58th Annual Meeting of the European Association for Animal Production, Dublin, Ireland, August 26th-29th 2007
Session 38 "Free communications on Animal Genetics", Abstract number 0541

Benefits of using phenotypic observations of dams for estimating QTL variance components and MA-BLUP EBV

S. Neuner[1*], R. Emmerling[1], G. Thaller[2] und K.-U. Götz[1]
[1]Bavarian State Research Centre for Agriculture, Institute of Animal Breeding, D-85580 Grub
[2]Christian-Albrechts-University, Institute of Animal Breeding and Husbandry, D-24098 Kiel

Introduction

Reliable estimates for variance components in QTL-models are important for fine mapping experiments and MA-BLUP evaluations in breeding programs using marker assisted selection (MAS). In cattle populations in many cases only a small fraction of the population will be genotyped at genetic markers. As only these genotyped animals provide information for QTL specific evaluations a 'two step approach' was used to estimate QTL variance components and MA-BLUP EBV, e.g. by Liu et al. (2004) or Druet et al. (2006). At first a polygenic animal model evaluation is conducted for the entire population to estimate pre-corrected phenotypes (step 1). These estimates are further used as observations in an MA-BLUP model for the genotyped animals (step 2). Since the pre-corrected phenotypes may represent different amounts of information, the problem of weighting the phenotypes in the second step arises.

This study examines the use of daughter yield deviations (DYD) of bulls and yield deviations (YD) of cows as observations in MA-BLUP models. Various models were calculated to detect the best combination of phenotypic measures (DYD, YD) and weighting factors for the estimation of QTL variance components and MA-BLUP EBV.

In some situations estimated QTL effects may be small and accuracies of animals of interest will not increase in MA-BLUP Models compared to conventional animal models. Therefore it was analyzed if MA-BLUP offers a better selection among paternal half-sibs.

Material and Methods

A stochastic simulation model was applied to evaluate different alternatives for estimating QTL variance components and MA-BLUP EBV in a two step approach. Each simulation cycle is divided into two phases: data generation and analysis of the simulated data sets.

Data Generation

In the simulation a dairy cattle breeding scheme with progeny testing and use of bulls for the generation of second crop daughters was simulated to generate the data. The time horizon for the simulation was 16 years which equals approximately the time spans for collection of genotypic data in real research projects.

In the current simulation study a single trait model for 305-day milk yield with a heritability of 0.36 was assumed. The overall breeding value of each animal is the sum of a 'residual polygenic breeding value' and a 'QTL breeding value'. One biallelic QTL with an allele frequency for the favourable allele of 0.5 is linked to a genetic marker. Recombination rate between QTL and maker locus was fixed at 0.01 and also the other assumptions the marker locus were very optimistic: 100 different alleles following a uniform distribution (PIC=0.9899 and PFIM=0.9799). Calculations were done for different proportions of the QTL-variance as compared to the overall genetic variance (QVR) in the trait investigated: 0, 5, 10, 20, 30, 40, 50 and 90%.

---

*Corresponding author: Stefan.Neuner@LfL.bayern.de

Analysis of simulated data sets

QTL variance components and MA-BLUP EBV were estimated in a two step approach. In the first step a classical animal model (AM) evaluation was done for the entire population to estimate DYD for progeny tested bulls and YD for cows in milk. Observations in this step were phenotypic records of cows. Usually all animals are included in genetic evaluations of dairy cattle, however, only a small fraction of animals might be genotyped at genetic markers. These animals are most likely proven bulls, bull dams and selection candidates for progeny testing. As only genotyped animals can provide information for the estimation of QTL variance components and MA-BLUP EBV, the second evaluation step was only applied to this genotyped subset of the population. Observations in step 2 are DYD and YD calculated in step 1.

The MA-BLUP model used for evaluations equals that of Fernando and Grossmann (1989):

$$y_i = \mu + u_i + v_i^p + v_i^m + e_i$$

where y_i is the record (YD for dams and DYD for sires) of individual i, u_i is the residual polygenic effect of individual i, v_i^p and v_i^m are the paternal and maternal gametic effects of individual i and e_i is the residual.

