Mate Allocation And Genomic Selection

M. A. Toro[*] and L. Varona[†]

Introduction

Mate allocation has been used in animal breeding with several purposes: a) to control inbreeding; b) in situations where economic merit is not linear; c) when there is an intermediate optimum (or restricted traits); d) to increase connection among herds, and, e) to profit from dominance genetic effects. With respect to the last point it is well known that any methodology that pretends to use non-additive effects it must contemplate two types of matings (Toro, 1996): a) matings from which the population will be propagated; b) matings to obtain commercial animals. Moreover, selection should be applied not to individuals but to matings. Though it is usually thought that application of the above ideas requires two separate lines as in the classical crossbreeding programmes or in the so called reciprocal recurrent selection it can be carried out in a single population. Optimal mating allocation to profit from dominance relies on the idea that although selection should be carried out on estimated additive breeding values, animals used for commercial production should be the product of planned matings which maximize the overall (additive plus dominance effects) genetic merit of the offspring.

With the recent availability of very dense panels of SNPs and the advent of gnomic selection that promises high accuracies and responses to selection (Meuwissen et al., 2001), it is seems natural that methods proposed to use dominance variation should be revisited. From all methodologies aimed to profit from dominance (Maki-Tanila, 2007) mating allocation could be the easier option. The aim of this study is to quantify the efficiency of mating allocation under a whole genome evaluation scenario in terms of genetic response to selection in the first and subsequent generations.

Methods

Simulation. Parameters of simulation were those as in Meuwissen et al. (2001). Thus, the population was simulated for 1000 generations with an effective size of 100. After 1000 generations, the actual size of the population was increased up to 500 males and 500 females for three consecutive generations. During all the process any individual was generated taking one gamete from a random father and another gamete from a random mother. The data set available for the estimation of marker effects consisted of the 3000 individuals from the last three generations (generation 1001, 1002 and 1003). These 3000 individuals were genotyped

[*] ETS Ingenieros Agrónomos, 28040 Madrid, Spain
[†] Facultad de Veterinaria, Universidad de Zaragoza, 50013 Zaragoza, Spain

and phenotyped and used as the training population to estimate additive and dominance effects of SNPs.

The genome was assumed to consist of 10 chromosomes of 100 cM each with 1000 loci per chromosome (9000 SNPs and 1000 QTLs in total) located at random map positions. Both SNPs and QTLs were biallelic. Mutations were generated at rates of 2.5×10^{-3} and 2.5×10^{-5} per locus and generation for the marker and the QTL loci respectively. Both the additive and the dominance effects were sampled from a standard normal distribution and scaled to get the desired values of h^2 (V_A/V_P) and d^2 (V_D/V_P) being V_A, V_D and V_P the additive, dominance and phenotypic variance respectively. In generation 1 about half of the loci were fixed for allele 1 and the other half were fixed for allele 2.

Model of analysis. Estimation of marker effects was carried out using the Bayes A method of Meuwissen et al. (2001). The model assumed was for the phenotypic value of individual j (j =1, …N) is

$$y_j = \mu + \sum_i^p x_{ij} g_i + \sum_i^p w_{ij} d_i + e_j$$

where p is the number of SNPs and x_{ij} and w_{ij} are parameters that take the values 1, 0 and -1 and 0,1 and 0 for the SNPs genotipes AA, Aa and aa, respectively. The assumed distributions for the additive (g_i), dominance (d_i) and residual (e_j) components were

$$g_i \sim N(0, \sigma^2_{gi}) \qquad d_i \sim N(0, \sigma^2_{di}) \qquad e_j \sim N(0, \sigma^2_e)$$

The prior distribution of the variances was the scaled inverted chi-square distribution

$$\sigma^2_{gi} \sim \chi^{-2}(\nu, S) \qquad \sigma^2_{di} \sim \chi^{-2}(\nu, S) \qquad \sigma^2_e \sim \chi^{-2}(-2, 0)$$

where S is a scale parameter and υ is the number of degrees of freedom of the prior distribution, that takes the values of 4.012 and 0.020, respectively (Meuwissen et al.,2001). Gibbs sampling based on posterior distributions conditional on other effects and results were obtained by averaging the samples from 10,000 cycles, after discarding the first 1,000. Generation 1004 was formed from 25 sires and 250 dams selected from generation 1003. Two strategies of selection, carried out during five generations, were compared. In the first (GS) 25 males and 250 females were selected from 500 males and 500 females based on genomic estimated breeding values. Afterwards individuals were mated randomly (10 dams per sire) and 4 sibs were obtaining from each mating. Finally, the true genotypic of value of the offspring was calculated. In the second strategy (GS+MA) the best 250 matings out of 6250 (25 x 250) possible matings were chosen based on the dominance prediction of the mating, and 4 new individuals for each mate were obtained. The true genotypic value of the offspring was also calculated. The algorithm of searching used was the simulated annealing.

Finally, phenotypic selection was also carried out as a control. Fifty replicates of each method were performed.

