Livestock Studies, 2022; 62 (2): 58-64
http://doi.org/10.46897/livestockstudies.1209084

LIVESTOCK STUDIES

RESEARCH PAPER

Effects of Different Methods and Relationship Matrices on Reliabilities of Genomic Selection in Dairy Cattle

Anıl KASAKOLU[1,*], Seyrani KONCAGÜL[1]

[1]Ankara University Faculty of Agriculture, Department of Animal Science, Ankara, Türkiye
*Corresponding Author

Article History
Received: 11 November 2022
Accepted: 23 November 2022
First Online: 23 November 2022

Corresponding Author*
Tel.: +90 505 015 91 20
E-mail: kasakoluanil1@gmail.com

Keywords
Dairy cattle
Genomic relationship matrix
ssGBLUP
Reliability

Abstract
Since genomic prediction is widely used in dairy cattle, we aimed to evaluate the performance of pedigree based (ABLUP), SNP based (GBLUP) and single-step GBLUP (ss-GBLUP) methods with different sets of information in terms of reliability of genomic prediction. Four different methods were evaluated: (Method 1) ABLUP with all available phenotypes and pedigree; (Method 2) GBLUP with SNP genotypes and phenotypes of genotyped cows; (Method 3) single-step GBLUP with SNP genotypes, phenotypes of genotyped cows and all pedigree and (Method 4) single-step GBLUP with SNP genotypes, all phenotypes of both genotyped and nongenotyped cows and all pedigree. SNP based methods also used different genomic relationship matrices (GRMs) formed by different approaches: ***vanRaden***, ***Astle***, ***Yang*** and ***Endelman***. The simulated dataset replicates a common dairy cattle population.
A significant increase in reliability of prediction was observed in ss-GBLUP with all phenotypes and pedigree beside genotyped cows. This increase was apparent for both first lactation milk yield (LMY) and milk fat percentage (Fat%). Combining all available information with ss-GBLUP gave about 1.6 and 1.2 times higher reliabilities for LMY and Fat%, respectively, compared to those obtained from the other three methods.

Introduction

Livestock breeding, with the increase of today's nutritional problems, has brought with it the need for rapid progress in production processes. The genetic gain of the population subject to selection must be achieved in a shorter time than with conventional methods such as progeny testing (PT) in dairy cattle. Genomic selection (GS) may shorten the generation interval (GI), an important factor in terms of genetic gain for a given time and GI in cattle is equivalent to an average of 5.5 years in classical progeny testing program. However, this period can be reduced by almost half with GS methods (Schaeffer 2006).

Genomic selection has been used since the beginning of the 21st century, but it can be observed that the number of genotyped animals has become more and more economical in the last decade in the countries that have access to the genotyping process (Wiggans et al. 2017). Although this is a preferred point, computational times seem to be a more restrictive problem in terms of efficiency than the times in the earlier methods. In other words, the technological increase in the available information has also created the need for advanced technology in calculations (computations/computational hardwares)(Tsurata et al. 2021).

One of the problems encountered in GS in some countries is that when the number of SNP is less than the number of genotyped animals, the genomic relationship matrices (GRM) turn into a singular (non-invertable) structure and they have to be combined with the A matrix and included in the analysis (Misztal et al. 2020). Genomic selection has come to a point where it can be preferred despite all the difficulties in the field of animal breeding, since it has the convenience of making a decision by taking blood even from the fetus or embrios.

Several methods have been developed for the formation of GRMs, but GRMs have been obtained mainly by using SNP information. In the context of the Human Genome Project (IHGSC 2001). Due to the fact that these SNPs are scattered throughout the genome, GS has been used in breeding value estimation as a useful method in today's conditions, as it is consistent with the

basic theories of quantitative genetics (Meuwissen et al. 2001).

This study aimed to investigate the effects of four methods with GRMs formed by different approaches on the reliabilities of genomic estimated breeding values (GEBV).

