



Comparison of different selection methods for improving litter size in sheep using computer simulation

Meysam Latifi (Latifi, M)¹, Amir Rashidi (Rashidi, A)¹, Rostam Abdollahi-Arpanahi (Abdollahi-Arpanahi, R)² and Mohammad Razmkabir (Razmkabir, M)¹

¹ University of Kurdistan, Faculty of Agriculture, Dept. of Animal Science, Sanandaj, Iran ² University of Tehran, College of Aburaihan, Dept. of Animal Science, Tehran, Pakdasht, Iran

Abstract

Aim of study: To assess selection methods via introgression to improve litter size in native and synthetic sheep breeds.

Area of study: Sanandaj, Kurdistan, Iran.

Material and methods: Selection approaches were performed using classical, genomic, gene-assisted classical (GasClassical) and gene-assisted genomic (GasGenomic) selection. Litter size trait with heritability of 0.1 including two chromosomes was simulated. On chromosome 1, a single QTL as the major gene was created to explain 40% of the total additive genetic variance. After simulation of a historical population, the animals from the last historical population were split into two populations. For the next 7 generations, animals were selected for favorable or unfavorable alleles to create distinct breeds of A or B, respectively. Then from the last generation, both males and females from breed B were selected to create a native population. On the other hand, males from breed A and females from breed B were mated to simulate a synthetic population. Finally, intra-population selections were carried out based on high breeding values during the last five generations.

Main results: The genetic gain in the synthetic breed was higher than that of the native breed under all selection methods. The frequencies of favorable alleles after five generations in the classical, genomic, GasClassical and GasGenomic selection approaches in the synthetic breed were 0.623, 0.730, 0.850 and 0.848, respectively.

Research highlights: Combining gene-assisted selection with classical or genomic selection has the potential to improve genetic gain and increase the frequencies of favorable allele for litter size in sheep.

Additional key words: major gene; synthetic breed; genomic evaluation

Abbreviations used: CBLUP (conventional or best linear unbiased prediction); GasClassical (gene-assisted classical selection); GasGenomic (gene-assisted genomic selection); GPBV (prediction of genomic breeding values); LD (linkage disequilibrium); LS (litter size); MCMC (Monte Carlo Markov Chain); PBV (prediction of breeding value); QTL (quantitative trait loci); SNP (single nucleotide polymorphisms); TBV (true breeding value)

Authors' contributions: Designed the study: ML and AR. Analyzed the data: ML. Wrote the paper: ML, AR. and RAA. Supervised the work: AR. All authors read the paper and approved the final manuscript.

Citation: Latifi, M; Rashidi, A; Abdollahi-Arpanahi, R; Razmkabir, M (2020). Comparison of different selection methods for improving litter size in sheep using computer simulation. Spanish Journal of Agricultural Research, Volume 18, Issue 1, e0403. <https://doi.org/10.5424/sjar/2020181-15459>

Supplementary material (Fig. S1) accompanies the paper on SJAR's website.

Received: 13 Jul 2019. **Accepted:** 02 Apr 2020.

Copyright © 2020 INIA. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC-by 4.0) License.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Correspondence should be addressed to Amir Rashidi: arashidi@uok.ac.ir

Introduction

Litter size (LS) is defined as the number of lambs born per ewe lambing. In sheep, a large variation in litter size has been observed among and within breeds. A trait such as LS can be genetically affected by many genes with small effects as well as major genes with large effects (Drouilhet *et al.*, 2009). The major genes frequency is one of the main factors in

improving breeding program efficiency for LS in sheep (Elsen *et al.*, 1994). Since the heritability of LS is low (Hanford *et al.*, 2005; Mokhtari *et al.*, 2010), the selection based on the ewe's own performance might induce a slow genetic gain. In addition, DNA testing of major genes and learning about inheritance patterns shown that major genes could lead to improvement of ovulation rates and LS in sheep (Davis, 2005).

When a major gene with desirable characteristics is detected in a particular breed, introduction and introgression of this gene into other breeds might be desirable. The cross between two or more breeds and subsequent mating among crossbred animals is considered a synthetic breed. Almost 418 synthetic breeds have been developed from the combination of two and more breeds (Rasali *et al.*, 2006). One of the main purposes of developing synthetic breeds has been to increase prolificacy and reproduction efficiency (Hulet *et al.*, 1984; Fahmy, 1990). Hence, crossbreeding between high and low prolific breeds for increasing favorable allele frequencies in a synthetic breed appear to be useful for improving LS. Previous studies have shown that breeding programs for introduction and introgression of favorable alleles such as the Booroola gene (FecB) into other sheep breeds is highly successful (Gootwine *et al.*, 2008; Mishra *et al.*, 2009).

Conventional or best linear unbiased prediction (CBLUP) selection is based on pedigree and phenotypic information (Henderson, 1975). Genomic selection, described by Meuwissen *et al.* (2001) is a method for improving quantitative traits in plant and animal breeding. This technique relies on the segmentation of the genome in thousands of intervals bracketed by contiguous markers and effects of all markers on the whole genome are estimated (Gaspa *et al.*, 2015).

Several major genes (casual mutations) have been identified for LS in sheep (Galloway *et al.*, 2000; Souza *et al.*, 2001; Hanrahan *et al.*, 2004; Martinez-Royo *et al.*, 2008; Drouilhet *et al.*, 2009). The genomic selection method (Duchemin *et al.*, 2012) has opened its way into sheep breeding programs. Combining genomic or classical selection with gene-assisted selection could be used for creating a synthetic sheep breed. Therefore, the objectives of this study were: (1) to compare classical, genomic, gene-assisted classical (GasClassical) and gene-assisted genomic (GasGenomic) selection methods in introducing a favorable gene for the creation of a synthetic breed to improve LS; and (2) to track genetic gain between a native and a synthetic sheep breed via classical and genomic selection over five generations in a sheep population by simulation.

