The Effect of Using Genomic Breeding Values to Manage the Loss in Response to Selection Caused by Genotype by Environment Interactions

S. A. Clark[1], D. Brown[2] and J. H. J. van der Werf[1].
[1]University of New England, Armidale, NSW Australia [2]Animal Genetics and Breeding Unit, Armidale, NSW, Australia

ABSTRACT: Genotype by environment (GxE) interaction can reduce genetic gain because there is often insufficient information for accurate selection in each environment. Genomic selection may be a useful method to reduce the impact of GxE interaction as measurement in different environments may be propagated over more selection candidates. We compared the value of testing the performance of relatives and/or genomic breeding values across environments to increase genetic gain when GxE was present. Selection based on genomic breeding values (GEBVs), that were calibrated in the commercial environment and 60% accurate, increased genetic gain by 14% over recording 20 half siblings in that environment. When GEBVs were 80% accurate they were equivalent to testing 20 Progeny in the commercial environment. Even when the breeding objective was for performance in two environments, it was more beneficial to use a GEBV based on information from the commercial environment.

Keywords: Genotype; Environment; Interaction; Genomics

Introduction

In livestock breeding programs it is important to utilize information from various environments. Often animals are recorded and selected in an environment that is different to where final commercial production occurs. While for most traits there is little evidence for genotype by environment interaction (G x E), for some traits, animals may perform differently in each environment and therefore selection in one may not result in the optimum response to selection in the other. This G x E can have varying degrees of impact on the response achieved within the breeding program (Mulder and Bijma 2005).

The extent of which G x E effects response will depend on the genetic correlation between performance in different environments. Given this correlation, Falconer (1952) showed that selection could be based on one or both environments using multi-trait selection principles.

Mulder and Bijma (2005) suggested that various recording strategies could be used to limit the amount of genetic gain that is lost due to significant G x E interactions. They showed that the impact of G x E on response to selection could be minimized if performance records on siblings or progeny were obtained across alternative environments. They concluded that the loss of genetic gain (due to G x E) was mainly due to the loss of selection accuracy.

Genomic selection provides the opportunity to increase selection accuracy (Meuwissen et al., 2001) and therefore may be a useful tool to manage the impact of G x E interaction on the breeding program. It also presents an opportunity to predict breeding value based on a wider range of animals, possibly located in different environments, as performance information can be used to inform many individuals, not just the close pedigree relatives of selection candidates.

The objective of this study was to examine the effect of genomic selection on a breeding program where significant G x E interactions exists.

Materials and Methods

Deterministic simulations. A number of simulations were carried out to examine the impact of using genomic information in a livestock breeding program that was effected by various degrees of G x E. A population consisting of 200 females and 20 males were selected from 400 male and 400 female selection candidates. Therefore 50% of females and 5% males were selected. Each female had 4 offspring per generation, which is similar to a dairy/beef cattle or sheep breeding program.

Two environments were used within the simulation. This included a nucleus environment (Env1), where all selection decisions were made and a production environment (Env2), were commercial production occurred. Two main breeding objectives were tested: the first breeding objective was for response in a single trait (which could also represent an index of multiple traits) in Env2 alone. The second breeding objective was for a varying amount of selection emphasis on improvement in each of the two environments and could represent a breeding program where commercial production occurred in two different environments i.e. intensive and extensive production. To observe the effect of different recording strategies, alternative scenarios were developed which varied in the amount and sources of information used to derive the selection criteria. These included: Own phenotype in Env1 (OP), phenotypic records from 20 half siblings (HS) in Env1 or Env2 (or both), phenotypic records from 20 progeny (Pr) in Env1 or Env2 (or both) and GEBV2 for Env2 or GEBV1 for Env1.

Four different traits were simulated to test a range of G x E interactions. These were a high heritability (0.5) trait and a low heritability (0.1) trait, with either a high (0.8) or moderate (0.5) genetic correlation between performances

in each environment. Genomic breeding values were assumed to be an extra correlated trait (Dekkers 2007). The correlation between GEBV and performance in each environment was dependent on the accuracy of the GEBV (as shown in Dekkers 2007).

Following Dekkers (2007), response to selection was predicted by using the pseudo-BLUP selection index that is implemented in the program SelAction (Rutten et al., 2002).

In summary, for breeding objective 1 (response in Env2 only), the effect of using a GEBV calibrated in Env2 for a trait with high heritability and low correlation between environments and a trait with low heritability and high correlation between environments was tested. For breeding objective 2 (when production in each environment was equally important), the effect of using either a GEBV calibrated in Env1 or Env2 was examined.

