OPTIMISING MULTISTAGE DAIRY CATTLE BREEDING PROGRAMS WITH REGARD TO GENOMIC SELECTION

Vinzent Börner[1] and Norbert Reinsch

Leibniz-Institute for Farm Animal Biology, FBN, Germany
[1]present address: Animal Genetics and Breeding Unit, University of New England, Armidale, NSW 2351

SUMMARY

Multistage dairy cattle breeding schemes consisting of 4 selection paths were optimised in order to maximise the genetic gain per year with regard to genomic selection on 2 genomically estimated breeding values differing in costs and accuracy. Results clearly show that the selection of bull dams is the major field of application for low-density genotyping but also emphasise the selection of sires to be of continuously highest importance for the generation of the genetic gain irrespective of increasing costs for high-density genotyping.

INTRODUCTION

Genomic selection (**GS**) will increase the genetic gain per year (Δ**Ga**) in dairy cattle breeding because of a tremendously decreased generation interval due to abandoning progeny testing of potential sires and own performance testing of potential bull dams in conjunction with a sufficiently accurate genomically estimated breeding value (**GEBV**) (Schaeffer 2006; Hayes et al. 2009). The economic feasibility of the selection of females on GEBVs is enhanced due to the possible usage of cheap low-density SNP-chips and imputation algorithms (Weigel *et al*. 2010). However, the scope of decision for breeding organisations concerning the detailed structure of breeding schemes has become tremendously complex because GS and the availability of different SNP-chips allow for a variety of one-, two- or multistage breeding schemes in every single selection path, whereas economic resources are limited. Furthermore, the actually realised accuracy of GEBVs (r_{GEBV}) may change in practical breeding programs, and the future development of genotyping costs (C_{GEBV}) is still unclear. The aim of this study was to find optimum multistage dairy cattle breeding schemes with regard to GS on (a) a GEBV derived from an expensive high-density chip (**GS$_{HD}$, GEBV$_{HD}$, r$_{GEBV,HD}$, C$_{GEBV,HD}$**), and on (b) a GEBV derived from a less expensive low-density chip (**GS$_{LD}$, GEBV$_{LD}$, r$_{GEBV,LD}$, C$_{GEBV,LD}$**) under the constraint of limited economic resources. Additionally, the sensitivity of ΔGa to changes in C_{GEBV} and r_{GEBV} could be examined. Therefore, multistage breeding schemes were evaluated using a grid search and varying $r_{GEBV,HD}$, $r_{GEBV,LD}$, $C_{GEBV,HD}$ and $C_{GEBV,LD}$ within a semi-continuous range. The outcomes provide answers to questions concerning the sensitivity of ΔGa to a change in r_{GEBV} and C_{GEBV} as well as the sensitivity of ΔGa as a function of the stage selection intensities and the interrelation between selection paths.

METHODS

For the model calculation the structure of a cooperative Holstein dairy cattle breeding program with 4 selection paths with following numbers of initial selection candidates and finally selected individuals was used: "sire-sire" (**SS**, 10, 5), "sire-dam" (**SD**, 500, 10), "dam-sire" (**DS**, 50000, 1000) and "dam-dam" (**DD**, 50000, 50000). For the sake of simplicity, only one trait was in the aggregate genotype (kg milk, h^2 =0.25, σ_p = 700). Possible selection stages within the paths SS, SD and DS are 1) selection on pedigree information (performances and/or GEBVs), 2) selection on GEBV$_{LD}$ and 3) selection on GEBV$_{HD}$. In path DD no selection was applied. The accuracies of the breeding values of successive selection stages were derived via selection index methodology using a pedigree backward to grand parents level. The GEBVs were modelled as traits with a heritability of

Table 1: Parameter range of costs and accuracies of genomically estimated breeding values

parameter	start	end	step width
$C_{GEBV,HD}$	150 €	250 €	20 €
$C_{GEBV,LD}$	20 €	100 €	20 €
$r_{GEBV,HD}$	0.7	0.9	0.05
$r_{GEBV,LD}$	0.4	0.65	0.05