Materials and Methods

Data sets

The data considered in this study were simulated for 3 discrete generations. The structure of data is given in Table 1 and has been generated according to the optimized scenario for the first lactation of Holstein dairy cattle by the steps below and the simulation study was performed in R environment (R Core Team 2022). The assumptions were:

- First lactation milk yield and fat percentage distributed as:

$$MVN \sim \left(\begin{bmatrix} 11000 \\ 3.4 \end{bmatrix}, \begin{bmatrix} 7954975 & -37.7904 \\ -37.7904 & 0.0067 \end{bmatrix}\right)$$

- Heritabilities of the traits are 0.17 and 0.30, respectively,

Table 1. The structure of simulated data

	Pedigree Information	SNP-effect	Phenotype	#Cow	#Sire*
Gen 0[1]	Yes	No	No	2200	200
Gen 1	Yes	No	Yes	2200	200
Gen 2	Yes	Yes	Yes	2000	-

*No records, just have pedigree

[1]Gen 0-2: generation 0 to 2

- Genetic correlation between traits is -0.80,
- 3 generations data were simulated:
  - Base population: 2200 dams without phenotypes and without SNP-genotypes
  - First generation: 2200 cows with phenotypes, without SNP-genotypes, and they were the progeny of the dams in the base population,
  - Second generation: 2000 cows with phenotypes and with SNP-genotypes, they were the progeny of the dams(cows) in the first generation,
- Each dam has one and female progeny (cow),
- Sires have average of 10 progeny and no phenotypic record,
- 54K SNP were simulated and each SNP has only additive effect,
- SNPs were mapped with reference to the bovine genome,

Method

In this study, the reliabilities of genomic predictions using different methods with different GRMs (**Henderson**, **vanRaden**, **Astle**, **Yang**, **Endelman**) developed by Henderson (1976), vanRaden (2008), Astle and Balding (2009), Yang et al. (2010) and Endelman and Jannink (2012) were compared. GRMs are based on pedigrees and/or SNP markers and each matrix calculation has the property of being idenditical by descent or state with additive relations. The Henderson relationship matrix (A) was used in classical ABLUP and single-step-genomic-BLUP (ss-GBLUP) analyses, while the other matrices were used in genomic-BLUP (GBLUP) and ss-GBLUP analyses.

Henderson relationship matrix (A) performs calculations based on the probability of gametic identity. It takes into account that each animal inherits a set of chromosomes from its ancestors and performs calculations for sets of chromosomes from common ancestor(s). The other GRMs, on the other hand, encode the markers as “0” for major homozygotes, “1” for heterozygotes and “2” for minor homozygotes as the basic framework, even though they are specifically separated. This is obtained by subtracting the mean from the coded matrix and dividing the variance of the SNPs by the total variance of the SNPs according to the equal/different number condition.

The vanRaden-GRM has been used most widely in genomic selection studies. This matrix subtracts the marker matrix from the expected values. The expected values computed by each calculated SNP frequencies from the sample or assuming the base population are known SNP frequencies, and divides it by 2pq, considering that the variance of the SNPs is equal and multiplied by the 2pq coefficient (vanRaden 2008):

$$GRM = \frac{ZZ'}{2\sum_i^m p_i(1-p_i)}$$

The Z matrix is the centered marker matrix and is the substraction of the marker matrix from a matrix of expected values of the SNPs. In this GRM, the diagonal elements correspond to 1+f. This is indeed inclined and consistent with the theory of inbreeding, measure of homozygosity in an individual. Additionally, the mean of the diagonal elements corresponding to 1+f is calculated to be 1. The correlation of inbreeding coefficients obtained with this matrix and inbreeding coefficients obtained by the classical method was calculated as 0.63 (vanRaden 2008)

The Endelmen-GRM has almost the same structure as the vanRaden-GRM. The only difference is that instead of taking the expected value the matrix, it assumes that the expected values of the SNPs are 0.5. The operations used in a GRM of this structure are no different from those in the vanRaden-GRM. These GRMs appear as the most appropriate predictors when the mean squared error is considered under one condition, which is when the number of markers is greater than the number of genotyped animals (Endelman and Jannink 2012).

