Theor Appl Genet (2015) 128:693–703
DOI 10.1007/s00122-015-2464-6

ORIGINAL PAPER

Shrinkage estimation of the genomic relationship matrix can improve genomic estimated breeding values in the training set

Dominik Müller · Frank Technow · Albrecht E. Melchinger

Received: 17 September 2014 / Accepted: 10 January 2015 / Published online: 4 March 2015

Abstract
Key message **We evaluated several methods for computing shrinkage estimates of the genomic relationship matrix and demonstrated their potential to enhance the reliability of genomic estimated breeding values of training set individuals**.
Abstract In genomic prediction in plant breeding, the training set constitutes a large fraction of the total number of genotypes assayed and is itself subject to selection. The objective of our study was to investigate whether genomic estimated breeding values (GEBVs) of individuals in the training set can be enhanced by shrinkage estimation of the genomic relationship matrix. We simulated two different population types: a diversity panel of unrelated individuals and a biparental family of doubled haploid lines. For different training set sizes (50, 100, 200), number of markers (50, 100, 200, 500, 2,500) and heritabilities (0.25, 0.5, 0.75), shrinkage coefficients were computed by four different methods. Two of these methods are novel and based on measures of LD, the other two were previously described in the literature, one of which was extended by us. Our results showed that shrinkage estimation of the genomic relationship matrix can significantly improve the reliability of the GEBVs of training set individuals, especially for a low number of markers. We demonstrate that the number of markers is the primary determinant of the optimum shrinkage coefficient maximizing the reliability and we recommend methods eligible for routine usage in practical applications.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-015-2464-6) contains supplementary material, which is available to authorized users.

Communicated by Hiroyoshi Iwata.

D. Müller (✉) · F. Technow · A. E. Melchinger
University of Hohenheim, Stuttgart, Germany
e-mail: Dominik_Mueller@uni-hohenheim.de

Present Address:
F. Technow
DuPont Pioneer, Johnston, IA, USA

Introduction

Since genomic prediction was first proposed by Meuwissen et al. (2001), it has proven to be a promising approach for numerous applications in both animal (e.g., Hayes et al. 2009; Hayes and Goddard 2010) and plant breeding (e.g., Bernardo and Yu 2007; Riedelsheimer et al. 2012). In the literature, the focus has so far been on the reliability of GEBVs for unobserved genotypes, whereas the training set (TS) of individuals used for calibrating the prediction model has received only little attention. However, in applied plant breeding programs, the TS individuals constitute a considerable fraction of the total breeding population and are usually themselves selection candidates. For TS individuals, both their phenotypic values and their GEBVs are available.

One of the most popular methods for genomic prediction is genomic best linear unbiased prediction (GBLUP), which has proven to be simple and efficient with performance that compares well with more sophisticated prediction methods (de Los Campos et al. 2013). It is based on the animal model (Lynch and Walsh 1998) that has been widely used by animal breeders for decades. The difference lies in the definition of the relationship matrix **A**. While in the classical animal breeding literature, **A** is calculated from pedigree data (e.g., Lynch and Walsh 1998), the principal innovation of GBLUP was to calculate **A**

from genome-wide marker data (Habier et al. 2007; VanRaden 2008; Goddard et al. 2009), often referred to as the genomic relationship matrix (GRM).

The elements of the GRM are estimates of the genetic correlation between alleles taken from pairs of individuals and can be conveniently computed with reference to the current population (Powell et al. 2010). As such, they can be interpreted as deviations from expected allele sharing between individuals, given the allele frequencies of the current population (Astle and Balding 2009). These deviations are a result of Mendelian sampling and linkage during the segregation of loci (Hill and Weir 2011).

Estimating genetic covariances from marker data allows for defining relationships among individuals of unknown ancestry, which would classically be treated as unrelated. An example in plant breeding would be a diversity panel of lines. Furthermore, it enables to identify additive-genetic variation within groups of individuals having identical pedigree relationships, for instance full-sib families.

Endelman and Jannink (2012) examined genomic prediction using GBLUP in the TS and demonstrated that the reliability of GEBVs of TS individuals can be substantially increased by shrinking the GRM towards a less complex target matrix that can be estimated from the data with higher precision. The problem was also addressed by Riedelsheimer and Melchinger (2013), who applied selection index theory to construct a selection index that aims to optimally combine GEBVs and phenotypic values of TS individuals. Apart from those previous studies, the importance of genomic prediction in the TS has not been appropriately recognized in the literature so far. Our study aims to alleviate this neglect by comparing the performance of several alternative shrinkage methods as well as the method of Riedelsheimer and Melchinger (2013). Besides two novel shrinkage methods that are based on measures of linkage disequilibrium between marker loci, we applied a regression approach similar to the proposal of Yang et al. (2010) and Goddard et al. (2011) and also used the method presented by Endelman and Jannink (2012). The objective of our study was to compare the alternative shrinkage methods in terms of reliabilities of GEBVs for different population types and marker densities.

Material and methods

Statistical model

The GEBVs were computed by GBLUP with the basic linear mixed model

$$y_i = \mu + a_i + e_i, \tag{1}$$

where the phenotypic value y_i of the ith individual is decomposed into a common intercept μ (fixed), a true genetic value a_i (random), and a residual term e_i. Using vector notation, the model assumes that $\mathbf{a} \sim \mathcal{N}\left(0, \mathbf{A}\sigma_a^2\right)$ and $\mathbf{e} \sim \mathcal{N}\left(0, \mathbf{I}\sigma_e^2\right)$, where σ_a^2 and σ_e^2 are the genetic and residual variance components, respectively. The matrix $\mathbf{A}$ is the GRM and its computation will be detailed later. The genetic values were predicted using the standard BLUP formulas (Lynch and Walsh 1998)

$$\hat{\mu} = \left(\mathbf{1}^T\mathbf{V}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}^T\mathbf{V}^{-1}\mathbf{y} \tag{2}$$

$$\hat{\mathbf{a}} = \sigma_a^2\mathbf{A}\mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{1}\hat{\mu}\right), \tag{3}$$

where $\mathbf{V} = \mathbf{A}\sigma_a^2 + \mathbf{I}\sigma_e^2$. Variance components and heritabilities were estimated using the spectral decomposition algorithm of Kang et al. (2008) as implemented in the *R* package *rrBLUP* (Endelman 2011).

Simulation

We simulated two different population types, a population of unrelated lines (UR) and a biparental family of lines (BP). The UR population was simulated by sampling genotypes from a joint distribution as described in Montana (2005) using allele frequencies sampled from the interval $[0.35, 0.65]$ and LD modelled following the exponential decay function $\mathrm{LD}(d) = 0.8 \times e^{-20d}$, where d is the genetic distance in Morgan. The BP population was generated by recombining the genomes of two divergent parental lines (i.e., lines that were generated by randomly assigning SNP alleles to one or the other parent with equal probability) using the *R* package *hypred* (Technow 2013). In both populations, haplotypes were doubled to obtain fully homozygous doubled haploid lines. We simulated ten chromosomes, the lengths of which were taken from the Genetics (2008) Composite Map of Maize (http://www.maizegdb.org) with a total map length of $\sim$18 Morgan. We used a constant number of 200 QTL, such that the QTL density amounted to about 11 QTL per Morgan. In both scenarios, we used different TS sizes $N \in \{50, 100, 200\}$ and heritabilities $h^2 \in \{0.25, 0.5, 0.75\}$. The size of the prediction set (PS) was held constant at 200 individuals. TS sizes were chosen to reflect the numbers used in practical plant breeding programs.

