Multiple-trait single-step genomic prediction of udder health traits in Nordic Red and Jersey cattle applying alternative SNP weighting

A. Chegini[1], I. Strandén[1], E. Karaman[2], T. Iso-Tuoru[1], M. Lidauer[1]

[1]Natural Resources Institute Finland (Luke)
[2]Aarhus University

BovReg Final Conference - Brussels
(14-15 February 2024)

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 815668

Introduction

- Clinical Mastitis ---- > high incidence.
- Improving resistance to CM -> higher profit and sustainability.
- Mastitis has relatively low heritability.
- Strategies to accelerate the rate of improvement: recording of clinical mastitis, multiple-trait and genomic selection.

Genomic selection

1. Single nucleotide polymorphism best linear unbiased prediction (**SNPBLUP**)
2. Genomic best linear unbiased prediction (**GBLUP)**
3. Single-step
   - **a) GBLUP**
   - **b) SNPBLUP**

- All above-mentioned methods assign equal weights to all SNPs, assuming that each marker contributes equally to the genetic variation.
- Some markers are in the range of influential genes or in proximity with them.

- Hypothesis: If it is not true, specific weight allocation = improvement.

Data

- Clinical mastitis, test-day somatic cell score, fore udder attachment and udder depth for Nordic Red cattle (RDC) and Jersey (JER) dairy cows.
- Calved from 1980s in Denmark, Finland and Sweden.
- 17 million records for 0.9 million JER cows and
- 74.5 million records for to 5.6 million RDC cows.

Data preparation and trait definition

Clinical mastitis (CM) traits:

CM11: -15 to 50 days in milk, first lactation.

CM12: 51 to 305 days in milk, first lactation.

CM2: -15 to 150 days in milk, second lactation.

CM3: -15 to 150 days in milk, third lactation.

Correlated traits:

- 3 somatic cell score (SCS) traits: first, second, third lactation test-day observations (were transformed to logarithmic scale)
- fore udder attachment (UA): first lactation.
- udder depth (UD): first lactation.

Table 2. Summary statistics of studied traits.

Trait	JER			RDC		
	n	Mean	SD	n	Mean	SD
CM11	601988	0.141	0.349	4791842	0.065	0.246
CM12	590198	0.106	0.308	4641064	0.062	0.242
CM2	427327	0.138	0.346	3452089	0.111	0.316
CM3	287409	0.160	0.368	2246733	0.142	0.351
TDSCS1	7317581	1.481	0.369	29944467	1.349	0.402
TDSCS2	5100038	1.558	0.405	20997798	1.469	0.433
TDSCS3	3301183	1.616	0.417	12978293	1.543	0.442
UA	307635	1.367	0.305	1161944	1.396	0.345
UD	307634	1.349	0.297	1160324	1.404	0.441

Genomic data

- Genotyped JER and RDC 136,562 and 249,223, respectively.
- Only genotyped individual born after 2008 were used for the analyses, leaving 64,777 and 125,789 genotyped JER and RDC, respectively.
- 41,897 and 46,914 SNPs for JER and RDC, respectively.

Multi-trait random regression test-day model

$$\boldsymbol{y} = \mathbf{Xb} + \mathbf{Qa} + \mathbf{Wpe} + \mathbf{e}$$

Fixed effects: herd × 5-yr production period (herd × yr of production for SCS traits); year × month of production × parity class; fixed linear regression on total heterosis of the cow; linear and quadratic regression coefficients on calving age × breed interaction.

Random effects: herd × year (herd × test-date for SCS traits); additive genetic effect of animal; permanent environmental effect of animal; residual effect.

- Fixed regression function on DIM nested within country × parity × year-season of calving × calving age.
- All breeding values of an animal were modelled by covariance functions

SNP weighting scenarios for ssSNPBLUP:

1. **Nonlinear**: VanRaden (2008) and Cole et al. (2009)

   in $\frac{ZDZ'}{\sum_{j=1}^{m} 2p_j q_j}$ D was calculated as $1.25^{\frac{|\hat{u}_i|}{s\,d(\hat{u})}-2}$

2. **2pqa^2**: (Falconer and Mackay, 1996; Zhang et al., 2010)

3. **20SNP_window**: The average weights (calculated by $2pqa^2$) of every 20 adjacent SNPs

4. SNP variances → (using annotations + NextGP)

Preparation of Input files

Allele frequencies → personal shell script.

Inbreeding coeff. → **RelaX2** (Strandén and Vuori, 2006).

T48eig_make_v 0.710 program → to calculate C^{-1} and Z_c (Strandén and Mäntysaari, 2023).

RPG = 0.1.

For scaling, we multiplied GRM by **trace (A_{22}) / trace (G)**.

Candidate selection and validation method

- Last four years of data were removed to create a reduced data.
- Genomic breeding values:
    - CM combined (0.15, 0.15, 0.25, 0.45)
    - SCS combined (0.50, 0.30, 0.20)
- Reliabilities for combined traits were calculated using an appropriate method for multiple-trait (Tier and Meyer, 2004).
- ERCs by reverse reliability approximation (Mäntysaari, 2013).
- **Males**: $ERC_full >= 2.0 && $ERC_reduced =< 0.001
- **Females**: $ERC_full >= 0.9 && $ERC_reduced =< 0.001
- **Validation method**: Legarra and Reverter, 2018)

Results of forward validation for Clinical Mastitis

Breed	Gender	n	Weight	Reg.full/red.	b_1	R^2	%gain
JER	Male	115	regular	G	0.78	0.65	
			Nonlinear	G	0.80	0.67	3.2
			$2pqa^2$	G	0.83	**0.76**	**17.0**
			20SNP_window	G	0.80	0.67	3.2
			annotation		**0.85**	0.73	12.4
RDC	Male	86	regular	G	0.75	0.50	
			Nonlinear	G	0.77	0.52	5.4
			$2pqa^2$	G	**0.87**	**0.67**	**35.9**
			20SNP_window	G	0.75	0.52	5.4

Results of forward validation for SCS

Breed	Gender	n	Weight	Reg.full/red.	b_1	R^2	%gain
JER	Male	119	regular	G	0.82	0.61	
			Nonlinear	G	0.83	0.65	6.0
			2pqa^2	G	0.82	**0.77**	**26.1**
			20SNP_window	G	0.83	0.68	10.3
			annotation		**0.90**	0.71	15.5
RDC	Male	125	regular	G	0.87	0.59	
			Nonlinear	G	0.89	0.63	6.2
			2pqa^2	G	**0.92**	**0.79**	**32.4**
			20SNP_window	G	0.89	0.65	9.3

BovReg *PARTNERS*

Thank you for your attention

www.bovregproject.eu

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 815668