Acta Universitatis Agriculturae Sueciae

Doctoral Thesis No. 2024:6
Faculty of Veterinary Medicine and Animal Science

Dairy cattle mating plans at herd level using genomic information

Christian Bengtsson

Dairy cattle mating plans at herd level using genomic information

Christian Bengtsson
Faculty of Veterinary Medicine and Animal Science
Department of Animal Breeding and Genetics
Uppsala

SWEDISH UNIVERSITY
OF AGRICULTURAL
SCIENCES

DOCTORAL THESIS
Uppsala 2024

Acta Universitatis Agriculturae Sueciae
2024:6

Cover: Red cow and DNA strain.
Photo: Elisabeth Theodorsson. Design: Johanna Vuori

ISSN 1652-6880
ISBN (print version) 978-91-8046-276-1
ISBN (electronic version) 978-91-8046-277-8
https://doi.org/10.54612/a.6ipjv4r8nv

Swedish University of Agricultural Sciences, Department of Department of Animal Breeding and Genetics, Uppsala, Sweden

Print: SLU Grafisk service, Uppsala 2024

Dairy cattle mating plans at herd level using genomic information

Abstract

Genotyping of dairy cattle can benefit farmers by increasing the accuracy of breeding values and improving mating plans at herd level. Validation studies on breeding values in this thesis revealed that for the vast majority of traits analysed in Holstein, Jersey and Nordic Red Dairy Cattle, genomically enhanced breeding values for virgin heifers were able to predict cow performance significantly more accurately than parent average breeding values. Linear programming was used to optimise matings based on economic scores for Red Dairy Cattle and Holstein, considering genetic level, semen cost, recessive genetic defects, and genetic relationship. For Holstein, we also studied polledness and beta-casein genotype. The mating results for Red Dairy Cattle showed that it was possible to reduce genetic relatedness between parents and eliminate expression of genetic defects with minimal effect on genetic level. Similar results were achieved for Holstein cattle, in which it was also possible to increase the frequency of polled and beta-casein genotype A2A2 offspring without negatively impacting other comparison criteria. Evaluation of the long-term impact of genomic mating allocations in a simulation study revealed that planning matings with genomic information at herd level involves important risk management decisions, e.g. a trade-off between using fewer bulls to increase the polled allele frequency more quickly and using more bulls to reduce the rate of inbreeding and the variation in carrier frequency for genetic defects.

Keywords: genomic relationship, pedigree relationship, mating program, linear programming, genomic breeding value, genotyping, dairy cow

Användning av genomisk information vid parningsplanering av mjölkkor på besättningsnivå

Sammanfattning

I denna avhandling undersökte vi fördelarna med genomisk analys av mjölkkor, i form av högre säkerhet för genomiska avelsvärden och förbättrad parningsplanering. Vi validerade avelsvärden och fann att genomiska avelsvärden förutsåg kvigornas framtida egenskaper betydligt bättre än härstamningsindex för majoriteten av analyserade egenskaper hos nordiska röda raser, holstein och jersey. Vi använde linjär programmering för att optimera parningar med hjälp av ekonomiska poängsummor för nordiska röda raser och holstein, och tog hänsyn till genetisk nivå, seminkostnad, recessiva genetiska defekter och genetiskt släktskap. För holstein undersökte vi även anlag för kullighet och Beta-kasein-genotyp. Parningsresultaten för nordiska röda raser visade att det var möjligt att minska genetiskt släktskap mellan föräldrar och eliminera uttryck av genetiska defekter med minimal påverkan på genetisk nivå. För holstein fann vi liknande resultat och det var också möjligt att öka andelen som var kulliga och hade önskad Beta-kasein-genotyp utan att negativt påverka andra jämförelsekriterier. Slutligen undersökte vi de långsiktiga effekterna av genomiska parningar genom en simuleringsstudie. Resultaten visade att parningar med genomisk information på besättningsnivå innebär viktiga riskhanteringsbeslut, såsom avvägningen mellan att använda färre tjurar för att snabbare öka frekvensen av hornlöshet och att använda fler tjurar för att minska inavelstakten och variationen i bärarfrekvens för genetiska defekter.

Nyckelord: Genomiskt släktskap, släktskap baserat på stamtavla, parningsprogram, linjärprogrammering, genomiska avelsvärden, DNA-analys, mjölkko, valideringsstudie

Dedication

To dairy farmers around the world: this thesis is dedicated to you amazing and passionate people who produce nutritious foods for humanity. You overcome many challenges and uncertainties in your daily work, and you always pursue your noble mission.

Contents

Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I. Bengtsson, C., Stålhammar, H., Strandberg, E., Eriksson, S. and Fikse, W.F. (2020). Association of genomically enhanced and parent average breeding values with cow performance in Nordic dairy cattle. *Journal of Dairy Science* 103, 6383-6391.

II. Bengtsson, C., Stålhammar, H., Thomasen, J.R., Eriksson, S., Fikse, W.F. and Strandberg, E. (2022). Mating allocations in Nordic Red Dairy Cattle using genomic information. *Journal of Dairy Science* 105(2), 1281-1297.