In order to vary linkage disequilibrium between markers and QTL, we used increasing numbers of markers $M \in \{50, 100, 500, 1,000, 2,500\}$. To place QTL and markers on the genome, first their number per chromosome was sampled from a multinomial distribution with class probabilities equal to the relative chromosome lengths. Subsequently, QTL and markers were uniformly distributed along

the respective chromosomes. QTL effects were drawn from a gamma distribution (Meuwissen et al. 2001) with shape 1.0 and rate 2.0. The signs of the effects were sampled from a Bernoulli distribution with $p = 0.5$. The QTL effects were then scaled to achieve an overall genetic variance equal to 1.0. Phenotypes were simulated by adding an independent Gaussian error term with $\sigma_e^2 = \frac{1-h^2}{h^2}$, depending on the heritability h^2. The reliability of GEBVs was calculated as the squared correlation coefficient between GEBVs and the simulated true genetic values and is denoted by ρ^2.

All of our results were obtained from 500 independent simulation runs. In order to determine the maximum reliability $\rho^2_{\max}$ in the TS and the corresponding optimum shrinkage coefficient δ_{opt} to be used in Eq. 5 described below, we computed the reliability of the resulting GEBVs in the TS at a sequence of 100 shrinkage coefficients equally spaced between 0 and 0.9 for each simulation run. Averages across all runs were calculated for each position in the sequence and $\rho^2_{\max}$ and δ_{opt} were determined numerically. The reliability of the phenotypic values, i.e., the squared correlation coefficient between phenotypic values and true genetic values corresponded to the heritability h^2. All computations were performed within the statistical computing environment *R* (R Core Team 2014).

Shrinkage methods

As a starting point and reference for all methods, the GRM was computed according to the first method of VanRaden (2008), which we refer to as Method VR1. As shown by Endelman and Jannink (2012), this method is also suitable for populations of inbred lines and the GRM is computed according to the following formula:

$$\widehat{\mathbf{A}} = \frac{\mathbf{W}\mathbf{W}^T}{2\sum_k p_k(1-p_k)}, \tag{4}$$

(Habier et al. 2007; VanRaden 2008; Endelman and Jannink 2012), where $\mathbf{W}$ is the column-centered genotype matrix with $w_{ik} = x_{ik} - 2p_k$; here $x_{ik} \in \{0, 1, 2\}$ codes the number of major alleles at the kth locus in the ith individual and p_k is the sample allele frequency at the kth locus. Under the infinitesimal model, the genetic value is determined by an infinitely large number of unlinked loci each of which contributes a small effect (Hill 2010). Given these assumptions, the genomic relationship matrix can be optimally estimated from the observed marker loci by Eq. 4 (Endelman and Jannink 2012).

In the following, we describe four methods that are based on the principle of imposing shrinkage on $\widehat{\mathbf{A}}$ to obtain a modified relationship matrix that can be written as

$$\widehat{\mathbf{A}}^* = \delta\mathbf{T} + (1-\delta)\widehat{\mathbf{A}}, \tag{5}$$

where $\mathbf{T}$ is a target matrix toward which $\widehat{\mathbf{A}}$ is shrunken. The shrinkage coefficient δ specifies the strength of shrinkage imposed on $\widehat{\mathbf{A}}$. Methods 1 and 2 are novel, Method 3 is based on Yang et al. (2010) and Goddard et al. (2011) and further developed by us, and method 4 was presented by Endelman and Jannink (2012). In Methods 1–3, the target matrix toward which $\widehat{\mathbf{A}}$ is shrunken is a diagonal matrix with elements equal to the average of the diagonal elements of $\widehat{\mathbf{A}}$, which is equal to $1 + \widehat{f}$. Here $\widehat{f}$ is the average inbreeding coefficient in the population, which equals 2 for fully inbred lines as used in the present study.

Method 1: adjLD

In preliminary analyses we observed that the optimum shrinkage coefficient is in a strong relationship with LD. We, therefore, developed a heuristic method in which the LD between adjacent marker loci (LD_{adj}) was used to compute the shrinkage coefficient as $\delta_{\text{adjLD}} = 1 - \text{LD}_{\text{adj}}$. The LD between adjacent markers was obtained as the average of the squared correlation between all pairs of neighboring markers across the genome (Hill and Robertson 1968).

Method 2: effLD

Because LD_{adj} only captures LD between adjacent loci, we devised a measure for effective LD (LD_{eff}) between a single hypothetical QTL and its surrounding markers. In short, LD_{eff} measures the amount of variation in the genotype of a single locus that is simultaneously explained by the genotypes of several surrounding loci. The shrinkage coefficient δ is then analogously computed as $\delta_{\text{effLD}} = 1 - \text{LD}_{\text{eff}}$. A detailed description of the method is provided in the "Appendix".

Method 3: RG

The third method extends the regression approach described by Yang et al. (2010) and Goddard et al. (2011). Here, the rationale is to regress relationship coefficients computed with QTL on those computed with markers and use the slope β for shrinkage to obtain an unbiased estimate of the GRM . In practice, β has to be estimated based on marker data alone, because the QTL are unkown. In Yang et al. (2010), β is estimated by randomly splitting markers into two equally sized sets for different numbers of markers and subsequently treating one set as proxies for QTL. The regression coefficient β is obtained by regressing the elements of $(\mathbf{A} - \mathbf{I})$ on the elements of $(\widehat{\mathbf{A}} - \mathbf{I})$, where $\mathbf{A}$ is the GRM computed with the (pseudo-) QTL and $\widehat{\mathbf{A}}$ the GRM computed with the markers. In our study, we estimated β by randomly splitting the total number of markers into two distinct sets. Because the number of QTL is relevant for the estimation of β, we

varied the set size of the pseudo-QTL starting from 5 up to half the number of all markers. Then we performed separate regressions for each set size with 25 replications, where we regressed the elements of $(\mathbf{A} - \mathbf{T}^{\mathrm{QTL}})$ on the elements of $(\widehat{\mathbf{A}} - \mathbf{T})$, including the diagonal. Here, $\mathbf{T}$ and $\mathbf{T}^{\mathrm{QTL}}$ are the diagonal matrices that contain the average of the diagonal elements of $\widehat{\mathbf{A}}$ and $\mathbf{A}$, respectively. The mean of all regression coefficients was used as an estimate $\widehat{\beta}$ and the corresponding shrinkage coefficient was obtained as $\delta_{\mathrm{RG}} = 1 - \widehat{\beta}$. In addition, we computed the shrinkage coefficient of Method RG using the true QTL genotypes to calculate $\mathbf{A}$, denoted $\delta_{\mathrm{RG}}^{\mathrm{QTL}}$, for comparison.

