Across-breed genomic evaluation based on BovineHD genotypes, and phenotypes of bulls and cows

Chris Schrooten¹), Ghyslaine Schopen¹), Phil Beatson²)

1) CRV BV, The Netherlands
2) CRV Ambreed, New Zealand

EAAP 2013, Nantes | Schrooten et al | 27 August 2013 |
Genomic evaluation New Zealand

- Currently, no national genomic evaluation
  - Performed in-house, by breeding organisations
  - CRV Ambreed, LIC
- Genomic evaluation CRV Ambreed
  - Started in 2007/2008
  - Using 50k SNP data
    - customCRV v1 and v2, BovineSNP50 v1 and v2
  - Single breed evaluations
    - Friesians
    - Jerseys
Reference populations CRV Ambreed

- Size by Feb. 2008 (protein):
  - Friesians: 1,050
  - Jerseys: 420

- Size by Dec. 2012 (protein):
  - Friesians: 2,350
  - Jerseys: 1,200

- Growth through
  - Yearly expansion: waiting bulls becoming daughter proven bulls
  - Exchange between CRV Ambreed and LIC

- Note: Eurogenomics reference population: > 25,000
How to increase size of reference?

- Multi-breed evaluation
  - Combine reference populations for Friesians, Jerseys and Crossbreds into one genomic evaluation
    - LD between SNP and mutations affecting the trait
    - Linkage phase the same in each breed
  - At least 300k SNP needed (de Roos et al, 2008)
  - Use BovineHD

- Expand reference population with cow phenotypes
  - Phenotypic info of cows is less reliable than phenotypic info of bulls
  - Therefore, relatively more cows are needed
Objective

to study the effect on the reliability of genomic EBV when combining single-breed reference populations into one multi-breed reference population, and using HD genotypes and cow genotypes and phenotypes
Material & Methods

- BovineHD (777k) genotypes
  - 463 Friesians
  - 229 Jerseys
  - 57 crossbreds
- Custom CRV 50K chip / BovineSNP 50
  - 3,550 reference bulls
  - 4,500 cows, 10 herds
  - removed cows with pedigree inconsistencies
    - approx. 3,600 cows left
Processing genotypes

- Imputation of 50k genotypes to BovineHD
  - approx. 623k SNP
  - Beagle 3.3
  - allelic imputation error rate dependent on chip and breed
    - 0.44% for Jersey, BovineSNP50 v2
    - 1.13% for Friesian, custom CRV chip v1
- Determine haplotype id based on Beagle output
  - at each BovineHD locus
- Select 1 out of 10 loci for further analyses
  - reduce computer requirements
  - omit redundant information
Validation study

- Subset of 7 traits in NZMI with moderate to high reliability
  - Prot., milk, livew., somatic cells, capacity, rump angle, udder
- Phenotype: deregressed proofs
  - of bulls
  - of cows, reliability of proof > threshold
    - #cows: 2,000 – 2,700
- weights: EDC
- youngest bull cohort considered as validation bulls
  - Friesians: 350
  - Jerseys: 160
  - Crossbred: 60
  - phenotype omitted from analyses
  - phenotype of their daughters omitted from analyses
Genomic evaluation model

- Estimation of GBV with Bayes SSVS (Calus et al, 2008)

\[ y_i = \mu + u_i + \sum_{j}^{n} (q_{ij1} + q_{ij2})v_j + e_i \]

- where \( y_i \): deregressed proof
- \( \mu \): overall mean
- \( u_i \): random polygenic effect of animal \( i \)
- \( q_{ij1(2)} \): size of effect for haplotype 1 (2) of animal \( i \) at locus \( j \)
- \( v_j \): direction vector of effects at locus \( j \)
- \( e_i \): residual

- Separate runs on same data without genomic component: PBV

- 4 replicates per trait
Validation

- Compare genomic (GBV) and polygenic breeding values (PBV) with daughter based phenotype (DRP)

\[
\Delta R^2 = \frac{R^2_{GBV, DRP} - R^2_{PBV, DRP}}{REL_{DRP}}
\]
Results – average $\Delta R^2$ across 7 traits

