

Breeding value estimation combining QTL and polygenic information

H.A. Mulder, M.P.L. Calus, R.F. Veerkamp

INTRODUCTION

- Challenge is to deal with genotyped and ungenotyped animals in marker-assisted breeding value estimation.
- Two possible solutions: 1) include ungenotyped animals into marker-assisted breeding value estimation using predicted genotypes/haplotypes or 2) blend for genotyped animals conventional EBV with QTL effects using selection index.

OBJECTIVES

- To compare accuracy of marker-assisted breeding value estimation using predicted and observed haplotypes (Mablup) with blending conventional EBV and QTL-effects for genotyped animals using selection index.

MATERIAL AND METHODS

Prediction of haplotypes

- Translation of 4 SNP-markers into haplotypes, e.g. 1111, 1112, 2111, 2112, etc.
- Prediction of number of haplotype copies (n_{hc}) for genotyped and ungenotyped animals using mixed model equations:

$$\begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\mathbf{M} \\ \mathbf{M}'\mathbf{1} & \mathbf{M}'\mathbf{M} + \lambda\mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \mu_{n_{hc}} \\ \mathbf{d}_y \\ \mathbf{d}_x \end{bmatrix} = \begin{bmatrix} \mathbf{1}'\mathbf{n}_{hc,y} \\ \mathbf{M}'\mathbf{n}_{hc,y} \end{bmatrix} \quad (1)$$

where $\mathbf{1}$ is a vector of ones, $\mathbf{M}$ is a design matrix linking $\mathbf{d}$ with $\mathbf{n}_{hc,y}$, $\mathbf{A}^{-1}$ is the inverse additive genetic relationship matrix, λ is the variance ratio of residual variance and additive genetic variance for n_{hc} allowing for a small proportion of genotyping errors $\lambda = \sigma_{e_{n_{hc}}}^2 / \sigma_{a_{n_{hc}}}^2 = 0.01/0.99$, $\mathbf{d}$ is a vector with the EBV for n_{hc} with $\mathbf{d}_y$ for genotyped animals and $\mathbf{d}_x$ for ungenotyped animals, $\mathbf{n}_{hc,y}$ is a vector with observed n_{hc} of genotyped animals and is set to missing for ungenotyped animals.

Marker-assisted breeding value estimation

$$y = \mu + u_{pol} + \sum_i^n (n_{hc_i} \times h_i) + e \quad (2)$$

where y is the phenotype, μ is the overall mean and modeled as a fixed effect, u_{pol} is the random polygenic EBV, $n_{hc_i} = \hat{\mu}_{n_{hc_i}} + \hat{d}_i$, which is the predicted number of haplotype copies for haplotype i , h_i is the random regression coefficient for haplotype i and e is the residual. The variances used are $\sigma_{h_i}^2 = 0.5\sigma_{A_{qtl}}^2$ and $\sigma_{u_{pol}}^2 = \sigma_{A_{pol}}^2$. MiX99 was used to perform breeding value estimation.

Selection index to blend conventional EBV with estimated haplotype effects.

- Conventional EBV are calculated ($\hat{u}_{con}$) using equation 2 without haplotype effects.
- For genotyped animals QTL-EBV effects ($\hat{u}_{QTL}$) are taken from the full model in equation 2.
- The total EBV is calculated as: $\hat{u}_{tot} = b_{con}\hat{u}_{con} + b_{QTL}\hat{u}_{QTL}$. Selection index weights were calculated as $\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}$ with $\mathbf{P} = \begin{bmatrix} r_{con}^2 & r_{con}^2 q^2 \\ r_{con}^2 q^2 & q^2 \end{bmatrix}$ and $\mathbf{G} = \begin{bmatrix} r_{con}^2 \\ q^2 \end{bmatrix}$, where r_{con}^2 is the reliability of $\hat{u}_{con}$ and q^2 is the proportion of genetic variance explained by the haplotype (Ma-index pract). In another scenario, all covariances and variances in $\mathbf{P}$ and $\mathbf{G}$ were estimated from the simulated data (Ma-index theor).

Simulation

- One bi-allelic additive QTL on a chromosome with 4 bi-allelic flanking markers (0.5 CM) explaining 15% of genetic variance; rest of genetic variance is additive polygenic.
- 100 generations of random mating with 50 sires and 50 dams to create linkage disequilibrium ($N_e = 100$).
- Generation 101-104 consist of 1000 males and 1000 females all with a phenotypic record. Generation 105 are juveniles and do not have a phenotypic record yet.
- Conventional BLUP selection up to generation 104; 50 sires and 250 dams are selected each generation.
- Marker-assisted breeding value estimation in generation 105; sires in generation 101-104 are genotyped and males in generation 105.

RESULTS

- The accuracy of total EBV with practical index weights (Ma-index pract) yielded lower accuracy than marker-assisted breeding value estimation (Mablup) due to too much emphasis on QTL-EBV (Table 1 and 2).
- With theoretical weights (Ma-index theor) the accuracy was slightly higher than with marker-assisted breeding value estimation (Mablup), but this is hard to achieve without knowing the true breeding values (Table 1).

Table 1. Summary statistics of accuracy of total EBV for genotyped juveniles with conventional breeding value estimation (Conblup), marker-assisted breeding value estimation (Mablup) or with blending of conventional EBV and QTL-EBV based on 4-marker haplotypes.

	Conblup	Mablup	Ma-index theor	Ma-index pract
Mean	0.184	0.275	0.276	0.247
SD	0.057	0.082	0.077	0.086
Min	0.021	0.024	0.022	-0.016
Max	0.309	0.439	0.422	0.434
Proportion of replicates with lower accuracy than Conblup		0.12	0.00	0.18

Table 2. Selection index weights (standard error between brackets) for conventional EBV and QTL-EBV using approximated (co)variances (Ma-index pract) or estimated (covariances) (Ma-index theor).

Method	Selection index weights			
	Conventional EBV		QTL-EBV	
Ma-index theor	0.93	(0.03)	0.62	(0.04)
Ma-index pract	0.92	(0.01)	0.97	(0.00)

CONCLUSION

- Blending QTL-EBV and conventional EBV requires careful calculation of selection index weights, whereas marker-assisted breeding value estimation does not require selection index weights.
- Marker-assisted breeding value estimation yields higher accuracy than blending conventional EBV with QTL-effects using a selection index based on realistic weights calculated from reliabilities and proportion of genetic variance explained by the haplotype.

Affiliation

Animal Sciences Group,
Wageningen University and Research Centre,
Animal Breeding and Genomics Centre,
PO Box 65, 8200 AB Lelystad,
The Netherlands
Tel. : +31-320-238363

Email: Herman.Mulder@wur.nl

www.asg.wur.nl



ANIMAL SCIENCES GROUP
WAGENINGEN UR