Efficiency of multi-breed genomic selection for dairy cattle breeds with different sizes of reference population

C. Hozé,*† S. Fritz,† F. Phocas,* D. Boichard,* V. Ducrocq,* and P. Croiseau*

*Institut National de la Recherche Agronomique (INRA), UMR 1313, Génétique Animale et Biologie Intégrative (GABI), 78350 Jouy-en-Josas, France
†Union Nationales des Coopératives d’Élevages et d’Insémination Animales (UNCEIA), 149 rue de Bercy, 75012 Paris, France

ABSTRACT

Single-breed genomic selection (GS) based on medium single nucleotide polymorphism (SNP) density (~50,000; 50K) is now routinely implemented in several large cattle breeds. However, building large enough reference populations remains a challenge for many medium or small breeds. The high-density BovineHD BeadChip (HD chip; Illumina Inc., San Diego, CA) containing 777,609 SNP developed in 2010 is characterized by short-distance linkage disequilibrium expected to be maintained across breeds. Therefore, combining reference populations can be envisioned. A population of 1,869 influential ancestors from 3 dairy breeds (Holstein, Montbéliarde, and Normande) was genotyped with the HD chip. Using this sample, 50K genotypes were imputed within breed to high-density genotypes, leading to a large HD reference population. This population was used to develop a multi-breed genomic evaluation. The goal of this paper was to investigate the gain of multi-breed genomic evaluation for a small breed. The advantage of using a large breed (Normande in the present study) to mimic a small breed is the large potential validation population to compare alternative genomic selection approaches more reliably. In the Normande breed, 3 training sets were defined with 1,597, 404, and 198 bulls, and a unique validation set included the 394 youngest bulls. For each training set, estimated breeding values (EBV) were computed using pedigree-based BLUP, single-breed BayesC, or multi-breed BayesC for which the reference population was formed by any of the Normande training data sets and 4,989 Holstein and 1,788 Montbéliarde bulls. Phenotypes were standardized by within-breed genetic standard deviation, the proportion of polygenic variance was set to 30%, and the estimated number of SNP with a nonzero effect was about 7,000. The 2 genomic selection (GS) approaches were performed using either the 50K or HD genotypes. The correlations between EBV and observed daughter yield deviations (DYD) were computed for 6 traits and using the different prediction approaches. Compared with pedigree-based BLUP, the average gain in accuracy with GS in small populations was 0.057 for the single-breed and 0.086 for multi-breed approach. This gain was up to 0.193 and 0.209, respectively, with the large reference population. Improvement of EBV prediction due to the multi-breed evaluation was higher for animals not closely related to the reference population. In the case of a breed with a small reference population size, the increase in correlation due to multi-breed GS was 0.141 for bulls without their sire in reference population compared with 0.016 for bulls with their sire in reference population. These results demonstrate that multi-breed GS can contribute to increase genomic evaluation accuracy in small breeds.

Key words: multi-breed genomic selection, dairy cattle, high-density chip

INTRODUCTION

Genomic selection has been implemented in many countries. To date, more than 16 countries apply genomic information for dairy cattle breeding (Nilforooshan et al., 2010; Eggen, 2012). However, the accuracy of genomic breeding values depends mainly on the size of the reference population (Hayes and Goddard, 2008; Goddard, 2009) and therefore, genomic evaluations are mainly implemented in large breeds. In small breeds, only a small number of progeny-tested bulls is available and assembling a large reference population consisting of animals with accurate phenotypes is challenging. In this context, a potential solution is to combine reference populations from different breeds and develop multi-breed genomic selection (GS).

Such an approach requires conserved linkage disequilibrium (LD) across breeds to maintain the association between SNP and QTL. Studies on real data show that the association between marker alleles is maintained for SNP <10 kb apart (Gautier et al., 2007; de Roos et al., 2008), a condition not fulfilled with the classically used BovineSNP50 BeadChip (50K, ~50,000 SNP; Illumina...
Mc0129444/14

MULTI-BREED GENOMIC SELECTION FOR DAIRY CATTLE

3919

programs and allow setting a proportion (\(\pi\)) of SNP with a zero effect. Therefore, these approaches were chosen here to compare accuracy of single-breed and multi-breed GS for a small reference population.

In small breeds, reference population size is limiting. Not only is the training set small but assessing the achieved accuracy is difficult because of the small validation population. If the validation population is enlarged, it would be at the expense of the training set and, therefore, at the expense of prediction accuracy (Erbe et al., 2010). Therefore, we chose here to use a large dairy breed to mimic a small breed and develop a multi-breed GS method. This strategy offers the opportunity to study several training population sizes to mimic either small or large breeds while using a unique reasonably large validation set. We first investigated the phenotypes used, the proportions of SNP with a nonzero effect, and the proportion of residual polygenic variance on a 50K basis. Then, we compared the predictive ability of genomic evaluation based on single-breed and multi-breed reference populations using the HD data set. Then, we assessed the benefit of using the HD chip for multi-breed analysis and investigated the effect of population structure on GS accuracy.

