossbred cattle whereas it monomorphic in local cattle because of growth hormone gene polymorphism was detected only in Madura crossed cattle population and there is no relation between differences of body growth and polymorphism in Madura crossbred cattle. The present study was the first report on growth hormone genotyping in Madura crossbred cattle and has to be considered as a preliminary study. A larger number of observation is needed to establish or deny the existence of an association between growth hormone genotypes and quantitative traits in those cattle also to evaluate crossbreeding program of Madura cattle in Indonesia.

REFERENCES


ACKNOWLEDGEMENTS

This study was supported by Universitas Gadjah Mada-World Class Research University Grant (2010), and thanks to 4th ICGC 2012 teams for this chance to presenting this research.
Certificate of Attendance

This is to certify that

Mifta Pratiwi Rachman

attended

4th International Conference on Quantitative Genetics: Understanding Variation in Complex Traits

17th - 22th June 2012
Edinburgh International Conference Centre, Scotland

Bill Hill (Chair)
The University of Edinburgh
(on behalf of the Local Organising Committee)