Material and methods

Simulation

Simulations were performed (Fig. 1) based on the forward-in-time process using the QMSim software (Sargolzaei & Schenkel, 2009). The first step involved the historical population being generated and ran-

domly mated for 1000 generations with an effective population size of 1000 animals (500 males and 500 females). At the second step, two random samples consisting of 330 animals as founders (30 males and 300 females for each breed) were drawn out of the last generation of historical population to simulate two different breeds, hereafter called A and B breeds. For the next 7 generations, using a random mating design, animals from breed A and B were selected based on favorable and unfavorable alleles of the major gene, respectively. Hence, two breeds were generated and fixed for favorable and unfavorable alleles. After fixation of favorable and unfavorable alleles, the breeding values of animals for LS were predicted using a threshold model (Model 2 below). At the third step, 30 males and 300 females from the last generation of Breed B were selected based on high breeding values and treated as founders to generate the native breed. Furthermore, to generate the synthetic breed, 30 males and 300 females from the last generation of breed A and B, respectively, were selected based on high breeding values and treated as founders. The founder animals expanded over five generations and were selected and mated based on high breeding values and minimizing inbreeding. In order to have a constant population size across generations, each dam produced 5 offsprings with an equal probability of each sex. Replacement rates for males and females were 40% and 20%, respectively. In generations G1 to G5, the native and the synthetic breeds were kept at a constant size of 1500 breeding candidates per each breed and 30 males and 300 females were selected in each generation as parents of the next generation.

Genome structure

The LS trait in sheep was simulated with heritability of 0.1. The genome consisted of 2 chromosomes, each 100 cM in length. For each chromosome, 10000 markers and 100 QTLs were simulated (20000 SNPs and 200 QTLs in total). Due to the limitation in computational requirements, only two chromosomes were considered. One of the QTLs at position 25.7 cM on chromosome 1 was considered as the major gene which explained 40% of the additive genetic variance for LS trait. This fraction of genetic variance for the major gene was similar to that reported in the Lacaune sheep by Bodin *et al.* (2014). Other QTLs were randomly distributed over the chromosomes. The rest of the additive genetic variance (60%) was allocated to the remaining 199 QTLs. The QTLs effects were sampled from a gamma distribution with shape of parameter $\gamma = 0.4$. Biallelic SNP markers were randomly distrib-

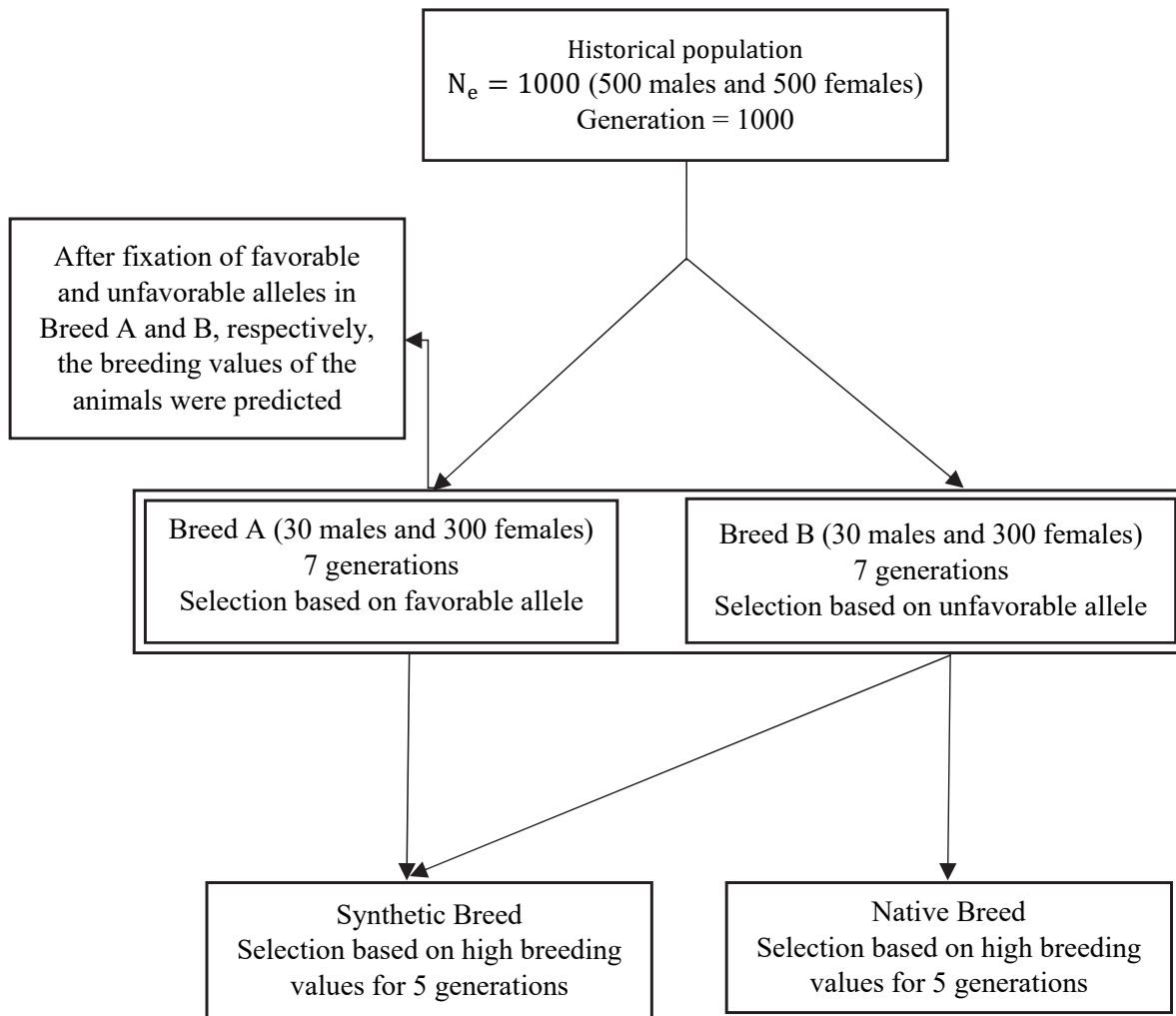


Figure 1. Structure of the simulation schemes.

uted along the two chromosomes. The mutation rate for both markers and QTL was assumed 2.5×10^{-4} per locus per generation (Esfandyari *et al.*, 2015). The true breeding value (TBV) for an individual was calculated by multiplying the genotype codes by the QTL effects. Finally, the phenotypic value of each individual i (y_i), was created by adding a normally distributed residual $e_i \sim N(\mathbf{0}, \sigma_e^2)$ to the sum over QTLs of genetic values as shown below:

$$y_i = \sum_{k=1}^m x_{ik} \alpha_k + e_i \quad (1)$$

Above, x_{ik} ($i=1, \dots, n; k=1, \dots, m$) is an element of the incidence marker matrix for additive genetic effects (α_k) and e_i is a random residual [$\mathbf{e} \sim N(\mathbf{0}, \mathbf{I}\sigma_e^2)$], where σ_e^2 is the residual variance. To convert a continuous trait into a threshold trait in each generation, 20% of the high-level phenotypes were considered to be 2, and the rest were considered as 1 to simulate a categorical

trait such as LS. All parameters used for the simulation are summarized in Table 1.

