EAAP 2009, Barcelona, Session 4, Abstract 4024

Breeding for a global dairy market using genomic selection

A. P. W. de Roos,[1*] C. Schrooten,[1] R. F. Veerkamp,[2] and J. A. M. van Arendonk,[3]
*[1]CRV, PO Box 454, 6800 AL Arnhem, Netherlands, [2]Wageningen University and Research Centre, Animal Breeding and Genomics Centre, PO Box 65, 8200 AB Lelystad, Netherlands, [3]Wageningen University and Research Centre, Animal Breeding and Genomics Centre, PO Box 338, 6700 AH Wageningen, Netherlands, *sander.de.roos@crv4all.com*

Abstract

From the simulation of a closed nucleus breeding program for dairy cattle it was concluded that the introduction of genomic selection and the use of young animals as parents increased the rate of genetic gain by a factor 2.4 when genetic markers explained 50% of the genetic variance. In this situation, all bulls in the top 100 EBV list were young bulls. While genomic selection reduced the rate of inbreeding, the actual rate of inbreeding per year was increased by a factor 1.6 because of the use of young animals as parents. When a reference population was available in environment A but not in environment B, selection based on the average EBV in environment A and B was the most effective strategy when the genetic correlation between A and B was ≥0.90. When the genetic correlation between A and B was ≤0.75 the rate of genetic gain was lower across all strategies. Splitting the population gave the highest rate of genetic gain but also the highest rate of inbreeding.

Introduction

Global dairy producers may prefer different bulls as a result of different breeding objectives or genotype by environment interaction (GxE). To meet global market demands, breeding organisations can use one general purpose breeding objective or multiple specialised breeding objectives in their breeding program (Mulder et al., 2006). Using multiple breeding objectives will create specialists for each market segment and is advantageous if the genetic correlation between the breeding objectives or environments is low. On the other hand, it can induce a split within the breeding program which would result in a higher rate of inbreeding within each part of the breeding program, or a lower rate of genetic gain if the rate of inbreeding is constrained (Mulder et al., 2006).

Recently, dairy cattle breeding organisations have started to use genomic selection in their breeding programs and scientists have reported reliabilities of genomic breeding values (GBV) of >0.60 for important dairy traits (VanRaden et al., 2009). Genomic selection is expected to increase the rate of genetic gain because young bulls can be selected for progeny testing with higher accuracy and superior young bulls and heifers will be used as parents, which shortens the generation interval (Schaeffer, 2006). Several breeding organisations have already commercialised young bulls with high GBV. Because reliabilities of GBV are expected to increase further over time and farmers will get more used to genomically selected young bulls, it may be expected that young bulls rather than proven bulls will dominate the semen market within some years. It is clear that these developments have major consequences for the optimal design of breeding programs. The optimal design of a genomic selection breeding program therefore needs further study, especially if the aim is to meet global market demands. One specific situation is where a reference population is established in one environment but the aim is to breed bulls for multiple environments, in the presence of GxE.

The objectives of this study were (1) to quantify the effects of using genomic selection and using young bulls and heifers as parents on the rate of genetic improvement and the rate of inbreeding, and (2) to compare alternative strategies to breed for multiple environments using genomic selection based on a reference population in one environment.

Material and methods

Breeding program design

A closed nucleus breeding program was simulated in which 1000 heifer and 1000 bull calves were born and genotyped annually, out of which 800 heifers and 800 bulls were culled after birth based on their EBV. The remaining 200 heifers and 200 bulls obtained a phenotype when they were 3 and 5 years old, respectively, and were culled when they were 6 and 8 years old. Among the non-culled animals, 200 dams and 40 sires were selected based on their EBV which produced 10 progeny per dam and 50 progeny per sire. In the first 20 years, only animals with a phenotype were selected as parents and EBV did not include genomic information. After that, genomic selection was performed for 30 years and animals of at least 1 year old could be selected as parents. Each scenario was replicated 100 times.