Results and discussion

Results of selection response in the first generation are presented in Table 1. The observed genetic response was always higher when mating selection was practiced and the advantage of GS+MA over GS ranked from 6% (h^2=0.40, d^2=0.05) to 22% (h^2=0.20, d^2=0.10). In general, this superiority increases as the ratio of dominance variance increases and as the heritability decrease. Finally, as expected, the superiority of GS and GS+MA over mass selection was clear in each scenario.

Table 1: Comparison of selection response in phenotypic standard deviations with different methods in the first generation[α]

h^2	d^2	MS	GS	GS+MA	Ratio
0.20	0.05	0.282±0.066	0.471±0.054	0.527±0.048	1.118
0.20	0.10	0.267±0.045	0.470±0.045	0.575±0.060	1.223
0.40	0.05	0.562±0.056	0.771±0.062	0.815±0.058	1.057
0.40	0.10	0.557±0.050	0.754±0.052	0.875±0.066	1.160

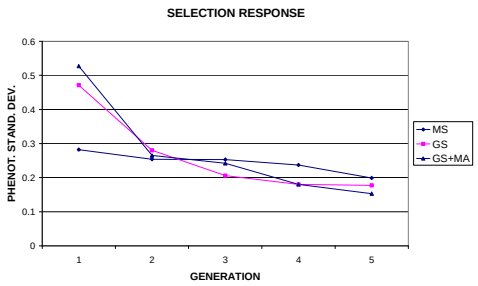
[α] Mass Selection (MS), Genomic selection with Bayes A (GS) and genomic selection with Bayes A and optimal mate allocation (GS+MA) and ratio between GS and GS+MA.

Results in the medium term for each scenrio are presented in Figure 1. The advantage of GS and GS+MA over MS observed in the first generation dissappear in subsequent generations due to a reduction in linkage disequilibrium between markers and QTL. However, it should be noted than MS requires to phenotype the candidates of selection at each generation.

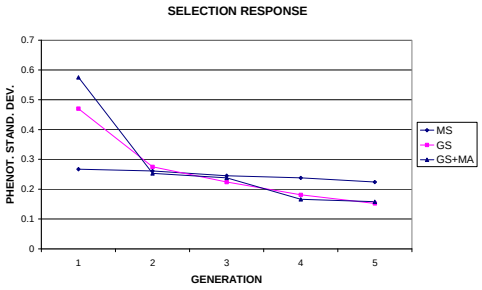
In addition, it is remarkable that the increase of response with GS+MA when compared with GS is observed only in the first generation, being similar to GS in generations two to five. Thus, the advantage in terms of selection response that it is obtained in the first generation is mantained in the subsequent generations. However, a single generation of random mating eliminates this superiorty (results not shown).

The genomic evaluation procedure used here was Bayes A, but results were confirmed with Bayes B. When both procedures were compared under a model with additive and dominance effects Bayes A gave higher accuracy than Bayes B (0.728 vs 0.700, h^2=0.20 and d^2=0.10) but when both procedures were compared under an additive model Bayes B gave higher accuracy than Bayes A (0.705 vs. 0.699, h^2=0.20 and d^2=0.05). A possible explanation of these results could be related with prior specifications of Bayes A and B, that it should be adapted to a model with additive and dominance effects.

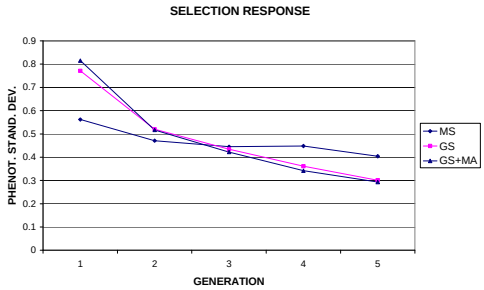
a) h^2=0.20, d^2=0.05



b) h^2=0.20, d^2=0.05



c) h^2=0.40, d^2=0.05



d) h^2=0.40, d^2=0.10

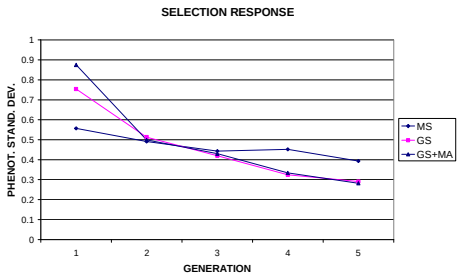


Figure 1: Comparison of selection response in the first five generations with different methods: Mass selection (MS); Genomic selection with Bayes A (GS); Genomic selection (Bayes A) and optimal mate allocation (GS+MA), measured in phenotypic standard deviations.

Conclusion

Mate allocation can improve the expected genetic response on the first generation, but its advantage is not improved over subsequent cycles of selection. Morover, genomic selection procedures loss their efficiency in the second generation when compared with simple mass selection. The need of a frequent re-estimation of marker effects should, therefore, be emphasized.

References

Maki-Tanila, A. (2007). *Agricultural and food science*, 16: 188-198.

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

Toro, M.A.(1996). *Genet. Sel. Evol.*,30: 339-349.