Prediction of expected variation in progeny groups and its application in mating programs

EAAP, 2013, Nantes, session 10:
How can farmer benefit from genomic information

D. Segelke*, F. Reinhardt* & G. Thaller§
Dierck.Segelke@vit.de
*Vereinigte Informationssysteme Tierhaltung w.V. (vit), Verden/Germany
§Institute of Animal Breeding and Husbandry Christian-Albrechts-University
Olshausenstraße 40 24098 Kiel/Germany
Background

- Different variation of breeding values of sire progeny groups

- AI company: Sires showing high variation
  - Increased probability to find extreme positive candidates

- Production herds: Sires showing less variation
  - Uniform progeny groups are better to manage

- Objective: Predict the variability of future offspring groups using SNP data

<table>
<thead>
<tr>
<th></th>
<th>Cassano</th>
<th>Alexander</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>32 kg</td>
<td>33 kg</td>
</tr>
<tr>
<td>Variation</td>
<td>16.5 kg</td>
<td>19.5 kg</td>
</tr>
<tr>
<td>p(&gt;70kg)</td>
<td>1.1%</td>
<td>2.8%</td>
</tr>
</tbody>
</table>
Background

Haplotype DGV +10

Haplotype DGV = 0

Sire A: DGV +50

Sperm 1       DGV =  0

Sperm 2       DGV = 10

Sperm17    DGV = 20

Sperm278  DGV = 30

Sperm 391  DGV = 40

Sperm 632  DGV = 50

Mean & variation of sperms

Haplotype DGV = 0

Sire B: DGV +70

Sperm 1        DGV =  20

Sperm 2       DGV = 30

Sperm17    DGV = 30

Sperm278  DGV = 40

Sperm 391  DGV = 40

Sperm 632  DGV = 50

Chrom 1                                        Chrom 2                                        Chrom 3                                        Chrom 4                                        Chrom 5

Chrom 1                                        Chrom 2                                        Chrom 3

Sire B: DGV +70
Materials & Methods

Simulated mean and variation (n = 100,000)

Sire strand A:
Start: 50% strand A, 50% strand B

Sire strand B:

→ Calculate sperm mean and sperm variation of 100,000 simulated sperms
Materials & Methods

- 58,035 Holstein animals (50K chip)
- Phasing via Beagle (v 3.3; Browning and Browning, 2010)
- Traits: Protein and fat yield

\[ h_{ij} = \sum_{k=1}^{n} z_{kj} \alpha_k \]

\[ \text{MGBV} = \frac{1}{N} \sum_{j=1}^{N} \sum_{i=1}^{H} h_{ij} \]

\[ \text{VGBV} = \sqrt{\frac{1}{N - 1} \left[ \left( \sum_{j=1}^{N} \left( \sum_{i=1}^{H} h_{ij} \right)^2 \right) - \frac{1}{N} \left( \sum_{j=1}^{N} \sum_{i=1}^{H} h_{ij} \right)^2 \right]} \]

- \( h_{ij} \): i-th parental or maternal haplotype breeding value
- \( z \): indicator of marker k
- \( \alpha_k \): k-th half SNP effect
- \( n \): number of SNP belonging to i-th haplotype
- \( \text{MGBV} \): mean gamete breeding value
- \( N \): number of repetitions of the simulation
- \( H \): number of haplotypes
- \( \text{VGBV} \): Variation of the gamete breeding values
Results

Relation between mean (MGBV) & variation of gamete breeding values (VGBV)
Results

Q-Q Plots for the variation of the gamete breeding values (VGBV)
Results

Distribution of VGBV for fat yield with & without the DGAT1 region

![Graph showing distribution of VGBV for fat yield with and without the DGAT1 region.](image)
## Results

### Examples for specific matings (protein yield)

<table>
<thead>
<tr>
<th>Sperm ($\sigma_a$)</th>
<th>Ovar ($\sigma_a$)</th>
<th>Offspring ($\sigma_a$)</th>
<th>$&gt;1\sigma_a$</th>
<th>$&gt;2\sigma_a$</th>
<th>$&gt;3\sigma_a$</th>
<th>$&gt;4\sigma_a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGBV    VGBV</td>
<td>MGBV    VGBV</td>
<td>E(mBV)</td>
<td>E(vBV)</td>
<td>p(%)</td>
<td>N</td>
<td>p(%)</td>
</tr>
<tr>
<td>1.81 0.29</td>
<td>0.55 0.39</td>
<td>2.36</td>
<td>0.68</td>
<td>98</td>
<td>2</td>
<td>70</td>
</tr>
<tr>
<td>1.68 0.52</td>
<td></td>
<td>2.23</td>
<td>0.91</td>
<td>91</td>
<td>3</td>
<td>60</td>
</tr>
</tbody>
</table>

![Distribution of the offspring breeding values](image)

- Animal 1
- Animal 2

05 October 2013
gBAP: A genomic mating software for breeding associations to produce future candidates

- **Input:**
  - MGBV (DGV/2) and VGBV monthly estimated for all genotyped bulls & cows
  - Imputed genetic defect state for all animal

- **Output:**
  - Recommendation of optimized mating combinations:
    - \(E(GEBV)\), Prob. of extreme offspring
    - Benefit: Low computation time for costumer
Conclusions

- Offspring groups of sires vary and prediction of the distribution is possible
  - Useful information to increase/decrease the probability to get extreme offspring

- Farmers benefit from genomic selection:
  - Imputing LD to 50K: reduce costs and screening of the whole herds is possible
  - Screening the herds for genetic properties (polled, red factor)
  - Screening and managing recessive alleles/haplotypes (HH1, HH2, HH3)

- New tools (i.e. gBAP) to handle all genomic information and to find the most suitable mating partners are needed
  - gEBV
  - Variation
  - Genetic defects (HH1, HH2, HH3, ...)
  - Genetic properties (polled, coat colour..)
  - ....
Acknowledgements

German national organization FBF is thanked for financial support

Thank you for attendance!
## Results

Correlation between variation of gamete breeding values for different traits and inbreeding coefficients

<table>
<thead>
<tr>
<th></th>
<th>$VGBV_{PY}$</th>
<th>$VGBV_{FY}$</th>
<th>$F_G$</th>
<th>$F_P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$VGBV_{PY}$</td>
<td>1</td>
<td>0.41</td>
<td>-0.19</td>
<td>-0.09</td>
</tr>
<tr>
<td>$VGBV_{FY}$</td>
<td></td>
<td>1</td>
<td>-0.10</td>
<td>-0.06</td>
</tr>
<tr>
<td>$F_G$</td>
<td></td>
<td></td>
<td>1</td>
<td>0.52</td>
</tr>
<tr>
<td>$F_P$</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

$VGBV_{PY}$: Variation gamete breeding values protein yield  
$VGBV_{FY}$: Variation gamete breeding values fat yield  
$F_G$: genomic inbreeding coefficient  
$F_P$: pedigree inbreeding coefficient