1 and a correlation to the aggregate genotype equal to its accuracies (Dekkers 2007). Thus sires in the pedigree had an own performance record. In an extension of the approach of Dekkers (2007) the correlation between $GEBV_{LD}$ and $GEBV_{HD}$ was derived from the ratio of their standard deviations. A double layer grid search was applied, where in the first layer r_{GEBV} and C_{GEBV} were varied as given in table 1. In the second layer, for each set of C_{GEBV} and r_{GEBV}, the breeding scheme maximising ΔG_a was searched for by varying the proportion selected at the pedigree stage, at the GS_{LD} stage and at the GS_{HD} stage in the paths SS, SD and DS between 0.05 and 1 in steps of 0.05, where 1 was equal to excluding the stage. The proportion selected at the last used stage was calculated as a dependent variable to account for the fixed number of finally selected individuals in each path. The selection intensities after selecting at the final stage in each path were derived via multidimensional integration merging integration algorithms of Genz (1992) and maximisation techniques of Brent (1973). The overall breeding costs included C_{GEBV} (including laboratory analysis and calculation of GEBVs), purchasing costs for male calves after the final selection stage, compensation payments to breeders for keeping finally non-selected selection candidates as long as the final selection has not taken place, and husbandry costs for finally selected males until maturity. The maximum breeding costs were derived from the cost structure of a progeny performance scheme testing 50 bull per year, but only purchasing costs, husbandry costs until maturity, husbandry costs from maturity to proven sire age and compensation payments for test bull insemination were regarded. A total of 146 million breeding schemes were included in the cost calculation process, where 6.7 million fulfilled the cost constraint and were evaluated in terms of ΔG_a.

RESULTS

Table 2 summarises the results concerning the genetic gain per year and per generation in different selection paths, and the proportion of genotyped initial selection candidates in the paths SD and DS. Independent of r_{GEBV} and C_{GEBV}, bull sires were always selected from cow sires by taking the best without gathering any additional information. Furthermore, as one kind of GS was always applied in the path DS, this path caused the highest proportion of overall breeding costs. The contributions of the different selection paths to the overall genetic gain were in the following order: SS > SD > DS > DD. The achievable ΔG_a varied between 0.46 and 0.62 genetic standard deviations and was mainly generated due to the selection of males, whereas the path DS never contributed more than 31 % to ΔG_a. GS_{HD} was always used to select males. The proportion of high-density genotyped initial male selection candidates (**$PG_{HD,SD}$**) was ≤ 1, independent of $r_{GEBV,HD}$, if the difference between $r_{GEBV,HD}$ and $r_{GEBV,LD}$ was ≥ 0.15 and no GS_{LD} was applied. Dependent on this difference GS_{HD} was also combined with GS_{LD} to select males. On the contrary, for the path DS breeding schemes were found suggesting selection on both GS_{LD} and GS_{HD} or excluding one of these. Furthermore, combined selection of females on $GEBV_{HD}$ and $GEBV_{LD}$ was found to produce a higher genetic gain than extending the proportion of low-density genotyped initial female selection candidates (**$PG_{LD,DS}$**). However, such combination was only useful if the difference in r_{GEBV} was ≤ 0.35. In other cases, a selection of females only on pedigree data and $GEBV_{HD}$ was found to be more

Table 2: Results across accuracies and costs of genomic estimated breeding values.

	$\bar{x}$	max	min	reference[1]
$\Delta G_a(\%)$[2]	191.21 (55 %)	223.97 (64 %)	161.38 (46 %)	178.50 (51 %)
$\Delta G_{SS}(\%)$[2]	745.21 (39 %)	839.83 (41 %)	647.30 (36 %)	701.52 (39 %)
$\Delta G_{SD}(\%)$[2]	675.21 (35 %)	762.83 (37 %)	582.35 (33 %)	637.18 (36 %)
$\Delta G_{DS}(\%)$[2]	491.70 (26 %)	639.65 (31 %)	373.02 (21 %)	446.27 (25 %)
$PG_{HD,SD}$[3]	0.46	1	0.06	0.475
$PG_{HD,DS}$[4]	0.03	0.05	0	0
$PG_{LD,SD}$[5]	0.56	0.95	0	0.95
$PG_{LD,DS}$[6]	0.1	0.55	0	0.10
absolute and relative total breeding costs	705,091 (98 %)	719,050 (100 %)	566,675 (79 %)	717,800 (99 %)

1: calculation results for a parameter combination of $r_{GEBV,HD} = 0.75$, $r_{GEBV,LD} = 0.6$, $C_{GEBV,HD} = 210 €$ and $C_{GEBV,LD} = 100$ €, 2: genetic gain per year and of different selection path in kg milk and as proportion of the additive genetic standard deviation 3: proportion of the initial selection candidates in the path "sire-dam" being genotyped with a high-density SNP-chip, 4: proportion of the initial selection candidates in the path "dam-sire" being genotyped with a high-density SNP-chip, 5: proportion of the initial selection candidates in the path "sire-dam" being genotyped with a low-density SNP-chip, 6: proportion of the initial selection candidates in the path "dam-sire" being genotyped with a low-density SNP-chip

rewarding. For a selection of females only on pedigree data and GEBV$_{LD}$, an $r_{GEBV,LD} \geq 0.55$ was necessary. As show in figure 1, ΔG_a was positively affected by an increasing r_{GEBV}, where $r_{GEBV,HD}$ had a higher effect than $r_{GEBV,LD}$. Not surprisingly, increasing C_{GEBV} decreased ΔG_a, but $C_{GEBV,LD}$ had a stronger effect compared to $C_{GEBV,HD}$. The effects of the variation of these parameters on ΔG_{DS} were similar to those on ΔG_a, whereas C_{GEBV} in general, and $r_{GEBV,LD}$ had no effect on ΔG_{SD}, and an increasing $r_{GEBV,HD}$ sharply increased ΔG_{SD}(results not shown). The line in figure 1 reflects the developments for a **reference scenario** ($r_{GEBV,HD} = 0.75$, $r_{GEBV,LD} = 0.6$, $C_{GEBV,HD} = 210 €$ and $C_{GEBV,LD} = 100 €$) if the abscissa parameter was varied and all other were kept constant.

