Pure

Scotland's Rural College

Estimation of genomic breeding values for milk yield in UK dairy goats

Mucha, S; Mrode, R; MacLaren-Lee, I; Coffey, MP; Conington, JE

Link to publication

Estimation of genomic breeding values for milk yield in UK dairy goats

S. Mucha,[*][1] **R. Mrode,**[*][†] **I. MacLaren-Lee,**[‡] **M. Coffey,**[*] **and J. Conington**[*]
*Animal and Veterinary Sciences, Scotland's Rural College, Easter Bush, Midlothian EH25 9RG, Scotland, United Kingdom
†Animal Biosciences, International Livestock Institute, Nairobi 00100, Kenya
‡Illumina Cambridge Ltd., Great Abington, Cambridge CB21 6GP, United Kingdom

ABSTRACT

The objective of this study was to estimate genomic breeding values for milk yield in crossbred dairy goats. The research was based on data provided by 2 commercial goat farms in the UK comprising 590,409 milk yield records on 14,453 dairy goats kidding between 1987 and 2013. The population was created by crossing 3 breeds: Alpine, Saanen, and Toggenburg. In each generation the best performing animals were selected for breeding, and as a result, a synthetic breed was created. The pedigree file contained 30,139 individuals, of which 2,799 were founders. The data set contained test-day records of milk yield, lactation number, farm, age at kidding, and year and season of kidding. Data on milk composition was unavailable. In total 1,960 animals were genotyped with the Illumina 50K caprine chip. Two methods for estimation of genomic breeding value were compared—BLUP at the single nucleotide polymorphism level (BLUP-SNP) and single-step BLUP. The highest accuracy of 0.61 was obtained with single-step BLUP, and the lowest (0.36) with BLUP-SNP. Linkage disequilibrium (r^2, the squared correlation of the alleles at 2 loci) at 50 kb (distance between 2 SNP) was 0.18. This is the first attempt to implement genomic selection in UK dairy goats. Results indicate that the single-step method provides the highest accuracy for populations with a small number of genotyped individuals, where the number of genotyped males is low and females are predominant in the reference population.
Key words: genomic selection, single step, milk yield, dairy goat

INTRODUCTION

Genomic selection has become routine in many farmed livestock species such as dairy and beef cattle. This is mainly due to exchange of genotypes between countries, and reference populations for those species are now large, consisting of thousands of bulls with high reliability breeding values. This allows the prediction of genomic breeding values for young animals, which have no phenotypic records, with acceptably high accuracy. In the case of dairy goats, the breeding industry is not so well developed worldwide. Routine breeding value estimation is carried out in several countries such as Canada, France, United States, and Norway (Bélichon et al., 1999; Montaldo and Manfredi, 2002). Recently, because of the introduction of the Illumina Caprine 50K BeadChip (Illumina Inc., San Diego, CA; Tosser-Klopp et al., 2012), genomic tools became available for genetic improvement of goats. Currently, genomic selection in dairy goats has been introduced only in France (Carillier et al., 2013), with 2,810 genotyped Saanen and Alpine goats. In the UK, the number of genotyped goats is also relatively small, which poses certain restrictions with respect to the estimation of genomic breeding values. Accuracy of methods that use only phenotypes of the genotyped animals and ignore records of the nongenotyped part of the population (e.g., genomic BLUP, BLUP-SNP) is limited when the reference population is small. Therefore, an alternative approach was considered that integrates all of the available phenotypic, pedigree, and genomic information in a single-step procedure (Legarra et al., 2009; Misztal et al., 2009; Christensen and Lund, 2010). This approach has been regarded as computationally demanding in cases of large data sets with hundreds of thousands of genotyped animals. However, in goats the amount of data used in genetic evaluations is considerably lower than that of dairy cattle. Moreover, the method is easy to implement because it can use raw phenotypic records without the need to calculate deregressed proofs (**DRP**). It also allows evaluation of all animals (with and without genotypes) simultaneously.

The objective of this study was to evaluate BLUP-SNP and single-step approach for estimation of genomic breeding values in dairy goats. Additionally, level of

linkage disequilibrium in the reference population was investigated.