Method 4: EJ

This method was devised by Endelman and Jannink (2012) and differs from the previous ones in that a different target for shrinkage is used. In the original presentation of Endelman and Jannink (2012), the shrunken GRM is computed as

$$\widehat{\mathbf{A}}^{*} = \frac{\delta_{\mathrm{EJ}}\ \langle \mathbf{S}_{ii} \rangle \mathbf{I} + (1 - \delta_{\mathrm{EJ}})\mathbf{S} + \langle \mathbf{W}_{\cdot k} \rangle \langle \mathbf{W}_{\cdot k} \rangle^{T}}{2\langle p_k q_k \rangle},$$

where $\langle \mathbf{S}_{ii} \rangle$ is the mean of the diagonal elements of $\mathbf{S}$ with $\mathbf{S} = M^{-1}\mathbf{W}\mathbf{W}^{T} - \langle \mathbf{W}_{\cdot k} \rangle \langle \mathbf{W}_{\cdot k} \rangle^{T}$ being the sample covariance matrix, $\langle \mathbf{W}_{\cdot k} \rangle$ is a column vector containing the row means of $\mathbf{W}$, and $\langle p_k q_k \rangle$ is the average of the product between allele frequencies across all loci. This can be rearranged to

$$\widehat{\mathbf{A}}^{*} = \delta_{\mathrm{EJ}} \left(\frac{\langle \mathbf{S}_{ii} \rangle \mathbf{I}}{2\langle p_k q_k \rangle} + \frac{\langle \mathbf{W}_{\cdot k} \rangle \langle \mathbf{W}_{\cdot k} \rangle^{T}}{2\langle p_k q_k \rangle} \right) + (1 - \delta_{\mathrm{EJ}})\widehat{\mathbf{A}}. \quad (6)$$

Hence, Endelman and Jannink (2012) use a similar target matrix as we do, which has the same diagonal elements as $\mathbf{T}$, but has in addition non-zero off-diagonal elements determined by the second term in the first parenthesis in Eq. 6. The computation of the shrinkage coefficient δ_{EJ} was described in Endelman and Jannink (2012).

Method 5: RM

In the context of ressource optimization for a single breeding cycle with genomic selection, Riedelsheimer and Melchinger (2013) proposed a selection index that combines GEBVs with phenotypic data for individuals in the training set. Their index is based on the theory presented in Lande and Thompson (1990) originally developed for marker-assisted selection. Although this method is not based on shrinkage estimation of the GRM, we included it in our analyses because it was originally constructed with the objective to improve the reliability of GEBVs of training set individuals, which is also the ultimate goal of the shrinkage methods presented earlier. Moreover, shrinkage estimation of the GRM effectively leads to an up-weighting of the own phenotypic value of an individual, while down-weighting the information of related individuals. Thus, the shrinkage coefficient can be conceptually regarded as a selection index combining a phenotype's own value with its GEBVs, estimated by using a non-shrunken GRM. In the "Appendix", we provide a detailed derivation of the formulas presented in Riedelsheimer and Melchinger (2013) and point out that some key assumptions implicitly made are violated.

Results

Reliability of method VR1 in the TS and PS

For the same size of the training set N, heritability h^2, and number of markers M, reliabilities for both TS and PS using Method VR1 were always higher in the BP population than in the UR population (Table 1) . In general, reliabilities increased with increasing N, h^2 and M. In the BP population, reliabilities in the PS amounted to 51–61 % of those observed in the TS for $N = 50$ and to 81–88 % for $N = 200$, with increasing percentage value for increasing number of markers. On the other hand, in the UR population reliabilities in the PS amounted to 11–25 % of those in the TS for $N = 50$ and 37–57 % for $N = 200$. While the reliabilities for $N = 50$ were above 0.17 and thus reasonably high in the BP population, they were lower than 0.17 in the UR population. In the UR population, the reliability in the TS decreased for increasing TS size when the number of markers was <500, but increased for $M \geq 500$ (Online Resource 1, Table S2) . Moreover, the reliability in the TS of the UR population only surpassed h^2 when $M > 200$, for all levels of N and h^2.

Reliabilities in the BP and UR population

The relative performance of the methods was similar for all levels of N. We, therefore, limit our presentation of results to those obtained for $N = 200$, for the sake of brevity. Results for $N = 50$ and $N = 100$ are shown in Online Resource 1. The performance of the various methods in the UR population for a training set size of 200 showed a strong dependency on the heritability h^2 and the number of markers M (Fig. 1). The difference between Method VR1 and the maximum reliability ρ^2_{max} was largest for high h^2 and low M and smallest vice versa. For $M = 100$, the methods adjLD, effLD, and EJ performed equally well, whereas RG showed slightly lower performance, especially for high h^2. Method RM led to the lowest reliability of GEBVs compared to all the other methods and was hardly better than Method VR1. For $M = 500$, effLD and RG were superior, followed by EJ and RM, which had comparable reliabilities. The reliability

Table 1 Reliability in the training and prediction set using the standard GRM after VanRaden (2008) (Method VR1) for different training set sizes ($N = 50, 100, 200$), heritabilities ($h^2 = 0.25, 0.50, 0.75$) and number of markers ($M = 100, 500, 2,500$) uniformly distributed on 10 chromosomes with a total length of about 18 Morgans

N	h^2	M	BP		UR	
			TS	PS	TS	PS
50	0.25	100	0.335	0.177	0.237	0.033
		500	0.357	0.201	0.278	0.055
		2,500	0.362	0.215	0.295	0.072
	0.5	100	0.540	0.281	0.417	0.052
		500	0.575	0.332	0.503	0.100
		2,500	0.589	0.351	0.523	0.119
	0.75	100	0.729	0.372	0.598	0.071
		500	0.779	0.459	0.728	0.138
		2,500	0.780	0.475	0.736	0.163
200	0.25	100	0.477	0.383	0.219	0.088
		500	0.545	0.465	0.328	0.160
		2,500	0.559	0.488	0.363	0.200
	0.5	100	0.648	0.527	0.366	0.138
		500	0.725	0.630	0.551	0.262
		2,500	0.733	0.641	0.590	0.323
	0.75	100	0.747	0.608	0.476	0.168
		500	0.849	0.735	0.731	0.337
		2,500	0.861	0.763	0.778	0.413

of Method adjLD was lowest. For $M = 2{,}500$, the reliability of VR1 was already almost identical with the optimum ρ^2_{opt}. Here, the best methods were RG, EJ, and RM, whereas effLD and adjLD showed the lowest reliability.

In the BP population, for $M = 100$, Method RG and effLD had the highest reliability. Method RM showed comparable performance to VR1, whereas methods adjLD and EJ were only marginally better than VR1 for $h^2 = 0.75$ and otherwise worse. For $M \geq 500$, however, the differences between the methods and VR1 were very small. However, for $M = 2{,}500$ and $h^2 = 0.75$, Method effLD showed a distinctly lower performance than the other methods.

