Multi-breed multi-trait single-step genomic predictions for Holstein and Jersey including crossbred animals

Renzo Bonifazi, Stella Aivazidou, Jan ten Napel, Matias Schrauf, Gerben de Jong, Jeremie Vandenplas
Crossbreeding and genomic predictions

Advantages

- Heterosis
- Breed Complementarity
- ↓ Inbreeding
- ↑ Fertility
- ↑ Longevity

- Reference population difficult to establish

- SNP and QTL effects may differ across breeds due to differences in:
  - genetic backgrounds
  - LD between QTL-SNPs
  - environment
Aim of the study

- Identify the best scenario for multi-breed single-step evaluations
  - Two purebred populations (PBD)
  - Derived crossbred population (XBD)
- Impact of modelling missing pedigree information using genetic groups or meta-founders
Data available

- **Pedigree (~1,1 M)**
  - Purebred HOL & JER, and XBD

- **Genotypes ~127K, 80K SNPs**

- **Phenotypes**
  - Milk yield (~538K) $h^2 = 0.31$
  - CSD (~370K) $h^2 = 0.05$
  - $r_g_{\text{milk, CSD}} = 0.22$

---

**Actual calving date**

- $-N$ days
- $+N$ days

**Planned calving date for given HY**
Scenarios

- **Single-breed reference population (SINGLE_POP)**
  - 3 separate evaluations (with 255 GG)

- **Multi-breed reference population using:**
  - 255 genetic groups : ALL
  - 4 genetic groups : ALL_4GG
  - 255 metafounders : ALL_255MF
  - 4 metafounders : ALL_4MF

- **Multi-breed multi-trait approach (MBMT): breeds as traits**
Validation

- LR method (Legarra & Reverter, 2018)
- **Forward-in-time**: whole vs partial dataset
- **Cut-off**: last 6 years phenotypes removed

- **Validation groups, by breed:**
  - **Cows**: with genotyped & phenotype for milk and/or CSD, born after cut-off
  - **Bulls**: genotyped & ≥20 daughters with phenotypes for milk and/or CSD born after cut-off and no daughters with records born before cut-off

<table>
<thead>
<tr>
<th>Breed</th>
<th>Cows</th>
<th>Bulls</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOL</td>
<td>~ 9K</td>
<td>~ 350</td>
</tr>
<tr>
<td>JER</td>
<td>~ 4K</td>
<td>~ 170</td>
</tr>
<tr>
<td>XBD</td>
<td>~ 13K</td>
<td>~ 130</td>
</tr>
<tr>
<td>Total</td>
<td>~ 27K</td>
<td>~ 650</td>
</tr>
</tbody>
</table>
Software

- GEBVs estimation using ssSNPBLUP
- (pedigree-based) VCE for MTMB scenario: BLUPF90 (gibbs & postgibbs)

Validation: R functions
<table>
<thead>
<tr>
<th></th>
<th>CSD</th>
<th></th>
<th></th>
<th>Milk</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HOL</td>
<td>JER</td>
<td>XBD</td>
<td>HOL</td>
<td>JER</td>
<td>XBD</td>
</tr>
<tr>
<td><strong>CSD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOL</td>
<td>0.03</td>
<td>0.59</td>
<td>0.93</td>
<td></td>
<td>0.41</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.00)</td>
<td>(0.09)</td>
<td></td>
<td>(0.05)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>JER</td>
<td></td>
<td>0.04</td>
<td>0.66</td>
<td>0.24</td>
<td>-0.02</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td>(0.06)</td>
<td></td>
<td>(0.10)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>XBD</td>
<td></td>
<td>0.03</td>
<td>0.03</td>
<td>0.70</td>
<td>0.47</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td></td>
<td>(0.05)</td>
<td>(0.01)</td>
</tr>
<tr>
<td><strong>Milk</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOL</td>
<td></td>
<td>0.34</td>
<td>0.34</td>
<td>0.87</td>
<td>0.24</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.04)</td>
<td>(0.04)</td>
<td></td>
<td>(0.07)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>JER</td>
<td></td>
<td>-0.03</td>
<td>-0.03</td>
<td>0.70</td>
<td>0.82</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.10)</td>
<td>(0.10)</td>
<td></td>
<td>(0.06)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>XBD</td>
<td></td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.04)</td>
<td>(0.04)</td>
<td></td>
<td>(0.05)</td>
<td>(0.03)</td>
</tr>
</tbody>
</table>

**Diagonal:** heritability, **Off-diagonal:** genetic correlations, **Brackets:** Standard Errors
**Metafounders**

- **ALL_4MF**

  **Matrix**

<table>
<thead>
<tr>
<th></th>
<th>HOL</th>
<th>JER</th>
<th>XBD</th>
<th>OTHER</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOL</td>
<td>0.93</td>
<td>0.78</td>
<td><strong>0.83</strong></td>
<td>0.57</td>
</tr>
<tr>
<td>JER</td>
<td>0.72</td>
<td>0.75</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>XBD</td>
<td></td>
<td></td>
<td>0.78</td>
<td>0.56</td>
</tr>
<tr>
<td>OTHER</td>
<td></td>
<td></td>
<td></td>
<td><strong>0.77</strong></td>
</tr>
</tbody>
</table>
Accuracy partial EBV

- **MTMB**: highest accuracy
- **ALL, ALL_4MF**: increased accuracy compared to the **SINGLE_POP** (2nd lowest)
- **SINGLE_POP**: lowest accuracy
- **MTMB**: lowest accuracy

- Similar results for cows (smaller SE)
- Similar results between: ALL, ALL_255MF, ALL_4GG, ALL_4MF
Overall, **dispersion within ±15%** (except milk-XBD bulls, CSD-HOL bulls).

- Cows have less dispersion and smaller SE
- Milk: both **over/under-dispersion**. CSD: mainly **over-dispersion** (except JER bulls).
- Similar results for **level bias**: overall smaller for Milk (within ±0.05 \( \sigma_A \)) than CSD (within ±0.1 \( \sigma_A \))
Conclusions

- Benefit of combining data in multi-breed evaluations for PBD and XBD
- Reduce n. GG (similar results with 255 and 4)
- Limited added benefit of using metafounders instead of GG
- MTMB accuracy highest for milk - lowest for CSD
- Further investigations using conformation traits
Acknowledgments
Conclusions

- Benefit of combining data in multi-breed evaluations for PBD and XBD
- Reduce n. GG (similar results with 255 and 4)
- Limited added benefit of using metafounders instead of GG
- MTMB accuracy highest for milk - lowest for CSD

Thanks for your attention

renzo.bonifazi@wur.nl  @RenzoBonifazi