MATERIALS AND METHODS

Phenotypic Data

Lactation data were from 2 separate farm units in the UK owned by a single farming business. The data set comprised 590,409 records on 14,453 dairy goats kidding between 1987 and 2013. The population was created in 1985 by crossing 3 breeds: Alpine, Saanen, and Toggenburg. No particular crossing strategy existed. In each generation the best performing animals were selected for breeding, and as a result, a synthetic breed was created. The pedigree file contained 30,139 individuals, of which 2,799 were considered as founders. A total of 296 sires and 12,468 dams were in the pedigree. The data set contained test-day records of milk yield, along with information about lactation number (1 to 6), farm (2 farms), age at kidding (12 to 90 mo), and year (1987 to 2013) and season of kidding [summer (June to August), autumn (September to November), winter (December to February), and spring (March to May)]. Fat and protein content was not included in the analysis because it was not recorded on either of the farms contributing data. Only goats with more than 3 test-day observations were used for analysis. Additionally, the data set was restricted to groups having at least 10 records per level of herd test day, year-season and age at kidding. Test-day milk records below 0.2 and above 12.5 kg were removed from the data as error records. Lactation length was restricted to between 4 and 520 DIM because goats from the 2 farms are routinely milked for long lactations. Heritability of milk yield in the analyzed population was 0.56. For a detailed description of the analyzed population, see Mucha et al. (2014).

Pedigree Analysis

Pedigree completeness was assessed with complete generations equivalent. The number of equivalent generations traced was computed as the sum over all known ancestors of the terms $(1/2)^t$, where t is the ancestor's generation number, which is equal to 1 for the parents, 2 for the grandparents, and so on (Maignel et al., 1996).

Effective population size (realized N_e) based on individual increase in inbreeding (ΔF_i) was calculated following the approach proposed by Gutiérrez et al. (2009). The ΔF_i coefficients were computed as

$$\Delta F_i = 1 - {}^{t-1}\!\sqrt{1 - F_i},$$

where F_i is the individual coefficient of inbreeding and t is the complete generations equivalent (Maignel et al., 1996). The effective population size (N_e) was obtained from $\overline{\Delta F}$, which was computed by averaging the ΔF_i of the n individuals included in a given reference subpopulation as $N_e = \dfrac{1}{2\overline{\Delta F}}$.

Genotypes

In total 1,960 animals were selected for genotyping. All of the available sires (150 individuals) were sampled, and subsequently, this set was supplemented by females. Selection of females for genotyping was based on 2 criteria: average daily lifetime yield and genetic relationship between the animals. The process was optimized in a way to select animals from the upper (group 1) and lower (group 2) tail of the distribution of average daily lifetime yield. Animals had been selected so that the relationship within the 2 groups was minimized and the relationship between the 2 groups was maximized. This was done with the software package Corona produced by Brian Kinghorn (University of New England, Armidale, Australia, personal communication).

Animals were genotyped commercially with a 50K Caprine Illumina SNP chip at Edinburgh Genomics (Edinburgh, UK). After filtering out SNP that were not in Hardy-Weinberg equilibrium, had minor allele frequency below 0.05, were monomorphic, had a call rate below 0.95, or had an Illumina GenCall (GC score) below 0.6, the data set contained 47,306 markers. Additionally, animals with a call rate below 0.9 were removed from further analyses. This resulted in 1,902 genotyped animals born between 2003 and 2012 being used. The SNP information was also used to correct the pedigree by means of sire and dam verification for animals with known parents and for parent discovery for animals with unknown pedigree. Finally, the genotype matrix was used for clustering based on principal component analysis, performed with SNP & Variation Suite v7.7.8 (Golden Helix Inc., Bozeman, MT). This analysis was done to investigate potential population stratification due to historical crossbreeding in the population.

Linkage Disequilibrium

Linkage disequilibrium (**LD**) was measured as r^2, which is the squared correlation of the alleles at 2 loci (Hill and Robertson, 1968):

$$r^2 = \frac{\left[f(AB) - f(A)f(B)\right]^2}{f(A)f(a)f(B)f(b)},$$

where $f(AB)$, $f(A)$, $f(a)$, $f(B)$, and $f(b)$ are observed frequencies of haplotype AB and of alleles A, a, B, and b, respectively. Linkage disequilibrium was calculated for all syntenic marker pairs (markers from the same chromosome). That SNP markers that could not be mapped to any chromosome were excluded from this analysis. Average LD was calculated as an arithmetic mean of r^2 values for SNP pairs in 1-kb windows from all chromosomes. Linkage disequilibrium based on the marker data was compared with an approximate expectation (E) of r^2 (Sved, 1971):

$$E(r^2) = \frac{1}{(4N_e c + 1)},$$

where N_e is the effective population size and c is the recombination distance in Morgans (we assumed 100 Mb = 1 Morgan) between SNP. Assuming that LD at short distances is dependent on long-term population history (Hayes et al., 2003; Hill, 1981), the historic effective population size was estimated as

$$N_t = \frac{1 - r^2}{4cr^2},$$

where N_t is the effective population size t generations ago, where $t = 1/(2c)$ (Hayes et al., 2003). The N_e from 5 generations ago (110 individuals) was considered as the most recent with $c = 0.1$ Morgans.