Simulation of true breeding values

True breeding values were calculated as: $\mathbf{u}_i = \mathbf{u}_{i,M} + \mathbf{u}_{i,P}$ where $\mathbf{u}_i$ is a vector of size 2 comprising the total breeding values of animal *i* for trait A and B, $\mathbf{u}_{i,M}$ is the marker part of the breeding value which can be explained without error by genetic markers and $\mathbf{u}_{i,P}$ is the polygenic part of the breeding value which cannot be traced by the markers and which was assumed to be independent of $\mathbf{u}_{i,M}$. The marker breeding values were calculated as: $\mathbf{u}_{i,M} = \frac{1}{2}\mathbf{u}_{i,M,sire} + \frac{1}{2}\mathbf{u}_{i,M,dam} + \mathbf{u}_{i,M,MS}$ where $\mathbf{u}_{i,M,sire}$ and $\mathbf{u}_{i,M,dam}$ are the marker breeding values of the sire and dam of animal *i* and $\mathbf{u}_{i,M,MS}$ is the marker Mendelian sampling effect, which was drawn from a multivariate normal distribution: $\mathbf{u}_{i,M,MS} \sim N\left(0, \frac{1}{2}\left(1 - \frac{1}{2}\left(F_{i,sire} + F_{i,dam}\right)\right)\mathbf{K}_M\right)$ where $F_{i,sire}$ and $F_{i,dam}$ are the pedigree inbreeding coefficients of the sire and dam of animal *i* (Meuwissen and Luo, 1992), respectively, and $\mathbf{K}_M$ is the marker covariance matrix, assuming a reference population for trait A only: $\mathbf{K}_M = \begin{pmatrix} 1 & \rho_{AB}^2 \\ \rho_{AB}^2 & \rho_{AB}^2 \end{pmatrix}\sigma_{M,A}^2$ where ρ_{AB} is the genetic correlation between trait A and B and $\sigma_{M,A}^2$ is the marker variance of trait A. Similarly, polygenic breeding values were calculated as: $\mathbf{u}_{i,P} = \frac{1}{2}\mathbf{u}_{i,P,sire} + \frac{1}{2}\mathbf{u}_{i,P,dam} + \mathbf{u}_{i,P,MS}$, $\mathbf{u}_{i,P,MS} \sim N\left(0, \frac{1}{2}\left(1 - \frac{1}{2}\left(F_{i,sire} + F_{i,dam}\right)\right)\mathbf{K}_P\right)$ and $\mathbf{K}_P = \begin{pmatrix} 1 & \rho_{AB} \\ \rho_{AB} & 1 \end{pmatrix} - \mathbf{K}_M$. Values of $\sigma_{M,A}^2$ and ρ_{AB} were varied from 0 to 1.

Breeding value estimation

Cows and bulls were assigned to one environment (trait), A or B, based on their EBV after birth, and obtained a phenotype for that trait at 3 and 5 years of age, respectively, with $h^2 = 0.30$ for cows $h^2 = 0.90$ for bulls. Bulls assigned to environment A that were used as sire for cows in environment B, also obtained a phenotype in environment B at the same time as their daughter, and vice versa. This corresponds to internationally used sires that provide genetic links between the environments. The phenotype of animal *i* for trait *j* was calculated as $y_{ij} = u_{ij} + e_{ij}$, where e_{ij} was drawn from a normal distribution with mean 0 and variance $(1-h^2)/h^2$. EBV were calculated as $\hat{\mathbf{u}} = \mathbf{u}_M + \hat{\mathbf{u}}_P$, assuming the marker breeding value was known without error. Polygenic breeding values ($\hat{\mathbf{u}}_P$) were estimated with a BLUP model: $\mathbf{y}_P = \mathbf{u}_P + \mathbf{e}$ where $\mathbf{y}_P = \mathbf{y} - \mathbf{u}_M$. The variance of $\mathbf{u}_P$ and $\mathbf{e}$ were assumed to be known: $\text{var}(\mathbf{u}_P) = \mathbf{A} \otimes \mathbf{K}_P$ where $\mathbf{A}$ is

the additive genetic relationship matrix derived from the full pedigree, and $\text{var}(\mathbf{e}) = \mathbf{I} \otimes \left(1 - h^2\right) / h^2$. The model was solved using iteration on data with a preconditioned conjugate gradient algorithm (Strandén and Lidauer, 1999).

