

Mating allocations in Nordic Red Dairy Cattle using genomic information

C. Bengtsson^{1,2*}, H. Stålhammar¹, J.R. Thomasen¹, S. Eriksson², W.F. Fikse³ and E. Strandberg²

¹VikingGenetics, VikingGenetics Sweden AB, 53294 Skara, Sweden ² Dept. of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Box 7023, 75007 Uppsala, Sweden ³ Växa Sverige, Box 288, 75105 Uppsala, Sweden; *chben@vikinggenetics.com

Abstract

In this study, we compared mating allocations in Nordic Red Dairy Cattle using genomic or pedigree information. We used linear programming to optimise different economic scores within each herd, considering genetic level, semen cost, the economic impact of recessive genetic defects, and genetic relationships. We found high correlations (≥ 0.83) between the pedigree and genomic relationship measures. The mating results showed that it was possible to reduce the different genetic relationships between parents with minimal effect on genetic level. Including the cost of known recessive genetic defects eliminated expression of genetic defects. It was possible to reduce genomic relationships between parents with pedigree information, but it was best done with genomic measures. Linear programming maximised the economic score for all herds studied within seconds, which means that it is suitable for implementation in mating software to be used by advisors and farmers.

Introduction

Mating programs are an important support tool for livestock breeders, helping them to identify the best parental matings to maximise genetic level and avoid mating between closely related individuals, preventing excessive inbreeding (Carthy et al., 2019; Bérodier et al., 2021). New genetic insights at single nucleotide level can be used in mating programs. Single nucleotide polymorphism (SNP) markers can give information about major genes and genetic defects. Minimising the probability of obtaining offspring homozygous for a lethal recessive genetic defect is of economic importance for farmers (Pryce et al., 2012). SNP markers also offer the possibility to reduce genomic relationships between parents when optimising mating plans. Various methods have been proposed for calculation of genomic relationships, including SNP-by-SNP relationships as described by e.g., VanRaden (2008). Further, there are methods using shared genomic segments, as described by e.g., Cara et al. (2013), which aim to reduce the number of runs of homozygosity in the offspring.

The objective of this study was to investigate the ability of different mating scenarios to maximise expected genetic level in the next generation, limiting parent relationship and the probability of expression of genetic defects. We investigated all scenarios at herd level using real data. We used linear programming to optimise economic scores within each herd, considering genetic level, semen cost, the economic impact of recessive genetic defects, and five different measures of relationships (two pedigree-based and three genomic-based).

Materials & Methods

Animals. Breeding values, pedigree data, SNP data, and data on the carrier status of genetic defects were obtained from the Nordic Cattle Genetic Evaluation. Genotype information for a total of six genetic defects were available. We selected 9,841 genotyped Red Dairy Cattle (RDC) females born in Denmark, Finland, or Sweden in 2019 to simulate matings. All females included belonged to herds with 20 or more genotyped females. The females were potentially mated to 50 genotyped RDC bulls from the Nordic breeding cooperative VikingGenetics.

Relationship measures. We used two different pedigree relationship coefficients. The first relationship coefficient traced the pedigree three generations back from the parents of the potential mating ($a_{3\text{Gen}}$). The second pedigree relationship coefficient was based on all available pedigree information (a_{AllGen}). We used three different genomic relationship coefficients, one SNP-by-SNP genomic relationship coefficient (g_{SNP}) according to VanRaden (2008). The two segment-based genomic relationship coefficients were based on different minimum lengths of segments: 1 centimorgan (cM) (g_{SEG1}) and 4 cM (g_{SEG4}).