MATERIALS AND METHODS

Reference Population

Genotypes used in this study came from the French genomic evaluation. The biological tissues, either cryopreserved semen or blood samples, were provided by various commercial AI companies and breeder organizations in the framework of their breeding program activities. Therefore, no ethical approval was required for this study.

In total, 535, 527, and 773 influential bulls from the Normande (NO), Montbéliarde (MO), and Holstein (HO) breeds were genotyped with the Illumina Bovine HD BeadChip and were used to impute HD genotypes for animals of the French reference population genotyped with the 50K chip. Quality control was performed within breed on the HD and 50K genotypes, using the same criteria for both chips. Genotyped animals with a call rate <0.95 were removed from the analysis. Only markers mapped on the UMD3.1 assembly (http://bovinegenome.org/cgi-bin/gbrowse/bovine_UMD31/) covering the 29 bovine autosomes were used. Any SNP showing departure from Hardy-Weinberg equilibrium (\(P < 0.0001\)) or with more than 10% missing genotypes were removed. In addition, genotype consistency was checked using 1,838 animals that were discordant for more than 1% of these animals were excluded.
Imputation was performed using Beagle 3.3.0 software (Browning and Browning, 2007) with estimated allelic imputation error rates of 0.33, 0.51, and 0.71% in NO, MO, and HO, respectively (Hozé et al., 2013). The imputation study revealed 1,980 SNP with high imputation error rates in most breeds (Hozé et al., 2013). These SNP were also discarded for the analysis leading to an HD data set of 706,791 SNP. The 37,364 SNP in common between the HD and 50K panels were used to mimic 50K genotypes and compare 50K and HD panels for genomic evaluation.

Among the complete French reference population, only bulls with daughters in France and an equivalent daughter contribution (EDC; VanRaden and Wiggans, 1991) >20 were retained, leading to an overall reference population of 8,768 animals (Table 1). Birth year distribution per breed is shown in Figure 1. Phenotypes used were daughter yield deviation (DYD; VanRaden and Wiggans, 1991) for milk yield, fat yield (FatY), protein yield (ProY), fat content (Fat%), protein content (Pro%), and SCS.

### Scenarios Tested

A validation scheme was used to study predictive ability of the genomic selection approaches. In a small breed, where the number of animals in the reference population is low, having a reasonably large validation population to correctly interpret accuracy is nearly impossible. To circumvent this problem, we chose here to use the Normande reference population. The youngest 20% bulls (i.e., those born after November 2004) formed the validation population and the 80% oldest animals were used to form several training populations of different sizes (Figure 1).

The first scenario (A) mimicked a large-breed situation: all 1,597 available NO bulls formed the training population. Most (96.5%) of validation bulls had their sire and maternal grandsires in training set A. In the second scenario (B), only the 404 HD-genotyped NO bulls from the 1,597 were kept in the training population. The 404 animals correspond to the major contributors of the breeds with, whenever possible,
2 to 4 siblings per sire to keep a familial structure. Major contributors were selected based on their marginal contributions to the population (Boichard et al., 1997) computed using the PEDIG software (Boichard, 2002). This training set B mimicked a regional breed. In this case, 83% of validation bulls had their sire and maternal grandsires in training set. A third scenario (C), with 198 NO bulls in the training set, was used to mimic an even smaller breed. It was built by selecting the ancestors of the validation bulls among the HD-genotyped bulls. Because of the selection procedure, the proportion of validation bulls with their sire and maternal grandsires in training set was the same as in scenario B (83%). Finally, scenario D mimicked an extreme case with no animal from the NO breed in the reference population; therefore, no sires of validation animals were part of the training population.

For animals in the validation population, phenotypes were erased to mimic candidates to selection. Then, the training population was used to estimate EBV for the validation animals.

**Evaluation Model**

The EBV were estimated using 3 approaches: pedigree-based BLUP performed with average information (AI)-REML approach (Jensen et al., 1996) and single-breed and multi-breed genomic evaluations performed using the BayesC approach (Kizilkaya et al., 2010) implemented in GS3 software (Legarra et al., 2013).