Selection

Two selection strategies were used from G0 to G5. The first was creation of the synthetic breed to introduce a favorable allele by classical, genomic, GasClassical and GasGenomic selection and the second was the selection of superior animals in the native breed based on classical and genomic selection.

Prediction of breeding values

—*Classical selection.* The Bayesian method was used for prediction of breeding values (PBVs) based on pedigree and phenotypic information. For Bayesian method, the MCMCglimm R package developed by

Table 1. Parameters of the simulation process.

Historical and current populations				
No. of generations (effective population size)	1000 (1000)			
Breed	A	B	Native	Synthetic
No. of founder sires (dams)	30 (300)	30 (300)	30 (300)	30 (300)
Criteria for selection/culling	Favorable allele/ age	Unfavorable allele/age	High breeding value/age	High breeding value/age
No. of generations for each breed	7	7	5	5
Mating system	Random	Random	Minimizes inbreeding	Minimizes inbreeding
No. of offspring per dam	5			
Replacement ratio for males (females)	0.4 (0.2)			
Sex probability for offspring	0.5			
Genome				
No. of chromosomes	2			
Total length of chromosome (M)	2			
Position of markers and QTLs	Random			
No. of QTL	200			
No. of markers	20000			
Marker and QTL mutation rate	2.5×10^{-4}			
Heritability	0.1			
Variance of QTL as the major gene (%)	40			
Position of major gene (cM)	25.7			

Hadfield (2010) was applied to analyze classical selection using BLUP methodology. The following animal model was used:

$$\mathbf{l} = \mathbf{1}\mu + \mathbf{Z}\mathbf{a} + \mathbf{e} \quad (2)$$

where \mathbf{l} is the vector of underlying latent variable for trait (one threshold and two categories of response), μ is the overall mean, $\mathbf{1}$ is a vector of ones, \mathbf{Z} is the incidence matrix relating phenotypic records to the animals, \mathbf{a} is a vector of random additive genetic effects with distribution $\mathbf{a} \sim N(\mathbf{0}, \mathbf{A}\sigma_a^2)$ and \mathbf{e} is a vector of random residuals with $\mathbf{e} \sim N(\mathbf{0}, \mathbf{I}\sigma_e^2)$, where σ_a^2 and σ_e^2 are additive genetic and residual variances, respectively. \mathbf{A} and \mathbf{I} are the additive genetic relationship and the identity matrices, respectively. A total of 300,000 rounds of Gibbs sampler were considered and the first 30,000 rounds were discarded as a burn-in period. The thinning interval was set to 100 cycles. In each generation, pedigree and phenotypic records of new generation were added to the recent pedigree and data file. Therefore, PBVs for every round of selection were re-calculated and the predicted breeding values were used for selection.

—*Genomic selection.* The Bayes B method was used to estimate marker effects with Monte Carlo Markov Chain (MCMC). In Bayes B, it is assumed that each SNP has either an effect of zero or non-zero with probabilities of π and $1-\pi$, respectively, and for those with

non-zero effect, it is assumed that each SNP has a different variance. The BGLR (Bayesian Generalized Linear Regression) package of R software (Perez & Campos, 2014) was used for prediction of genomic breeding values (GPBVs). The Gibbs sampler was run for 30,000 rounds with a 3,000 burn-in period. The thinning interval was set to 10 cycles. The following model was used to estimate the genomic breeding values:

$$\mathbf{l} = \mu + \sum_{j=1}^n \mathbf{x}_{ij} \mathbf{m}_j + \mathbf{e} \quad (3)$$

where \mathbf{l} is the vector of underlying latent variable for trait (one threshold and two categories of response), μ is the overall mean, $\mathbf{1}$ is a vector of ones, \mathbf{x}_{ij} is an incidence matrix for marker j and individual i , \mathbf{m}_j is a random effect for marker j and \mathbf{e} is the random residual error with distribution of $\mathbf{e} \sim N(\mathbf{0}, \mathbf{I}\sigma_e^2)$. In each generation, the effects of markers were estimated using female phenotypes. GPBVs for male and females were calculated as the sum of all marker effects according to their marker genotypes:

$$\text{GPBV}_i = \sum_{j=1}^n \mathbf{x}_{ij} \widehat{\mathbf{m}}_j \quad (4)$$

where \mathbf{x}_{ij} is the genotype of animal i at locus j , and $\widehat{\mathbf{m}}_j$ is the estimated substitution effect of marker j .

Combining gene-assisted selection with classical and genomic selection

In the present study, selection and culling of animals in each generation was based on (G)PBVs and age, respectively. For combining gene-assisted selection with classical and genomic selection (GasClassical and GasGenomic), we developed an alternative approach that allows selection of animals with favorable allele and the highest (G)PBVs as the parents of the next generation. According to this approach: 1) the PBV or GPBV for each animal was predicted using model 2 and 3, respectively; 2) homozygosity or heterozygosity of each animal was determined based on the favorable allele, and 3) based on the equation below, the weighted predicted breeding value of each animal was obtained:

$$(G)PBV_w = (G)PBV_p + NFA \quad (5)$$

where $(G)PBV_w$ is the weighted PBV or GPBV, $(G)PBV_p$ is the predicted PBV or GPBV for each animal based on model 2 and 3, respectively, and NFA is number of favorable alleles in each animal. If an animal had 2, 1 or zero copies of the favorable allele, NFA was 2, 1 and 0, respectively (Fig. S1 [suppl.]).

Linkage disequilibrium (LD)

The LD in the native and the synthetic breeds from G0 to G5 was assessed by r^2 among all pairs of markers using QMSim software (Sargolzaei & Schenkel, 2009):

$$r^2 = \frac{D^2}{\text{freq}(A) \times \text{freq}(a) \times \text{freq}(B) \times \text{freq}(b)} \quad (6)$$

where $D = \text{freq}(AB) - (\text{freq}(A) \times \text{freq}(B))$; $\text{freq}(AB)$ is frequency of observed haplotype; and $\text{freq}(A)$ and $\text{freq}(B)$ are frequencies of alleles A and B, respectively.