Results and Discussion

Breeding objective 1 - Selection for response in Env2. Selection based on genomic information calibrated in Env2 enabled increased response to selection. Although, the extra value added by genomic information was dependent on the accuracy of the GEBV2 (as shown by Dekkers 2007), the heritability of the trait, the correlation between the environments and the amount of information that was otherwise available on relatives of the selection candidate.
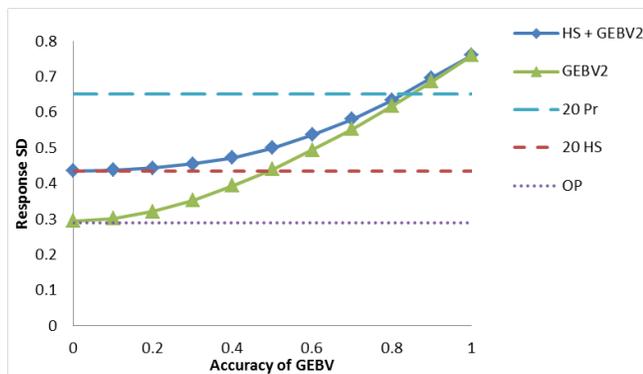
Figure 1: The effect of various sources of information[*] on response to selection[1] for a trait with a high heritability (0.5) and a moderate correlation between environments (0.5).
[1] The breeding objective was for performance in ENV2 only. [*] OP- Own Performance in Env1, GEBV2- GEBV for Env2, 20 HS- records on 20 half siblings in Env2, HS+GEBV2 - records on 20 HS and a GEBV in Env2, 20 Pr- 20 Progeny

When there was a high heritability and a moderate G x E correlation (0.5), a GEBV2 with an accuracy of 0.5 was equivalent to 20 HS being testing in the target environment (Env2) (Figure 1). Figure 1 also shows that GEBV2 accuracies higher than 0.8 were required to give as much value as 20 progeny. When there was a lower herita-

bility (0.1) and a high G x E correlation (0.8), a GEBV2 with an accuracy of 0.4 was required to give as much information as 20 HS (Figure 2). Similarly the GEBV provided more information than 20 progeny at accuracies of greater than 0.6.
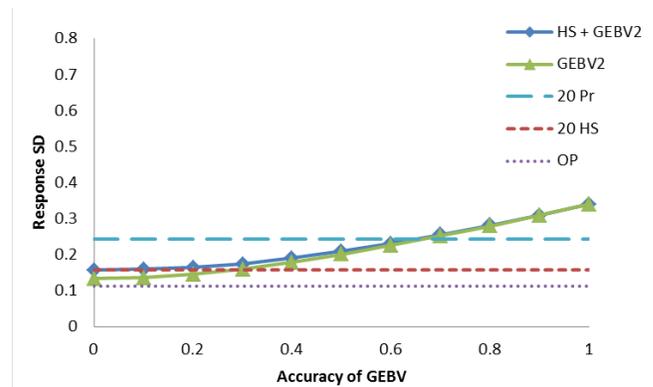
Figure 2: The effect of various sources of information[*] on response to selection for a trait with a low heritability (0.1) and a high correlation between environments (0.8).
[*] The breeding objective and information sources were the same as used in Figure 1.

When the genetic correlation between environments was high extra gains due to using GEBV2 were lower. This was also the case for other sources of information such as HS and progeny information. Furthermore when the genetic correlation between environments was high, the overall value of recording in Env2 reduced. Although, even at a correlation of 0.8 (which is regarded as a high correlation) there were significant additional gains to be made by using GEBV2 based on animals in the production environment (Env2).

As shown by Mulder and Bijma (2005) managing the loss in response due to G x E was highly dependent on the accuracy of selection and here we show that increases in accuracy, whether it was from measuring the performance of relatives (HS or Progeny) or using GEBV2 from Env2, resulted in proportionally greater response to selection.

Breeding Objective 2 - Selection for response in both environments. The breeding objective was changed to target response in both environments and when the importance of each environment became more targeted towards the nucleus environment (Env1) (>70% of the index weight), there was little gain in testing in Env2. When the index had a weight of >30% on Env2 then it was always better to have a GEBV2 that was relevant to Env2. However, this assumed that some level of recording was still occurring in Env1 (i.e. own performance and half sibling performance measures) and therefore the addition of the GEBV1 for Env1 added little extra information (Figure 3), whereas the additional information for Env2 was much greater. Furthermore, the same trend was observed for recording additional HS and progeny as it was also more useful to progeny test in Env2 (Figure 4). Even when there is

an equal importance of each environment it is more important to collect additional information from Env2 than Env1 (Figure 3 and 4), provided some recording already occurred in the Env1.
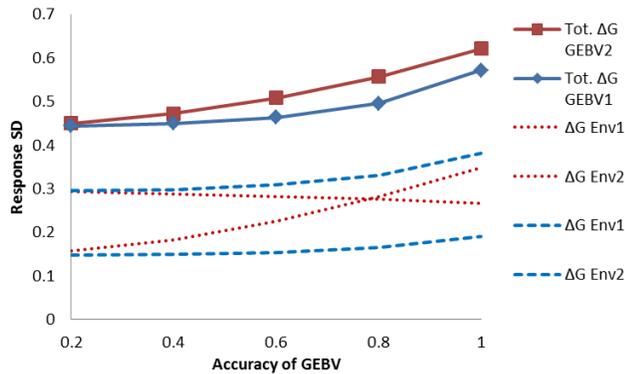
Figure 3: The effect of using a GEBV trained in alternative environments[*] for a trait with a high heritability (0.5) and a moderate correlation (0.5) between environments and equal weighting on each environment in the breeding objective.