EAAP 2013, Nantes | Schrooten et al | 27 August 2013 |
# Results: $\Delta R^2$ per trait per breed per ref. set

<table>
<thead>
<tr>
<th>Trait</th>
<th>Friesian</th>
<th></th>
<th></th>
<th>Jersey</th>
<th></th>
<th></th>
<th>Crossbred</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>sb50k</td>
<td>mbb</td>
<td>mbbc</td>
<td>sb50k</td>
<td>mbb</td>
<td>mbbc</td>
<td>mbb</td>
<td>mbbc</td>
</tr>
<tr>
<td>Protein</td>
<td>.117</td>
<td>.104</td>
<td>.054</td>
<td>.047</td>
<td>.092</td>
<td>.070</td>
<td>.047</td>
<td>.122</td>
</tr>
<tr>
<td>Milk</td>
<td>.154</td>
<td>.158</td>
<td>.116</td>
<td>.204</td>
<td>.135</td>
<td>.070</td>
<td>.023</td>
<td>.098</td>
</tr>
<tr>
<td>Livew.</td>
<td>.035</td>
<td>.068</td>
<td>.078</td>
<td>.049</td>
<td>.116</td>
<td>.164</td>
<td>.075</td>
<td>.077</td>
</tr>
<tr>
<td>Som. Cells</td>
<td>.069</td>
<td>.024</td>
<td>.088</td>
<td>.051</td>
<td>.106</td>
<td>.185</td>
<td>.224</td>
<td>.125</td>
</tr>
<tr>
<td>Capac.</td>
<td>.095</td>
<td>.130</td>
<td>.103</td>
<td>.040</td>
<td>.036</td>
<td>.087</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rump angle</td>
<td>.119</td>
<td>.108</td>
<td>.129</td>
<td>.013</td>
<td>.061</td>
<td>.112</td>
<td>.121</td>
<td>.111</td>
</tr>
<tr>
<td>Udder</td>
<td>.021</td>
<td>.062</td>
<td>.093</td>
<td>.084</td>
<td>.051</td>
<td>.072</td>
<td>.055</td>
<td>.085</td>
</tr>
<tr>
<td>Avg.</td>
<td>.087</td>
<td>.093</td>
<td>.094</td>
<td>.070</td>
<td>.085</td>
<td>.109</td>
<td>.091</td>
<td>.103</td>
</tr>
</tbody>
</table>
Results: $\Delta R^2$ per trait per breed per ref. set

<table>
<thead>
<tr>
<th>Trait</th>
<th>Friesian</th>
<th>Jersey</th>
<th>Crossbred</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>sb50k</td>
<td>mbb</td>
<td>mbbc</td>
</tr>
<tr>
<td>Protein</td>
<td>.117</td>
<td>.104</td>
<td>.054</td>
</tr>
<tr>
<td>Milk</td>
<td>.154</td>
<td>.158</td>
<td>.116</td>
</tr>
<tr>
<td>Livew.</td>
<td>.035</td>
<td>.068</td>
<td>.078</td>
</tr>
<tr>
<td>Som. Cells</td>
<td>.069</td>
<td>.024</td>
<td>.088</td>
</tr>
<tr>
<td>Capac.</td>
<td>.095</td>
<td>.130</td>
<td>.103</td>
</tr>
<tr>
<td>Rump angle</td>
<td>.119</td>
<td>.108</td>
<td>.129</td>
</tr>
<tr>
<td>Udder</td>
<td>.021</td>
<td>.062</td>
<td>.093</td>
</tr>
<tr>
<td>Avg.</td>
<td>.087</td>
<td>.093</td>
<td>.094</td>
</tr>
</tbody>
</table>
Summary / implications

- Squared correlations between genomic prediction and daughter based breeding value increased by using HD-genotypes, cow phenotypes in addition to bull phenotypes, and multi-breed evaluation.
- Increase in $R^2$ varied from 0.007 (Friesian) to 0.039 (Jersey).
  - $R^2$ did not increase for all trait-breed combinations.
- No need for further HD-genotyping.

Recommendation:
- Make use of HD- and cow-info for traits where it is beneficial.
- Genotype more females if genotyping cost are sufficiently low.
Thank you for your attention