III. Bengtsson, C., Stålhammar, H., Thomasen, J.R., Fikse, W.F., Strandberg, E. and Eriksson, S. (2023). Mating allocations in Holstein combining genomic information and linear programming optimization at the herd level. *Journal of Dairy Science* 106(5), 3359-3375.

IV. Bengtsson, C., Stålhammar, H., Thomasen, J.R., Fikse, W.F., Strandberg, E., Eriksson S. and Johnsson, M. (2023). Simulation of long-term impact of dairy cattle mating programs using genomic information. Manuscript.

The contribution of Christian Bengtsson to the papers included in this thesis was as follows:

I. Planned the study in collaboration with the supervisors and retrieved data. Used and modified scripts from his Master's thesis, where a pilot study on Red Dairy Cattle was performed under the supervision of W.F. Fikse. Analysed the results and drafted the manuscript in collaboration with the co-authors.

II. Participated in planning the study together with the supervisors. Used the mating R script provided by Bérodier et al. (2021), which was modified to fit the Nordic data. Investigated suitable programs to calculate genetic relationships. Analysed results, drafted and revised the manuscript together with the co-authors.

III. Had the main responsibility for planning the study and retrieved additional data. Modified the mating R scripts further and analysed the results. Drafted the manuscript, on which the co-authors provided feedback.

IV. Had the main responsibility for planning the study and finding external expertise. Set up the simulation together with the co-authors. Ran the simulations and analysed the results. Drafted the manuscript, on which the co-authors provided feedback.

Abbreviations

AI	Artificial insemination
BLUP	Best linear unbiased prediction
DFS	Denmark, Finland and Sweden
EBV	Estimated breeding value
GEBV	Genomically enhanced breeding value
NTM	Nordic Total Merit index
MOET	Multiple ovulation and embryo transfer
OCS	Optimum contribution selection
OPU	Ovum pick-up
PA	Parent average breeding value
QTL	Quantitative trait locus/loci
RDC	Red Dairy Cattle
ROH	Runs of homozygosity
SNP	Single nucleotide polymorphism
TMI	Total merit index

1. Introduction

Breeding work on dairy cattle has helped to improve the productivity and profitability of the dairy sector. In the past, such work focused on milk yield, but health and fertility have now become important breeding objectives (Oltenacu & Broom, 2010). Recent animal breeding research has also included traits related to sustainability and climate impact, *e.g.* efficiency, methane emissions and resilience (Løvendahl *et al.*, 2018; Bengtsson *et al.*, 2022).

A technological breakthrough that has transformed dairy cattle breeding in the past 15 years is the development of genomic selection (Meuwissen *et al.*, 2001; Schaeffer, 2006). It offers many advantages for dairy cattle breeding, such as higher accuracy of breeding values for young animals (which enables shorter generation intervals) and enhanced use of reproductive technologies such as multiple ovulation and embryo transfer (MOET) and ovum pick-up (OPU) (Thomasen *et al.*, 2016). Genomic selection makes it more cost-effective to breed for novel traits, as it reduces the need for large daughter groups from each bull (Henryon *et al.*, 2014).

Genotyping, which was once costly and used mainly for artificial insemination (AI) bulls or candidates, has now become more affordable and accessible for female selection and management. At present (2023 values), the average cost of genotyping per animal is around €20-25, making it a worthwhile investment for dairy farmers if the results are used actively (Hjortø *et al.*, 2015; Newton & Berry, 2020). At herd level, genomic selection is often combined with sexed dairy bull semen and beef bull semen, where the best heifers are inseminated with dairy bull semen and older cows are inseminated with beef bull semen (Hjortø *et al.*, 2015; Clasen *et al.*, 2021). Genomic data can also help with mating plans. For instance, single nucleotide polymorphisms (SNPs) can be used to estimate genomic

relationships and monogenic traits can be detected and considered in mating choices (Carthy *et al*., 2019; Bérodier *et al*., 2021). This thesis explored how to use genomic data effectively at herd level, particularly when planning matings.

2. Background – Dairy cattle breeding in the genomic era

Genomic selection is a method that predicts phenotypic traits based on many genetic markers that cover the whole genome, as all quantitative trait loci (QTL) are assumed to be in linkage disequilibrium with at least one marker. Genetic markers are DNA sequences that vary between individuals and can be detected by different technologies, such as SNP arrays or whole genome sequencing. The estimated marker effects are then used to calculate genomically enhanced breeding values (GEBV) for selection of candidates of interest (Meuwissen *et al*., 2001; Schaeffer, 2006; Rajora, 2019). In practice, this is achieved by first estimating the combined genetic effects for each individual of a reference population and then using the information obtained to infer GEBV for the selection candidates (Figure 1).

By enabling higher breeding value accuracies for young animals, genomic selection makes it possible to shorten the generation interval. Dairy cattle breeding programmes have widely adopted genomic selection since it was first implemented in 2008 (Hayes *et al*., 2009). Compared with progeny testing schemes in dairy cattle, genomic selection has been shown to increase the annual genetic gain by up to around 100%, mainly due to shorter generation interval (Schaeffer, 2006; Hayes *et al*., 2009). It also enhances use of reproductive technologies, such as MOET and OPU (Thomasen *et al*., 2016).