The general Bayesian mixture model (Meuwissen, 2009) is

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

where \( y \) is the vector of DYD, \( 1 \) is a vector of 1s, \( q \) is the vector of genotype effects, \( M \) is the incidence matrix for markers genotypes, \( u \) is the vector of polygenic effect, \( Z \) is the incidence matrix for the polygenic effect, and \( e \) is a vector of uncorrelated residuals normally distributed, whose variance is inversely proportional to the EDC of each DYD. The model for multi-breed genomic evaluation was the same as in equation [1], except that a breed effect was included.

The model assumes that a small proportion (\( \pi \)) of SNP has a nonzero effect and divides total additive variance into the sum of the variances due to each SNP effect and the residual polygenic variance. In our data sets, convergence was not obtained and we had to assume \( \pi \) known. A strong prior on the variances was used to constrain the proportion of residual polygenic component to a constant.

**Phenotypes and Parameters Used in Genomic Evaluations**

Total genetic variances were estimated using the average information (AI)-REML approach (Jensen et al., 1996). Genetic standard deviations (SD) vary widely among breeds (Table 2), in particular for fat and protein contents. Therefore, we tested 2 types of phenotypes for genomic evaluation: DYD and DYD standardized by the genetic SD of the breed.

Four values ranging from 10 to 40% were tested for the proportion of variance due to residual polygenic effect. This parameter was used to divide total genetic variance into the sum of the variance due to SNP effect and that due to the residual polygenic effect. Two values were tested for the proportion of SNP with a nonzero effect expressed as an equivalent number of expected SNP with a nonzero effect (Su et al., 2012): a value of about 700 SNP (\( \pi = 0.001 \) on the HD panel and \( \pi = 0.02 \) on the 50K) corresponding to a scenario with a small number of QTL influencing the trait, and a value of about 7,000 SNP (\( \pi = 0.01 \) on the HD panel and \( \pi = 0.2 \) on the 50K) corresponding to a scenario with a larger number of QTL.

We tested several sets of parameters using training set B and 50K genotypes as listed above and summarized in Table 3. The parameter settings for the analysis on HD data values were those resulting in the highest average weighted Pearson correlation between EBV and DYD across the 6 traits for validation animals. Because computational requirements for an HD analysis were high, we assumed that the parameters were stable among SNP panels and used them for the analysis on the HD chip. This assumption is supported by Erbe et al. (2012) and Haile-Mariam et al. (2013), where it was shown that the average number of SNP with a nonzero effect and the proportion of genetic variance unaccounted by SNP markers were relatively similar between the 50K and the HD panel.

**Comparison of Evaluation Methods**

After defining the parameters for genomic evaluation, we compared the accuracy of evaluation methods. The criterion to assess the accuracy of prediction was the weighted Pearson correlation (Peers, 1996) between EBV and DYD for validation animals and the weights being the EDC. Data sets A, B, and C were used to compare accuracies of EBV prediction using pedigree-based BLUP, single-breed BayesC, and multi-breed BayesC. For scenarios A and B, we also compared the achieved accuracy of GS using a 50K or an HD panel. In scenario D, no animal from the NO breed was used.
in training; therefore, only multi-breed GS was used to estimate EBV. Tested scenarios and evaluation methods are summarized in Table 3.

In our scenarios, most validation animals had their sire and grandsires in the training set. We assumed that the accuracy achieved with pedigree-based BLUP was close to that achieved with a classical pedigree-based evaluation except that dam performances were not taken into account. Therefore, we compared the gain in accuracy of genomic evaluation methods to the accuracy achieved with pedigree-based BLUP.

RESULTS

Parameters for Genomic Evaluations

Several sets of parameters were tested for BayesC on the 50K data set: 2 numbers of SNP with a nonzero effect, 4 values of proportions of residual polygenic variance ranging from 10 to 40%, and 2 types of phenotypes (raw or standardized DYD) were tested (Table 3). Retained values were based on the highest average accuracy over the 6 traits for training set B (containing 404 NO bulls) and for a single-breed or multi-breed situation.

In the single-breed situation, no clear differences were found between sets of parameters and the average accuracy over the 6 traits for the 394 animals in validation ranged from 0.377 to 0.390 depending on the chosen parameters (not shown). A larger influence of type of phenotypes, proportion of polygenic variance, and proportion of SNP with a nonzero effect was observed for multi-breed genomic evaluations (Figure 2), and average achieved accuracies ranged from 0.338 to 0.398. The largest differences were observed between the 2 tested proportions of SNP: moving from an expected number of SNP with an effect of 700 to 7,000 led to an increase in correlation of 0.03. The effect of the proportion of polygenic variance depended on the proportion of SNP with a nonzero effect. We observed that when a larger polygenic component was assumed (>20% of the genetic variance), fewer SNP were required to accurately predict phenotypes. This suggests that when fewer SNP are included in the model, a smaller part of genetic variance is explained by SNP and therefore a larger variance due to residual polygenic is required.