Young animals as parents

Corresponding to objective (1) of the study, one scenario was evaluated that included only trait A and in which the minimum age of cows and bulls to be selected as parents during the genomic selection phase (years 21 – 50) was either 3 and 5, 1 and 5, or 1 and 1, respectively.

Selection and mating strategy

Corresponding to objective (2) of the study, three methods of selection were compared:

AVE = selection candidates with the highest average EBV for trait A and B were selected.

SPL = the first generation of animals were randomly assigned to either environment A or B, and the breeding program was completely split afterwards, i.e. in environment A the selection candidates with the highest EBV for trait A were selected and in environment B the selection candidates with the highest EBV for trait B were selected.

EXT = selection candidates with the highest EBVs for trait A were selected as well as the selection candidates with the highest EBVs for trait B.

Across all three methods the number of selection candidates and the number of animals selected was the same. In strategy SPL and EXT dams were mated to sires that were selected for the same trait, whereas in AVE any dam could be mated to any sire.

Rate of genetic gain and inbreeding

To evaluate the breeding programs, the average EBV and inbreeding coefficients of the 100 highest ranked bulls of at least 1 year old was computed for each trait. The rate of genetic gain was computed as the average increase per year over years 20 to 50 of the average EBV of these top 100 bulls. The rate of inbreeding was computed as the average increase per year over years 20 to 50 of the inbreeding coefficients of the same group of bulls.

Results

Young animals as parents

When only older animals were used as parents (cows ≥3 yr, bulls ≥5 yr), genomic selection increased the rate of genetic gain from 0.21 to 0.32 σ_G per year when $\sigma^2_{M,A}$ increased from 0 to 100% (Figure 1). When also young animals were used as parents the rate of genetic gain was 0.26, 0.51, and 0.67 σ_G per year for $\sigma^2_{M,A}$ equal to 0, 50, and 100%, respectively (Figure 1). Genomic selection reduced the rate of inbreeding from 0.15% to 0.05% per year when $\sigma^2_{M,A}$ increased from 0 to 100%, but across all $\sigma^2_{M,A}$ the use of young animals as parents resulted in 3 to 5 times higher rate of inbreeding (Figure 2). Genomic selection and using also young animals as parents substantially reduced the percentage of proven bulls among the top 100 bull EBV within very few years, from which it remained constant until year 50 (Figure 3).