Estimation of Genomic Breeding Values

Two methods were used to estimate genomic breeding values (**GEBV**). The first method was BLUP at the SNP level (BLUP-SNP), where deregressed sire and dam proofs were used as phenotypes. The software package MIX99 (Lidauer et al., 2011) was used for the deregression using a full animal pedigree with effective offspring contributions (EOC) used as the weighting factor. The EOC were calculated in the same way for males and females as

$$\text{EOC}_i = \frac{\text{rel}_i \cdot \text{kdau}}{1 - \text{rel}_i},$$

$$\text{kdau} = \frac{4 - h^2}{h^2},$$

where rel_i is the reliability of EBV for animal i; kdau is a function of heritability as defined by Fikse and Banos (2001); and h^2 is the heritability of milk yield. The

SNP effects were estimated with the following statistical model:

$$y_i = \mu + \nu_i + \sum_{j=1}^{m} z_{ij} u_j + e_i,$$

where y_i is the DRP; μ is the overall mean; v_i is the residual polygenic effect of ith goat (10% of additive genetic variance); m is the number of SNP markers used in the analysis; z_{ij} is the genotype value coded as 0, 1, or 2 for homozygote, heterozygote, and the opposing homozygote; u_j is the random regression coefficient for jth SNP; and e_i is the residual effect. The residual polygenic effect was set as 10% of additive genetic variance to avoid high variance of direct genomic values (**DGV**) and to minimize prediction bias of GEBV (Liu et al., 2011).

The second approach to calculate GEBV was based on the single-step method (Legarra et al., 2009; Misztal et al., 2009). The software package BLUPf90 (Misztal et al., 2002) was used to fit the following random regression model:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp} + \mathbf{e},$$

where $\mathbf{y}$ is the vector of test-day observations; $\mathbf{b}$ the vector of fixed effects consisting of herd test day, year-season, age at kidding, and fixed lactation curves modeled by fitting Legendre polynomials (Kirkpatrick et al., 1990) of fourth order; $\mathbf{a}$ is a 1×3 vector of random regression coefficients (Legendre polynomials of second order) for the animal effect; $\mathbf{p}$ is the 1×3 vector of random regression coefficients (Legendre polynomials of second order) for the permanent environment effect; and $\mathbf{e}$ is the vector of random residual effects. The matrix $\mathbf{X}$ is the incidence matrix for fixed effects; $\mathbf{Z}$ and $\mathbf{W}$ are matrices of Legendre polynomials of DIM of second order for random animal and permanent environment effect, respectively.

Random effects were assumed to be normally distributed with zero means and the following covariance structure:

$$\text{Var} \begin{bmatrix} \mathbf{a} \\ \mathbf{p} \\ e \end{bmatrix} = \begin{bmatrix} \mathbf{H} \otimes \mathbf{U} & 0 & 0 \\ 0 & \mathbf{I} \otimes \mathbf{P} & 0 \\ 0 & 0 & \mathbf{I}\sigma_e^2 \end{bmatrix},$$

where $\mathbf{U}$ and $\mathbf{P}$ are 3×3 (co)variance matrices of random regression coefficients for the animal and permanent environment effects, respectively, $\mathbf{I}$ are identity matrices, and $\mathbf{H}$ is the relationship matrix calculated

using Van Raden's (2008) genomic relationship matrix $\mathbf{G}$ and pedigree relationship matrix $\mathbf{A}$ as

$$\mathbf{G} = 0.95 \frac{\mathbf{SS'}}{2\sum_{i=1}^{n} p_i \left(1 - p_i\right)} + 0.05\mathbf{A},$$

where $\mathbf{S}$ is a centered incidence matrix of SNP genotypes, n is the number of SNP markers, and p_i is allele frequency of marker i. The inverse of $\mathbf{H}$ (Aguilar et al., 2010; Christensen and Lund, 2010) is

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix},$$

where $\mathbf{A}_{22}^{-1}$ is the inverse of pedigree relationship matrix for the genotyped animals.