The optimal proportion of variance due to polygenic components seemed to be related to the heritability of the traits. A higher proportion of residual polygenic variances gave slightly higher results for low heritability traits (SCS), whereas lower proportions were required for fat and protein contents. A proportion of polygenic variance of 30% was the optimal trade-off across the 6 traits. Standardization of phenotypes allowed an aver-
age gain of 0.01 in accuracy. Standardization had little effect on accuracy when the proportion of SNP with a nonzero effect was high. This suggests that standardization reduces the residual noise associated with raw phenotypes and avoids selecting SNP that explain differences in genetic variance between breeds.

The highest accuracy was obtained when the phenotypes were standardized, the proportion of polygenic variance was set to 30%, and an estimated number of SNP with a nonzero effect was around 7,000. Therefore, these parameters were used in the rest of this study.

**Multi-Breed Versus Single-Breed GS**

Multi-breed and single-breed genomic evaluations were compared for training sets A, B, and C. Results based on the HD panel are presented in Table 4. Regardless of the method, accuracies were higher in training set A than in the smaller training sets, which was at least partly due to our design. Training set A included 97% of the validation bulls sires and maternal grandsires. This proportion decreased to 83% with training sets B and C. This led to a lower accuracy of pedigree-based BLUP and a poorer estimation of the polygenic component in GS. For the same reasons, regression slopes were closer to 1 in training set A.

Comparing pedigree-based and genomic evaluations, we notice that the use of DNA information increased the accuracy of EBV prediction in both data sets (Table 4). However, the gain was 2.5 times higher in training set A (+0.193) than in training sets B (+0.077) and C (+0.057). The larger reference population allowed a better estimation of SNP effects and therefore higher accuracies. The increase in correlation due to the GS was trait-dependent (Table 4). The largest gains were observed for highly heritable traits influenced by large QTL such as milk content and the lowest were observed for SCS.

In all data sets, moving from single-breed to multi-breed GS allowed a slight increase in the quality of prediction (Table 4). The average increase in accuracy compared with pedigree-based BLUP varied from 0.193 for single-breed GS to 0.209 with multi-breed GS in Table 4.

### Table 3. Overview of the main features of the scenarios tested for Montbéliarde (MO), Normande (NO), and Holstein (HO) breeds

<table>
<thead>
<tr>
<th>Feature</th>
<th>Training A</th>
<th>Training B</th>
<th>Training C</th>
<th>Training D</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLUP</td>
<td>1,597 NO</td>
<td>404 NO</td>
<td>198 NO</td>
<td>4,989 HO</td>
</tr>
<tr>
<td>Single-breed GS(^1)</td>
<td>1,597 NO</td>
<td>404 NO</td>
<td>198 NO</td>
<td>4,989 HO</td>
</tr>
<tr>
<td>Multi-breed GS</td>
<td>1,597 NO</td>
<td>404 NO</td>
<td>198 NO</td>
<td>4,989 HO</td>
</tr>
<tr>
<td>Phenotypes(^2)</td>
<td>DYD/sDYD</td>
<td>DYD/sDYD</td>
<td>sDYD</td>
<td>sDYD</td>
</tr>
<tr>
<td>SNP panel(^3)</td>
<td>50K/HD</td>
<td>50K/HD</td>
<td>HD</td>
<td>HD</td>
</tr>
<tr>
<td>Eff_SNP(^4)</td>
<td>7,000</td>
<td>700/7,000</td>
<td>7,000</td>
<td>7,000</td>
</tr>
<tr>
<td>Polygenic (%)</td>
<td>30</td>
<td>10/20/30/40</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Validation (with sire + mgs)(^5)</td>
<td>394 NO (380)</td>
<td>39 NO (326)</td>
<td>394 NO (326)</td>
<td>394 NO (0)</td>
</tr>
</tbody>
</table>

\(^1\)GS = genomic selection.

\(^2\)DYD = daughter yield deviation; sDYD = standardized DYD.

\(^3\)50K = ~50,000 SNP panel; HD = high-density SNP chip.

\(^4\)Eff_SNP = expected number of markers with a nonzero effect.

\(^5\)Number of validation bulls with sire and maternal grandsire (mgs) in the training set.
training set A, from 0.079 to 0.106 in training set B, and from 0.057 to 0.086 in training set C. It represents extra gain of, respectively, 8, 39, and 50% with multi-breed GS compared with single-breed GS. The use of information from MO and HO breeds allowed a gain for the 6 traits studied (Table 4) but the extra gain was trait- and data set-dependent. The largest increase was observed for Fat% in training set B (+0.08) and the lowest for SCS in training set A (+0.007). Fat content is influenced by large QTL, some of which are probably common across breeds. This could explain why the largest increases were observed for this trait. Common QTL may be more difficult to identify in traits influenced by small QTL (SCS, ProY). The benefit of multi-breed GS is also dependent on the accuracy of prediction of single-breed GS. For training set A, where the NO reference population was large, the accuracy of single-breed GS was already high and adding information from other breeds had a limited effect. Regression slopes of multi-breed GS were in the same range as those from single-breed GS, meaning that adding information from other breeds did not affect the bias.