Additionally, a classical animal model was calculated for this reduced data set. In this model there was only one predictor for the overall animal effect and it is henceforth denoted as 'animal model on MA-BLUP records' (AM-MA). For the AM-MA and the MA-BLUP model various combinations of phenotypic measures (DYD, YD) and weighting factors are calculated. The evaluations are divided into blocks A and B that are characterised by the phenotypic measures used. Block A is similar to the German MA-BLUP system (Liu et al., 2004) where only DYD (DYD Models) are used, whereas in block B DYD and YD (DYD-YD Model) are used together as in the French MA-BLUP system (Boichard et al., 2002). Within the blocks different weighting factors as described in the literature were applied to DYD: no weighting, variance of DYD (Bennewitz et al., 2004), effective daughter contributions (EDC)(Fikse and Banos, 2001; Liu et al., 2004; Szyda et al., 2005) and daughter equivalents (DE) (Van Raden and Wiggans, 1991, Druet et al., 2006). YD were not weighted since each cow had only one record in the current study (following Druet, 2006, personal communication).

The results for the different models are presented in terms of deviations and correlations between estimated and simulated parameters.

Analyzing the benefit of MA-BLUP for a better selection among paternal half-sibs was done for a simulated data set with a QVR of 0.10. Progenies of heterozygous sires were divided into two groups: those inheriting the favourable allele, and those inheriting the unfavourable allele. Then their simulated BV, their AM EBV and MA-BLUP EBV were compared. The same was done for homozygous sires, if there is a difference which allele of the sire, the first or the second, was inherited to the progenies.

Results

Variance Component Estimation

Criteria to evaluate the variance component estimation are the absolute values for the estimated components and the estimated QVR. The results are consistent for all simulated QVR. As an example, the results of a variance component estimation for a simulated QTL ratio of 30 % are shown in table 1. Presented results are the averages for 25 replicates. Results clearly show that the use of DYD-Models is superior to that of DYD-YD models. The best results were obtained by using 'DE' or 'EDC' as weighting factors. In DYD-YD MA-BLUP models estimated variance components were less accurate. Best results were obtained when the weighting factor 'variance of DYD' was applied. Weighting of information in block A and B was essential. No weighting of daughter information led to a gross overestimation of variance components.

Table 1: Simulated and estimated values of variance components for a QVR of 30 %

30 % QTL		Additive Genetic Variance		Residual Polygenic Variance		QTL Variance		QTL Variance Ratio (QVR)	
Block	**weighting factor**	simulated	MA-BLUP model	simulated	MA-BLUP model	simulated	MA-BLUP model	simulated	MA-BLUP model
A (DYD)	no	260100	310276	182070	226130	78030	84146	0,30	0,271
	DE	260100	267523	182070	189122	78030	78401	0,30	0,293
	var(DYD)	260100	285015	182070	205123	78030	79891	0,30	0,280
	EDC	260100	266468	182070	187311	78030	79157	0,30	0,297
B (DYD-YD)	no	260100	332366	182070	240387	78030	91979	0,30	0,277
	DE	260100	296924	182070	221844	78030	75080	0,30	0,253
	var(DYD)	260100	251005	182070	169885	78030	81120	0,30	0,323
	EDC	260100	295643	182070	220385	78030	75258	0,30	0,255

Accuracy of MA-BLUP EBV

For each group of animals (bulls, cows and young bulls) correlations between simulated and estimated breeding values were calculated for all combinations of block x weight x model. As for the estimation of variance components, the findings are consistent for all simulated QVR ratios. Correlations between EBV and simulated BV for a QTL ratio of 30% are shown in table 2. Presented results are the averages of 25 replicates.

Table 2: Correlations between simulated and estimated breeding values for various models using different phenotypic measures and weighting factors for QVR of 30%

30 % QTL		Progeny Tested Bulls			Cows			Young Bulls		
Block	**weighting factor**	AM	AM-MA	MA-BLUP	AM	AM-MA	MA-BLUP	AM	AM-MA	MA-BLUP
A (DYD)	no	0,919	0,907	0,910	0,744	0,467	0,491	0,551	0,488	0,547
	DE	0,919	0,908	0,911	0,744	0,468	0,491	0,551	0,488	0,547
	var(DYD)	0,919	0,907	0,910	0,744	0,467	0,491	0,551	0,488	0,547
	EDC	0,919	0,908	0,911	0,744	0,468	0,491	0,551	0,488	0,547
B (DYD-YD)	no	0,919	0,838	0,843	0,744	0,682	0,69	0,551	0,507	0,567
	DE	0,919	0,910	0,914	0,744	0,708	0,718	0,551	0,543	0,605
	var(DYD)	0,919	0,911	0,915	0,744	0,708	0,718	0,551	0,546	0,608
	EDC	0,919	0,910	0,914	0,744	0,708	0,718	0,551	0,543	0,605