Accuracy of Genomic Breeding Values

The genotyped animals were divided into a training and a validation set consisting of 1,474 (1,410 females and 64 males) and 305 animals (302 females and 3 males), respectively. Animals in the training set were born between 2003 and 2010 and had a minimum reliability of EBV of 0.76. Validation animals were born in 2011 and had a minimum reliability of EBV of 0.69. Validation animals had no phenotypic records because the data set was restricted to animals born up to and including 2010. Females in the reference and validation set had 1 to 6 and 1 to 2 lactations, respectively. The third group of animals consisted of 123 animals that did not meet the criteria for entering the reference or validation set. These animals were not used in the current analysis but will be included in the future GEBV evaluations once they have an appropriate reliability of their EBV. In total 59 of the validation animals had both parents in the reference set, 116 had only a sire in the reference set, and 25 animals had only their dam in the reference set. The accuracies of genomic predictions were calculated as a correlation between DRP and GEBV for the single-step method (**HBLUP**), or DGV for BLUP-SNP in the validation animals. Additionally, the accuracy of pedigree-based predictions (**PBLUP** and parent average) was calculated as a correlation of DRP with EBV and parent average of the validation animals. The EBV from PBLUP were obtained from the same model as for HBLUP, but the H matrix was replaced with the pedigree-based A matrix. Parent average was calculated as an average of parent EBV based on the validation data set. The gain of using SNP information was calculated as a difference between the accuracy of the evaluated methods (BLUP-SNP, HBLUP) and PBLUP.

RESULTS

Population Structure

The breed composition of the animals was not recorded and, thus, could not be included in the analysis. To mitigate this problem SNP information was used to assess breed composition of the animals. Figure 1 illustrates population structure by plotting the first 2 principal components of the genotype matrix for all genotyped individuals (training and validation). Clustering based on principal component analysis did not reveal any major distinct groups. The proportion of variance explained by the first 10 principal components is illustrated in Figure 2. The first 2 principal components explained only 16 and 13% of variance, respectively. The remaining principal components explained between 12 and 6% of variance. This suggests that the population is mostly homogenous, and therefore, breed was not included as a factor in further analysis.

LD Structure

All possible SNP pairs with a distance of $\leq 2,000$ kb from the 29 caprine chromosomes produced 1,654,288 pairwise r^2. Average r^2 among syntenic markers (within a 1-kb window) as a function of marker distance is presented in Figure 2. The largest decline of LD was for distances below 100 kb. In the studied population the mean r^2 at 50 kb (distance between 2 SNP) was 0.18. Linkage disequilibrium declined from 0.14 at 100 kb to 0.09 at 1,000 kb, and 0.07 at 2000 kb. Compared with the LD expected for a population with $N_e = 110$, the r^2 in the analyzed population was lower for small SNP distances and similar at large distances above 2,000 kb (Figure 3).

Relationship Between and Within Reference and Validation Animals

The pedigree contained 30,139 individuals over 16 generations with a mean pedigree completeness of 2.52 complete generation equivalents. Mean genetic relationship was 2.32 and 4.15% within the reference and validation populations, respectively. Mean relationship between the validation and reference animals was 2.49%. Inbreeding coefficient was 0.88 and 1.35% in the reference and validation populations, respectively.
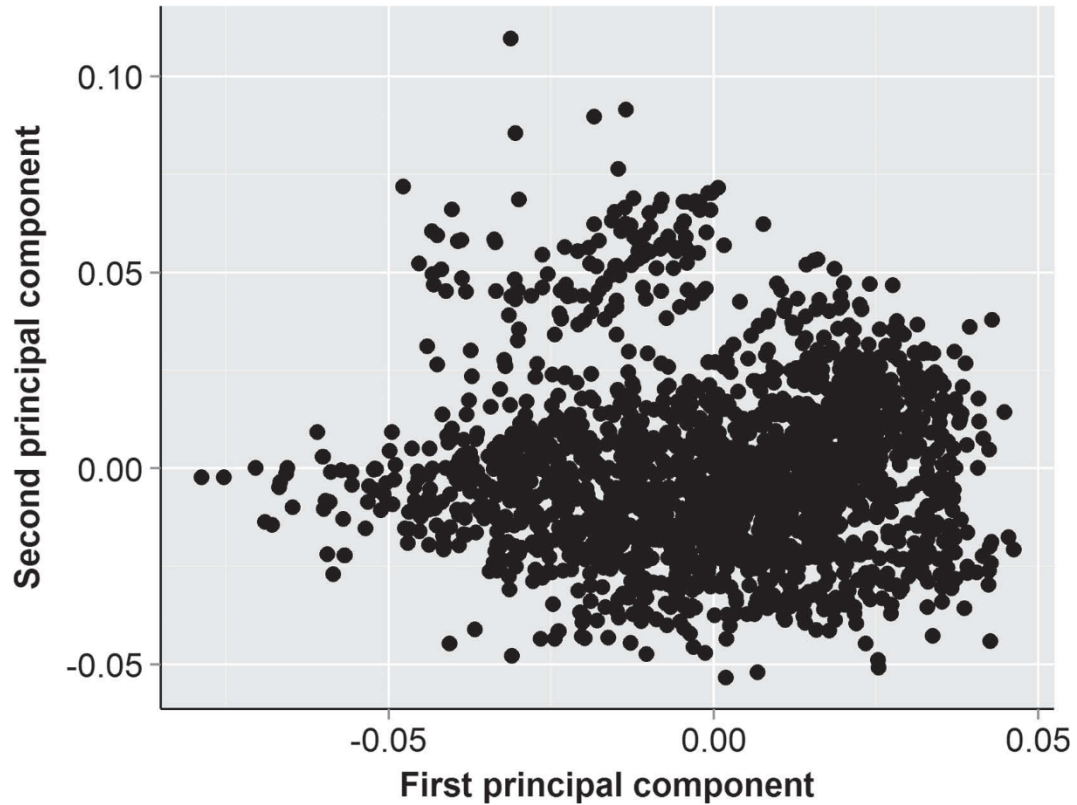
Figure 1. Plot of the first 2 principal components of the genotype matrix for the goat reference population.

