Admixture and Genomic Predictions in Danish Jersey

J.R. Thomasen\textsuperscript{1,2}, A.C. S\o{}rensen\textsuperscript{1}, M.S. Lund\textsuperscript{1} and B. Guldbbrandtsen

\textsuperscript{1}Aarhus University, Department of Molecular Biology and Genetics, Denmark
\textsuperscript{2}VikingGenetics, Denmark

Internal Seminar
November 8\textsuperscript{th} 2011
Park Inn, Copenhagen
Genomic selection in a Small Breed

- Small breed → Small reference population

- Low reliabilities of genomic prediction
  - 0.12 – 0.47 for the traits included in NTM
  - Limited gain compared to parent index selection

- Need for improvement of genomic reliabilities

- Assume homogeneous breed

- Admixed population → Origin specific LD → Origin specific marker effect?
Objectives

- Study the subpopulation structure in Danish Jersey based on pedigree and markers
- Reconstruct the breed proportions based on the markers
- Evaluate the genomic predictions using the population structure in the model
Admixture exists in Danish Jersey Breed proportion for genotyped bulls
Division into Subpopulations

- ALL - 1730 animals
- DNK - 231 animals >75% DNK pedigree
- US - 171 animals >75% US pedigree
LD only persists at short distances

<table>
<thead>
<tr>
<th></th>
<th>DNK/US</th>
<th>US/ALL</th>
<th>DNK/All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistence of LD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.87</td>
<td>0.95</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>Pairwise Marker Distance (kb)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>0.90</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>0.90</td>
<td>0.80</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>0.80</td>
<td>0.70</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>0.70</td>
<td>0.60</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>0.60</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Markers Reflect Breed Proportions

Proportion Danish Jersey

Markers (Structure)

R² = 0.798
Different Genetic Levels in DNK and US

<table>
<thead>
<tr>
<th></th>
<th>DNK</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk</td>
<td>106</td>
<td>114</td>
</tr>
<tr>
<td>Fat</td>
<td>107</td>
<td>98</td>
</tr>
<tr>
<td>Protein</td>
<td>109</td>
<td>108</td>
</tr>
<tr>
<td>Udder health</td>
<td>103</td>
<td>95</td>
</tr>
<tr>
<td>Fertility</td>
<td>100</td>
<td>96</td>
</tr>
</tbody>
</table>

*Interbull breeding values 2010 (NAV, 2011)*
Genomic reliabilities for Protein

DNK - $R^2 = 0.26$
USJ - $R^2 = 0.18$
MIX - $R^2 = 0.21$

Year adjusted DPP vs DGV for different categories and models.
Population Structure using Random Regression

\[ y_i = \mu + \beta w_i + a_{0i} + w_i a_{1i} + e_i \]

- DRP
- DGV for animal \( i \)
- Udder health

- \( w_i \): Breed proportion from *Structure* or pedigree

Covariance structure: Genomic or pedigree relationship
Dataset
- 1079 genotyped bulls
- DRP for udder health

Predictions
- 879 reference bulls with DRP
- Variance components
- 200 test bulls (DGV)

Reliability = $r^2_{DRP,DGV} / \text{Rel}_{DRP}$
#### Gains due to G-BLUP and structure

<table>
<thead>
<tr>
<th>Model</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic animal model</td>
<td>0.24</td>
</tr>
<tr>
<td>Basic genomic model</td>
<td>0.34</td>
</tr>
<tr>
<td>RR-genomic model – structure</td>
<td>0.33</td>
</tr>
<tr>
<td>RR-genomic model - pedigree</td>
<td>0.33</td>
</tr>
</tbody>
</table>
• Conclusions

• Population structure in Danish Jersey exist

• Population structure modeled in the random regression model did not improve predictions compared to a basic genomic model

• Why?

• The population structure modeled well by genomic relationship matrix

• Structure/pedigree information based on average marker information across the entire genome

• Current improvement

• Cluster markers based on effect size and origin