Cow genotyping strategies for genomic selection in small dairy cattle population

Janez Jenko


EAAP 2016
67th Annual Meeting of the European Federation of Animal Science
Belfast UK, 29 Aug – 2 Sept 2016
Overview of the talk

• Guernsey cattle breed
• Genotypes
• Quality check (QC) and imputation
• Phenotypes
• Prediction of genomic breeding values (GEBV)
• Scenarios of cow genotyping
• Conclusions
Guernsey cattle breed

- Small dairy cattle breed population
- Approximately 14,000 recorded individuals worldwide
- Milk from Guernsey cow has unique qualities
Genomic evaluations

- Small number of bulls with conventional proofs
  - 197 bulls genotyped (born from 1957 to 2013)
    - 168 on 75K
    - 29 on 777K

- Genotyping cows to improve accuracies
  - 1,440 cows (present in herds at the end of 2014)
    - All genotyped on 25K
QC & Imputation

- QC for genotypes and SNPs
  - Call rate, minor allele frequency, heterozygosity rate, Hardy-Weinberg equilibrium, Mendelian inconsistencies

- Imputation

Legend

- Bulls existing SNPs
- Cows existing SNPs
- Imputed SNPs

Initial genotypes

Imputation step 1

Imputation step 2
Imputation accuracy

Correlation between true and imputed SNP genotypes using 10-fold cross-validation on:

1,333 cow genotypes with 11,983 SNP existing on both: 75K and 25K chip

<table>
<thead>
<tr>
<th>Between individuals</th>
<th></th>
<th>Between SNP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.952</td>
<td>Mean</td>
</tr>
<tr>
<td>SD</td>
<td>0.033</td>
<td>SD</td>
</tr>
</tbody>
</table>
### Phenotypes

<table>
<thead>
<tr>
<th></th>
<th>Bulls</th>
<th>Cows</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Phenotypes”</td>
<td>PTA from MACE</td>
<td>Milk records</td>
</tr>
<tr>
<td>Adjustment of</td>
<td>Deregression</td>
<td>Adjusted lactation records</td>
</tr>
<tr>
<td>“phenotypes”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of milk, fat,</td>
<td>185</td>
<td>1,307</td>
</tr>
<tr>
<td>and protein yield</td>
<td></td>
<td></td>
</tr>
<tr>
<td>records</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Model to predict GEBV

**Model**

\[ y = 1\mu + Zu + e \]

- **Phenotypes**
- **Mean**
- **Inference matrix linking records from vector** \( y \) **to vector** \( u \)
- **Vector of random genetic effects**
- **Vector of errors**

- \( \text{Var}(u) = G\sigma_g^2 \)
- \( \text{Var}(e) = \sigma_e^2 W^{-1} \)
- \( G = \frac{MM'}{2 \sum_{j}^{N_{snp}} p_j(1-p_j)} \)

\( W^{-1} \) - weights the contribution from each individual
Cow genotyping strategies

- Bull genotypes
- Bull and cow genotypes
- Bull and 50% of cow genotypes selected at random
- Bull and 50% of cow genotypes with top performance
- Bull and 50% of cow genotypes with extreme phenotypes
Metrics of comparison

10-fold cross-validation

Accuracy
Correlation between phenotypes and GEBV

Bias
Slope of the regression of phenotypes on GEBV
• > 1 underestimation
• < 1 overestimation
Accuracy

Cow genotyping strategies

Accuracy

Milk yield

Fat yield

Protein yield
Cow genotyping strategies
Conclusions

• Genotyping cows improves the accuracy of GEBV in small dairy cattle population

• Genotyping half of cows (those with extreme adjusted phenotypes) yields only slightly lower accuracies as compared to the scenario when genotyping all cows

• Selective genotyping introduces bias - GEBV is overestimated
Further information

Manuscript in revision

• Results for Calving interval
• Genotyping only 40% or 30% of cows from both tails
• Selection of cow genotypes based on indexes (Profit Lifetime Index, Guernsey Merit Index)
• Predicted model
Acknowledgement

Financial support

The Royal Guernsey Agricultural & Horticultural Society

Thank you for your attention