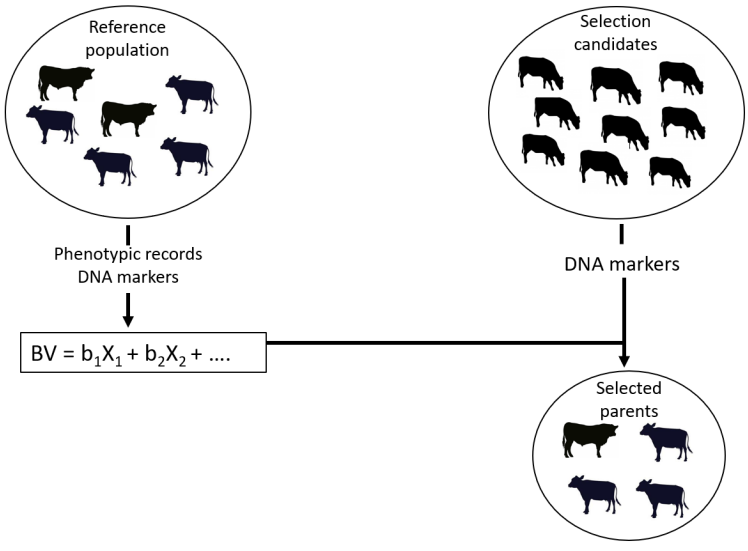


Figure 1. Illustration of the concept of genomic selection, which involves a reference population with both genotype and phenotype data and a pool of selection candidates with genotype data. Information on breeding values (BV) is used to choose parents from the selection candidate pool.

Genomic evaluation is a large research field within animal breeding that aims to improve the accuracy of GEBV while overcoming computational and bias problems (*e.g*. Legarra *et al*., 2009; Christensen & Lund, 2010; Misztal *et al*., 2020). Different methods have been developed to combine genomic information with pedigree data and phenotypes, often classified into multi-step and single-step approaches. In multi-step approaches, SNP effects are estimated based on phenotypes or de-regressed proofs. GEBV are then composed of an index with parent average, direct genomic value and a deduction of parent average to eliminate double counting. Single-step methods combine genomic and pedigree relationships to automatically create an index with all sources of information (Misztal *et al*., 2020).

Analysis of SNP marker data for dairy cattle has enabled new discoveries on the functional role of single nucleotide variations. One of these discoveries is the occurrence of missing genotypes within populations. This suggests that, if absent, a certain haplotype (which may comprise several adjacent SNPs) is a recessive deleterious (lethal) allele. Such alleles have previously been difficult to detect, except as reduced fertility of a specific bull or daughter group (because the conceptus often dies early *in utero*).

Reducing the chance of producing offspring that are homozygous for recessive genetic defects is beneficial for farm finances (Pryce *et al*., 2012) and also for animal health and welfare (EFFAB, 2020). Moreover, SNP arrays can include desirable monogenic traits such as polledness. Cattle dehorning has been common practice for many years and is done for various reasons, such as preventing injury to other cattle and increasing safety for animal handlers. However, dehorning has been proven to cause changes in animal behaviour, neuroendocrine responses and physiology, which may indicate that it is a stressful and painful procedure (Stock *et al*., 2013). Since 2022, organic farms in the European Union need to obtain a permit if they want to dehorn their cattle (EU Commission Regulation No. 889/2008; EU (European Union), 2008). Another example of a monogenic trait is beta-casein variant. Casein represents about 80% of bovine milk proteins, with beta-casein comprising around 35% of casein protein (Indyk *et al*., 2021). The most widespread variants of beta-casein in bovine milk are A1 and A2. Animals that are homozygous for the A2 allele produce so-called A2 milk, which is often advertised as a healthier alternative than regular cow milk, although the human health advantages of drinking A2 milk are still under debate (Summer *et al*., 2020). Despite this lack of confirmed benefits, some countries are seeking to increase consumption of A2 milk and some dairies pay extra for A2 milk (Bisutti *et al*., 2022).

Use of SNP markers also offers new opportunities for measuring and managing inbreeding at genomic level (VanRaden, 2008; de Cara *et al*., 2013). There are two main types of genomic measures: SNP-based measures (*e.g*. VanRaden, 2008) and runs of homozygosity (ROH) (*e.g*. Purfield *et al*., 2012). Inbreeding leads to an increase in autozygosity throughout the genome in the form of ROH. In a meta-analysis on the effects of inbreeding in livestock, Doekes *et al*. (2021) concluded that genomic measures are better than pedigree measures for indicating inbreeding depression, but did not find any difference between SNP-based measures and ROH. However, those authors highlighted the limited number of studies investigating ROH and inbreeding depression, and the arbitrary definitions of ROH. In principle, ROH would contribute to inbreeding depression if they contain recessive deleterious alleles (Charlesworth & Willis, 2009). Long ROH reflect new inbreeding and are expected to contain more deleterious alleles than short ROH, due to purging and recombination through the generations (Stoffel *et al*., 2021).

