Influence of Dominance Effects on the Estimation of Direct Genomic Values in a Crossbreeding Design

W. Mekkawy[*], T. Roughsedge[†], G. Banos[††], E.A. Navajas[†], C-A. Duthie[†], R.M. Sawalha[†] and R. Roehe[†]

Introduction

Many of the commercial breeding programmes of beef cattle, sheep, pigs and poultry are using crossbreeding systems of two or more breeds. At present, mainly purebred information is used in genetic improvement programmes. Wei (1992) showed that inclusion of crossbreeding information will increase the overall selection response in the breeding programme. Genomic selection based on a dense single nucleotide polymorphisms (SNPs) panel has been found to substantially increase selection response (Solberg et al. 2008). The use of genomic selection in a crossbreeding scheme has been studied by Ibáñez-Escriche et al. (2009). However, in their study only additive genetic effects were considered. But crossbreeding performances are expected to be influenced by dominance effects. Therefore, the main objective of this study was to assess the effect of ignoring dominance effects on the estimation of the direct genomic values (DGV) when utilising crossbred information.

Material and methods

Simulation. Genomes of animal were simulated consisting of 10 chromosomes of equal length (1 Morgan each). Along each chromosome, 1,000 markers and 100 biallelic QTL were randomly distributed. For both markers and QTL, random mutations were sampled from a Poisson distribution with rate of 2.5×10^{-5} per locus per meiosis. A mutation switched the allele 1 to 2 and vice versa. Recombination events were also sampled from a Poisson distribution according to the different genetic distance between each pairs of SNPs. Both additive and dominance effects were sampled from a gamma distribution with shape parameter 4.2 and scale parameter 1.4. Each QTL effect was given a positive or a negative sign with probability 0.5. Two different breeds were simulated in which the base populations of the two breeds were homozygous in the opposite allele for all markers and QTL loci. For each breed a base population of 200 unrelated animals (100 males and 100 females) was simulated. Sires and dams were mated randomly for 1,000 generations to achieve a mutation drift balance. Three different levels of relatedness between breeds had been used (unrelated, distantly related or closely related). Unrelated breeds had no common origin, whereas distantly and closely related breeds were crossed 500 generations ago, and 50 generations ago of the 1,000 historical generations, respectively, and these crossbred animals were used

[*] Animal Production Department, Faculty of Agriculture, Ain Shams University, Hadaeq Shubra, Cairo 11241, Egypt.

[†] Scottish Agricultural College, West Mains Road, Edinburgh EH9 3 JG, United Kingdom.

[††] Aristotle University of Thessaloniki, GR-54124 Thessaloniki , Greece.

The authors would like to thank Defra and RERAD for funding the project.

as the common origin for the subsequent generations of the two separated lines. In generation 1,000, the population size of each breed was increased to 2,000 animals and the two breeds were crossed to form a training population (crossbred animals, generation 1,001). Thereafter, eight generations of purebred animals with population size of 2,000 for each line were simulated. Random mating assumed for all generations. The purebred animals in generation 1,008 were used for validation. For the training population phenotypes and genotypes for all simulated markers were available, whereas the validation populations had only genotypes of the simulated markers available.

Three different levels of direct heritabilities (h^2_a= 0.1, 0.3 or 0.5) and four different levels of dominance heritabilities (h^2_d = 0.0, 0.1, 0.3 or 0.5) were assumed. Residual effects were simulated from a standard normal distribution. Both residual and QTL effects were rescaled in the training population according to the desired heritabilities. 20 replicates for each case were used for the study.