Accuracy of Genomic Breeding Values

The accuracy of the genomic predictions was 0.36 and 0.61 for BLUP-SNP and HBLUP, respectively (Table 1). The gain of using SNP information expressed as the difference between accuracy of PBLUP and HBLUP was 5.2%. The BLUP-SNP not only had a considerably lower accuracy, but also resulted in a low regression
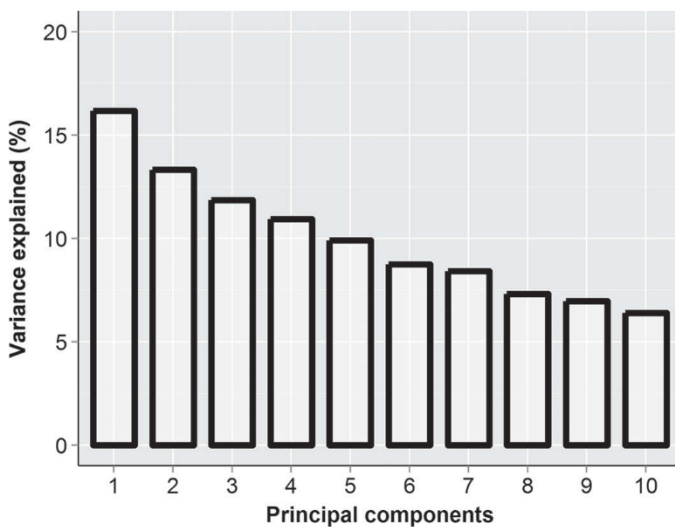

Figure 2. Percentage of variance explained by the first 10 principal components of the genotype matrix for the animals in the reference population.
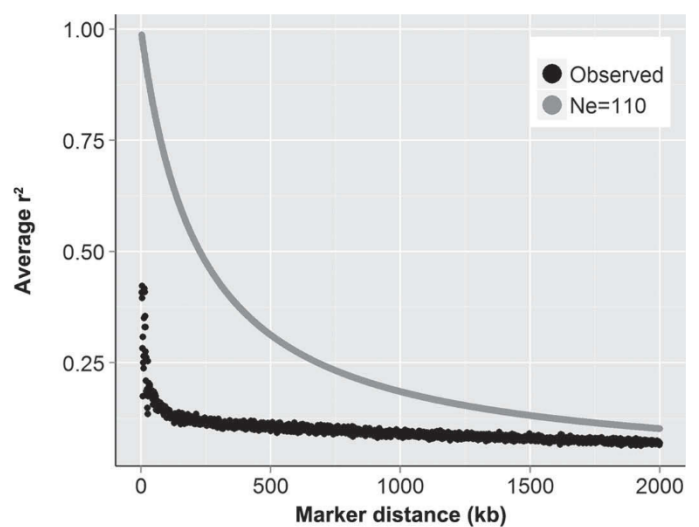

Figure 3. Linkage disequilibrium observed in British crossbred goats and expected for a population with effective population size (Ne) of 110.

Table 1. Correlations (r) and regression coefficients (b[1]) between deregressed proofs and parent average (PA), EBV from pedigree-based BLUP (PBLUP), direct genomic values from BLUP at the SNP level (BLUP-SNP), and GEBV from single-step BLUP (HBLUP) for animals in the validation population

Method	r	b[1]	Gain[1] (%)
PA	0.45	1.08	−22.4
PBLUP	0.58	1.27	0.0
BLUP-SNP	0.36	0.29	−37.9
HBLUP	0.61	0.99	5.2

[1]Gain in accuracy relative to PBLUP.

coefficient of 0.29. This suggests that DGV obtained from this method overpredicts the DRP. The GEBV from HBLUP appear to be less biased, as the regression coefficient was 0.99.