Shrinkage coefficients

In our simulations, we numerically determined the optimum shrinkage coefficient δ_{opt} that maximized the reliability in the TS. To assess the relative importance of the number of markers M, heritability h^2, and training set size N on the variation in δ_{opt}, we used linear regression with scaled predictors (Table 2) .

In the UR and BP populations, the total variation in the optimum shrinkage coefficient δ_{opt} explained by the linear regression amounted to $R^2 = 0.633$ and $R^2 = 0.394$, respectively. In both population types, the number of markers M showed the largest regression coefficient, with -2.16 in UR and -0.095 in BP. Compared to M, heritability h^2 and training set size N had only a small influence on δ_{opt} in both population types.

Because of this, we computed δ_{opt} for different numbers of markers, averaging over heritability and training set size and compared it to the shrinkage coefficients obtained by Methods adjLD, effLD, and RG (Table 3) , which do not vary with h^2 and N by definition. In addition, we calculated the shrinkage coefficient for Method RG using the true QTL ($\delta^{\text{QTL}}_{\text{RG}}$).

In the UR (BP) population, δ_{opt} was 0.81 (0.39) for $M = 50$ and was reduced to 0.05 (0.01) for $M = 2{,}500$. Across both population types, $\delta^{\text{QTL}}_{\text{RG}}$ was remarkably close to δ_{opt}, with a correlation of 0.98. For Method RG, δ_{RG} was considerably lower than δ_{opt} for $M \leq 100$, but in good agreement otherwise. The shrinkage coefficient δ_{adjLD} was generally higher than δ_{opt} in both population types for all levels of M and decreased only to 0.37 for $M = 2{,}500$ in the UR population. For Method effLD, δ_{effLD} was close to δ_{opt} for $M \leq 200$, but its value stayed almost constant for $M \geq 500$ in the UR population and even increased in the BP population. We found that the optimum shrinkage coefficient δ_{opt} and $\delta^{\text{QTL}}_{\text{RG}}$ were almost identical. The estimate δ_{RG} matched δ_{opt} for $M = 100$ and upward.

Discussion

Shrinkage estimation of the GRM

Best linear unbiased prediction (BLUP) is equivalent to a selection index when fixed effects are first estimated using

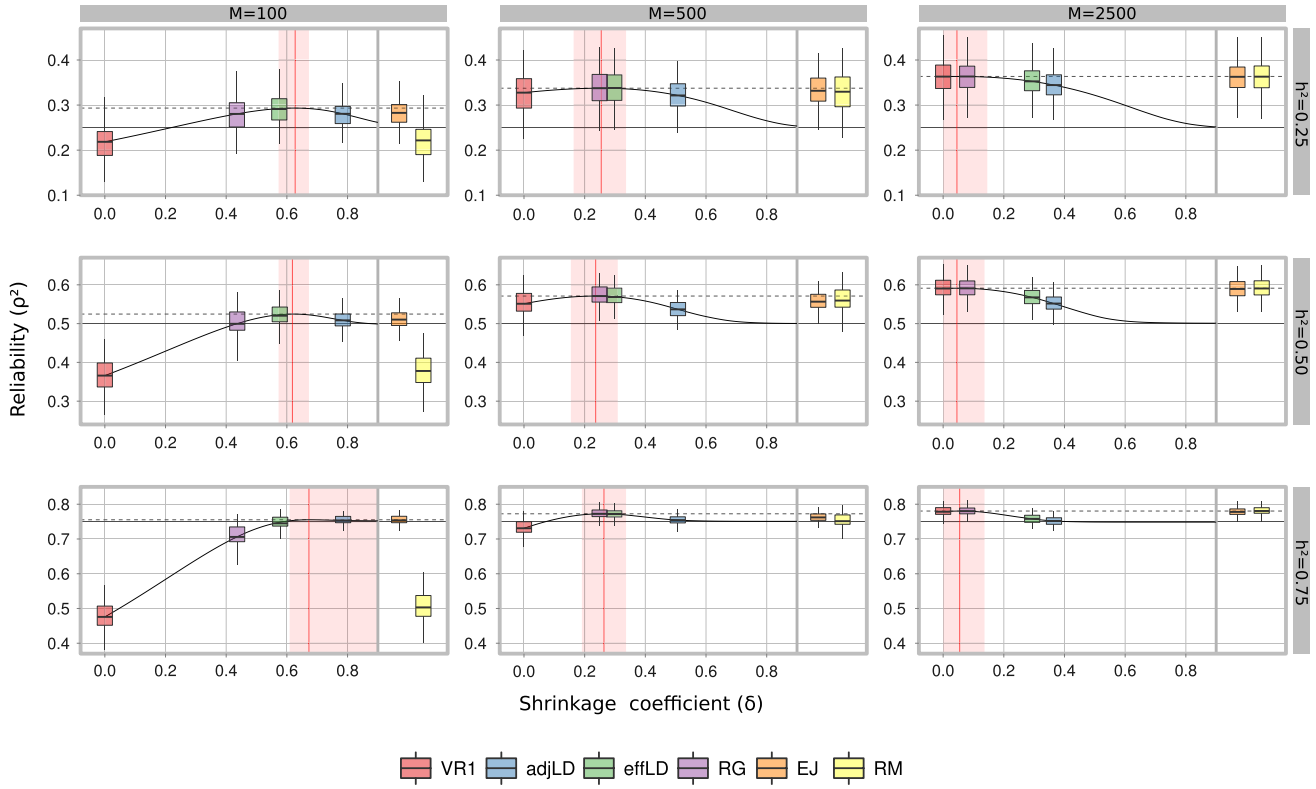

Fig. 1 Reliability (ρ^2) in the UR population for a training set size of $N = 200$ for different numbers of markers ($M = 100, 500, 2,500$) and heritabilities ($h^2 = 0.25, 0.50, 0.75$). The *solid black curve* shows the reliability when the shrinkage coefficient δ is systematically varied between 0 and 0.9. The maximum of this *curve*, $\rho^2_{\max}$, is indicated by the *dashed horizontal black line*, and the value of δ for which $\rho^2_{\max}$ was achieved, which is δ_{opt}, is shown by the *vertical red line*, surrounded by a *shaded region* where the reliability is not < 99.5 % of the $\rho^2_{\max}$. The *boxplots* shows the mean and the 0.90, 0.65, 0.35, and 0.10 quantiles for the different methods, centered at the average shrinkage coefficient of the respective method. The *boxplots* for Methods EJ and RM are drawn without a scale on the *x-axis* in a separate section within *each panel*, because they cannot be compared to the other methods based on their shrinkage coefficient

generalized least-squares and subsequently used to correct phenotypic values (Henderson 1973). This index optimally combines the available phenotypic information of related individuals and maximizes the correlation between predicted and true genetic values (Searle et al. 1992). However, this property depends on the correct specification of the covariance structure, i.e., the GRM and the variance components. If markers are not in sufficient LD with QTL, the relationships derived from marker genotypes deviate from the actual relationships at the QTL (Yang et al. 2010), resulting in a misrepresentation of the true QTL relationships in the GRM. This leads to spurious signals coming from the phenotypic values of other individuals and, as a consequence, the reliability of the GEBVs is impaired and can even be significantly lower than the heritability (Figs. 1, 2). A similar phenomenon was observed by Habier et al. (2013), where they showed that increasing the TS size can even lead to reduced reliability of individuals in the PS because of 'relationship noise' due to the misrepresentation of the actual pedigree relationships in the GRM. Shrinkage estimation of the GRM can then recover some of the lost reliability when a proportionally larger amount of 'noise' due to incomplete LD is shrunken to zero compared to actual QTL relationships traced by markers. In terms of the BLUP selection index, shrinkage leads to an up-weighting of the own phenotypic value of an individual and down-weighting of phenotypic values of other individuals and by this reduces the negative impact of spurious signals from misrepresented relationships.