**Table 4.** Weighted Pearson correlations between observed and predicted daughter yield deviations (DYD) for the 6 traits using pedigree-based BLUP, single-breed genomic selection (GS) and multi-breed GS using training sets A (1,597 Normande bulls), B (404 Normande bulls), and C (198 Normande bulls) with the high-density panel

<table>
<thead>
<tr>
<th>Training set</th>
<th>Model</th>
<th>Milk</th>
<th>FatY</th>
<th>ProY</th>
<th>Fat%</th>
<th>Pro%</th>
<th>SCS</th>
<th>Average corr.</th>
<th>Average slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Multi-breed GS</td>
<td>0.505</td>
<td>0.518</td>
<td>0.514</td>
<td>0.647</td>
<td>0.644</td>
<td>0.568</td>
<td>0.566</td>
<td>0.875</td>
</tr>
<tr>
<td></td>
<td>Single-breed GS</td>
<td>0.484</td>
<td>0.489</td>
<td>0.495</td>
<td>0.638</td>
<td>0.632</td>
<td>0.561</td>
<td>0.550</td>
<td>0.863</td>
</tr>
<tr>
<td></td>
<td>BLUP</td>
<td>0.316</td>
<td>0.346</td>
<td>0.395</td>
<td>0.450</td>
<td>0.397</td>
<td>0.430</td>
<td>0.357</td>
<td>0.760</td>
</tr>
<tr>
<td>B</td>
<td>Multi-breed GS</td>
<td>0.345</td>
<td>0.401</td>
<td>0.353</td>
<td>0.478</td>
<td>0.494</td>
<td>0.427</td>
<td>0.416</td>
<td>0.726</td>
</tr>
<tr>
<td></td>
<td>Single-breed GS</td>
<td>0.312</td>
<td>0.389</td>
<td>0.333</td>
<td>0.398</td>
<td>0.470</td>
<td>0.418</td>
<td>0.387</td>
<td>0.710</td>
</tr>
<tr>
<td></td>
<td>BLUP</td>
<td>0.234</td>
<td>0.299</td>
<td>0.244</td>
<td>0.305</td>
<td>0.382</td>
<td>0.399</td>
<td>0.311</td>
<td>0.663</td>
</tr>
<tr>
<td>C</td>
<td>Multi-breed GS</td>
<td>0.347</td>
<td>0.401</td>
<td>0.368</td>
<td>0.434</td>
<td>0.444</td>
<td>0.389</td>
<td>0.397</td>
<td>0.717</td>
</tr>
<tr>
<td></td>
<td>Single-breed GS</td>
<td>0.319</td>
<td>0.396</td>
<td>0.340</td>
<td>0.368</td>
<td>0.420</td>
<td>0.368</td>
<td>0.368</td>
<td>0.729</td>
</tr>
<tr>
<td></td>
<td>BLUP</td>
<td>0.243</td>
<td>0.306</td>
<td>0.249</td>
<td>0.305</td>
<td>0.373</td>
<td>0.393</td>
<td>0.311</td>
<td>0.685</td>
</tr>
</tbody>
</table>

1Genomic evaluations were based on a BayesC approach. Phenotypes used were standardized within breed; the proportion of residual polygenic component was set to 30% and the number of expected SNP with a nonzero effect to 7,000.

2Milk yield (Milk), fat yield (FatY), protein yield (ProY), fat content (Fat%), protein content (Pro%), and SCS.

3Average correlation between DYD and EBV over the 6 traits.

4Average regression slope over the 6 traits.

The comparison between the 50K study and the HD study allowed us to assess the effect of SNP density on the accuracy of prediction. Results for single-breed and multi-breed GS are presented in Figure 3 for training sets A and B.

As seen before, accuracies were higher for training set A than for training set B and were slightly higher for multi-breed GS than for single-breed GS. However, the effect of SNP density depended on the evaluation method. Slight or no differences were observed between the 50K and HD panels for single-breed GS for training set A (+0.006) and training set B (+0.002). Larger differences were observed for multi-breed GS. The average gain of HD data set was 0.022 for training set A and 0.018 for training set B. This could be explained by the higher linkage disequilibrium on the HD chip allowing a SNP-QTL association better conserved across breeds.

