

Accuracy of prediction for a genomic evaluation in rotational crossbreeding scheme (Montbéliarde x Holstein x Red Danish)

R. Saintilan, P. Croiseau, A. Baur, I. Croue, V. Ducrocq, J.R. Thomasen, E. Karaman, D. Boichard, H. Leclerc and B.C.D. Cuyabano

Plan of the presentation

- **Context**
- **Methods**
- **Dataset**
- **Results**
- **Discussions and perspectives**

Context

- **Crossbreeding strategy increase in dairy cattle farms**
- **Opportunity to combine strengths of many pure breeds, compensate weakness and benefit from Heterosis effect**
- **In France, few evaluations exist for crossbred animals and only for terminal purpose**
- **International Gentore project and French Evagenoc project to implement genomic evaluation for different crosses**
- **This study present the results of a genomic evaluation conducted on a Procross population on a two steps procedures: polygenic evaluation to get YD and a genomic evaluation for production traits**

Methods

- Estimation of Breed of Origins (BOA) (Eirikson et al., 2021):
 - Purpose is the recodification of allele of crossbred in function to which purebred give the allele
 - Comparison of the frequency of haplotypes of 16 consecutive SNP with the ones in the purebreds to identify the breed of origin (65,356 haplotypic combinations)
 - Sliding windows moving on SNP at time
 - If indetermination, divide the size of haplotype by 2
 - BreedOrigin fortran in house software
 - $BOA_{(i)}$ if $\frac{f_i}{\sum_{j=1}^{nbreed} f_j}$ higher than 0.90
- Compare BOA with breed proportions based on pedigree

Breed of Origin	Original Allele	Recodified Allele
A	1	1
A	2	2
B	1	3
B	2	4
C	1	5
C	2	6

Methods

- **Phenotypes:** YD of genotyped females were obtained from a polygenic model to estimate all non-genetic effects, heterosis and recombination losses and heterogenous variances
- **Crossbreeding parameters:** Computed for each crossbred animal from proportions of breed origins based on pedigree. Values computed as described in Dechow et al. (2007)

$$H = 1 - \sum_{i=1}^{nb_{breed}} s_i d_i \qquad R = 1 - \sum_{i=1}^{nb_{breed}} \frac{(s_i^2 + d_i^2)}{2}$$

with s_i and d_i the proportions of sire and dam genes from breed i , respectively

Methods

- **Genomic Evaluation:**

- **extension of the SNP BLUP model where a SNP effect is estimated for each breed**

$$y_i = \sum_{b=1}^{nb_breed} (p_{i,b} \mu_b) + \sum_{b=1}^{nb_breed} \left(\sum_{j=1}^{nb_SNP} (\beta_{i,j,b} X_{i,j,b}) \right) + e_i$$

With y_i is the YD of the animal i , μ is a vector of means defined within each breed, $p_{i,b}$ is the proportion of breed b in the genome of individual i , estimated with the BOA approach.

$X_{i,j,b}$ is the allele content of SNP j that originates from the breed b for animal i , centered for the allelic frequency of the SNP in breed b :

$$X_{i,j,b} = (k_{i,j} - n_{i,j,b} f_{j,b})$$

where $k_{i,j,b}$ and $n_{i,j,b}$ are the number of “2” alleles and the total number of alleles of the SNP j that originates from breed b for the animal i , respectively; f_b is the frequency of allele “2” of the SNP j in breed b .

Data

- Population of 5,238 genotypes of crossbred animals genotyped or imputed on a MD chips using FImpute (Sargolzaei et al., 2014). Two batches of imputed, purebred one and inclusion of crossbred animals in the second one
- Females with YD as a result from the polygenic evaluation
- Validation study: 2,000 youngest crossbred cows without progeny in the validation population

Results: Breed of Origins

Breed of Origin	Breed composition with BOA		Breed composition based on Pedigree		
	<i>Mean</i>	<i>Sd</i>	<i>Mean</i>	<i>Sd</i>	<i>Correlation</i>
Montbéliarde (Mo)	13.60%	20.00%	11.80%	19.20%	0.99
Holstein (Ho)	48.00%	19.40%	56.30%	20.20%	0.97
Red Danish (RD)	34.00%	17.80%	29.80%	17.80%	0.95

- For genotyped crossbred cows, 94% of the alleles were assigned to a purebred
- 48% of markers were from a Holstein origin, 34% a Red Danish origin and 14% a Montbéliarde origin
- Corresponding origins on pedigree were 56%, 30% and 12%, respectively
- Very high correlation between both breed composition ranging from 0.95 to 0.99

Results: Across bred genomic evaluation

Traits	performances without Heterosis effect				performances adjusted for Heterosis effect			
	Training Population		Validation Population		Training Population		Validation Population	
	<i>Corr</i>	<i>Slope</i>	<i>Corr</i>	<i>Slope</i>	<i>Corr</i>	<i>Slope</i>	<i>Corr</i>	<i>Slope</i>
milk yield	0.83	1.17	0.40	1.03	0.82	1.19	0.41	1.07
fat yield	0.80	1.20	0.36	1.09	0.78	1.22	0.38	1.15
protein yield	0.77	1.23	0.35	1.11	0.76	1.26	0.36	1.16
fat content	0.92	1.15	0.65	1.05	0.92	1.15	0.65	1.05
rotein conter	0.93	1.12	0.62	1.00	0.93	1.12	0.62	1.00

- On the training population, the correlations ranged from 0.76 for protein yield and 0.93 for protein content and fat content. The associated slopes are higher than 1 for all traits. Very few impact of heterosis effect
- On the validation population, correlations ranged from 0.35 to 0.65 an slopes are closer to 1. The effect of heterosis is marginal

Discussions and perspectives

- Application on real dataset of a genomic evaluation based on a SNP BLUP accounting for BOA
- Need imputation and phasing for genotyped crossbred animals to predict BOAs (around 1.5% average error in the prediction)
- Good adequation between BOA and pedigree-based breed composition
- Genomic evaluation showed honourable accuracy in the prediction of production traits, 0.30 for traits with $h^2=0.30$ and 0.60 for traits with $h^2=0.50$
- Few impacts of heterosis here
- Following steps:
 - include DYD in the reference population
 - Other traits
 - Other populations

Acknowledgments