2.1 Mating plans

Selection and mating are two key aspects of animal breeding that influence genetic improvement and variation in populations (Jansen & Wilton, 1985; Weigel & Lin, 2000). Non-random mating has most likely been around in animal breeding for as long as selection. Proper mating plans can help breeders control inbreeding, which can be managed at two levels: (i) population level, where the inbreeding rate can be limited to a desired level while maximising the genetic gain by optimising the long-term contributions of a selected number of breeding animals, and (ii) individual level, where large inbreeding coefficients in offspring can be avoided to reduce the impact of inbreeding depression (Pryce *et al.*, 2012; Liu *et al*., 2017). Farmers have previously tried to control inbreeding by avoiding matings of genetically related animals. However, as relationships within the breed increase, it becomes difficult to avoid such matings without the aid of a computer. Hence, mating programmes have become 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, thus preventing excessive inbreeding (Carthy *et al*., 2019; Bérodier *et al*., 2021). Various methods for calculation of genomic relationships have been proposed, including SNP-by-SNP relationships as described by *e.g*. VanRaden (2008). Further, methods using shared genomic segments, as described by *e.g*. de Cara *et al*. (2013), aim to reduce the number of ROH in the offspring.

The new genetic possibilities require updated methods that combine relevant information based on their economic value when setting up mating plans. Several studies have created economic scoring systems to rank each potential mating (Pryce *et al*., 2012; Carthy *et al*., 2019; Bérodier *et al*., 2021). The economic score often includes genetic level, expected inbreeding, the probability of conceiving an offspring homozygous for a genetic defect and semen price (Bérodier *et al*., 2021). The economic score is flexible and can be adjusted to match economic conditions on a specific farm, such as a price premium for A2 milk and/or polled animals. Using linear programming to maximise the mean economic score of every herd, subject to necessary constraints, is a fast and effective method (Carthy *et al*., 2019; Bérodier *et al*., 2021). Linear programming has also been shown to outperform other mating methods, such as sequential mate allocation (Sun *et al*., 2013; Carthy *et al*., 2019; Bérodier *et al*., 2021).

2.2 The net benefit of genotyping

Besides using the genomic information in mating plans, genotyping at herd level can provide several other benefits for dairy farmers. For example, it can help them select the best females for breeding and replacement or identify females for embryo transfer or *in vitro* fertilisation (Newton & Berry, 2020). The net benefit of genotyping females at herd level depends on many factors, such as genotyping price, the accuracy difference between parent average breeding values (PA) and GEBV, the proportion of females kept as replacement, parentage errors, age when first progeny is born, the standard deviation of the breeding goal, the value of better mating advice through genomic-assisted mating plans, and the value of identifying elite females and combining genotyping with state-of-the-art reproductive technologies (*e.g*. MOET, OPU and sexed semen) (Calus *et al*., 2015; Hjortø *et al.*, 2015; Newton & Berry, 2020). Hence, estimating the overall value of genotyping is challenging because it is influenced by multiple factors that may vary depending on country, region and herd. Several studies have tried to quantify the net benefit of genotyping candidate females for replacement using different methods and assumptions (Calus *et al*., 2015; Hjortø *et al*., 2015; Newton & Berry, 2020). Hjortø *et al*. (2015) found that in Denmark, Finland and Sweden (DFS), genomic testing was profitable in most cases at a genotyping price of €30 when combined with sexed semen and beef semen. The price of genotyping varies slightly between the Nordic countries and is currently around €25 in Sweden (Växa) and around €20 in Denmark (VikingDenmark).

2.3 Validation of genomically enhanced breeding values

To increase confidence in genomic technology among farmers, a clear illustration of the validity of the relationship between genomic predictions and future phenotypes is needed (Pryce *et al*., 2012). GEBV can be validated in different ways. One way is cross-validation, which involves dividing the available dataset into validation and training sets. By masking observations of all individuals in the validation set and predicting the observations or estimated breeding value (EBV) with a model based on individuals in the training set only, the correlation between masked phenotypes or EBV and predicted values for the validation individuals can be estimated. This correlation then reflects the accuracy of prediction (de Roos *et al*., 2009). A

disadvantage of validating GEBV against conventional EBV is that training and validation sets are rarely strictly independent (Su *et al*., 2010). Thus Yao *et al.* (2015) used genotypes and health data to predict future phenotypes, taking correlations between predicted values and phenotypes as measurements of accuracy. To illustrate the accuracy of GEBV compared with PA, Weigel *et al*. (2015) divided cows into quartiles based on their virgin heifer GEBV and sire predicted transmitting ability, and thereafter calculated actual cow performances for each quartile.