First of all, our results show that weighting is also essential for the estimation of MA-BLUP EBV. Furthermore the differences of the correlations are pretty small when a weighting is applied. Secondly, correlations of proven bulls are nearly unaffected whether YD are included in the evaluation models or not. In contrast to proven bulls correlations for bull dams and young bulls in the AM and the MA-BLUP model highly depend on whether YD are included in the phenotypic measures (block B) or not (block A). If YD are not included, there are no observations for bull dams and their EBV are only based on pedigree information in the MA-BLUP data. Since EBV of young bulls are calculated based on their parents' EBV, the advantage of using YD is obvious. The difference in correlations between DYD-YD models and DYD models is about 0.06 for young bulls and 0.25 for the cows. Thirdly, if only DYD are used in MA-BLUP evaluations (block A) a QTL of at least 30% QVR is needed to get accuracies for young bulls in MA-BLUP models that are close to the values found from the classical AM. In DYD-YD MA-BLUP models (block B), QTL ratios ≥10% (results not shown) are required to obtain higher accuracies for young bulls than in ordinary genetic evaluations.

Lower correlations for all animal groups in AM-MA than for AM for all scenarios (table 2) indicate that there is a loss of information if two step approaches are applied.

Half-sib selection

Following the results for a QVR of 0.10 (results not shown) accuracies for young bulls in the AM and the DYD-YD MA-BLUP are approximately the same. In this situation the question arises if MA-BLUP is also reasonable for low to moderate QTL effects. Therefore progenies of heterozygous sires were analysed. Progenies inheriting the favourable QTL allele had on average 216kg higher real breeding values, 30kg higher estimated breeding values in DYD MA-BLUP models, and 46kg higher estimated breeding values in DYD-YD MA-BLUP models. To control the results we also looked at progenies of non segregating families. The expectation of no differences in their BV and EBV due to inheritance at the QTL was fulfilled.

Discussion

Our results show the importance of weighting the daughter information in DYD and DYD-YD models. If there is no weighting different numbers of daughters per bull are not accounted for. This leads to inaccurate estimates of variance components and lower accuracies of MA-BLUP EBV. Following the results of the current study, the correct choice of the weighting factor is more important for the estimation of variance components than for the estimation of MA-BLUP EBV. For the latter, the ratio of the applied variances is more important than their absolute values if only accuracies are considered. The fact that the highest accuracies are always obtained if the estimated variance components and ratios are as close as possible to the simulated parameters shows the importance of correctly chosen weights. Weighting is especially difficult in DYD-YD models because the scales of the two types of information are not identical.

Applying two step approaches for MA-BLUP Models always causes a loss of information. As only a small fraction of the population is included in the model, several relationships among animals get lost in MA-BLUP data sets. The information content for proven bulls decreases only marginally because DYD are estimated very accurately from many daughters. For cows only relationship information and, in case of DYD-YD models, their own phenotypic records are taken as sources of information. As a consequence, this leads to a higher impact of missing relationships for cows as compared to bulls. Therefore, even in AM-MA DYD-YD models accuracies of cows are about 0.04 lower than in AM. In consequence the loss of information due to the 2-step-approach and to missing phenotypes of bull dams in DYD models has to be overcome by an additional source of information: QTL information.

Analyses of MAS applied to practical breeding programs describe the increase in accuracies of young bulls (Liu et al., 2004; Druet et al., 2005). Liu et al. (2004) described the increase in accuracy of young German Holstein bulls if two QTL are included as random effects and DGAT1 as a fixed effect in MA-BLUP evaluation. Correlations increased from 0.45 in the AM-MA to 0.65 in the MA-BLUP model, but the main effect was due to DGAT1. More important than comparing results of AM-MA and MA-BLUP Models is the superiority of accuracies from MA-BLUP models over traditional AM. Druet et al. (2005) investigated this for the French MAS program. In the French MAS system between 40-50% of the variance for all milk traits is explained by 3 to 5 QTL. Accuracies for milk yield EBV of young bulls increased from 0.47 to 0.55. Results of our analysis for a 40 % QTL show an increase from 0.58 to 0.68. Differences in the level can be explained by different heritabilities and different designs for the French breeding program and the one assumed in the simulation. A higher benefit through MAS in the simulation could be expected because only one QTL with advantageous properties was simulated.