Astle-GRM approach takes the kinship coefficients into account in the calculations of the GRM and handles the loci one by one. Although this matrix layout is not considered very suitable in animal breeding, it is used in plant breeding or human genetics studies. One more iteration to make unbiased of p at the expected value 2p and the p's can be recalculated. Negative values in the diagonal elements of the G indicate that less alleles are shared than the expected in line with the given p's (Astle and Balding 2009):

$$GRM_{i,j} = \frac{1}{L}\sum_{l}^{L}\frac{(x_l - 2p_l 1)(x_l - 2p_l 1)^T}{4p_l(1-p_l)}$$

Yang-GRM follows an approach that is very similar to the Astle method. The difference between them is that for an unbiased estimate of the inbreeding coefficient in the diagonal elements, the variance of SNPs is considered to be different, so that each SNP variance is affected by p (Yang et al. 2010). They divided the matrix into two elements, diagonal and off-diagonal, the expected value of the off-diagonal and diagonal elements will be zero and 1, respectively:

$$\frac{1}{N}\sum\nolimits_i G_{ijk} = \begin{cases} \frac{1}{N}\sum_i \frac{(x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2p_i(1-p_i)}, j \neq k \\ \frac{1}{N}\sum_i \frac{{x_{ij}}^2(1+2p_i)x_{ij} + 2{p_i}^2}{2p_i(1-p_i)}, j = k \end{cases}$$

In our study, the simulated data were subjected to multi-trait-BLUP analyzes.

The mixed model equation used is as follows (Mrode, 2014):

Here it is assumed that u's are u~MVN(0, g), and e's are e~MVN(0, R). R and g matrices are covariance matrices and the elements of the matrices are calculated with REML algorithms. It is also assumed that there is no correlation among genetic effects and the environmental effects:

$$cov\begin{bmatrix} g_1 \\ g_2 \\ e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} g_{11}G & g_{12}G & 0 & 0 \\ g_{21}G & g_{22}G & 0 & 0 \\ 0 & 0 & r_{11} & r_{12} \\ 0 & 0 & r_{21} & r_{22} \end{bmatrix}$$

The analyzes were carried out using the rrBLUP (Endelman 2011) and Sommer (Covarrubias-Pazaran 2016) packages. Genomic parameters were estimated with four different methods:

Method 1: ABLUP: used full pedigree and all available records in the data,

Method 2: GBLUP: used only SNP-genotypes and the phenotypes of genotyped cows,

Method 3: ss-GBLUP: used full pedigree, SNP-genotypes and phenotypes of genotyped cows,

Method 4: ss-GBLUP: used full pedigree, SNP-genotypes and all available phenotypes.

Heritability and genetic correlation were estimated and the reliabilities of breeding value estimates were compared. The reliability formula is given below and the accuracies are the square root of the reliability of the predictions:

$$REL_i = 1 - \frac{PEV}{Vg}$$

where,

REL_i = reliability of estimated breeding value of ith cow,

PEV = prediction error variance,

V_g = additive genomic variance.

$$\begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{u}_1 \\ \hat{u}_2 \end{bmatrix} = \begin{bmatrix} X_1^T R^{11} X_1 & X_1^T R^{12} X_2 & X_1^T R^{11} Z_1 & X_1^T R^{12} Z_2 \\ X_2^T R^{21} X_1 & X_2^T R^{22} X_1 & X_2^T R^{21} Z_1 & X_2^T R^{22} Z_2 \\ Z_1^T R^{11} X_1 & Z_1^T R^{12} X_2 & Z_1^T R^{11} Z_1 + G^{-1} g^{11} & Z_1^T R^{12} Z_1 + G^{-1} g^{12} \\ Z_2^T R^{21} X_1 & Z_2^T R^{22} X_1 & Z_2^T R^{21} Z_1 + G^{-1} g^{21} & Z_2^T R^{22} Z_2 + G^{-1} g^{22} \end{bmatrix}^{-1} \begin{bmatrix} X_1^T R^{11} y_1 + X_1^T R^{12} y_2 \\ X_2^T R^{21} y_1 + X_2^T R^{22} y_2 \\ Z_1^T R^{11} y_1 + Z_1^T R^{12} y_2 \\ Z_1^T R^{21} y_1 + Z_1^T R^{22} y_2 \end{bmatrix}$$

Results

Descriptive statistics of first lactation milk yields (LMY) and milk fat percentage (Fat%) of 4200 dairy cattles from 2200 and 2000 cows in generation 1 (Gen1) and generation 2 (Gen2), respectively, are shown in Table 2.

