Development of high density (600K) chicken genotyping array

Almas Gheyas
Roslin Institute, University of Edinburgh
almas.gheyas@roslin.ed.ac.uk

15th Jan, 2012
Plant and Animal Genomes (PAG) XX Conference
Steps in array development

- **Major steps:**
  - Re-sequencing of many chickens and detection of SNPs
  - Selection of SNPs for validation
  - Validation of selected SNPs
  - Final selection of 600K SNPs for array

Fig: Affymetrix Axiom™ genotyping array
Re-sequencing

243 individuals from 24 lines

- 4 commercial broiler lines (Aviagen) → 40 individuals
- 11 commercial layer lines (Hyline & Synbreed) → 113 individuals
- 9 experimental layer lines (IAH & RI) → 90 individuals

- Illumina high throughput sequencing
- 10-15 samples pooled within line
- Average depth of coverage: 7-17x per line
SNP detection

- Sequence reads were aligned to chicken reference genome (Gallus_gallus-4.0)
- SNP detection performed using Samtools (v.0.1.7a)
- SNP detection done by
  - Within line analyses: 78M SNPs
  - All lines together: 139M SNPs
SNP selection

78M
SNP quality score ≥ 20

24M
SNP quality score ≥ 60

10M
No other SNPs within 10bp on at least one side, uniformly spaced by physical distance

1.8M
Predicted reproducibility of SNPs, uniformly spaced by genetic distance; 50:50 broiler & layer SNPs
1.8M SNPs placed in three Affymetrix arrays for validation

- 288 individuals genotyped:
  - 3 broiler lines
  - 9 layer lines
  - 26 outgroup individuals
SNP validation criteria

- SNP must be polymorphic with $\geq 3$ copies of minor allele
- SNPs must be robust
  - Genotype call rate $\geq 98\%$
  - Good cluster separation
  - Reproducibility
- SNPs have stable Mendelian inheritance
- SNPs must meet performance criteria across all lines

<table>
<thead>
<tr>
<th>Polymorphic and robust SNPs</th>
<th>1,187,482</th>
<th>65%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stable Mendelian inheritance</td>
<td>1,176,808</td>
<td>99%</td>
</tr>
</tbody>
</table>
Selection of final 600k

- Backbone was created by taking all validated SNPs within exonic gene regions (n=21,534)
- SNPs with extreme breach of HWE (P<0.00001) were removed.
- LD between adjacent pairs of SNPs were calculated (Golden Helix SVS) and the average broiler and layer LDs were used to decide the optimum balance of broiler and layer SNPs (3:2).
- Broiler and layer SNPs were selected to get as even spacing as possible in terms of map distance.
Annotation of 600K SNPs

- Intergenic
- Intronic
- Exonic
- 1kb Upstream
- 1kb Downstream
- UTR3 & UTR5
- Splicing & ncRNA

- Synonymous: 56.14%
- Nonsynonymous: 43.31%
- Stopgain/loss: 0.56%
Fig 3: Distribution of SNPs across chromosomes in unit physical (mb) and map distance (cM)
Mean gap size is $1,748 \pm 5,274$ bases.
PCA analysis was performed to see if the markers can group the closely related individuals.
Uses of high density SNP array

- Genomic selection
- Selection signature analysis
- Genome wide association studies
- For fine mapping of QTLs
Acknowledgements

Funding: BBSRC/Defra LINK grant; Aviagen Ltd, Affymetrix Ltd; Hyline International; German Federal Ministry of Education and Research

Roslin Institute: David Burt; John A. Woolliams; Chris Haley; Almas Gheyas; Clarissa Boschiero; Frances Turner; Andy Law; Le Yu; Peter Kaiser; Paul Hocking

Aviagen: Kellie A. Watson; Andreas Kranis

Hyline: Janet E. Fulton

ARK genomics: Richard Talbot; Frances Turner; Sarah Smith; Alison Downing

Affymerix: Fiona Brew; Ali Pirani; Lucy Raynold; J. Montgomery; J. Bart

Synbreed: Henner Simianer; Ruedi Fries; Rudolf Preisinger; T.M. Strom; Steffen Weigend; Klaus Meyer; George Haberer