2.4 Genotyping in the Nordic countries

The establishment of the Nordic Cattle Genetic Evaluation in 2002 enabled greater cooperation between AI organisations in the Nordic countries. The differences between the Nordic countries were small at that time, according to a study on genotype × environment interactions by Kolmodin *et al*. (2002), and a joint breeding programme was established. Three breeds in the DFS countries have genomic breeding schemes: Red Dairy Cattle (RDC), Holstein and Jersey. Farmers in DFS quickly adopted the new technology and 75% of all inseminations in 2007 were from progeny-tested bulls, but by 2017 this proportion had decreased to less than 5% (Hans Stålhammar, Emeritus/Senior Researcher, VikingGenetics, personal communication, June 20, 2023). During the same period, the genetic improvement in the Nordic Total Merit (NTM) index per year in the two major breeds (RDC and Holstein) almost doubled (NAV, 2023).

Almost all progeny-tested bulls in DFS were genotyped some years after the introduction of SNP genotyping. Genotyping of cows and virgin heifers in DFS started on a large scale in 2012 with the VikingGenetics genotyping project. Initially, the main objective was to include genotyped females in the reference population and thereby increase the accuracy of GEBV. This was especially important for RDC and Jersey, which had more limited reference populations based on bulls than the Holstein breed. In 2022, around 107,500 females born in DFS were genomically tested (corresponding to approximately 20-25% of purebred born females), compared with just under 25,000 females in 2014 (Ulrik Sander Nielsen, Senior Researcher, SEGES, personal communication June 20, 2023) (Figure 2).

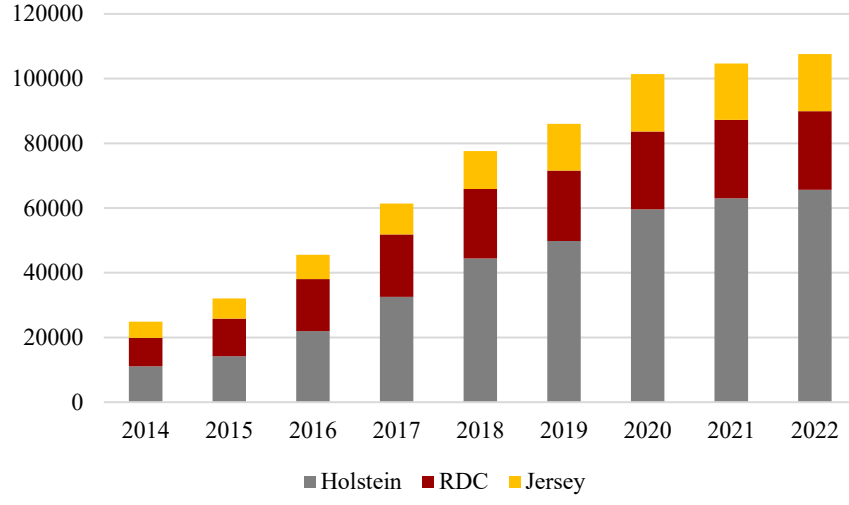


Figure 2. Number of female Holstein, Nordic Red Dairy Cattle (RDC) and Jersey genotyped per year in Denmark, Finland and Sweden in the period 2014-2022. Source: Ulrik Sander Nielsen, Senior Researcher, SEGES, personal communication June 20, 2023.

2.5 Description of knowledge gaps

In DFS, GEBV has not been validated on phenotypes with validation results published in scientific journals. This represents a knowledge gap in current research, as validating breeding values on phenotypes can help farmers understand how their animals' breeding values work in practice, especially when the validation is based on their own farm records. This could increase confidence in genomic technology among farmers.

Use of genomic relationships for matings is a novel research topic in DFS. Moreover, most international research to date has focused mainly on SNP-by-SNP relationships (*e.g.* Bérodier *et al*., 2021; Carthy *et al*., 2019) and the effects of using a segment-based relationship are not well studied, especially for herd-level matings.

At present, there are more than 20 monogenic traits available for mating programmes in DFS (A1 in Appendix 1), including recessive genetic defects, polledness and casein traits. However, these traits have only been included recently and farmers and advisors need guidance on how to handle them when information is available for both males and females. Trait data on

polledness and beta-casein status in females have been available since 2021 in DFS, but few Nordic or international studies have investigated mating planning at herd level with both desired and undesired monogenic traits.

In some studies using real data, genomic information and linear programming have been applied to mating programmes (Carthy *et al*., 2019; Bérodier *et al*., 2021). However, these studies have mainly focused on the current breeding population and the immediate offspring, while the long-term consequences of these mating strategies could not be evaluated. Dairy cattle breeding is a long-term process with cumulative effects, and it is important to understand the consequences of different breeding strategies (Doublet *et al*., 2019).

