Genomic evaluation combining different French dairy cattle breeds

S. Karoui, MJ. Carabaño, C. Díaz, A. Legarra

INRA

INIA
Instituto Nacional de Investigación
y Tecnología Agraria y Alimentaria

Introduction

- Accuracy of genomic breeding values (GEBVs) depends on the size of the reference population (e.g. VanRaden et al., 2008; Hayes et al., 2009).

- The use of a multiple breed reference population might be a way to increase accuracy of genomic breeding values (GEBV's) in small breeds.

- Interesting for traits with low h^2 (Brondum et al., 2011).

Methods

- SNP effects might potentially differ across breeds (different LD, different QTLs, G x E) (Varona et al.,2010).

- In an equivalent model, breeding values of each bull in different breeds are different (but correlated) and are assumed to follow a multivariate distribution.

$$\mathbf{u} = \begin{cases} \mathbf{u}_{breed\,1} \\ \mathbf{u}_{breed\,2} \end{cases} \sim N\begin{pmatrix} \mathbf{0} & \mathbf{G}\sigma_{u1}^2 & \mathbf{G}\sigma_{u1,2} \\ \mathbf{0}' & \mathbf{G}\sigma_{u2,1} & \mathbf{G}\sigma_{u2}^2 \end{pmatrix} = N\left(\mathbf{0}, \mathbf{G}_0 \otimes \mathbf{G}\right)$$

Objective

To evaluate the impact of

1. using a large multi-breed dairy cattle reference population
2. and an across-breed genetic correlation model

on the accuracies of GEBVs

Genotypes

- All bulls were genotyped for 50k SNPs using the Illumina Bovine array.

- Genotypes of all breeds were merged including only SNPs which were polymorphic in all three breeds.

- SNPs were filtered by extreme HWD, Mendelian inconsistencies.

- Finally, 43852 SNPs were used.

Phenotypes

- Daughter yield deviations (2DYD 's) for :

  - ✓ Milk yield.
  - ✓ Fat content.
  - ✓ Fertility (Non return rate at 56 days).

- Weighted by "equivalent daughter contributions"(EDC's).

- Training set included bulls from Holstein (H), Montbéliarde (M) and Normande (N) breeds.

Populations

French Holstein bulls
Reference : 2976
Validation :   964

French Montbéliarde bulls
Reference : 950
Validation : 222

French Normande bulls
Reference : 970
Validation : 248

Computations

- Multiple trait version of **G**BLUP (VanRaden, 2008; Hayes and Goddard,2008)

➤ A single **G** includes all 3 breeds

➤ In all analyses, one trait (i.e., milk) is considered in $\mathbf{G}_0$ as a different trait for each breed (e.g., as in across-country MACE).

$$\begin{pmatrix} \mathbf{y}_H \\ \mathbf{y}_M \\ \mathbf{y}_N \end{pmatrix} = ... + \begin{bmatrix} \mathbf{Z}_H & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_M & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Z}_N \end{bmatrix} \begin{pmatrix} \mathbf{u}_H \\ \mathbf{u}_M \\ \mathbf{u}_N \end{pmatrix} + \mathbf{e}$$

$$Var \begin{pmatrix} \mathbf{u}_H \\ \mathbf{u}_M \\ \mathbf{u}_N \end{pmatrix} = \mathbf{G}_0 \otimes \mathbf{G}$$

➤ Bulls:
  ➤ are pure bred but connected through genomic relationships in **G**.
  ➤ Have EBVs in the three scales (H, M, N)

Computations

➤ Genetic parameters (genetic variances, <u>genetic correlations between different breeds</u>) were estimated using Bayesian procedures in `gibbs2f90 (Misztal et al., 2002)`.

➤ Accuracy of GEBVs by cross-validation = R^2 (2DYD , GEBV) weighted by EDC.

Scenarios

Reference population	Size	Correlation between breeds
Multibreed		
H+M+N	4896	Correlation = estimated
H+M+N	4896	Correlation = 0.95
Uni-breed		
H	2976	
M	970	Correlation = 0
N	950	

breeds assumed ~identical

Three separate single breed analysis

Validation : 964 H + 222 M + 248 N

Results

Table 1. Genetic correlation (r_g) between breeds

Trait	Breeds Montbéliarde-Normande	Montbéliarde-Holstein	Normande-Holstein
Milk	0.46 [0.26 ; 0.65]	0.79 [0.63 ; 0.93]	0.38 [0.19 ; 0.55]
%Fat	0.35 [0.07 ; 0.64]	0.66 [0.50 ; 0.84]	0.56 [0.34 ; 0.76]
Fertility	-0.01 [-0.50, 0.54]	0.39 [-0.05; 0.73]	0.22 [-0.15 ,0.54]

Results

Table 2. Reliabilities (R²) by cross-validation

Trait		M	N	H
Milk	r_g=estimated	0.21	0.13	0.31
	r_g= 0	0.19	0.12	0.30
	r_g=0.95	0.21	0.14	0.31
%Fat	r_g =estimated	0.33	0.39	0.52
	r_g=0	0.27	0.39	0.51
	r_g=0.95	0.33	0.39	0.52
Fertility	r_g=estimated	0.19	0.07	0.11
	r_g=0	0.19	0.07	0.11
	r_g=0.95	0.19	0.07	0.11

Conclusions

➢ Multi-breed evaluation is equal or more accurate than single breed evaluation.

➢ The use of the multi-breed reference population only helped to increase accuracy of GEBVs for traits and populations that showed large correlations and in the breed with the smallest data set.

➢ The genetic correlation between breeds could be a good indicator of the interest of a multi-breed reference population.

Acknowledgements

- Authors thank financing by ANR project AMASGEN and by APISGENE.

- Project partly supported by Toulouse Midi-Pyrénées bioinformatic platform.

- A. Legarra acknowledges financing from POCTEFA (www.poctefa.eu) project GENOMIA.

- Comments from V. Ducrocq, S. Fritz, D. Boichard, C. Robert-Granié, P. Croiseau and F. Guillaume are fully acknowledged.

- This work has been done during a stay of S. Karoui in INRA Toulouse, France, who acknowledges his INIA fellowship (project RTA2007-0071, co-funded by Ministerio de Economía y Competitividad and FEDER).

Results

Table 3. coefficient of regression of 2DYD on GEBV

Trait	Models	RRM		
		M	N	H
Milk	r_g=estimated	0.822	0.713	0.740
	r_g= 0	0.812	0.683	0.740
	r_g=0.95	0.811	0.698	0.739
%Fat	r_g =estimated	1.118	1.082	0.935
	r_g=0	1.023	1.039	0.930
	r_g=0.95	1.086	0.988	0.930
Fertility	r_g=estimated	1.517	0.996	0.741
	r_g=0	1.513	1.013	0.744
	r_g=0.95	1.484	0.913	0.731

Relationship Montbéliarde

Relationship between Montbéliarde and Holstein

Relationship Normande

Relationship between Montbéliarde and Normande

Relationship Holstein

Relationship between Normande and Holstein