Statistical analyses. The crossbred data sets were analyzed using BLUP method described by Meuwissen et al. (2001). This method assumes that all marker effects are sampled from the same normal distribution. Only the additive effects of the markers were included in the model as follow:

$$y = Xb + Zu + e$$

where y is the vector of observations, Xb is a vector of means, u is a vector of random additive genetic effects that associated with each marker genotype, Z is the incidence matrix of the markers additive genetic effects and e is the vector of random residual. The markers included in the analyses had minor allele frequency (MAF) $\geq$ 0.05. The estimated additive effects of the markers were used to calculate the estimated direct genomic values (DGV) of the animals in the validation populations. For each simulated data set we calculated the accuracy as the correlation between the simulated and estimated DGV of the animals in the validation populations and the bias of the estimation mean of the differences between simulated and estimated (DGV). Moreover, we calculated the expected (difference of mean of the estimated DGV of the selected animals from the mean of the estimated DGV of the whole population in the validation population) and actual selection response (difference of mean of the true DGV of the selected animals from the mean of the true DGV of the whole population in the validation population) of 10% of the parents selected by their DGV ignoring dominance.

Results and discussion

The results of the validation populations of the two breeds were almost identical so that we are presenting here only the results of one validation population (Table 1). The level of relatedness between the crossed breeds substantially influenced the accuracies i.e. accuracies increased with the increase in level of relatedness between breeds. Ignoring dominance effects decreased accuracy in unrelated and distantly related populations but only slightly when closely related breed were being crossed. With increase in dominance heritability, the accuracy was reduced, mainly in traits of low direct heritability, for which in the unrelated and distantly related population the accuracy was even around zero. In contrast, with high additive heritability (h^2_a=0.5) and close relationship between crossed breeds, the negative effect of ignoring dominance on the accuracies disappeared even in scenarios with high

dominance heritability. The bias was very close to zero when the two breeds were closely related, whereas for the other two levels of relatedness, higher biases were obtained. In most cases, the bias increased with the increase in the dominance heritability but without clear trend and mostly the bias was non-significantly different from zero. This increase in standard error of bias indicates that much more variation among replicates occurs if unrelated or distantly related populations are crossed.

Table 1: Estimates (± S.E.) of accuracy and bias for the different scenarios of the simulation

Scenario		Unrelated		Distantly related		Closely related	
h^2_a	h^2_d	Accuracy	Bias	Accuracy	Bias	Accuracy	Bias
0.1	0.0	0.34 ± 0.06	-0.04 ± 0.11	0.70 ± 0.03	-0.01 ± 0.06	0.86 ± 0.02	-0.01 ± 0.01
0.1	0.1	0.06 ± 0.08	-0.08 ± 0.10	0.47 ± 0.05	-0.31 ± 0.14	0.80 ± 0.03	0.00 ± 0.02
0.1	0.3	-0.01± 0.08	-0.03 ± 0.08	0.22 ± 0.08	-0.24 ± 0.11	0.82 ± 0.03	-0.01 ± 0.02
0.1	0.5	-0.13± 0.07	-0.11 ± 0.10	-0.01 ± 0.09	0.08 ± 0.18	0.76 ± 0.04	0.01 ± 0.03
0.3	0.0	0.46± 0.04	0.14 ± 0.14	0.81 ± 0.03	0.03 ± 0.09	0.92 ± 0.01	0.00 ± 0.02
0.3	0.1	0.42± 0.05	0.17 ± 0.17	0.72 ± 0.04	0.21 ± 0.16	0.91 ± 0.01	0.01 ± 0.02
0.3	0.3	0.21± 0.05	-0.10 ± 0.20	0.58 ± 0.06	0.00 ± 0.16	0.89 ± 0.02	0.01 ± 0.04
0.3	0.5	0.17± 0.07	-0.36 ± 0.20	0.40 ± 0.10	0.12 ± 0.15	0.86 ± 0.03	-0.04 ± 0.03
0.5	0.0	0.46± 0.03	0.44 ± 0.22	0.89 ± 0.01	-0.13 ± 0.12	0.91 ± 0.01	-0.04 ± 0.02
0.5	0.1	0.48± 0.04	0.23 ± 0.24	0.77 ± 0.02	0.11 ± 0.14	0.92 ± 0.01	-0.01 ± 0.03
0.5	0.3	0.36± 0.08	0.08 ± 0.20	0.66 ± 0.04	-0.06 ± 0.20	0.91 ± 0.01	-0.02 ± 0.03
0.5	0.5	0.25± 0.06	-0.19 ± 0.24	0.57 ± 0.07	0.62 ± 0.29	0.90 ± 0.01	0.01 ± 0.04