Optimum shrinkage coefficient

By using linear regression , we found that in both population types most of the variation in the optimum shrinkage coefficient δ_{opt} can be explained by the number of markers (Table 2). The number of markers is strongly related to LD, so that in turn, LD is an important influencing factor of δ_{opt}. Consequently, if a sufficient number of markers is present to ensure a high level of LD, relationships in the GRM are specified correctly and shrinkage is not required. This corroborates the notion that information about actual relationships conveyed by markers is tightly

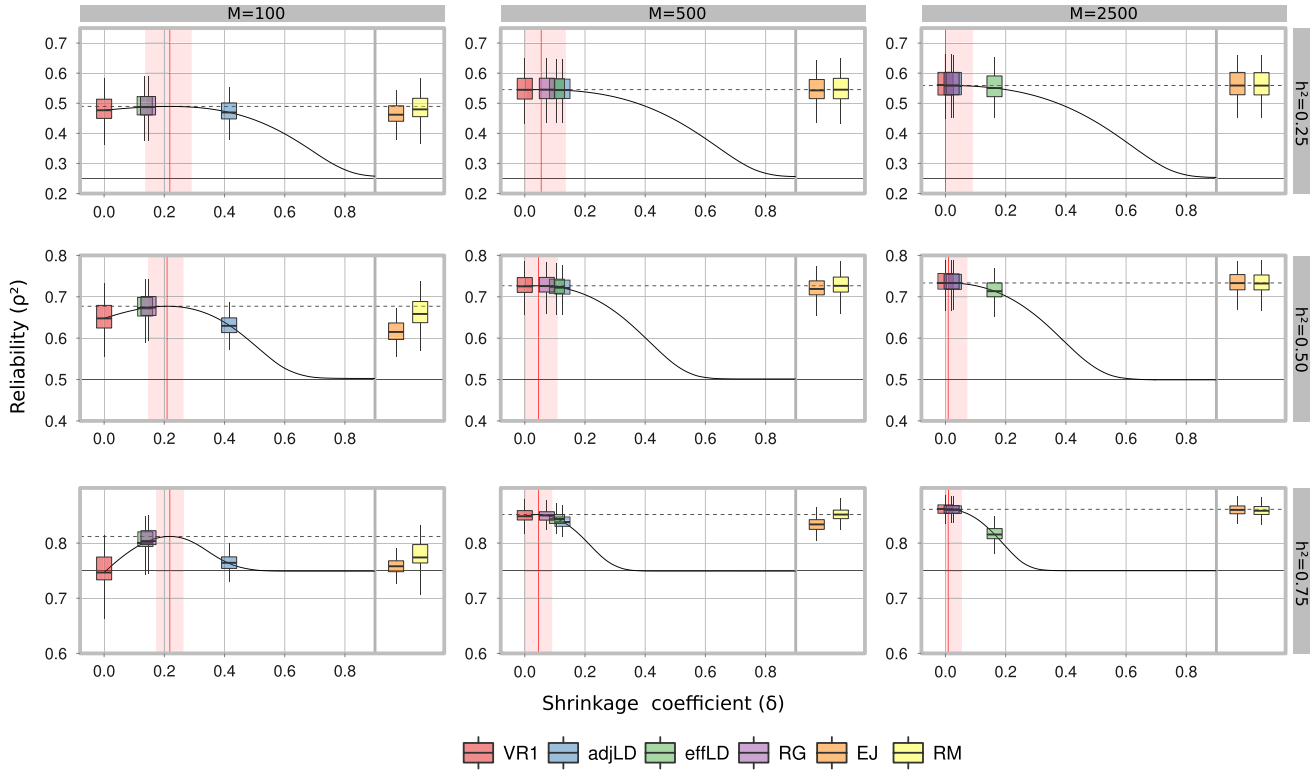

Fig. 2 Reliability (ρ^2) in the BP population for a training set size of $N = 200$ for different numbers of markers ($M = 100, 500, 2{,}500$) and heritabilities ($h^2 = 0.25, 0.50, 0.75$). The *solid black curve* shows the reliability when the shrinkage coefficient δ is systematically varied between 0 and 0.9. The maximum of this *curve*, $\rho^2_{\max}$, is indicated by the *dashed horizontal black line*, and the value of δ for which $\rho^2_{\max}$ was achieved, which is δ_{opt}, is shown by the *vertical red line*, surrounded by a *shaded region* where the reliability is not <99.5 % of the $\rho^2_{\max}$. The *boxplots* shows the mean and the 0.90, 0.65, 0.35, and 0.10 quantiles for the different methods, centered at the average shrinkage coefficient of the respective method. The *boxplots* for Methods EJ and RM are drawn without a scale on the *x-axis* in a separate section within *each panel*, because they cannot be compared to the other methods based on their shrinkage coefficient

Table 2 Linear Regression of the optimum shrinkage coefficient δ_{opt} on the number of markers (M), heritability (h^2) and training set size (N) as predictors scaled by subtracting the mean and dividing by the standard deviation

	UR	BP
M	−0.216 (0.025)***	−0.095 (0.017)***
h^2	0.073 (0.025)**	0.006 (0.017)
N	−0.051 (0.025)*	−0.006 (0.017)
R^2	0.633	0.394

Shown are the regression coefficient estimates, followed by the respective standard errors in parentheses

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

associated with LD (Yang et al. 2010). LD also strongly impacted the reliability of GEBVs. The lower LD in the UR compared to BP population can explain the generally lower reliability in both TS and PS in the former. The presence of extended linkage blocks due to cosegregation (Frisch and Melchinger 2007; Smith et al. 2008) in biparental populations of doubled haploid lines can explain the higher reliability in the BP compared to the UR population (Habier et al. 2013).