DISCUSSION

Linkage disequilibrium was analyzed because it has an influence on the accuracy of genomic selection (Habier et al., 2007). The extent of LD found in the current population is similar to that reported by Carillier et al. (2013) in the French purebred dairy goats (r^2 of 0.17 at 50 kb) but higher than the value obtained for the crossbred population (r^2 of 0.14 at 50 kb). This might indicate that the analyzed population is mostly homogeneous with respect to breed composition and can be treated as a synthetic pure breed. The average linkage disequilibrium in dairy goats here (0.18 at 50 kb) appears to be lower than that reported for dairy cattle (0.20–0.23 at 40 kb; de Roos et al., 2008; Khatkar et al., 2008; Habier et al., 2010) or pigs (0.47–0.49 at 30 kb; Uimari and Tapio, 2011). The shape of the LD decay curve was analyzed to infer information about population history with respect to the recent and historical effective population size. The LD for small SNP distances was smaller than that expected for a population with N_e of 110, which indicates that historically the analyzed population had higher levels of diversity, most likely resulting from crossbreeding of the 3 founder breeds. Linkage disequilibrium for large distances, which reflects recent history of the population (Hayes et al., 2003), was similar to that expected for a population with N_e of 110. This is in agreement with the pedigree-based estimates of effective population size, which were between 90 and 117 in the last 10 yr.

The kinship coefficient in the validation population was higher than that reported for French dairy goats (1.3%; Carillier et al., 2013). Kinship between the reference and validation in the current study was also higher than that reported in the French goats (1.3–1.4%; Carillier et al., 2013), which should have a positive effect on the accuracy of genomic breeding

values (Pszczola et al., 2012). Still, kinship found in the current study is lower than that reported in dairy cattle (Mrode et al., 2009; Habier et al., 2010). Inbreeding in the analyzed population appears to be lower compared with that reported in the French dairy goats (2.1–2.8% within reference population, and 2.1 within candidates; Carillier et al., 2013) and dairy cattle (Sørensen et al., 2005). Considering pedigree completeness in the analyzed population both the estimates of inbreeding and kinship are most likely to be underestimated (Oliehoek and Bijma, 2009). Selective genotyping based on the distribution of average daily lifetime yield and relationship between the animals was meant to maximize the predictive power of the reference population and thus maximize the accuracy in the validation population. Selected animals that were close in trait values should be distantly related, and animals that had very different trait values should be closely related. This ensures that the widest range of marker variants is captured. Two-tailed selection strategies have been recommended for small populations because they increase the accuracy of genomic predictions (Jiménez-Montero et al., 2012).

The accuracy of the genomic predictions was 0.36 and 0.61 for BLUP-SNP and HBLUP, respectively. Low accuracy of BLUP-SNP could be due to having mostly females in the validation set and therefore less precise information. Other studies with similarly small size of the reference population obtained higher accuracy ranging from 0.43 to 0.83 (Berry et al., 2009; Pintus et al., 2013). Low accuracy of BLUP-SNP in the current study can be explained by the fact that the reference population was composed mainly of females, which do not add much to the accuracy of genomic predictions. Effectively, most of the information is captured by the males present in the reference, as shown by Cooper et al. (2014), who confirmed that adding females to the reference does not contribute much to the accuracy of genomic predictions. We have verified this with an additional analysis, where instead of using the mixed reference population (males and females), only sires were left in the reference, whereas the validation population remained unchanged. This resulted in accuracy of 0.34, which was just slightly lower than the initial accuracy of 0.36 obtained with the mixed reference population. The BLUP-SNP not only had a considerably lower accuracy than HBLUP, but also resulted in a low regression coefficient of 0.29. This suggests that DGV obtained from this method overpredicts the DRP. The GEBV from HBLUP appear to be less biased, as the regression coefficient was 0.99.

The accuracy of genomic breeding values for milk yield obtained in the current project (HBLUP) was higher than that reported by Carillier et al. (2013, 2014) in the French dairy goats (0.39 and 0.43 with

a 2-step and single-step approach, respectively). However, the gain in accuracy when comparing PBLUP and HBLUP was very similar to the 5.1% reported in the French study (Carillier et al., 2013). When comparing traits with similarly high heritability in the French population (Carillier et al., 2014), such as fat and protein content (h^2 between 0.48 and 0.60), the reported accuracies (between 0.59 and 0.70) were similar to or higher than those reported in the current study. Relatively high accuracy of both PBLUP and HBLUP in the current study can be explained by high heritability (0.56) of milk yield in the evaluated population. This is much higher in comparison to the heritabilities of 0.30 observed in other goat and sheep populations (Bélichon et al., 1999; Baloche et al., 2014). Such high heritability can be explained by the fact that the data originate from only 2 farms with automated recording equipment, which reduces environmental "noise" (Mucha et al., 2014). Additionally, the relationship between the reference and validation animals was higher in the current population, which also contributed to higher accuracy of the GEBV. It is also worth mentioning that the reference population in the current study consisted of one breed in comparison to a mixture of Saanen and Alpine animals in the French population (Carillier et al., 2013).