Evaluation of scenarios

The genetic gain, frequency of favorable allele and mean of inbreeding coefficient were monitored across generations for all selection schemes. The genetic gain per generation was computed as the average TBV over time. In addition, the accuracy of prediction for both classical and genomic selection methods was obtained as the Pearson correlation between the TBVs and (G)PBVs in each generation. A total of 10 replicates were produced for each scenario and selection method. The 10 replicates were simulated separately, but the initial generation (G0) was identical for all scenarios within each replicate.

Results

Native breed

The results obtained using simulations performed for the native breed are summarized in Table 2. The results obtained under classical and genomic selection methods showed that the trend of true breeding values increased across generations. The mean of true breeding values for both methods were -0.22 at G0 and increased to 0.264 and 0.621 at G5 for classical and genomic selection, respectively. Furthermore, the mean of true breeding values obtained under genomic selec-

Table 2. Mean of true breeding values, accuracy of prediction (r) and mean of inbreeding coefficient (F) in each generation based on classical and genomic selection in the native breed.

Selection method	Gen	Mean _{TBV}	$r_{(PBV,TBV)}$	F
Classical	0	-0.220 ± 0.155	0.253 ± 0.056	0.000 ± 0.000
	1	-0.146 ± 0.154	0.309 ± 0.064	0.000 ± 0.000
	2	-0.064 ± 0.168	0.374 ± 0.066	0.000 ± 0.000
	3	0.042 ± 0.187	0.370 ± 0.102	0.000 ± 0.000
	4	0.146 ± 0.206	0.375 ± 0.055	0.004 ± 0.004
Genomic	0	-0.220 ± 0.155	0.669 ± 0.077	0.000 ± 0.000
	1	-0.145 ± 0.155	0.696 ± 0.067	0.000 ± 0.000
	2	0.017 ± 0.170	0.722 ± 0.061	0.000 ± 0.000
	3	0.209 ± 0.197	0.763 ± 0.047	0.000 ± 0.000
	4	0.416 ± 0.215	0.779 ± 0.047	0.000 ± 0.000
	5	0.621 ± 0.228	0.738 ± 0.046	0.002 ± 0.001

tion was 74% higher than that of the classical selection. Thus, the accuracy of prediction using genomic selection was higher compared to the classical selection. The highest accuracies for classical and genomic selection was obtained in G4 and afterwards, decreased in G5 (Table 2). At G5, the mean of inbreeding coefficients for the native breed under classical and genomic selection were 0.016 and 0.002, respectively.

Synthetic breed

The results obtained using simulations performed for the synthetic breed are summarized in Table 3. As expected, regardless of selection methods, mean of true breeding values increased across generations. Genetic gain obtained under GasGenomic (1.347) was higher than that of obtained using the GasClassical selection (1.007), and the value of 1.319 obtained for Genomic selection was higher than that of obtained for the classical selection method (0.884). Combining gene assisted-selection with classical and genomic selection led to higher genetic gain. Consequently, genetic gain obtained by GasGenomic and GasClassical selection methods was greater than those of genomic and clas-

sical selection (2% and 16% higher genetic gain from G0 to G5, respectively). As shown in Table 3, genomic and GasGenomic selection resulted in better accuracies of PBV prediction in comparison with classical and GasClassical selection methods. The highest accuracy for classical, genomic, GasClassical and GasGenomic selection in the synthetic breed were obtained at G4 which decreased slightly at G5 (Table 3). For the synthetic breed, a similar pattern to that of the native breed was observed for the mean of inbreeding coefficient. At G5, mean of inbreeding coefficients were 0.011, 0.002, 0.008 and 0.002 for classical, genomic, GasClassical and GasGenomic genomic selection, respectively. The mean of inbreeding coefficient in classical and GasClassical selection varied more compared to genomic and GasGenomic selection.

The frequency of the favorable allele in each generation under classical, genomic, GasClassical and GasGenomic selection in the synthetic breed for 10 replicates is shown in Fig. 2. The frequency of the favorable allele started with an initial value of 0.091 at G0 and reached to 0.623, 0.730, 0.850 and 0.848 at G5 for classical, genomic, GasClassical and GasGenomic selection methods, respectively (Table 3). The frequency of favorable allele under classical and genomic

Table 3. Mean of true breeding values, accuracy of prediction (r) and mean of inbreeding coefficient (F) in each generation based on classical, genomic, GasClassical and GasGenomic selection methods in the synthetic breed.

Selection method	Gen	Mean _{TBV}	$r_{(PBV,TBV)}$	Frequency ¹	F
Classical	0	-0.142 ± 0.151	0.263 ± 0.034	0.091 ± 0.000	0.000 ± 0.000
	1	0.282 ± 0.154	0.298 ± 0.066	0.500 ± 0.000	0.000 ± 0.000
	2	0.332 ± 0.149	0.361 ± 0.095	0.451 ± 0.005	0.000 ± 0.000
	3	0.384 ± 0.163	0.506 ± 0.057	0.429 ± 0.024	0.000 ± 0.000
	4	0.530 ± 0.168	0.514 ± 0.059	0.491 ± 0.051	0.002 ± 0.003
Genomic	0	-0.142 ± 0.151	0.581 ± 0.089	0.091 ± 0.000	0.000 ± 0.000
	1	0.281 ± 0.153	0.646 ± 0.056	0.500 ± 0.000	0.000 ± 0.000
	2	0.399 ± 0.144	0.720 ± 0.052	0.449 ± 0.003	0.000 ± 0.000
	3	0.579 ± 0.177	0.799 ± 0.043	0.442 ± 0.003	0.000 ± 0.000
	4	0.852 ± 0.198	0.834 ± 0.045	0.537 ± 0.004	0.001 ± 0.001
GasClassical	0	-0.142 ± 0.151	0.251 ± 0.044	0.091 ± 0.000	0.000 ± 0.000
	1	0.279 ± 0.152	0.303 ± 0.084	0.500 ± 0.000	0.000 ± 0.000
	2	0.304 ± 0.168	0.386 ± 0.053	0.449 ± 0.005	0.000 ± 0.000
	3	0.460 ± 0.184	0.510 ± 0.060	0.550 ± 0.033	0.000 ± 0.000
	4	0.666 ± 0.220	0.516 ± 0.049	0.700 ± 0.053	0.002 ± 0.003
GasGenomic	0	-0.142 ± 0.151	0.635 ± 0.089	0.091 ± 0.000	0.000 ± 0.000
	1	0.281 ± 0.151	0.672 ± 0.068	0.500 ± 0.000	0.000 ± 0.000
	2	0.407 ± 0.145	0.741 ± 0.052	0.449 ± 0.004	0.000 ± 0.000
	3	0.637 ± 0.150	0.784 ± 0.050	0.538 ± 0.027	0.000 ± 0.000
	4	0.915 ± 0.159	0.805 ± 0.041	0.698 ± 0.003	0.001 ± 0.000
	5	1.205 ± 0.184	0.793 ± 0.052	0.848 ± 0.005	0.002 ± 0.003