The difference between the maximum reliability $\rho^2_{\max}$ obtained using the optimum shrinkage coefficient δ_{opt} and the reliability obtained for Method VR1 can be regarded as the maximum achievable gain in reliability that can be brought about by shrinkage. This gain was generally highest for a low number of markers M and high heritability h^2, and vice versa (Figs. 1, 2). However, because the focus is on the reliability in the TS, for which phenotypic values are available, any gain in reliability due to shrinkage has to be set into relationship to h^2, which represents the reliability achieved when selecting on the phenotypic values directly. Therefore, although the gain in reliability went up with increasing h^2, the difference between $\rho^2_{\max}$ and h^2 went down. Hence, there is a range where h^2 is high enough to allow shrinkage to substantially improve the reliability of GEBVs in the TS relative to the one obtained with Methods VR1, but yet low enough to allow $\rho^2_{\max}$ to be appreciably higher than h^2. This range is precisely what was termed the "sweet spot" by Endelman and Jannink (2012). In their

Table 3 Shrinkage coefficients for different numbers of markers ($M = 50, 100, 200, 500, 1,000, 2,500$), averaged across heritability and training set size

M	M_d	δ_{adjLD}	δ_{effLD}	δ_{RG}	δ_{RG}^{QTL}	δ_{opt}
UR						
50	2.8	0.87	0.75	0.48	0.78	0.81
100	5.6	0.78	0.58	0.44	0.63	0.64
200	11.1	0.67	0.39	0.37	0.46	0.46
500	27.8	0.51	0.30	0.25	0.25	0.25
1000	55.7	0.42	0.29	0.16	0.14	0.14
2500	139.2	0.37	0.29	0.08	0.06	0.05
BP						
50	2.8	0.58	0.32	0.16	0.44	0.39
100	5.6	0.42	0.14	0.15	0.29	0.22
200	11.1	0.26	0.10	0.12	0.17	0.11
500	27.8	0.13	0.11	0.07	0.07	0.05
1000	55.7	0.07	0.12	0.04	0.04	0.02
2500	139.2	0.03	0.16	0.02	0.02	0.01

M_d marker density (number of markers per Morgan), δ_{adjLD}, δ_{effLD}, δ_{RG} shrinkage coefficients for Methods adjLD, effLD, and RG, δ_{RG}^{QTL} shrinkage coefficient for Method RG using QTL, δ_{opt} numerically determined optimum shrinkage coefficient

article, the showed that shrinkage estimation of the GRM using Methods EJ can improve the reliability of GEBVs in the TS in an "unstructured" population of 274 maize inbred lines genotyped for 384 markers, where by "unstructured" they implied that the first principal component explained only 5 % of the total variation.

In the PS, regardless of the combination of the parameters M, h^2 and N, shrinkage did not lead to any gain in reliability, *i.e*, the maximum achievable gain in reliability was essentially zero (Online Resource 1, Table S3). This result corroborates the findings of Endelman and Jannink (2012) that shrinkage did not improve the GEBV reliability for unphenotyped individuals, even for a low number of markers.

Comparison between methods

In our simulation study, the optimum shrinkage coefficient δ_{opt} could be identified because the true genetic values and the QTL were known. For real applications, however, the shrinkage coefficient must be estimated from the data. The regression methods RG would lead to a shrinkage coefficient δ_{RG}^{QTL} that closely matches δ_{opt} if the QTL were known, which demonstrates that Method RG is in principal the right approach. However, neither QTL nor their number is known in practice, which is the reason why markers have to be employed as a proxy for QTL. This poses the problem to decide on the proportion of the sets into which the markers are partitioned, which should best reflect the unknown true proportion between QTL and markers. Our strategy of assuming the number of QTL ranging from a minimum of 5 up to half the number of markers ensured that values δ_{RG} close to δ_{RG}^{QTL} were achieved for a high number of markers, but it causes δ_{RG} to have a pronounced downward bias relative to δ_{RG}^{QTL} when <200 markers were used (Table 3), which equals the number of QTL we used throughout our simulations. Consequently, Methods RG featured shrinkage coefficients close to δ_{opt} for $M \geq 200$ and thus was one of the best performing methods for both population types. The Methods effLD had a shrinkage coefficient in good agreement with δ_{opt} for $M \leq 500$, where it showed reliabilities close to ρ^2_{max}. However, for more than 500 markers, δ_{effLD} was considerably higher than δ_{opt}, which led to shrinkage that was too strong and consequently reliabilities were even lower than those obtained for Method VR1. The same trend was observed for Method LDadj with shrinkage coefficients δ_{adjLD} that were even more exaggerated for a large number of markers. Method EJ is also based on a shrinkage approach, but towards a slightly different target matrix than methods RG, effLD and adjLD, which is the reason why it cannot be compared to the other methods based on its shrinkage coefficient. The method showed superior performance in the UR population, especially for a low number of markers, but revealed deficiencies in the BP population for low to medium number of markers, where it can underperform Method VR1. The method RM is not based on shrinkage, but on a selection index approach (Riedelsheimer and Melchinger 2013). Although critical assumptions of the method are not fulfilled, it shows reasonable performance in both population types for $M \leq 500$, but is hardly better than Method VR1 for $M = 50$, particularly in the UR population.

In conclusion, our results demonstrate that shrinkage estimation of the GRM can substantially improve the reliability of GEBVs of TS individuals, in particular when the number of markers is low and the heritability is at intermediate values. Of the shrinkage methods evaluated, Method RG was the most promising with superior performance and reliabilities always as high as or higher than those obtained from VR1.

Author contribution statement Author contribution statement: DM conducted all simulations and analyses, devised Methods effLD, adjLD and RG, and wrote the manuscript. FT supported the development of the shrinkage concept, contributed software to conduct the simulations and revised the manuscript. AEM initiated and guided through the study, did the algebra of the 'Method RM' part of the manuscript and revised the manuscript.

Conflict of interest The authors declare no conflict of interest associated with this study.

Ethical standards The authors declare that ethical standards are met, and all the experiments comply with the current laws of the country in which they were performed.

Appendix

In this appendix, we describe Methods effLD and RM in detail.

Method effLD

In order to account for the genetic variance explained by markers beyond the ones immediately adjacent to QTL, we devised a measure for effective LD (LD_{eff}). Because QTL genotypes are generally unobservable, we use marker loci as a proxy.

Suppose that M biallelic markers are located on a chromosomal segment where p_i is the estimated allele frequency (of the major allele) at the ith marker. The LD between marker i and j can be computed according to Hill and Robertson (1968) as

$$r_{ij}^2 = \frac{\left(p_{ij} - p_i p_j\right)^2}{p_i p_j (1-p_i)\left(1-p_j\right)}, \tag{7}$$

where p_{ij} is the joint probability of the major allele occurring at both marker loci i and j. LD_{eff} is then calculated as follows. For each chromosome,