[*]Tot. ΔG GEBV1- total response when selecting on GEBV1 for Env1, Tot. ΔG GEBV2 - Total response when selecting on GEBV2 for Env2, --ΔG Env1 (or Env2) is the response in Env1 (or Env2), when selecting on GEBV for Env1. ··· ΔG Env1 (or Env2) is the response in Env1(or Env2), when selecting on GEBV for Env2.
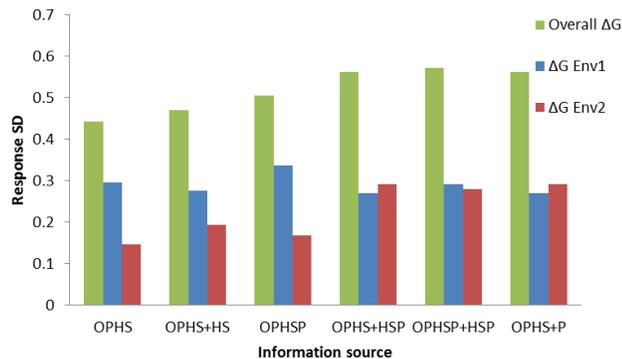
Figure 4: The effect of using different sources[1] of information from alternative environments [*] for a trait with a high heritability (0.5) and a moderate correlation (0.5) between environments.

[*]Overall ΔG- overall response, ΔG Env1 (or Env2) is the response in Env1 (or Env2). [1]OP= own performance in Env1, HS= 20 half sibs in Env1, +HS = HS performance in Env2, P =Progeny. i.e. OPHS+HS = own performance/ half sib performance in Env1 and HS performance in Env2.

Genomic breeding values not only had an effect on managing the loss of response to selection due to G x E, they also helped to reduce inbreeding within the breeding program (results not shown). This was regardless of whether they were trained in one or the other environment. This is not unexpected given GEBV have been shown to explain some information about variation due to Mendelian sampling (Daetwyler et al., 2010; Clark et al., 2013).

Implications for reference population design and performance recording. The results of this study may have important implications for the makeup of reference populations and the phenotypes that are used to create GEBVs. Knowledge of the structure of the breeding program and the importance of different environments can be used to determine the makeup of the reference. For example if the breeding objective is to improve production in a second environment, which is different to that of the nucleus, then creating GEBV2 based on the production environment data will be useful.

If the second environment is of little importance to the breeding objective then using GEBVs based on this data will give limited increases to overall genetic gain. In this case, GEBV1 based on nucleus data will have the greatest impact on response to selection. This has clear advantages as nucleus animals are already routinely recorded for performance and can then be simply genotyped to be used as part of the reference dataset for GEBVs. In practice, there has been a large amount of data shared across different environments (i.e. international collaborations, esp. dairy) to improve the accuracies of GEBVs. It is important to consider the impact of each environment when combining this data in regard to the importance of production in alternative environments and how these relate to the target population.

In this study, it was assumed that there are separate GEBVs for each environment; however, G x E interactions will also have an effect on estimating GEBVs. For example, if GEBVs were estimated from performance data from an unrelated environment, sub-optimal response to selection may occur. Furthermore, in practice there could be many more than two environments that are relevant to the breeding program. James (2009) noted that if the overall production system involves are large number of environments and management systems it would be important to test sires across a wide range of environments. This would also be the case for GEBVs, as it would be important to have GEBVs that was relevant to performance in many and not just a single environment.

The use of GEBVs to increase selection accuracy across environments may also have important implications when making decisions regarding which animals should be performance recorded. Recording the performance of a large number of relatives for each selection candidate, across all environments will often be impractical and therefore not occur. It may be more feasible to construct a reference population for genomic predictions that includes fewer animals and covers each environment. This is plausible as the information from GEBV can be propagated to many individuals not just the close direct relatives of few selection candidates as used in pedigree based systems. However, the added value relating to the construction of such a population would depend on the importance and significance of G x E to the breeding program.

Conclusion

Genomic breeding values can be effectively used to increase genetic gain when significant G x E exists. However, the added value over other recording strategies depends on the amount of performance information available at the time of selection. The impact of using GEBVs to manage losses due to G x E will be dependent on the reference data used to formulate the GEBVs.

Literature Cited

Mulder, H., Bijma, P. (2005). J. Anim. Sci. 63:27-33

Dekkers, J. (2007). J. Anim. Sci. 85; 9: 2104-2114

James J. (2009). *in* van der Werf et al.*,* Adapt. Fit. in Anim. Pop. 3:151-168

Faulkner, D.S. (1952). American Nat. 86:293-298

Rutten, M. J. M., Bijma, P., Woolliams, J. A. and van Arendonk, J. A. M. (2002). J. Hered. 93:456–458.

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

Clark, S.A., Hickey, J.M., Kinghorn, B.P. and van der Werf, J.H.J (2013). Genet. Sel. Evol. 45: 44

Daetwyler, H.D., Villanueva, B., Bijma, P. and Woolliams, J.A. (2007) J Anim. Breed. Genet. 124:369-376