¹Frequencies of favorable allele

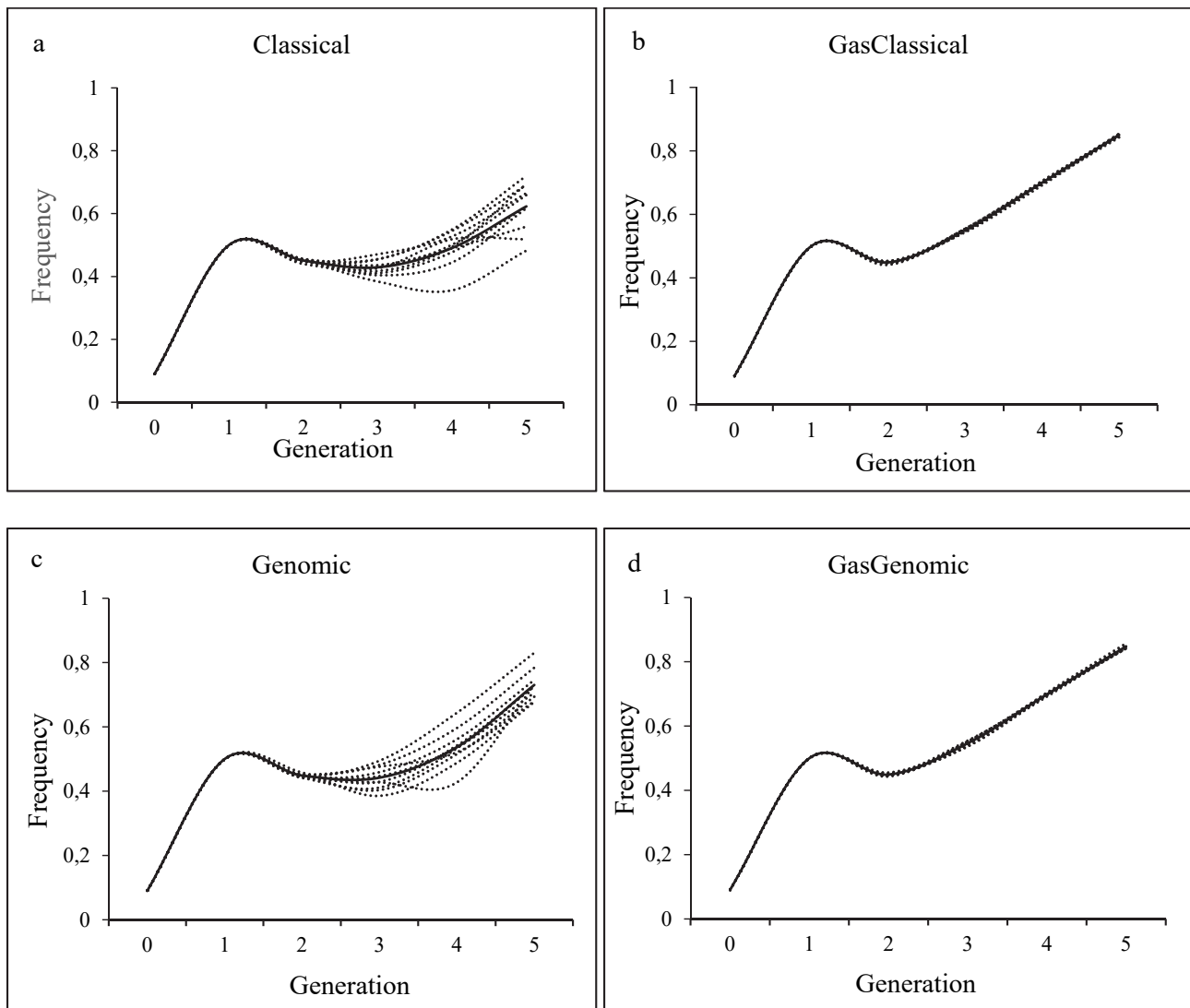


Figure 2. Pattern of allele frequencies of the favorable allele for the synthetic breed under a) Classical b) Gene-assisted classical (GasClassical), c) Genomic and d) Gene-assisted genomic (GasGenomic) selection across 10 replicates (the black line is the mean of 10 replicates).

selection (Figs. 2a and 2c) showed high fluctuations whereas when applying GasClassical and GasGenomic selections, no fluctuation was observed over 10 replicates (Figs. 2b and 2d).

Comparison between the native and the synthetic breeds

Mean of true breeding values for the native and the synthetic breeds obtained using different methods are shown in Fig. 3. As illustrated, genetic gains for LS in the synthetic breed under classical or genomic selections were higher than selections in the native breed. Genetic gain in the synthetic breed with GasClassical (117%) and classical (82%) selection methods was greater than that in the native breed under classical

selection. Moreover, genetic gain in the synthetic breed under GasGenomic (60%) and Genomic (57%) selection methods was higher than that obtained for the native breed using the genomic selection method.

Linkage disequilibrium in the native and the synthetic breeds

The average LD decay obtained using genomic selection method for all possible pairs of markers in the native and the synthetic breeds are shown in Fig. 4. As illustrated, the maximum average of r^2 for the native and the synthetic breeds (genomic and GasGenomic selection) at distance 0.0 to 50 kb at G0 were 0.19 and 0.18, respectively. However, the minimum average of r^2 at a distance of 400 to 500 kb for the native and the