Economic score. For each potential mating between female i and bull j , we calculated an economic score similarly to Bérodier et al. (2021) and Pryce et al. (2012):

$$Score_{ij} = \left(\frac{NTM_i + NTM_j}{2} + \lambda F_{ij} \right) \times \text{prob}(\text{♀}) - \sum_{r=1}^{n_r} p(aa)_r \times v_r - \text{semen cost}$$

where NTM_i and NTM_j are the value of Nordic Total Merit units in Euros (€) for female i and bull j , λ is the economic consequence of an 1% increase in inbreeding similarly to Cole (2015), F_{ij} is the pedigree or genomic based co-ancestry (Relationship/2), $\text{prob}(\text{♀})$ is the probability of producing a female conceptus, n_r is the number of recessive genetic defects considered, $p(aa)_r$ is the probability of expression of a genetic defect r , v_r is the economic cost associated with the recessive genetic defect r similarly to Cole (2015) and Bérodier et al. (2021), and semen cost is the average amount (€) spent on semen for a pregnancy. In addition, an economic score based on only NTM and semen cost was calculated (MaxNTM). Mate allocation was programmed in R version 3.6.3 (R Core Team, 2020). Linear programming optimisation was performed with the ‘Lp_solve’ package in R (Berkelaar et al., 2020). The objective of linear programming was to maximise the average economic score in the herd with a maximum of 5% of females per bull and herd. The mating R script was provided by Bérodier et al. (2021).

Results

The correlations between the genetic relationship coefficients were at least 0.83 or higher (Table 1). The mean value of the relationship coefficients between all possible combinations of females and males (Random) ranged from 0.009 to 0.188 (Table 2). The strongest correlation was between a_{AllGen} and $a_{3\text{Gen}}$ ($r=0.99$), and the second strongest was between g_{SEG1} and g_{SEG4} ($r=0.98$). The strongest correlation between pedigree and genomic relationships was between a_{AllGen} and g_{SEG4} ($r = 0.88$) (Table 1). The coefficients of regression of $a_{3\text{Gen}}$ and g_{SNP} on a_{AllGen} were close to 1, and the coefficients of regression of g_{SEG1} and g_{SEG4} on a_{AllGen} were both 0.94.

It was possible to maximise economic score with limited impact on the average NTM level (Table 2). Including the cost of the known recessive genetic defects when optimising mating strategies eliminated the risk of loss from a genetic defect, regardless of which genetic relationship was used. When just maximising NTM minus semen cost (MaxNTM), the NTM level improved compared with Random, but it resulted in higher average genetic relationship coefficients than Random and did not reduce the probability of loss from a genetic defect. Including a genomic relationship in the economic score also kept the other genomic relationship averages

at a low level. For example, with the constraint of 5% females per bull and herd, including g_{SNP} in the objective function (scenario GSNP) resulted in a g_{SEG1} of 0.148, compared with 0.143 with GSEG1 (Table 2). Using the pedigree relationships also reduced the genomic relationships compared with Random and MaxNTM, but not as much as using genomic relationships in the objective function.

Table 1. Correlations between the different relationship coefficients for all possible combinations of 9,841 females and 50 bulls.

Relationship ¹	aAllGen	g _{SNP}	g _{SEG1}	g _{SEG4}
a _{3Gen}	0.99	0.88	0.83	0.87
aAllGen		0.88	0.85	0.88
g _{SNP}			0.9	0.93
g _{SEG1}				0.98

¹a_{3Gen} = pedigree relationships using three generations of ancestors, aAllGen = pedigree relationships using all available pedigree information, g_{SNP} = genomic relationship calculated according to VanRaden (2008), g_{SEG1} (g_{SEG4}) = genomic segment-based relationship according to de Cara et al. (2013) with a minimum segment length of 1 (4) centimorgan

Table 2. Comparison of outcomes of planned matings of 9,841 females for 7 mating scenarios in Nordic Red Dairy Cattle using various comparison criteria¹.

Comparison criterion ¹	Ran-dom	5% females/bull scenarios ²					
		Max NTM	3Gen	AllGen	G _{SNP}	G _{SEG1}	G _{SEG4}
NTM	19.5	20.8	20.8	20.8	20.8	20.7	20.8
a _{3Gen}	0.028	0.033	0.007	0.009	0.014	0.014	0.014
aAllGen	0.066	0.070	0.046	0.043	0.050	0.050	0.050
g _{SNP}	0.009	0.014	-0.012	-0.016	-0.038	-0.034	-0.033
g _{SEG1}	0.188	0.191	0.167	0.163	0.148	0.143	0.146
g _{SEG4}	0.115	0.119	0.094	0.091	0.078	0.075	0.074
Expression of genetic defects (%)	0.4	0.4	0	0	0	0	0

¹Average Nordic total merit (NTM) level, average five different genetic relationships between parents, the probability of expression of genetic defects.