1. compute p_{ij} for all marker pairs as $p_{ij} = r_{ij}\sqrt{p_i p_j (1-p_i)\left(1-p_j\right)} + p_i p_j$
2. compute the covariance matrix $\Sigma = \left\{\Sigma_{ij}\right\}$ by solving the equations $\Phi\left(z(p_i), z(p_j); \Sigma_{ij}\right) = p_{ij}$ for Σ_{ij} for all marker pairs, where Φ is the cumulative distribution function of the standard bivariate normal distribution with mean zero and covariance Σ_{ij} and $z(p_i)$ refers to the p_ith quantile of the univariate standard normal distribution (Montana 2005).
3. compute the conditional variance for each locus i, given all others, as $\sigma_i = \Sigma_{\mathbf{i},\mathbf{i}} - \Sigma_{\mathbf{i},-\mathbf{i}}\Sigma^{-1}_{-\mathbf{i},-\mathbf{i}}\Sigma_{-\mathbf{i},\mathbf{i}}$. Here, the subscript $\mathbf{i}$ denotes the ith row or column, whereas $-\mathbf{i}$ denotes all but the ith row or column. Considering now the ith locus as a QTL, we imagine a hypothetical marker locus h in the proximity that would effectively lead to the same conditional variance at the ith locus.
4. compute $p^*_{ih} = \Phi\left(z(p_i), 0; \sqrt{1-\sigma_i}\right)$
5. compute the effective LD for each locus i of all loci (L) as

$$\mathrm{LD}_{\mathrm{eff}} = \sum_{i=1}^{L} \frac{\left(p^*_{ih} - 0.5p_i\right)^2}{0.5p_i(1-p_i)(1-0.5)} \tag{8}$$

6. take the average across all loci on the same chromosome

Finally, take the average across all chromosomes. Intuitively, LDeff would be the average coefficient of LD that would be observed between a QTL and a hypothetical marker with 0.5 allele frequency that would reduce the variance of the QTL genotype from $\Sigma_{\mathbf{i},\mathbf{i}}$ to σ_i.

Method RM

We use the model and notation of Dekkers (2007)

$$Y_i = G_i + E_i = \widehat{Q}_i + e_i + R_i + E_i, \tag{9}$$

where the phenotypic value Y_i of the ith individual is decomposed into its genetic value G_i and an environmental deviate E_i. The genetic value is further partitioned into QTL effect Q_i that is associated with marker through LD and effects R_i that is independent of markers. The effects Q_i can be further subdivided into a prediction $\widehat{Q}_i$ and a prediction error e_i, both being uncorrelated with one another.

A selection index combining phenotypic data and GEBVs can be constructed as $\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}$, e.g., Lande and Thompson (1990), where

$$\mathbf{G} = \begin{pmatrix} \mathrm{cov}(\hat{Q}_i, G_i) \\ \mathrm{cov}(Y_i, G_i) \end{pmatrix} \quad \text{and} \quad \mathbf{P} = \begin{pmatrix} \mathrm{var}(\hat{Q}_i); & \mathrm{cov}(\hat{Q}_i, Y_i) \\ \mathrm{cov}(Y_i, \hat{Q}_i); & \mathrm{var}(Y_i) \end{pmatrix}. \tag{10}$$

Without loss of generality, we assume $\sigma^2_G = \mathrm{var}(G_i) = 1$ and $\sigma^2_P = \mathrm{var}(P_i) = \frac{1}{h^2}$. Also, let $q^2 = \mathrm{var}(Q_i)$ be the proportion of variance contributed by QTL that are in LD with markers. Then

$$r\left(\widehat{Q}_i, Q_i\right) = r_{\widehat{Q}_i} = \frac{\mathrm{cov}(\widehat{Q}_i, Q_i)}{\sigma_{\widehat{Q}_i}\sigma_{Q_i}} = \frac{\sigma_{\widehat{Q}_i}}{\sigma_{Q_i}}, \tag{11}$$

where the last equality follows from the uncorrelatedness of the predictor $\widehat{Q}_i$ with the model residual e_i. Thus, $r_{\widehat{Q}_i} = \frac{\sigma^2_{\widehat{Q}_i}}{\sigma^2_{Q_i}}$

is the proportion of genetic variance contributed by Q_i that is explained by the GEBV $\widehat{Q}_i$. Assuming $r\left(\widehat{Q}_i, R_i\right) = 0$, we obtain

$$r\left(\widehat{Q}_i, G_i\right) = \frac{\text{cov}(\widehat{Q}_i, G_i)}{\sigma_{\widehat{Q}_i}\sigma_{G_i}} = \frac{\sigma_{Q_i}\text{cov}(\widehat{Q}_i, G_i)}{\sigma_{\widehat{Q}_i}\sigma_{Q_i}\sigma_{G_i}} = qr_{\widehat{Q}_i}. \quad (12)$$

With this, we obtain $\text{cov}(\widehat{Q}_i, G_i) = q^2 r^2_{\widehat{Q}_i}$. Since $\text{cov}(Y_i, G_i) = 1$, we have

$$\mathbf{G} = \begin{pmatrix} \text{cov}(\hat{Q}_i, G_i) \\ \text{cov}(Y_i, G_i) \end{pmatrix} = \begin{pmatrix} q^2 r^2_{\hat{Q}_i} \\ 1 \end{pmatrix} \quad (13)$$

Further, we have $\text{var}(\widehat{Q}_i) = q^2 r^2_{\widehat{Q}_i}$, $\text{var}(P_i) = \frac{1}{h^2}$. Assuming that $\widehat{Q}_i$ and E_i are uncorrelated, i.e., $r\left(\widehat{Q}_i, E_i\right) = 0$, we have $\text{cov}(\widehat{Q}_i, P_i) = \text{cov}(\widehat{Q}_i, G_i) = q^2 r^2_{\widehat{Q}_i}$. Hence,

$$\mathbf{P} = \begin{pmatrix} \text{var}(\hat{Q}_i); & \text{cov}(\hat{Q}_i, Y_i) \\ \text{cov}(Y_i, \hat{Q}_i); & \text{var}(Y_i) \end{pmatrix} = \begin{pmatrix} q^2 r^2_{\hat{Q}_i}; & q^2 r^2_{\hat{Q}_i} \\ q^2 r^2_{\hat{Q}_i}; & \frac{1}{h^2} \end{pmatrix} \quad (14)$$

By multiplying $\mathbf{P}^{-1}$ and $\mathbf{G}$, we obtain

$$b_1 = \frac{1 - h^2}{1 - h^2 q^2 r^2_{\widehat{Q}_i}} \quad \text{and} \quad b_2 = \frac{h^2 - h^2 q^2 r^2_{\widehat{Q}_i}}{1 - h^2 q^2 r^2_{\widehat{Q}_i}}, \quad (15)$$

In particular, we have

$$\frac{b_1}{b_2} = \frac{\frac{1}{h^2} - 1}{1 - q^2 r^2_{\widehat{Q}_i}} = \frac{\frac{1}{h^2} - 1}{1 - r^2_{MG}}. \quad (16)$$

This is equivalent to Eq. 3 in Lande and Thompson (1990). The quantity $q^2 r^2_{\widehat{Q}_i}$ is equal to r^2_{MG} in Dekkers (2007), which is the proportion of genetic variance that is explained by the GEBV. In practice, this parameter can be estimated using cross-validation as the squared predictive ability. In particular, we used fivefold cross-validation with five replications to estimate r^2_{MG} from the training set. The assumptions $r\left(\widehat{Q}_i, R_i\right) = 0$ and $r\left(\widehat{Q}_i, E_i\right) = 0$ are obviously not fulfilled with finite population sizes, as was validated by means of simulation.