In dairy cattle, half-sib family structure is very common because a sire has the ability to mate with more than one cow in a given time. The probability of these siblings being geneticaly related to each other is 25% if it is considered with the Henderson-A. Even though this possibility changes with crossing over, the general expectation is in this direction.

Table 2. The descriptive statistics of simulated traits

Trait	N	Mean	Median	Min	Max	SD	CV%
Fat%	4200	3.40	3.40	3.12	3.69	0.082	2.40
LMY	4200	10980.95	10950.77	1595.36	20457.09	2700.168	24.50

LMY: lactation milk yield; Fat%: milk fat percentage; N: number of observations used; Min: minimum value; Max: maximum value; SD: standard deviation; CV%: coefficient of variation

Since SNP-based GRMs are based on observations rather than probability, they also take crossingover into account and it allows one to make more valid predictions. Comparisons of coefficient in different GRMs of a half-sib family with a random individual within the family are given in Table 3. Genomic relationship among halfsibs varies around the classical expectation of 0.25.

The heritability and genomic correlations are given in Table 4. All methods slightly underestimated the heritabilities, however underestimation was more drastic for genomic correlation. It can be attributable to the number of observation simulated in this study.

Table 3. Genomic relationships based on different GRMs for the animal F10_3 with its some halfsibs

i^{th} animal	j^{th} halfsib	Henderson	Endelman	vanRaden	Astle	Yang
F10_3	F181_3	0.25	0.238	0.238	0.232	0.232
F10_3	F489_3	0.25	0.247	0.247	0.239	0.238
F10_3	F608_3	0.25	0.199	0.199	0.198	0.198
F10_3	F729_3	0.25	0.231	0.231	0.221	0.220
F10_3	F930_3	0.25	0.270	0.270	0.265	0.264
F10_3	F1370_3	0.25	0.169	0.169	0.164	0.163

Table 4. Heritabilities and genomic correlations

Method		LMY	Fat%	r^G
ABLUP	Henderson	0.152	0.272	-0.617
GBLUP[1] or ss-GBLUP	vanRaden	0.155	0.282	-0.533
	Astle	0.151	0.279	-0.518
	Yang	0.153	0.275	-0.523
	Endelman	0.155	0.282	-0.533

LMY: lactation milk yield; Fat%: milk fat percentage; h^2: heritability; r^G: eklemeli genomic correlation; ABLUP: used pedigree and phenotypes (from generation 0 to 2); GBLUP: used SNP-genotypes and phenotypes (in generaion 2 only)
[1]GBLUP and ss-GBLUP produced the same results

The reliabilities obtained from four methods with different GRMs for the first lactation milk yield (LMY) and milk fat percentage (Fat%) are given in Table 5 and Table 6, respectively. For LMY, reliability estimates were close but lower using Method 2 and 3 (varied from 0.126 to 0.315) than that of obtained from classical approach (Henderson ABLUP) (0.281). However, Method 4 gave more higher reliabity estimtes (varied from 0.340 to 0.447) coparing to the other three methods. For Fat%, reliability estimates were close but higher using Method 2 and 3 (varied from 0.271 to 0.414) than that of obtained from classical approach (Henderson ABLUP) (0.362). However, Method 4 provided much higher reliability (between 0.406 and 0.496) compared to the other three methods.

Phenotypes of nongenotyped cows in the first generation improved the reliabilities of the estimations for genotyped cows in the second generation dramaticaly in the ss-GBLUP analyses (Method 4). Presence of the phenotype and pedigree information of the dams of genotyped cows were contributed significantly to the reliability of the predictions for the both trait evaluated. Moreover, ss-GBLUP with full pedigree and all available phenotypes produced higher reliabily estimates for Fat% with smaller standard deviation than for LMY with larger standard deviation. Combining all available information with ss-GBLUP gave about 1.6 and 1.2 times higher reliabilities for LMY and Fat%, respectively, compared to those obtained from ABLUP (Method 1), GBLUP (Method 2) or ss-GBLUP (Method 3) using only phenotypes of genotyped cows.