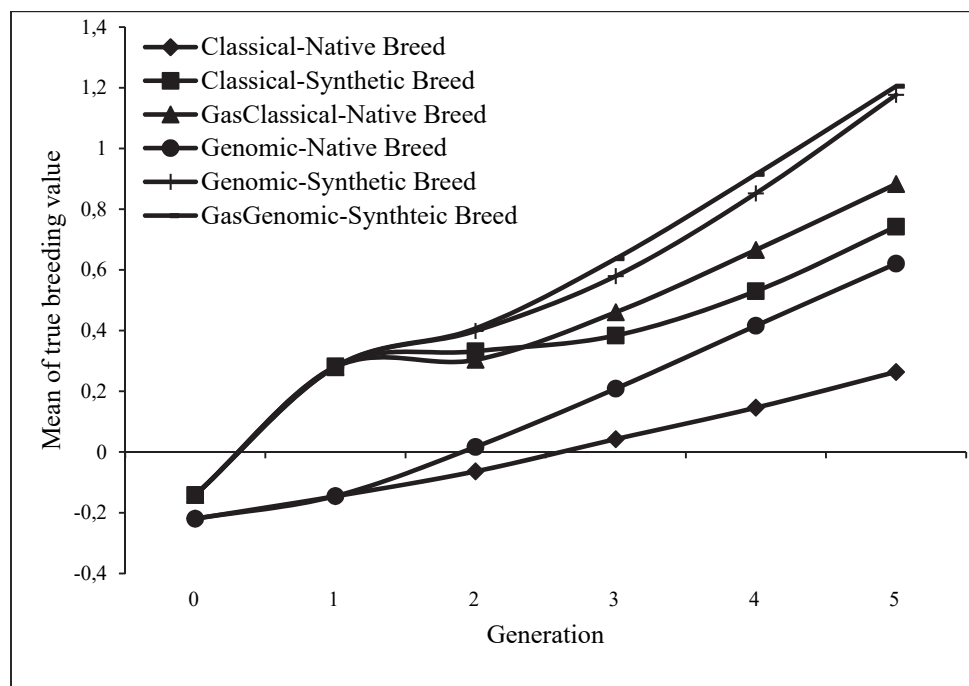


Figure 3. Mean of breeding values for the native and the synthetic breeds with Genomic and Classic selection methods across 10 replicates.

synthetic breeds (genomic and GasGenomic selection) at G0 were 0.11 and 0.10, respectively. For the native and the synthetic breeds (genomic and GasGenomic selection) at G5, the average maximum values of r^2 for distance of 0.0 to 50 kb were 0.30 and 0.26, respectively. In addition, the average minimum values of r^2 for the native and the synthetic breeds (genomic and GasGenomic selection) from 400 to 500 kb at G5 were 0.23 and 0.18, respectively. For the native and the synthetic breeds, by increasing marker pair distance, the values of r^2 decreased at G0 and G5 (Fig. 4). The native breed showed higher levels of LD in comparison with the synthetic breed at G0 and G5.

Discussion

The results of this study showed that employing genomic and GasGenomic selection methods rather than Classical and GasClassical selection lead to increase in genetic gain in the native and the synthetic breeds. Meuwissen *et al.* (2001) showed that using genomic selection method resulted in an increased accuracy of selection for traits with low heritability and female sex-limited such as LS. A reason for the high accuracy of prediction in genomic selection could be due to use of the marker information to capture the Mendelian sampling terms (Daetwyler *et al.*, 2007). The results reported by other researchers indicated that the accuracy of prediction depends on the extent of

linkage disequilibrium, the density of markers, statistical methods, heritability of trait and selection design (Wang *et al.*, 2013; Gowane *et al.*, 2018; Karimi *et al.*, 2019). Using simulated data, higher accuracies were reported for genomic selection compared to classical selection in dairy cattle (Gaspa *et al.*, 2015), swine (Putz *et al.*, 2018) and American mink (Karimi *et al.*, 2019). In the current study, the accuracies in all selection methods increased from G0 to G4. Afterward, a decrease was observed in G5, because individuals in this generation did not have any recorded offspring.

When the combination of gene-assisted selection with classical and genomic selection was carried out for improvement of a trait by introducing a major gene, genetic gain was higher than when classical and genomic selection were used alone. These findings indicate that if the major gene is controlling a large proportion of genetic variance in a population, the combination of classical and genomic selection with gene-assisted selection leads to an increase in genetic gain. Pedersen *et al.* (2009) indicated that using genotype information increased within-family selection and lead to an increase in genetic gain as the accuracy of breeding values increased. Generally, this result was obtained using classical and genomic selection methods in the synthetic breed which resulted in higher genetic gain compared to the native breed. High genetic gain for the synthetic breed led to an increase of heterozygosity at the QTL level and accuracy of selection method (Ødegard *et al.*, 2009b). Our results indicated that two breeds (A and B)

were separated based on the frequencies of the major gene after 7 generations. If the two breeds were considerably divergent, the relative advantage of crossing could be higher.

In both the native and the synthetic breeds, the rate of increase of inbreeding coefficient was lower in genomic selection compared to classical selection. Dae-twyler *et al.* (2007) stated that the low inbreeding rate in genomic selection compared to the classical selection method is due to an increase in prediction accuracy of the Mendelian sampling. In our simulation, the mating system was based on optimized mating design. Thus, pairs of mates were chosen based on minimizing genetic relationship to create the next generation (Sargolzaei & Schenkel, 2009). Hence, the rate of inbreeding coefficient in both genomic and classical selection was lower than the results reported in previous research (Ødegard *et al.*, 2009b; Gaspa *et al.*, 2015).

The frequency of favorable allele after five generations in classical, genomic, GasClassical and GasGenomic selection methods in the synthetic breed were 0.62, 0.73, 0.85 and 0.85, respectively (Table 3). Results obtained in our study showed that the frequency of favorable allele in genomic selection increased by more than 20% compared to the classical selection. These findings indicate that genomic selection is more effective in increasing the frequency of favorable allele than the classical selection without any knowledge of the major gene's position. Under GasClassical and GasGenomic selection methods, the frequency of the favorable allele was never declined in comparison with classical and genomic selections (Fig. 2). The frequency of the favorable allele was the same in the GasClassical and GasGenomic, but the rate of genetic gain in the GasGenomic was higher than that of the

GasClassical selection method (31%). These results are in agreement with Ødegard *et al.* (2009a), who stated that combining genomic selection with gene-assisted selection for the target QTL acted as an additional action against decline of the target QTL and gave surpass genetic gain.