The expected selection responses increased with the increase in the degree of relatedness between the two breeds and the actual responses were overestimated when the breeds were unrelated, whereas actual responses were accurately predicted when the breeds were closely related (Figure 1). Ignoring dominance effects decreased the actual response from selection except for the scenarios with high additive heritabilities of the closely related breeds.

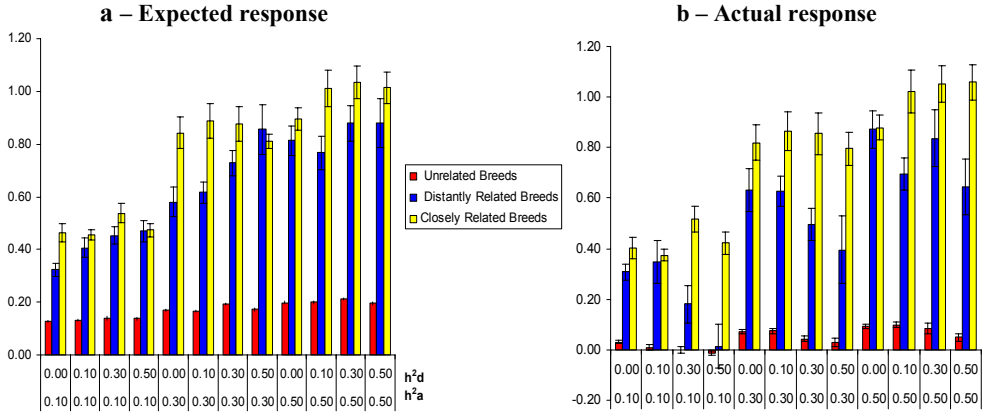

Figure 1: Expected and actual response to selection (± S.E.) for the different scenarios of the simulation

The very low accuracies and actual selection responses when the breeds were distantly related is at least partly due to the fact that the base populations of the two breeds were homozygous in the opposite allele for all markers and QTL and the low mutation rate limited the role of the recombination in breaking the linkage disequilibrium and after the 1,000 of historical generations most of the markers and QTL remained homozygous. Crossing the two different breeds created a high level of linkage disequilibrium but most of the markers and QTL were heterozygous (expression of the dominance) which may confound the estimation of the direct genetic effects of close markers.

Further studies are necessary to considered dominance effects in the genomic evaluation model for crossbred populations. In purebred populations, Toro and Varona (2009) showed that inclusion of dominance effects in the genomic evaluation could increase the selection response.

Conclusion

Crossing of unrelated populations substantially reduced the accuracy of the direct genomic breeding values and therefore specification of their relationship is necessary in order to use crossbred information for genomic selection efficiently. Inclusion of dominance effects in genomic evaluation models is important when crossbred breeds are unrelated or distantly related because in these cases ignoring dominance will substantially decrease the accuracies of DGV depending on the dominance heritability. In contrast, when the crossed breeds are closely related, ignoring dominance showed a low effect on the accuracy of DGV especially when the direct heritability is high.

References

Ibánẽz-Escriche, N., Fernando, R. L., Toosi, A. and Dekkers J.C.M. (2009) *Genet. Sel. Evol*., 41:12.

Meuwissen, T. H. E., Hayes, B. and Goddard, M.E. (2001). *Genetics* 157:1819–1829.

Solberg T.R., Sonesson, A.K., Woolliams, J.A. and Meuwissen T.H.E. (2008) *J. Anim*. Sci., 86:2447–2454.

Toro, M. A. and Varona, L. (2009) Conference on Next Generation Sequencing: Challenges and Opportunities

Wei, M. (1992) PhD thesis Wageningen Agri. Univ. Netherlands