**Effect of Population Structure**

Accuracies of prediction were compared for validation bulls whether they had their sire in the training sets B and C or not (Table 5). Differences in correlation across training sets and evaluation methods were smaller for bulls with their sire in reference population than for those without. In both training sets, the correlation was halved for bulls without their sire in the training population but the average gain in correlation with multi-breed GS was more than 10 times higher for bulls that did not have their sire in the training population (+0.16) than for those that did (+0.013): moving to a larger reference population or a multi-breed GS was more beneficial for validation animals that did not have their sire in training population. The increase in EBV prediction accuracy is likely due to a more accurate estimation of SNP effect. For bulls without their sire in training set B, the correlation between direct genomic breeding value and DYD increased from 0.218 to 0.375 when the reference population included more animals, regardless of the breed of the additional training animals. In the case of validation bulls with their sire in the reference population, accuracies were not improved.
When no animals from the Normande breed were in the reference population, accuracies of multi-breed GS were around 0.18 for all traits except for Pro% and SCS (Table 6). For the latter traits, the benefit of multi-breed GS compared with single-breed GS was already low in training sets B and C (Table 5). On average, the accuracy achieved was close to the one using data set C and single-breed GS for animals without their sire in the reference population. This suggests that multi-breed GS could be beneficial even for breeds with no animals at all in the reference population.

**DISCUSSION**

In this study, we assessed the benefit of combining reference populations from several breeds and performed multi-breed GS. A BayesC approach known to provide good results in genomic evaluation (e.g., Croiseau et al., 2012; Duchemin et al., 2012) was used.

The model used here requires some key parameters such as the proportion of SNP with a nonzero effect and the proportion of genetic variance due to a residual polygenic component. We looked for the optimal values for these parameters using data from the 50K chip and later used these for the analysis on the HD chip. When the population was small (<500 bulls) and HD genotypes were used, the average gain in accuracy compared with pedigree-based BLUP was 0.067 using single-breed GS and up to 0.105 using multi-breed GS. The benefit of multi-breed GS compared with single-breed GS was reduced when 50K genotypes were used and almost disappeared in the case of an already-large reference population.

Because of the low number of progeny-tested bulls, the current reference population for French regional breeds (Abondance, Brown Swiss, Tarentaise, and Simmental) consisted of fewer than 200 bulls (Hozé et al., 2013); therefore, using them for a validation study was questionable. The alternative implemented here was to use a national breed with a large reference population and to artificially reduce the training population size. The major advantage of this approach is to keep a reasonably large validation population, allowing for more robust conclusions and for a better estimation of the effect of reference population. This design also has some disadvantages. In particular, the choice of the animals kept in a reduced training set has a major influence on population structure and on genomic selection accuracy. Studies on simulated and real data also showed that for one given animal, the reliability of GEBV and direct genomic breeding value was linked to its relationship to reference population (Habier et al., 2010; Clark et al., 2012; Pszczola et al., 2012). Erbe et al. (2012) observed a loss of correlation between 0.05 and 0.09 when the sire of the animal considered was not in the reference population. Here, observed losses were dramatically larger (between 0.095 and 0.285) because reference population sizes considered in this study were much smaller. Accuracy of GS depends on the average relationship between individuals and LD between genotypes.

**Table 5.** Weighted-average Pearson correlations between observed daughter yield deviations and EBV for the 6 traits in relation to single-breed and multi-breed genomic selection (GS) in the training set, the evaluation methods on the high-density panel, and the presence of sires in the training set

<table>
<thead>
<tr>
<th>Group</th>
<th>Training set C (198 bulls)</th>
<th>Training set B (404 bulls)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single-breed GS</td>
<td>Multi-breed GS</td>
</tr>
<tr>
<td>Bulls with sire in training set (n = 355)</td>
<td>0.414 0.432</td>
<td>0.431 0.444</td>
</tr>
<tr>
<td>Bulls without sire in training set (n = 39)</td>
<td>0.129 0.251</td>
<td>0.189 0.349</td>
</tr>
</tbody>
</table>

1Milk yield, fat yield, protein yield, fat content, protein content, and SCS.

2Genomic evaluations were based on a BayesC approach. Phenotypes used were standardized within breed; the proportion of residual polygenic component was set to 30% and the number of expected SNP with a nonzero effect to 7,000.