The study presented here about the extent of the LD can be used to interpret the difference observed among the breeds. Observed LD is a function of the recombination rate between loci within a breed and the selection performed for specific quantitative or qualitative traits of interest (Prieur *et al.*, 2017). Each of the breeds showed the decrease in r^2 with increased distance between markers. In the present study, the levels of LD for the native and the synthetic breeds at a distance from 0.0 to 50 kb at G0 and G5 was in agreement with the results reported by Kijas *et al.* (2014) at a distance of 10 kb. For the native and synthetic breeds under selection, the level of LD increased over the generations. The high level of LD showed a high level of selection intensity over generations for these breeds. The extent of LD could be increased by several factors, including inbreeding, population structure, and bottlenecks (Pritchard & Przeworski, 2001). In this study, the extent of LD was found to be remarkably different between the native and the synthetic breeds (Fig. 4). This could be due to the creation of the synthetic breed from two cross-breeds (Breed A and B). Toosi *et al.* (2010) mentioned that individual animals are less related to each other in a crossbreed population; therefore, LD haplotypes in crossbred populations are narrower than in a purebred. Results of the current study showed that the LD decay was different between breeds, and similar results in sheep were reported by Al-Mamun *et al.* (2015).

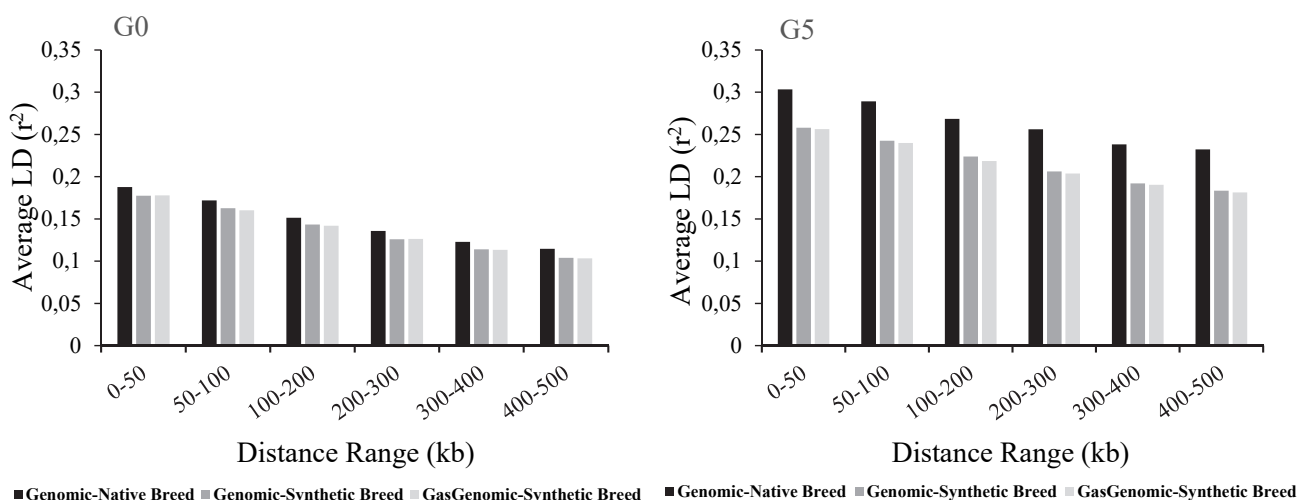


Figure 4. The average LD decay for the native and the synthetic breeds (genomic and GasGenomic selection) at generation zero (G0) and five (G5) across 10 replicates.

The results of this study confirmed that reproduction and sex-limited traits such as LS in sheep can be improved by using genomic selection methods. The results also demonstrated that inclusion of information about known causal mutations into genomic and classical selection methods can lead to increase in frequency of favorable allele and subsequently resulted in higher genetic gain in the synthetic breed.

Acknowledgments

The authors are sincerely grateful to the Dr. M Hossein Yazdi (Head of Technical Analysis Division, Akvaforsk Genetics Center AS, N-6600 Sunndalsøra, Norway) and Dr. Mehdi Bohlouli and Carsten Scheper (Institute of Animal Breeding and Genetics, Justus-Liebig-University Gießen, 35390 Gießen, Germany) for the valuable discussions.

References

- Al-Mamun HA, Clark SA, Kwan P, Gondro C, 2015. Genome-wide linkage disequilibrium and genetic diversity in five populations of Australian domestic sheep. *Genet Sel Evol* 47: 1-14. <https://doi.org/10.1186/s12711-015-0169-6>
- Bodin L, Martin PM, Raoul J, 2014. Effects of the FecL major gene on mean and variance of litter size in the Lacaune meat sheep population. *Proc 10th World Conf on genetics applied to livestock production (WCGALP)*, Vancouver (Canada), Aug 17-22.
- Daetwyler HD, Villanueva B, Bijma P, Woolliams JA, 2007. Inbreeding in genome-wide selection. *J Anim Breed Genet* 124: 369-376. <https://doi.org/10.1111/j.1439-0388.2007.00693.x>
- Davis GH, 2005. Major genes affecting ovulation rate in sheep. *Genet Sel Evol* 37: 11-23. <https://doi.org/10.1186/1297-9686-37-S1-S11>
- Drouilhet L, Lecerf F, Bodin L, Fabre S, Mulsant P, 2009. Fine mapping of the FecL locus influencing prolificacy in Lacaune sheep. *Anim Genet* 40: 804-812. <https://doi.org/10.1111/j.1365-2052.2009.01919.x>
- Duchemin S, Colombani C, Legarra A, Baloché G, Larroque H, Astruc JM, Barillet F, Robert-Granié C, Manfredi E, 2012. Genomic selection in the French Lacaune dairy sheep breed. *J Dairy Sci* 95: 2723-2733. <https://doi.org/10.3168/jds.2011-4980>
- Elsen JM, Bodin L, Francois D, Poivey JP, Teyssier J, 1994. Genetic improvement of litter size in sheep. *Proc 5th World Conf on genetics applied to livestock production (WCGALP)*. Guelph, Ontario, (Canada), pp: 237-244.
- Esfandyari H, Sorensen AC, Bijma P, 2015. Maximizing crossbred performance through purebred genomic selection. *Genet Sel Evol* 47: 16. <https://doi.org/10.1186/s12711-015-0099-3>
- Fahmy MH, 1990. Development of DLS breed of sheep: genetic and phenotypic parameters of date of lambing and litter size. *Can J Anim Sci* 70: 771-778. <https://doi.org/10.4141/cjas90-096>
- Galloway SM, McNatty KP, Cambridge LM, Laitinen MP, Juengel JL, Jokiranta TS, McLaren RJ, Luiro K, Dodds KG, Montgomery G, 2000. Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner. *Nat Genet* 25: 279-283. <https://doi.org/10.1038/77033>
- Gaspa G, Veerkamp RF, Calu MPL, Windig JJ, 2015. Assessment of genomic selection for introgression of polledness into Holstein Friesian cattle by simulation. *Livest Sci* 179: 86-95. <https://doi.org/10.1016/j.livsci.2015.05.020>
- Gootwine E, Reicher S, Rozov A, 2008. Prolificacy and lamb survival at birth in Awassi and Assaf sheep carrying the FecB (Booroola) mutation. *Anim Reprod Sci* 108: 402-411. <https://doi.org/10.1016/j.anireprosci.2007.09.009>
- Gowane GR, Lee SH, Clark S, Moghaddar N, Al-Mamun HA, van der Werf JHJ. 2018. Effect of selection on bias and accuracy in genomic prediction of breeding values. *bioRxiv* 298042. <https://doi.org/10.1101/298042>
- Hadfield JD, 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J Stat Softw* 33: 1-22.
- Hanford KJ, Van Vleck LD, Snowden GD, 2005. Estimates of genetic parameters and genetic change for reproduction, weight, and wool characteristics of Rambouillet sheep. *Small Rumin Res* 57: 175-186. <https://doi.org/10.1016/j.smallrumres.2004.07.003>
- Hanrahan JP, Gregan SM, Mulsant P, Mullen M, Davis GH, Powell R, Galloway SM, 2004. Mutations in the genes for oocyte-derived growth factors GDF9 and BMP15 are associated with both increased ovulation rate and sterility in Cambridge and Belclare sheep (*Ovis aries*). *Biol Reprod* 70: 900-909. <https://doi.org/10.1095/biol-reprod.103.023093>
- Henderson CR, 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31: 423-447. <https://doi.org/10.2307/2529430>
- Hulet CV, Ercanbrack SK, Knight AD, 1984. Development of the Polypay breed of sheep. *J Anim Sci* 58: 15-24. <https://doi.org/10.2527/jas1984.58115x>
- Karimi K, Sargolzaei M, Plastow GS, Wang Z, Miar Y, 2019. Opportunities for genomic selection in American mink: A simulation study. *PLOS ONE* 14: e0213873. <https://doi.org/10.1371/journal.pone.0213873>
- Kijas JW, Porto-Neto L, Dominik S, Reverter A, Bunch R, McCulloch R, Hayes BJ, Brauning R, McEwan J, 2014. Linkage disequilibrium over short physical distances measured in sheep using a high-density SNP chip. *Anim Genet* 45: 754-757. <https://doi.org/10.1111/age.12197>
- Martinez-Royo A, Jurado JJ, Smulders JP, Marti JI, Alabart JL, Roche A, Fantova E, Bodin L, Mulsant P, Serrano M, Folch J, Calvo JH, 2008. A deletion in the bone morphogenetic protein 15 gene causes sterility and increased prolificacy in Rasa Aragonesa sheep. *Anim Genet* 39: 294-297. <https://doi.org/10.1111/j.1365-2052.2008.01707.x>