References

Astle W, Balding DJ (2009) Population structure and cryptic relatedness in genetic association studies. Stat Sci 24(4):451–471. doi:10.1214/09-STS307. http://projecteuclid.org/euclid.ss/1271770342,arXiv:1010.4681v1

Bernardo R, Yu J (2007) Prospects for genomewide selection for quantitative traits in Maize. Crop Sci 47(3):1082. doi:10.2135/cropsci2006.11.0690. https://www.crops.org/publications/cs/abstracts/47/3/1082

de Los Campos G, Hickey JM, Pong-Wong R, Daetwyler HD, Calus MPL (2013) Whole-genome regression and prediction methods applied to plant and animal breeding. Genetics 193(2), pp. 327–45. doi:10.1534/genetics.112.143313. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3567727&tool=pmcentrez&rendertype=abstract

Dekkers JCM (2007) Prediction of response to marker-assisted and genomic selection using selection index theory. J Anim Breed Genet 124(6):331–41. doi:10.1111/j.1439-0388.2007.00701.x. http://www.ncbi.nlm.nih.gov/pubmed/18076470

Endelman JB (2011) Ridge regression and other kernels for genomic selection with R Package rrBLUP. Plant Genome J 4(3):250. doi:10.3835/plantgenome2011.08.0024. https://www.crops.org/publications/tpg/abstracts/4/3/250

Endelman JB, Jannink JL (2012) Shrinkage estimation of the realized relationship matrix. G3 2(11):1405–13. doi:10.1534/g3.112.004259. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3484671&tool=pmcentrez&rendertype=abstract

Frisch M, Melchinger AE (2007) Variance of the parental genome contribution to inbred lines derived from biparental crosses. Genetics 176(1):477–88, doi:10.1534/genetics.106.065433. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1893034&tool=pmcentrez&rendertype=abstract

Goddard ME, Wray NR, Verbyla K, Visscher PM (2009) Estimating effects and making predictions from genome-wide marker data. Stat Sci 24(4):517–529. doi:10.1214/09-STS306. http://projecteuclid.org/euclid.ss/1271770346,arXiv:1010.4710v1

Goddard ME, Hayes BJ, Meuwissen THE (2011) Using the genomic relationship matrix to predict the accuracy of genomic selection. J Anim Breed Genet 128(6):409–21, doi:10.1111/j.1439-0388.2011.00964.x. http://www.ncbi.nlm.nih.gov/pubmed/22059574

Habier D, Fernando RL, Dekkers JCM (2007) The impact of genetic relationship information on genome-assisted breeding values. Genetics 177(4):2389–97. doi:10.1534/genetics.107.081190. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2219482&tool=pmcentrez&rendertype=abstract

Habier D, Fernando RL, Garrick DJ (2013) Genomic BLUP decoded: a look into the black box of genomic prediction. Genetics 194(3):597–607. doi:10.1534/genetics.113.152207. http://www.ncbi.nlm.nih.gov/pubmed/23640517

Hayes B, Goddard M (2010) Genome-wide association and genomic selection in animal breeding. Genome 53(11): 876–83. doi:10.1139/G10-076. http://www.ncbi.nlm.nih.gov/pubmed/21076503

Hayes BJ, Bowman PJ, Chamberlaina J, Goddard ME (2009) Invited review: genomic selection in dairy cattle: progress and challenges. J Dairy Sci 92(2):433–43. doi:10.3168/jds.2008-1646. http://www.ncbi.nlm.nih.gov/pubmed/19164653

Henderson CR (1973) Sire evaluation and genetic trends. J Anim Sci, pp 10–41

Hill W, Robertson A (1968) Linkage disequilibrium in finite populations. Theor Appl Genet 38(6):226–231. http://link.springer.com/article/10.1007/BF01245622

Hill WG (2010) Understanding and using quantitative genetic variation. Philos Trans R Soc Lond Ser B Biol Sci 365(1537);73–85. doi:10.1098/rstb.2009.0203. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2842708&tool=pmcentrez&rendertype=abstract

Hill WG, Weir BS (2011) Variation in actual relationship as a consequence of Mendelian sampling and linkage. Genet Res 93(1):47–64. doi:10.1017/S0016672310000480. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3070763&tool=pmcentrez&rendertype=abstract

Kang HM, Zaitlen Na, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E (2008) Efficient control of population structure in model organism association mapping. Genetics 178(3):1709–23. doi:10.1534/genetics.107.080101. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2278096&tool=pmcentrez&rendertype=abstract

Lande R, Thompson R (1990) Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124(3):743–56. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1203965&tool=pmcentrez&rendertype=abstract

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits, 1st edn. Sinauer Associates, Sunderland

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157(4):1819–1829. http://www.genetics.org/content/157/4/1819.abstract

Montana G (2005) HapSim: a simulation tool for generating haplotype data with pre-specified allele frequencies and LD coefficients. Bioinformatics (Oxford, England) 21(23): 4309–11, doi:10.1093/bioinformatics/bti689. http://www.ncbi.nlm.nih.gov/pubmed/16188927

Powell JE, Visscher PM, Goddard ME (2010) Reconciling the analysis of IBD and IBS in complex trait studies. Nat Rev Genet 11(11): 800–5. doi:10.1038/nrg2865. http://www.ncbi.nlm.nih.gov/pubmed/20877324

R Core Team (2014) R: a language and environment for statistical computing. http://www.r-project.org/

Riedelsheimer C, Melchinger AE (2013) Optimizing the allocation of resources for genomic selection in one breeding cycle. TAG Theoret Appl Genet 126(11):2835–48. doi:10.1007/s00122-013-2175-9. http://www.ncbi.nlm.nih.gov/pubmed/23982591

Riedelsheimer C, Technow F, Melchinger AE (2012) Comparison of whole-genome prediction models for traits with contrasting genetic architecture in a diversity panel of maize inbred lines. BMC genomics 13(1):452. doi:10.1186/1471-2164-13-452. http://www.mendeley.com/research/comparison-of-whole-genome-prediction-models-for-traits-with-contrasting-genetic-architecture-in-a-d-1/

Searle SR, Casella G, McCulloch CE (1992) Variance components, 1st edn. Wiley-Interscience, Hoboken

Smith JSC, Hussain T, Jones ES, Graham G, Podlich D, Wall S, Williams M (2008) Use of doubled haploids in maize breeding: implications for intellectual property protection and genetic diversity in hybrid crops. Mol Breed 22(1):51–59. doi:10.1007/s11032-007-9155-1. http://link.springer.com/10.1007/s11032-007-9155-1

Technow F (2013) hypred: simulation of genomic data in applied genetics. http://cran.r-project.org/web/packages/hypred/

VanRaden PM (2008) Efficient methods to compute genomic predictions. J Dairy Sci 91(11):4414–23. doi:10.3168/jds.2007-0980. http://www.ncbi.nlm.nih.gov/pubmed/18946147

Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden Pa, Heath AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM (2010) Common SNPs explain a large proportion of the heritability for human height. Nat Genet 42(7):565–9. doi:10.1038/ng.608. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3232052&tool=pmcentrez&rendertype=abstract