The choice of whether DYD or DYD-YD models are used, depends on the intention of the research: fine mapping or estimation of MA-EBV for MAS. While fine mapping is especially interested in correct estimates for variance components, MAS requires an increase in accuracies for MA-EBV of animals without own phenotypic or progeny information. Therefore, MA-BLUP models for MAS should include both DYD and YD to ensure the highest possible efficiency of selection.

Following our results MA-BLUP also is a useful tool for the selection among half sibs, even if the QTL effects seem to be to small when we look at their accuracies. The accuracies of young bulls for a QVR of 0.10 were 0.542 in the AM, 0.486 in the MA-BLUP DYD model and 0.545 in the MA-BLUP DYD-YD model. That means there is only a little increase in accuracy if we move from the AM to the MA-BLUP DYD-YD model. But our results show that a better distinction between progenies in segregating families due to QTL inheritance is possible. The effect used is the reduction of the within family variance between half-sibs. But even if benefits are possible for small QTL effects the costs should be justified.

Conclusions

To estimate QTL variance components in MA-BLUP models the use of DYD models is the most appropriate strategy. Weighting is essential and best results are obtained by using 'DE' or 'EDC' as weighting factors. In DYD-YD MA-BLUP models estimated variance components are less accurate.

MA-BLUP evaluations that do not make use of phenotypic data for bull dams will only give benefits for QTL explaining more than 30% of the additive genetic variance. In DYD-YD models the data is still incomplete compared to conventional animal model evaluation. To outweigh the loss of information caused by the two step approach a 10% QTL is necessary. As a consequence of the results of this study MA-BLUP models used to estimate EBV for MAS should include DYD and YD to ensure that MAS improves selection efficiency even for moderate QTL effects.

Even if QTL effects are small MA-BLUP still offers benefits through a better selection among half-sibs in segregating families.

References

Bennewitz, J.; Reinsch, N.; Paul, S.; Looft, C.; Kaupe, B.; Weimann, C.; Erhardt, G.; Thaller, G.; Kuhn, Ch.; Schwerin, M.; Thomsen, H.; Reinhardt, F.; Reents, R. and Kalm, E. (2004): The DGAT1 K232A Mutation Is Not Solely Responsible for the Milk Production Quantitative Trait Locus on the Bovine Chromosome 14. J. Dairy Sci., 87: 431-442.

Boichard, D.; Fritz, S.; Rossignol, M. N.; Boscher, M. Y.; Malafosse, A. and Colleau, J. J. (2002): Implementation of Marker Assisted Selection in French Dairy Cattle Breeding. Proc. 7th World Congr. Genet. Appl. Livest. Prod., Montpellier, France, n° 22.03.

Druet, T.; Fritz, S.; Colleau, J.J.; Gautier, M.; Eggen, A.; Rossignol, M.N.; Boscher, M.Y.; Malafosse, A and Boichard, D. (2005): Genetic markers in breeding programs. 26th European Holstein and Red Holstein Conference, Prague 2005.

Druet, T.; Fritz, S.; Boichard, D. and Colleau, J. J., (2006): Estimation of Genetic Parameters for Quantitative Trait Loci for Dairy Traits in the French Holstein Population. J. Dairy Sci., 89: 4070-4076.

Fernando, R. L. and Grossman, M. (1989): Marker assisted selection using best linear unbiased prediction. Genet. Sel. Evol., 21: 467–477.

Fikse, W. F. and Banos, G. (2001): Weighting Factors of Sire Daughter Information in International Genetic Evaluations. J. Dairy Sci., 84: 1759-1767.

Liu, Z.; Reinhardt, F.; Szyda, J.; Thomsen, H. and Reents, R. (2004): A Marker Assisted Genetic Evaluation System for Dairy Cattle Using a Random QTL Model. Int. Bull Evaluation Service, Uppsala, Bulletin 32: 170-174.

Szyda, J.; Liu, Z.; Reinhardt, F. and Reents, R. (2005): Estimation of Quantitative Trait Loci Parameters for Milk Production Traits in German Holstein Dairy Cattle Population. J. Dairy Sci., 88: 356-367.

Van Raden, P. M. and Wiggans, G. R. (1991): Derivation, calculation, and use of national animal model information. J. Dairy Sci., 74: 2737–2746.