²Maximum percent of females per bull and herd set to 5%.

Random: All possible combinations of 9,841 females and 50 bulls

MaxNTM: mates were selected based on maximising an economic score including NTM, sexed semen, and semen cost.

3Gen, AllGen, G_{SNP}, G_{SEG1}, G_{SEG4}: mates were selected based on maximising an economic score including NTM, sexed semen, semen cost, a penalty for genetic defects and a pedigree relationship include three generations (a_{3Gen}) or all available ancestors (aAllGen), or a genomic relationship calculated according to VanRaden (2008) (g_{SNP}), according to Cara et al. (2013) with a minimum genomic segment length of 1 centimorgan (g_{SEG1}) or 4 centimorgan (g_{SEG4}).

Discussion

The results show that it is possible to reduce genetic relationships between RDC parents in herds with minimal effect on the genetic level. Including the cost of known recessive genetic defects when optimising mating strategies eliminated expression of known genetic defects, regardless of the genetic relationship used.

The strong correlation estimated between pedigree and genomic relationships confirms that dairy pedigrees are well documented in the Nordic countries. The correlations between the pedigree relationship and genomic relationship estimates were high: ≥ 0.83 for $a_{3\text{Gen}}$ and ≥ 0.85 for a_{AllGen} (Table 1). Carthy et al. (2019) reported a correlation of 0.57 between pedigree relationships and genomic relationship, which is lower than in other studies (0.67-0.88) (VanRaden et al., 2011; Pryce et al., 2012). Pryce et al. (2012) concluded that pedigree depth plays a major role for the strength of correlation between pedigree relationships and genomic relationships. They found that when the number of generations of recorded ancestry was 2, 4, 6, and 8, the correlations were 0.67, 0.73, 0.84, and 0.87, respectively (Pryce et al., 2012).

In general, genomic relationships were better at reducing pedigree relationships than pedigree relationships were at reducing genomic relationships. Furthermore, using GSEG1 reduced gSEG1 compared with Random from 0.188 to 0.143, and 3Gen reduced gSEG1 to 0.163, i.e., the relative difference was 55%. Hence, using genomic relationships could be an overall better and safer option than using pedigree relationships in keeping all average relationships studied low.

Conclusions

We studied mating allocations in Nordic Red Dairy Cattle and found that it was possible to reduce genetic relationships between parents with minimal effect on genetic level. Including the cost of known recessive genetic defects entirely eliminated the risk of loss due to the six known genetic defects. It was possible to reduce genomic relationships between parents with pedigree measures, but it was best done with genomic measures. Linear programming maximised the economic score for all herds studied within seconds, which means that the method is suitable for implementing in mating software to be used by advisors and farmers.

References

- Berkelaar and others. 2020. LpSolve: Interface to “Lp_solve” v. 5.5 - Available at: <https://CRAN.R-project.org/package=lpSolve>
- Bérodier M., Berg P., Meuwissen T., Boichard D., et al. 2021. Animal 15:100016. <https://doi.org/10.1016/j.animal.2020.100016>
- de Cara M.Á.R., Villanueva B., Toro M.Á., and Fernández J. 2013. Mol. Ecol. 22:6091–6099. <https://doi.org/10.1111/mec.12560>
- Cole J.B. 2015. Genet. Sel. Evol. 47:94. <https://doi.org/10.1186/s12711-015-0174-9>
- Carthy T.R., McCarthy J., and Berry D.P. 2019. J. Dairy Sci. 102:8210–8220. doi:10.3168/jds.2019-16283. <https://doi.org/10.3168/jds.2019-16283>
- Pryce J.E., Hayes B.J., and Goddard M.E. 2012. J. Dairy Sci. 95:377–388. <https://doi.org/10.3168/jds.2011-4254>
- R Core Team. 2020. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. - Available at: <https://www.r-project.org/>
- VanRaden P.M. 2008. J. Dairy Sci. 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- VanRaden P.M., Olson K.M., Null D.J., and Hutchison J.L. 2011. J. Dairy Sci. 94:6153–6161. <https://doi.org/10.3168/jds.2011-4624>