**Table 6.** Weighted Pearson correlations between observed and predicted daughter yield deviations for the 6 traits using multi-breed genomic selection (GS) with training set D (no Normande bulls) and a high-density panel

<table>
<thead>
<tr>
<th>Training set</th>
<th>Model</th>
<th>Milk</th>
<th>FatY</th>
<th>ProY</th>
<th>Fat%</th>
<th>Pro%</th>
<th>SCS</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>Multi-breed GS</td>
<td>0.172</td>
<td>0.168</td>
<td>0.182</td>
<td>0.174</td>
<td>0.097</td>
<td>0.045</td>
<td>0.140</td>
</tr>
</tbody>
</table>

1Milk yield (Milk), fat yield (FatY), protein yield (ProY), fat content (Fat%), protein content (Pro%), and SCS.

2Genomic evaluations were based on a BayesC approach. Phenotypes used were standardized within breed; the proportion of residual polygenic component was set to 30% and the number of expected SNP with a nonzero effect to 7,000.
markers and causative genes (Habier et al., 2011). The accuracy due to LD is considered as a lower bound of the global accuracy for animals with no close relatives in the reference population and was found to increase with the reference population size (Habier et al., 2011). Our results are consistent with this statement: the average increase in accuracy for animals without their sires in the reference population was higher with larger training population or with a multi-breed evaluation.

The accuracy achieved in GS depends on the familial structure and on the accuracy obtained with pedigree-based BLUP. For this reason, comparing the gain in accuracy due to GS may be more relevant than comparing the accuracies themselves. The accuracy of prediction dramatically increased with GS in training set A (+0.193) and increased less for training sets B (+0.077) and C (+0.057). This result confirms the effect of reference population size on the accuracy of GS (Daetwyler et al., 2008; Goddard, 2009). In contrast, the extra gain with multi-breed GS was higher with training sets B (+0.030) and C (+0.029) than with training set A (+0.016). Holstein-Jersey studies (Hayes et al., 2009; Erbe et al., 2012) and a Holstein-Jersey-Brown Swiss study (Olson et al., 2012) also showed that combining reference populations is more beneficial for small populations than for large populations. It suggests that when the single-breed reference population is large enough, estimation of SNP effects and genomic evaluation are already accurate and adding information from other breeds does not improve them.

Density of markers is also known to affect the accuracy of GS (Daetwyler et al., 2008; Goddard, 2009). In single-breed evaluations, simulation studies forecast high increases in accuracy when marker density is increased (de Roos et al., 2009; Meuwissen and Goddard, 2010; VanRaden et al., 2011) but results on real data found limited gain with the HD chip (Erbe et al., 2012; Su et al., 2012; VanRaden et al., 2013). In our study, we also found little or no improvement (+0.006 for training set A and +0.003 for training set B) when moving from the 50K panel to the HD chip. One reason for this limited gain is probably the already high proportion of genetic variance captured by SNP on production traits with the 50K panel. This was estimated at 80% by Haile-Mariam et al. (2013) and was only slightly higher with the HD chip (Erbe et al., 2012). The higher number of effects to estimate may also be a factor limiting the increase in correlation. This hypothesis is supported by the lower gain in correlation due to the use of the HD chip for training set B where the reference population was small. Erbe et al. (2012) even observed a decrease in correlation with the HD chip for a Jersey population of 540 bulls. The benefit of the HD chip was clearer when multi-breed GS was used. The increase in correlation observed with the HD chip was 0.022 for training set A and 0.018 for training set B. These results are in the same range as in Erbe et al. (2012), where the use of HD chip led to a 0.03 increase of multi-breed GS accuracy observed in the Jersey population. The higher density of SNP allowed a higher persistence of LD across breeds and a more conserved QTL-SNP phase, as stated by de Roos et al. (2009).

However, the achieved accuracies observed in multi-breed studies using the HD chip are lower than those observed in simulated data sets where QTL and SNP effects are assumed to be shared across breeds (de Roos et al., 2009; Erbe et al., 2012). This may suggest that marker-QTL associations across breeds are less conserved than expected or that QTL effects may differ among breeds. Applying SNP effects estimated in one breed to another breed led to poor correlations between EBV and observed phenotypes (Erbe et al., 2012; Olson et al., 2012). This confirmed that associations between QTL and SNP are at least partially breed specific and that SNP effects cannot be directly transposed from one breed to another. In this study, breeds were pooled.
together and an average SNP effect was estimated. This model may dilute the observed associations of markers with phenotypes as it does not take into account for possible differences in SNP effect among breeds or for the possible existence of breed-specific QTL. A possible improvement for multi-breed GEBV estimation would be to include a SNP × breed interaction in the model and to allow SNP effect to differ in variance, value, and sign between populations. Multi-trait models, where phenotypes of each breed are considered correlated traits, were proposed to account for the breed-QTL interaction. In Nordic Red dairy cattle where the populations are closely related, an increase in correlation of 2 to 3% was observed when comparing a breed-specific model and a genomic BLUP assuming a uniform population (Makgahlela et al., 2013). Considering Holstein, Jersey, and Brown Swiss data sets, Olson et al. (2012) showed that using phenotypes of each breed as correlated traits led to significantly higher reliabilities than just pooling the reference population over several breeds but they found limited gain compared with a single-breed genomic evaluation. In a study on the 3 main French dairy breeds, Karoui et al. (2012) observed an increase in correlation when comparing a multi-trait evaluation with each breed considered as a different trait to a single-breed genomic evaluation. However, using estimated genetic correlations between breeds did not improve EBV estimation compared with a situation where genetic correlations were set arbitrarily to 0.95. This situation is close to that where populations are pooled and suggests that for the main French breeds, using a multi-trait model would not improve genomic evaluation. An alternative to account for possible differences in SNP-QTL association was proposed by Brøndum et al. (2012). In their model, a first prediction of SNP effects is performed on the largest breeds and the posterior probability for a SNP to have a small or a large effect is used as a prior for the estimation of SNP effect in the second breed. Accuracies achieved with this approach were higher than those obtained with a single-breed GS but lower than those achieved when reference populations are pooled (Brøndum et al., 2012).