The accuracy of single-step GEBV obtained in the current study was also higher than that based on genomic BLUP reported for the New Zealand sheep (Auvray et al., 2014), especially when compared with breeds such as Coopworth with a similar size of the reference population (1,488–1,759 for live weight traits) and heritability of the traits (0.40–0.45), which had accuracies between 0.21 and 0.46. Considerably higher accuracies were reported when using single-step methodology in the French Lacaune sheep (Baloche et al., 2014), with values ranging from 0.47 (milk yield) to 0.71 (fat percentage). In this case the reference population consisted of 1,593 genotyped rams that were supplemented with 5,904 nongenotyped rams with progeny. Also in dairy cattle, single-step analysis has been reported to increase accuracy in comparison with genomic BLUP and BLUP-SNP (Gao et al., 2012; Koivula et al., 2012). This is mainly due to the fact that nongenotyped animals with progeny can provide additional information for the estimation of genomically enhanced breeding values.

CONCLUSIONS

This study is the first attempt to create a reference population and implement genomic selection in UK dairy goats. Linkage disequilibrium found in the current population is in agreement with the levels described in other goat populations. The single-step approach resulted in higher accuracy of genomic breeding values in comparison with BLUP-SNP. This method can be recommended for breeding programs with reference populations containing a small number of sires supplemented with females. The accuracy of genomic prediction obtained in the current study was higher than that reported in other goat studies but comparable to those observed in some sheep and cattle populations.

REFERENCES

Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy Sci. 93:743–752.

Auvray, B., J. C. Mcewan, S. N. Newman, M. Lee, and K. G. Dodds. 2014. Genomic prediction of breeding values in the New Zealand sheep industry using a 50K SNP chip. J. Anim. Sci. 92:4375–4389.

Baloche, G., A. Legarra, G. Sallé, H. Larroque, J.-M. Astruc, C. Robert-Granié, and F. Barillet. 2014. Assessment of accuracy of genomic prediction for French Lacaune dairy sheep. J. Dairy Sci. 97:1107–1116.

Bélichon, S., E. Manfredi, and A. Piacère. 1999. Genetic parameters of dairy traits in the Alpine and Saanen goat breeds. Genet. Sel. Evol. 31:529–534.

Berry, D. P., F. Kearney, and B. L. Harris. 2009. Genomic selection in Ireland. Interbull Bull. 39:29–34.

Carillier, C., H. Larroque, I. Palhière, V. Clément, R. Rupp, and C. Robert-Granié. 2013. A first step toward genomic selection in the multi-breed French dairy goat population. J. Dairy Sci. 96:7294–7305.

Carillier, C., H. Larroque, and C. Robert-Granié. 2014. Comparison of joint versus purebred genomic evaluation in the French multi-breed dairy goat population. Genet. Sel. Evol. 46:67.

Christensen, O. F., and M. S. Lund. 2010. Genomic prediction when some animals are not genotyped. Genet. Sel. Evol. 42:2.

Cooper, T. A., G. R. Wiggans, and P. M. VanRaden. 2014. Including cow information in genomic prediction of Holstein dairy cattle in the US. In Proc. 10th World Congr. Genetics Appl. Livest. Prod., Vancouver, Canada. American Society of Animal Science, Champaign, IL.

de Roos, A. P. W., B. J. Hayes, R. J. Spelman, and M. E. Goddard. 2008. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. Genetics 179:1503–1512.

Fikse, W., and G. Banos. 2001. Weighting factors of sire daughter information in international genetic evaluations. J. Dairy Sci. 84:1759–1767.

Gao, H., O. F. Christensen, P. Madsen, U. S. Nielsen, Y. Zhang, M. S. Lund, and G. Su. 2012. Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. Genet. Sel. Evol. 44:8.

Gutiérrez, J. P., I. Cervantes, and F. Goyache. 2009. Improving the estimation of realized effective population sizes in farm animals. J. Anim. Breed. Genet. 126:327–332.

Habier, D., R. L. Fernando, and J. C. M. Dekkers. 2007. The impact of genetic relationship information on genome-assisted breeding values. Genetics 177:2389–2397.

Habier, D., J. Tetens, F.-R. Seefried, P. Lichtner, and G. Thaller. 2010. The impact of genetic relationship information on genomic breeding values in German Holstein cattle. Genet. Sel. Evol. 42:5.