3. Aims of the thesis

The main aims of the work in this thesis were to compare the accuracy of genomically enhanced breeding values (GEBV) with parent average breeding values (PA) in prediction of cow performance and to provide guidance on how to best incorporate genomic information into mating programmes for Nordic dairy cattle. Specific objectives were:

- To compare the ability of virgin heifer genomically enhanced breeding values and parent average breeding values to predict future cow performance (Paper I)

- To evaluate the ability of different approaches for mating allocation in order to maximise expected genetic level, while limiting parent relatedness and minimising the probability of expression of genetic defects in the next generation (Paper II)

- To assess the ability of different approaches for mating allocation that also consider favourable monogenic traits (polledness and beta-casein) (Paper III)

- To investigate the long-term impact of mating programmes using genomic information on genetic gain, genetic diversity and monogenic traits (Paper IV)

4. Summary of Papers I-IV

Approximately 85% of farms in DFS are enrolled in a national milk recording scheme. This enables validation of GEBV with phenotype data on a large scale. The aim in Paper I was to compare the ability of virgin heifer GEBV and PA in prediction of future cow performance.

New genetic insights at single nucleotide level can be used in mating programmes. Single nucleotide polymorphism markers can give information about major genes and genetic defects and also offer the possibility to reduce genomic relationships between parents when making mating plans. The objective in Papers II and III was to investigate the ability of different approaches for mating allocation to maximise expected genetic level, limiting parent relatedness, while also considering monogenic traits. In both studies, all scenarios at herd level were investigated using real data and linear programming was used to optimise different economic scores within each herd. Paper II focused on RDC and evaluated the economic score for this breed, 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). Paper III focused on Holstein dairy cattle, with the economic score extended to also included polledness and beta-casein. In Paper IV, which was inspired by the work in Papers II and III, different scenarios were simulated and compared in order to investigate the long-term impact of mating programmes using genomic information and linear programming.

4.1 Association of genomically enhanced and parent average breeding values with cow performance in Nordic Red Dairy Cattle (Paper I)

4.1.1 Materials and methods

The ability of virgin heifer GEBV and PA to predict future cow performance was assessed in Paper I. Twelve different traits in first parity were analysed, including production, conformation, fertility and other functional traits. Phenotype data were obtained from national milk recording schemes and breeding values from the Nordic Cattle Genetic Evaluation (NAV, 2019). Direct genomic breeding values were calculated using genomic best linear unbiased prediction (BLUP) and combined with traditional breeding values, using bivariate blending (Mäntysaari & Strandén, 2010; Taskinen *et al*., 2013). The data covered 14,862 RDC, 17,145 Holstein and 7,330 Jersey genotyped virgin heifers born between 2013 and 2015 in DFS. Phenotypes adjusted for systematic environmental effects were used as measures of cow performance and were named according to the respective trait, *e.g*. cow adjusted milk yield was denoted $Milk_{Adj}$. Correlations between breeding values and adjusted phenotypes were calculated using Statistical Analysis Software (SAS) version 9.4 (SAS Institute Inc., Cary, NC). A 95% confidence interval using Fisher's Z transformation was applied to assess the significance of differences between correlations. The PROC RANKS procedure in SAS was used to rank cows into four quartiles across herds for GEBV or PA.

4.1.2 Results and comments

For RDC and Holstein, all correlations between breeding values and adjusted phenotypes were significantly stronger for GEBV than for PA (Table 1).

Table 1. Relative change[1] (%) in correlation of genomically enhanced breeding values (GEBV) and of parent average breeding values (PA) with cow adjusted (Adj) phenotypes for Red Dairy Cattle (RDC), Holstein and Jersey, n/a = not applicable

Traits	RDC	Holstein	Jersey
$Milk_{Adj}$	+63%	+49%	+59%
Fat_{Adj}	+46%	+46%	+60%
$Protein_{Adj}$	+52%	+44%	+63%
SCS_{Adj}	+63%	+65%	+67%
Clinical $mastitis_{Adj}$	+62%	+38%	+13%[2]
IFL_{Adj}[3]	+68%	+64%	+78%
$Udder_{Adj}$	+42%	+61%	+71%
Feet and $legs_{Adj}$	+73%	+56%	+32%[2]
Calving $ease_{Adj}$	+77%	+88%	+37%[2]
Claw $health_{Adj}$	+74%	+91%	n/a
General $health_{Adj}$	+194%	+94%	n/a
Survival $1\text{-}2_{Adj}$	+71%	+136%	+11%[2]

[1] $\frac{\text{Correlation with GEBV} - \text{Correlation with PA}}{\text{Correlation with PA}}$ x 100.

[2]No significant difference between GEBV correlation and PA correlation ($p<0.05$).

[3]Interval from first to last service, in days.

For Jersey, GEBV correlations were significantly stronger for all traits except clinical mastitis, calving ease and survival 1-2. The correlations with adjusted phenotypes for the different traits were 42-194% higher for GEBV than for PA in RDC, 38-136% higher for GEBV than for PA in Holstein, and 11-78% higher for GEBV than for PA in Jersey. However, it should be noted that the large relative percentage change between PA and GEBV for correlations was in many cases from initially low levels.

Traits with low heritability, such as interval from first to last insemination, clinical mastitis, calving ease, claw health, and general health, gained relatively more from inclusion of genomic information than did highly heritable traits such as production.