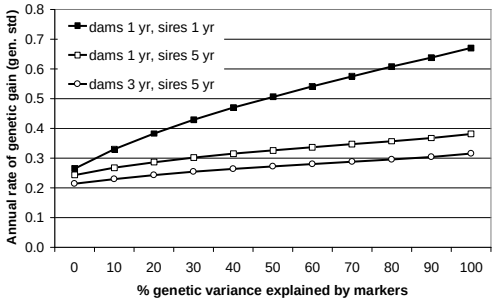

Figure 1. *Rate of genetic gain with different minimum age for sires and dams.*

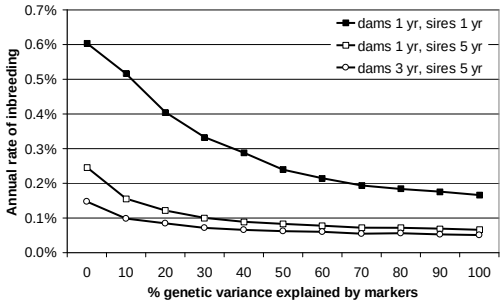

Figure 2. *Rate of inbreeding with different minimum age for sires and dams*

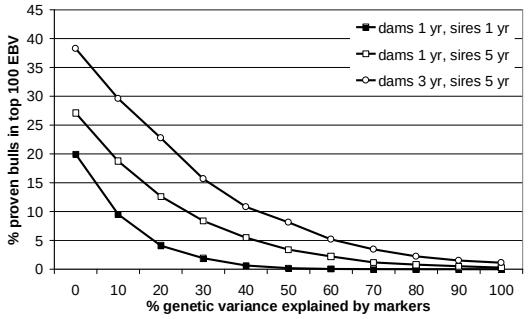

Figure 3. *Percentage proven bulls (age ≥5 yr) in top 100 EBV in year 50, with different minimum age for sires and dams*

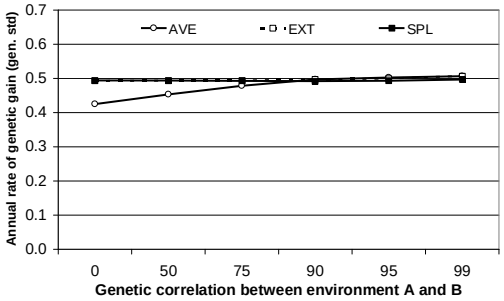

Figure 4. *Rate of genetic gain of top 100 bulls for trait A, for selection strategy AVE, SPL, and EXT.*

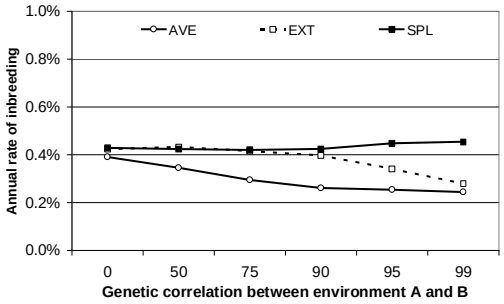

Figure 5. *Rate of inbreeding of top 100 bulls for trait A, for selection strategy AVE, SPL, and EXT.*

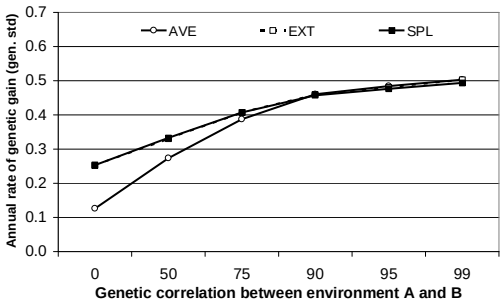

Figure 6. *Rate of genetic gain of top 100 bulls for trait B, for selection strategy AVE, SPL, and EXT.*

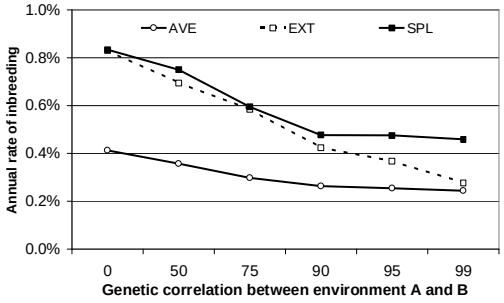

Figure 7. *Rate of inbreeding of top 100 bulls for trait B, for selection strategy AVE, SPL, and EXT.*