Table 5. Descriptive statistics of reliability estimates of genotyped cows from different methods for the first laktation milk yield (LMY)

	Mean	Min	Max	S.D.								
Henderson	0.281	0.206	0.302	0.009								
	Method 2				Method 3				Method 4			
	Mean	Min	Max	S.D.	Mean	Min	Max	S.D.	Mean	Min	Max	S.D.
vanRaden	0.272	0.145	0.315	0.013	0.272	0.145	0.315	0.013	0.417	0.348	0.447	0.009
Astle	0.262	0.126	0.309	0.016	0.262	0.126	0.309	0.016	0.412	0.340	0.449	0.010
Yang	0.264	0.139	0.302	0.012	0.264	0.139	0.302	0.012	0.412	0.346	0.439	0.008
Endelman	0.272	0.145	0.315	0.013	0.272	0.145	0.315	0.013	0.417	0.348	0.447	0.009

Table 6. Descriptive statistics of reliability estimates of genotyped cows from different methods for the milk fat percentage (Fat%)

	Mean	Min	Max	S.D.								
Henderson	0.362	0.305	0.383	0.009								
	Method 2				Method 3				Method 4			
	Mean	Min	Max	S.D.	Mean	Min	Max	S.D.	Mean	Min	Max	S.D.
vanRaden	0.373	0.284	0.414	0.012	0.373	0.284	0.414	0.012	0.464	0.413	0.495	0.009
Astle	0.367	0.271	0.410	0.014	0.367	0.271	0.410	0.014	0.461	0.406	0.496	0.010
Yang	0.363	0.273	0.399	0.012	0.363	0.273	0.399	0.012	0.461	0.411	0.487	0.008
Endelman	0.373	0.284	0.414	0.012	0.373	0.284	0.414	0.012	0.464	0.413	0.495	0.009

Using different GRMs, formed in four different approaches, did not make any significant effect on the reliabilities of genomic prediction. All four approaches (vanRaden, Astle, Yang, Endelman) yielded almost the same results.

Discussion

The genomic prediction reliability of the methods using GRM and only the phenotype of the genotyped cows was slightly higher for Fat% than that of the conventional ABLUP approach, but slightly lower than that for LMY. This could be due to the fact that Fat% was simulated with a low standard deviation and a high heritability, while LMY had a high standard deviation and a low heritability. Moreover, using only phenotypes of genotyped animals (GBLUP/Method 2) or all pedigree (ss-GBLUP/Method 3) did not make an noteworthy change in the results. However, a significant degree of superiority was found among the methods. Including phenotypes of nongenotyped animals into the analyses (ss-GBLUP/Method 4) dramatically improved the reliability estimates of genomic prediction. This results are consistant with previous studies (Christensen and Lund 2010; Gray et al. 2012).

Another important finding of this study is that the reliability of genomic prediction is almost the same regardless of which GRM (vanRaden, Astle, Yang, or Endelman) is used.

Conclusions

The ss-GBLUP method, which considers the entire pedigree, genomic information and phenotypes of genotyped and non-genotyped cows, provides higher realiability of genomic prediction compared to traditional BLUP (ABLUP) and the other methods that use only genotyped individuals.

Author contributions

All authors contributed equally to this study.

Conflicts of interest

The authors declares that they have no known competing financial or non-financial, professional, or personal conflicts that could have appeared to influence the work reported in this paper.

References

Astle, W., Balding, D. J. (2009). Population Structure and Cryptic Relatedness in Genetic Association Studies." Statistical Science, 24(4) 451-471. https://doi.org/10.1214/09-STS307

Covarrubias-Pazaran, G. (2016). “Genome assisted prediction of quantitative traits using the R package sommer.” PLoS ONE, 11, 1-15.

Christensen O. F., Lund M. S. (2010). Genomic prediction when some animals are not genotyped. Genetics Selection Evolution 2010, 42:2-8. http://www.gsejournal.org/content/42/1/2.

Endelman, J. B. (2011) Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP, The Plant Genome, 4(3), https://doi.org/10.3835/plantgenome2011.08.0024

Endelman, J. B. , Jannink, J. L. (2012) Shrinkage estimation of the realized relationship matrix. G3 (Bethesda).2(11):1405-1413. https://doi.org/10.1534/g3.112.004259

Forni, S. , Aguilar, I, Misztal, I. (2011) Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. Genet.Sel.Evol.,43(1):1 https://doi.org/10.1186/1297-9686-43-1

Fragomeni, B.O., Lourenco, D.A.L., Masuda, Y. et al.(2017) Incorporation of causative quantitative trait nucleotides in single-step GBLUP. Genet Sel Evol 49, 59 https://doi.org/10.1186/s12711-017-0335-0

Gaspa, G. , Veerkamp, R. F. ,Calus, M. P.L. , Windig, J. J. (2015) Assessment of genomic selection for introgression of polledness into Holstein Friesian cattle by simulation, Livestock Science, 179 86-95. https://doi.org/10.1016/j.livsci.2015.05.020.