One of the traits for which the correlations increased the most, cow adjusted phenotype for interval from first to last service (IFL_{Adj}) increased by over 64% for all three breeds when genomic information was included in the breeding values instead of PA. For Jersey, the correlation between breeding value and IFL_{Adj} increased by 78% when genomic information was included. Some selected results for quartiles of cows ranked by their GEBV or PA across different herds are presented in Figure 3 and 4, which provide clear illustrations of the increased accuracy of GEBV compared with PA and the implications for milk yield or days between the first and last service.

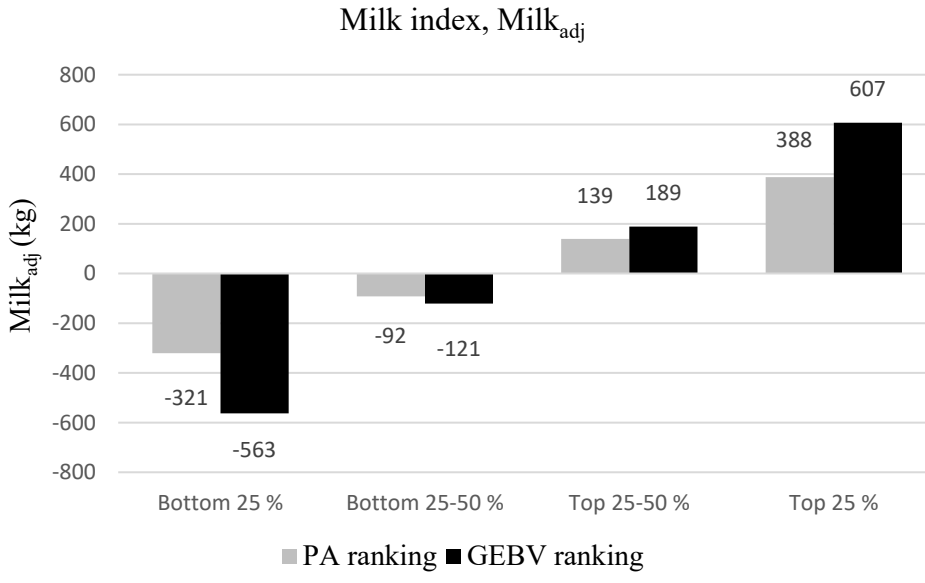


Figure 3. Results for 14,710 Red Dairy Cattle: Cow adjusted (Adj) phenotype for milk yield between quartiles ranked on heifer fertility index, parent average breeding values (PA) or genomically enhanced breeding values (GEBV).

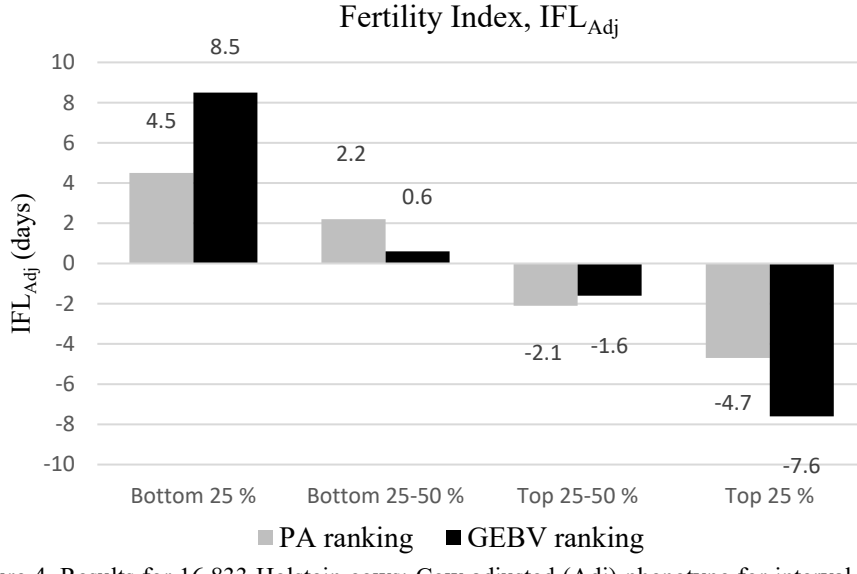


Figure 4. Results for 16,833 Holstein cows: Cow adjusted (Adj) phenotype for interval from first to last service (IFL$_{Adj}$) between quartiles ranked on heifer fertility index, parent average breeding values (PA) or genomically enhanced breeding values (GEBV).

The extracted virgin heifer GEBV and PA, estimated before on-farm information was recorded, reflected information available to farmers at the time of selection. The maximum age at which a breeding value for a heifer was taken was 14 months, to reflect the breeding values at first insemination for virgin heifers.

Based on the results, farmers in DFS can have confidence in using genomic technology on their herds for selection decisions.

4.2 Mating allocation in Nordic Red Dairy Cattle using genomic information (Paper II)

4.2.1 Materials and methods

In Paper II, different scenarios for mating allocations in Nordic Red Dairy Cattle using genomic information (Table 2) were compared. Linear programming was used to optimise different economic scores within each herd based on real data, considering genetic level, semen cost, the economic impact of recessive genetic defects and genetic relationships. In total, 9,841 genotyped RDC females born in Denmark, Finland or Sweden in 2019 were selected for mating allocations.