Hayes, B. J., P. M. Visscher, H. C. McPartlan, and M. E. Goddard. 2003. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. Genome Res. 13:635–643.

Hill, W. G. 1981. Estimation of effective population size from data on linkage disequilibrium. Genet. Res. 38:209–216.

Hill, W. G., and A. Robertson. 1968. Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38:226–231.

Jiménez-Montero, J. a., O. González-Recio, and R. Alenda. 2012. Genotyping strategies for genomic selection in small dairy cattle populations. Animal 6:1216–1224.

Khatkar, M. S., F. W. Nicholas, A. R. Collins, K. R. Zenger, J. A. Cavanagh, W. Barris, R. D. Schnabel, J. F. Taylor, and H. W. Raadsma. 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. BMC Genomics 9:187.

Kirkpatrick, M., D. Lofsvold, and M. Bulmer. 1990. Analysis of the inheritance, selection and evolution of growth trajectories. Genetics 124:979–993.

Koivula, M., I. Strandén, G. Su, and E. A. Mäntysaari. 2012. Different methods to calculate genomic predictions–comparisons of BLUP at the single nucleotide polymorphism level (SNP-BLUP), BLUP at the individual level (G-BLUP), and the one-step approach (H-BLUP). J. Dairy Sci. 95:4065–4073.

Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. J. Dairy Sci. 92:4656–4663.

Lidauer, M., K. Matilainen, E. Mantysaari, and I. Straden. 2011. MiX99 Manual. MTT, Jokioinen, Finland.

Liu, Z., F. R. Seefried, F. Reinhardt, S. Rensing, G. Thaller, and R. Reents. 2011. Impacts of both reference population size and inclusion of a residual polygenic effect on the accuracy of genomic prediction. Genet. Sel. Evol. 43:19.

Maignel, L., D. Boichard, and E. Verrier. 1996. Genetic variability of French dairy breeds estimated from pedigree information. Interbull Bull. 14:49–54.

Misztal, I., A. Legarra, and I. Aguilar. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. J. Dairy Sci. 92:4648–4655.

Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 and related programs (BGF90). Pages 743–745 in 7th World Congr. Genet. Appl. Livest. Prod., Montpellier, France. INRA, Paris, France.

Montaldo, H. H., and E. Manfredi. 2002. Organisation of selection programmes for dairy goats. Pages 1–8 in Proc. 7th World Congr. Genetics Appl. Livest. Prod. Institut National de la Recherche Agronomique (INRA), Montpellier, France.

Mrode, R., J. F. Kearney, S. Biffani, M. Coffey, and F. Canavesi. 2009. Short communication: Genetic relationships between the Holstein cow populations of three European dairy countries. J. Dairy Sci. 92:5760–5764.

Mucha, S., R. Mrode, M. Coffey, and J. Conington. 2014. Estimation of genetic parameters for milk yield across lactations in mixed-breed dairy goats. J. Dairy Sci. 97:2455–2461.

Oliehoek, P. A., and P. Bijma. 2009. Effects of pedigree errors on the efficiency of conservation decisions. Genet. Sel. Evol. 41:9.

Pintus, M. a., E. L. Nicolazzi, J. B. C. H. M. Van Kaam, S. Biffani, A. Stella, G. Gaspa, C. Dimauro, and N. P. P. Macciotta. 2013. Use of different statistical models to predict direct genomic values for productive and functional traits in Italian Holsteins. J. Anim. Breed. Genet. 130:32–40.

Pszczola, M., T. Strabel, H. A. Mulder, and M. P. L. Calus. 2012. Reliability of direct genomic values for animals with different relationships within and to the reference population. J. Dairy Sci. 95:389–400.

Sørensen, A. C., M. K. Sørensen, and P. Berg. 2005. Inbreeding in Danish dairy cattle breeds. J. Dairy Sci. 88:1865–1872.

Sved, J. A. 1971. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. Theor. Popul. Biol. 2:125–141.

Tosser-Klopp, G., P. Bardou, C. Cabau, A. Eggen, T. Faraut, H. Heuven, S. Jamli, C. Klopp, C. T. Lawley, J. McEwan, P. Martin, C. Moreno, P. Mulsant, I. Nabihoudine, E. Pailhoux, I. Palhière, R. Rupp, J. Sarry, B. Sayre, A. Tircazes, J. Wang, W. Wang, T.-P. Yu, and W. Zhang. 2012. Goat genome assembly, Availability of an international 50K SNP chip and RH panel: An update of the International Goat Genome Consortium projects. Pages 1–14 in Plant Anim. Genome Conf., San Diego, CA. International Plant & Animal Genome, San Diego, CA.

Uimari, P., and M. Tapio. 2011. Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds. J. Anim. Sci. 89:609–614.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423.