Goddard, M.E. , Hayes, B.J. (2007) Genomic selection, Journal of Animal Breeding and Genetics, 124(6) 323-330. https://doi.org/10.1111/j.1439-0388.2007.00702.x

Gray K. A., Cassady J. P., Huang Y., Maltecca C. (2012). Effectiveness of genomic prediction on milk flow traits in dairy cattle. Genetics Selection Evolution, 44:24-29. http://www.gsejournal.org/content/44/1/24.

Henderson, C. R. (1976) A Simple Method for Computing the Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values, Biometrics, 32(1) 69-83. https://doi.org/10.2307/2529339

International Human Genome Sequencing Consortium. (2001) Initial sequencing and analysis of the human genome. Nature 409, 860–921. https://doi.org/10.1038/35057062

Karaman, E., Lund, M. S. , Anche, M. T. , Janss, L. , Su, G. (2018) Genomic Prediction Using Multi-trait Weighted GBLUP Accounting for Heterogeneous Variances and Covariances Across the Genome, G3 Genes|Genomes|Genetics, 8 (11), 3549–3558. https://doi.org/10.1534/g3.118.200673.

Legarra A. , Christensen O. F. , Aguilar I. , Misztal I. (2014) Single Step, a general approach for genomic selection, Livestock Science 166 54-65.

https://doi.org/10.1016/j.livsci.2014.04.029.

Meuwissen, T. H. , Hayes, B. J. , Goddard, M. E. (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics.157(4):1819-1829. https://doi.org/10.1093/genetics/157.4.1819.

Misztal, I. , Lourenco, D. , Legarra, A. (2020) Current status of genomic evaluation, Journal of Animal Science, 98(4) , skaa101. https://doi.org/10.1093/jas/skaa101

Mrode, R. A. (2014) Linear models for the prediction of animal breeding values, CABI, Wallingford, Oxfordshire.

Pérez-Cabal, M. A. , Vazquez, A. I. , Gianola D. , Rosa G. J. , Weigel K. A. (2012) Accuracy of Genome-Enabled Prediction in a Dairy Cattle Population using Different Cross-Validation Layouts. Front Genet. 3(27). https://doi.org/10.3389/fgene.2012.00027

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Schaeffer, L.R. (2006) "Strategy for applying genome-wide selection in dairy cattle", Journal of Animal Breeding and Genetics, 123(4) 218-223. https://doi.org/10.1111/j.1439-0388.2006.00595.x.

Seno, L., Guidolin, D., Aspilcueta-Borquis, R., Nascimento, G., Silva, T., Oliveira, H., & Munari, D. (2018). Genomic selection in dairy cattle simulated populations. Journal of Dairy Research, 85(2), 125-132. https//doi.org/10.1017/S0022029918000304

Tsuruta, S. , Lourenco, D.A.L., Masuda, Y. , Lawlor, T.J. , Misztal, I. (2021) Reducing computational cost of large-scale genomic evaluation by using indirect genomic prediction, JDS Communications,2(6),356:360. https://doi.org/10.3168/jdsc.2021-0097

VanRaden, P.M. (2008) Efficient Methods to Compute Genomic Predictions, Journal of Dairy Science 91(11) 4414-4423. https://doi.org/10.3168/jds.2007-0980.

Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Health AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM.(2010) Common SNPs explain a large proportion of the heritability for human height. Nature Genet., 42,565–569. https://doi.org/10.1038/ng.608.

Wiggans, G. R. , Cole, J. B. , Hubbard, S. M., Sonstegard, T. S. (2017) Genomic Selection in Dairy Cattle: The USDA Experience, Annu Rev Anim Biosci.5:309-327. https://doi.org10.1146/annurev-animal-021815-111422