- Meuwissen THE, Hayes B, Goddard ME, 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157: 1819-1829.
- Mishra AK, Arora AL, Kumar S, Prince LLL, 2009. Studies on effect of Booroola (FecB) genotype on lifetime ewe's productivity efficiency, litter size and number of weaned lambs in Garole × Malpura sheep. *Anim Reprod Sci* 113: 293-298. <https://doi.org/10.1016/j.anireprosci.2008.06.002>
- Mokhtari MS, Rashidi A, Mohamadi Y, 2010. Estimation of genetic parameters for post-weaning traits of Kermani sheep. *J Anim Sci* 80: 22-27. <https://doi.org/10.1016/j.smallrumres.2008.08.002>
- Ødegård J, Sonesson Ak, Yazdi MH, Meuwissen THE, 2009a. Introgression of a major QTL from an inferior into a superior population using genomic selection. *Genet Sel Evol* 41: 1-10. <https://doi.org/10.1186/1297-9686-41-38>
- Ødegård J, Yazdi MH, Sonesson AK, Meuwissen THE, 2009b. Incorporating desirable genetic characteristics from an inferior into a superior population using genomic selection. *Genetics* 181: 737-745. <https://doi.org/10.1534/genetics.108.098160>
- Pedersen LD, Sørensen AC, Berg P, 2009. Marker-assisted selection can reduce true as well as pedigree-estimated inbreeding. *J Dairy Sci* 92: 2214-2223. <https://doi.org/10.3168/jds.2008-1616>
- Perez P, de los Campos G, 2014. Genome-wide regression and prediction with the BGLR statistical package. *Genetics* 198: 483-495. <https://doi.org/10.1534/genetics.114.164442>
- Prieur V, Clarke SM, Brito LF, McEwan JC, Lee MA, Brauning R, Dodds KG, Auvray B, 2017. Estimation of linkage disequilibrium and effective population size in New Zealand sheep using three different methods to create genetic maps. *BMC Genet* 18: 68. <https://doi.org/10.1186/s12863-017-0534-2>
- Pritchard JK, Przeworski M, 2001. Linkage disequilibrium in humans: models and data. *Am J Hum Genet* 69:1-14. <https://doi.org/10.1086/321275>
- Putz AM, Tiezzi F, Maltecca C, Gray KA, Knauer MT, 2018. A comparison of accuracy validation methods for genomic and pedigree-based predictions of swine litter size traits using Large White and simulated data. *J Anim Breed Genet* 135: 5-13. <https://doi.org/10.1111/jbg.12302>
- Rasali DP, Hrestha JNB, Crow GH, 2006. Development of composite sheep breeds in the world: A review. *Can J Anim Sci* 86: 1-24.
- Sargolzaei M, Schenkel FS, 2009. QMSim: a largescale genome simulator for livestock. *Bioinformatics* 25: 680-681. <https://doi.org/10.1093/bioinformatics/btp045>
- Souza C, MacDougall C, Campbell B, McNeilly A, Baird D, 2001. The Booroola (FecB) phenotype is associated with a mutation in the bone morphogenetic receptor type 1 B (BMPRI1B) gene. *J Endocrinol* 169: R1-R6. <https://doi.org/10.1677/joe.0.169r001>
- Toosi A, Fernando RL, Dekkers JCM, 2010. Genomic selection in admixed and crossbred populations. *J Anim Sci* 88: 32-46. <https://doi.org/10.2527/jas.2009-1975>
- Wang CL, Ding XD, Wang JY, Liu JF, Fu WX, Zhang Z, Yin ZJ, Zhang, 2013. Bayesian methods for estimating GEBVs of threshold traits. *Heredity* 110: 213-219. <https://doi.org/10.1038/hdy.2012.65>