DISCUSSION

The results clearly show that the applicability of GS for selecting females is enhanced when cheap low-density SNP-chips are used. Due to cost limitation the path DS was not found to generate the highest proportion of the genetic gain, which is in contrast to other deterministic calculations (Schaeffer 2006). The cost constraint also induced a strong interaction between selection strategies in different paths leading to the fact that $C_{GEBV,LD}$ had a stronger effect on ΔG_a compared to $C_{GEBV,HD}$, whereas this was vice versa for the accuracies. In many parameter combinations a combined selection of males and females on pedigree data, GS$_{LD}$ and GS$_{HD}$ was the favourable solution. Thus, as long as sufficient information from relatives are available and selection on GEBV$_{HD}$ is possible, the proportion of individuals being low-density genotyped should be optimised with regard to the diminishing marginal utility of the selection intensity on ΔG_a. Furthermore, in competitive markets an advantage can be achieved by generating the same result with lower costs. Since bull dams are selected from the total cow population, high-density genotypes will be available for the sires but not for the dams of selection candidates. Thus, population based algorithms have to be used for imputation, which might be a critical point for the implementation of GS$_{LD}$ because a minimum accuracy of GEBV$_{LD}$ has to be achieved to use it in conjunction or in favour of GS$_{HD}$ for selecting females.
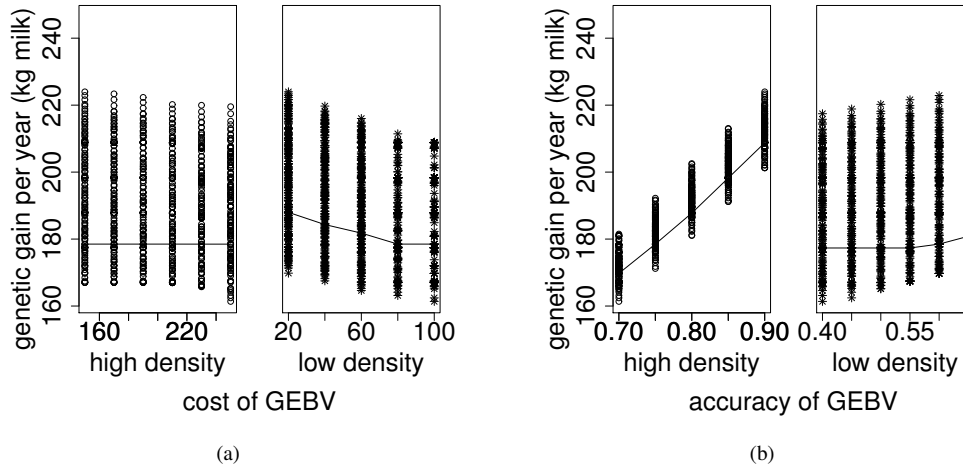
Figure 1: ΔG_a as a function of costs (1(a)) and accuracies (1(b)) of GEBVs.

Only those breeding schemes are plotted maximising ΔG_a for a given combination of $r_{GEBV,HD}$, $r_{GEBV,LD}$, $C_{GEBV,HD}$ and $C_{GEBV,LD}$. The values of the reference scenario, whereupon the abscissa parameter was varied, are given by the continuous line.

REFERENCES

Brent R. P. (1973) Algorithms for Minimisation without Derivatives, Prentice-Hall Series in Automatic Computation, Prentice-Hall, Englewood Cliffs, New Jersey.

Dekkers J. C. M. (2007) *J. Anim. Breed. Genet.* **124**(6):331.

Genz A. (1992) *J. Comput. Graph. Stat.* **1**:141 .

Hayes B. J., Bowman P. J., Chamberlain A. J. and Goddard M. E. (2009) *J. Dairy Sci.* **92**(2):433.

Schaeffer L. R. (2006) *J. Anim. Breed. Genet.* **123**(4):218.

Weigel K., de los Campos G., Vazquez A., Rosa C. V. T. G., Gianola D., O'Connell J., VanRaden P. and Wiggans G. (2010) in Proceedings of the 9th World Congress on Genetics applied to Livestock Production, Leipzig, Germany, August 1-6, 2010.