Selection for two environments

Figures 4 – 7 correspond to the scenario where selection is aimed at two environments, A and B, and where a reference population is present for trait A but not for trait B. The markers explain 50% of the genetic variance ($\sigma^2_{M,A} = 0.5$) and both young ald older animals were used as parents. When the genetic correlation between A and B (ρ_{AB}) approached 1 all selection strategies resulted in the same rate of genetic gain for both traits (Figure 4, 6), but strategy SPL, where the population was split, had a higher rate of inbreeding (0.45 vs. 0.24%) (Figure 5, 7). For ρ_{AB} = 0.90 all strategies still had the same rate of genetic gain, although slightly lower for trait B, but the rate of inbreeding increased substantially for EXT because the population started to split. For $\rho_{AB} \leq 0.75$, the rate of genetic gain for trait B was lower than for trait A because of the absence of a reference population. Furthermore, the rate of genetic gain was lower in AVE than in SPL and EXT for both traits when $\rho_{AB} \leq$ 0.75. In strategy EXT the population began to split when $\rho_{AB} \leq 0.90$ and the rates of genetic gain and inbreeding were similar to SPL. For SPL and EXT, when ρ_{AB} decreases from 0.90 to 0 the rates of genetic gain and inbreeding for trait A were not affected, whereas the rate of genetic gain for trait B was decreased and the rate of inbreeding was increased, corresponding to the situation where the markers explained less of the genetic variation (Figure 1, 2).

Discussion

An important assumption in this study was that $\sigma^2_{M,A}$ and ρ_{AB} were constant over time. This may be realistic if genetic markers trace many QTL simultaneously and their effects are updated regularly. Furthermore, the marker variance was assumed equal for all animals, which may not be the case if animals are unequally related to the reference population. Thirdly, ρ_{AB} was used as the correlation between marker breeding values of trait A and B, as well as between the polygenic breeding values of trait A and B, which may not necessarily be the same.

Genomic selection with $\sigma^2_{M,A} = 0.5$ in combination with using young animals as parents increased the rate of genetic gain from 0.21 to 0.51 σ_G per year (2.4 times higher), but also increased the rate of inbreeding from 0.15 to 0.24% per year (1.6 times higher). The percentage of proven bulls in the top 100 EBV list would reduce from 38 to 0%. Without genomic selection and without use of young animals as parents already 62% of the top 100 bulls are young bulls. This is also observed in practise, however, the reason that these young bulls with high parent average EBV were hardly used commercially or as parents lies in their low reliability (too much risk) and biases due to preferential treatment of bull dams. Furthermore it would dramatically increase the rate of inbreeding (Figure 2), with only a moderate effect on the rate of genetic gain (Figure 1). With genomic selection the reliabilities of young bull EBV are substantially higher and much less influenced by cow phenotypes, so it becomes much more attractive to utilise young animals in the breeding program and commercially. When the complete top 100 comprises only young bulls, as observed here, progeny testing may be obsolete and bulls may be culled after they have been used commercially as a young bull.

When a breeding organisation has a reference population in environment A it may very effectively breed superior bulls for environment B where it has no reference population when $\rho_{AB} \geq 0.90$. In these situations strategy AVE gave the lowest rate of inbreeding with the same rate of genetic gain. For $\rho_{AB} \leq 0.75$ the genetic gain in the environment B is substantially lower than in environment A. In these situations strategy EXT or SPL gave a higher rate of genetic gain than AVE but also more inbreeding. The best strategy for low ρ_{AB} therefore depends on the importance of genetic gain versus inbreeding, which is a complex dilemma.

References

Meuwissen, T.H.E. and Z. Luo, 1992. Computing inbreeding coefficients in large populations. *Genet. Sel. Evol.* 24:305-313.

Mulder, H.A., R.F. Veerkamp, B.J. Ducro, J.A.M. van Arendonk, and P. Bijma, 2006. Optimization of dairy cattle breeding programs for different environments with genotype by environment interaction. *J. Dairy Sci.* 89:1740-1752.

Schaeffer, L.R. 2006. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 123:218–223.

Strandén, I., and M. Lidauer. 1999. Solving large mixed linear models using preconditioned conjugate gradient iteration. *J. Dairy Sci.* 82:2779–2787

VanRaden, P.M., C.P. VanTassel, G.R. Wiggans, T.S. Sonstegard, R.D. Schnabel, J.F. Taylor and F.S. Schenkel, 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci. 92*, 16-24.