Until now, methods proposed to account for incompletely conserved LD between QTL and SNP resulted in limited gain in accuracy. This is probably explained by the higher number of effects to estimate with a breed-specific allele model, which requires a large reference population per breed (Ibáñez-Escriche et al., 2009). Use of haplotypes could also be useful to better detect conserved LD across breeds as it increases LD with the causal mutations (Hayes et al., 2007; Calus et al., 2008). With the reduction of genome resequencing cost and imputation from SNP chip to sequence data, GS based on sequence data can be envisioned (Druet et al., 2014). Sequence data are particularly appealing for multi-breed GS as the causal mutation is in the data set. Therefore, it is no longer needed to have a conserved marker-QTL association across breed. Again, the number of effects to estimate would be limiting and SNP preselection through QTL detection could probably be useful to better estimate SNP effect and take benefit of the higher density (Croiseau et al., 2011).

In this study, multi-breed genomic evaluations were used to increase reference population size and therefore to increase the accuracy of EBV predictions. A complementary solution is to genotype females and to include cow information in genomic evaluation. The benefit of inclusion of females on accuracy of genomic selection and increase in genetic gain has been assessed in simulations studies (Mc Hugh et al., 2011). On real data, adding 10,000 cows to a population of 3,000 bulls led to an increase in reliability of genomic prediction ranging from 4 to 8% (Pryce et al., 2012). For a smaller population and a highly heritable trait, Calus et al. (2013) showed that adding 1,609 cows to a population of 296 bulls increased prediction accuracy by 45%. Indeed, the increase in accuracy depends on the heritability of the traits studied, the number of cows added to the reference population, and the bull reference population size (Egger-Danner et al., 2012). Therefore, it is likely that, in a small breed where the number of progeny-tested bulls is limited, inclusion of cows in reference population is very beneficial.

CONCLUSIONS

This study compared GS approaches for small breeds of dairy cattle and 6 traits. It focused on the benefit of multi-breed GS for breeds with fewer than 500 animals in their reference population. We observed an average increase in accuracy over the 6 traits ranging from 0.057 and 0.209 with single-breed genomic evaluation compared with pedigree-based BLUP, and the lowest gain was observed for a training population of <200 animals. Compared with single-breed GS, multi-breed GS allows a gain in correlation between EBV and observed DYD that ranges between 0.016 and 0.029. The highest increase was observed for traits influenced by large QTL (e.g., fat content). Conserved linkage disequilibrium across breeds does exist and multi-breed GS may benefit breeds with a small reference population size. However, gain in correlation with multi-breed GS was low when the breed reference population was already large and when the validation animals had their sire in the reference population. Implementation of GS in small breeds is feasible provided that ancestors of selection candidates are in the reference population.
With the increase in reference population size under a genomic evaluation breeding scheme, female genotyping, and multi-breed GS, EBV prediction accuracy in small dairy cattle breeds could quickly increase in the near future.

ACKNOWLEDGMENTS

The GEMBAL project was funded by the Agence Nationale de la Recherche (Paris, France; ANR-10-GENM-0014), ApisGene (Paris, France), Races de France (Paris, France), and INRA AIP Bioressources (Paris, France). The EUROGENOMICS consortium provided most of the Holstein genotypes. Most 50K genotypes originated from the Cartofine-ANR-05-GEN-ANIMAL-007 project funded by ANR (French National Research Agency) and ApisGene, and from genomic selection activity generated by the French cattle breeding companies, with Labogena (Jouy-en-Josas, France) as main genotyping laboratory.

REFERENCES