Two different pedigree relationship coefficients were used, one tracing the pedigree three generations back from the parents of the potential mating (a_{3Gen}) and one based on all available pedigree information (a_{AllGen}). Three different genomic relationship coefficients were used, one SNP-by-SNP genomic relationship and two based on shared genomic segments. The SNP-by-SNP genomic relationship coefficient (g_{SNP}) between animal *i* and *j* was calculated according to VanRaden (2008):

$$g_{SNP_{ij}} = \frac{\sum_m (x_{im} - 2p_m) \times (x_{jm} - 2p_m)}{2\sum_m p_m (1 - p_m)} \quad \text{(eq. 1)}$$

where x_{im} and x_{jm} are the genotype scores of animal *i* and animal *j* at marker *m*, coded: 0 = homozygote, 1 = heterozygote, and 2 = alternative homozygote, and p_m is the frequency of the alternative allele of marker *m* in the founder population.

Because the founder population frequency was unknown, the allele frequency of all 149,943 genotyped RDC animals available in Paper II was

used, as is common practice in genomic evaluation (Wang *et al*., 2014). The software SNP1101 was used to calculate g_{SNP} (Sargolzaei, 2014).

The two genomic relationship coefficients based on shared genomic segments ($g_{SEG_{ij}}$) were calculated following de Cara *et al*. (2013):

$$g_{SEG_{ij}} = \frac{\sum_k \sum_{ai=1}^{2} \sum_{bj=1}^{2} (L_{SEGk}(a_i b_j))}{2L_{AUTO}} \quad \text{(eq. 2)}$$

where L_{SEGk} is the length (in base pairs) of the *k*th shared segment measured over homolog *a* of animal *i* and homolog *b* of animal *j*, and L_{AUTO} is the total length of the autosomes covered by the SNP in base pairs (bp). Two different L_{SEGk} values were used: 1 centimorgan (cM) (g_{SEG1}) and 4 cM (g_{SEG4}), assuming 1 cM = 1,000,000 bp (Gautier *et al*., 2007).

The lengths of segments were chosen to represent short and long segments, as done in other studies (Zhang *et al.*, 2015; Forutan *et al*., 2018; Makanjuola *et al*., 2020). Phasing of genotypes was carried out in Beagle 4.1 with default settings (Browning & Browning, 2007), and segments of minimum desired length were extracted in RefineIBD (Browning & Browning, 2013).

For each potential mating between female *i* and bull *j*, an economic score was calculated as done by Bérodier *et al*. (2021) and Pryce *et al*. (2012):

$$Score_{ij} = \left(\frac{NTM_i + NTM_j}{2} + \lambda F_{ij}\right) \times \text{prob}(♀) - \sum_{r=1}^{n_r} p(aa)_r \times v_r - \text{semen cost} \quad \text{(eq. 3)}$$

where NTM_i and NTM_j are the value of Nordic Total Merit units in Euros (€) for female *i* and bull *j*, $\boldsymbol{\lambda}$ is the economic consequence of a 1% increase in inbreeding, F_{ij} is the pedigree or genomic based co-ancestry (Relationship/2), prob(♀) is the probability of producing a female conceptus, n_r is the number of recessive genetic defects considered (A1 in Appendix 1), $p(aa)_r$ is the probability of expression of genetic defect *r*, v_r is the economic cost associated with recessive genetic defect *r*, and semen cost is the average amount (€) spent on semen for a pregnancy.

The value of 1 index unit of NTM was set at €24.8, based on the value per NTM unit and year (€9.2) and the average production lifetime (2.7 years) (Fikse & Kargo, 2020). Only sexed semen was considered and the probability of producing a female conceptus was assumed to be 0.9, which is the minimum expected sexing rate for most sexing technologies (Burnell, 2019). The economic consequence of a 1% increase in inbreeding was set to

€24.8. The carrier frequencies considered for genetic defects are shown in Table 3 (for details, see Table A1 in Appendix 1). The cost of an early abortion (genetic defect at BTA12, PIRM/AH1, AH2) was assumed to be €80, based on the resulting longer calving interval (€30-40/month) and the cost of extra insemination(s) (€30). The cost of a later abortion or an early calf death was assumed to be €160 (genetic defect SMA, BH2, BTA23). Prices for sexed semen set by VikingGenetics in 2020 were used, where a semen dose for a bull with NTM of 30 or more cost €26, that for a bull with NTM of 25-30 cost €22.5, and that for a bull with NTM of 20-25 cost €19 (Jakob Lykke Voergaard, product manager, VikingRed, VikingGenetics, personal communication, January 11, 2021). The semen price was multiplied by 1.8, which is the average number of inseminations needed for a pregnancy in RDC (Sørensen *et al*., 2018).

Table 2. Description of the mating scenarios considered for Nordic Red Dairy Cattle in Paper II

Scenario	Economic score includes:		
	Nordic Total Merit, NTM	Relationship[1]	Genetic defect value
MaxNTM	Yes	No	No
3Gen	Yes	a_{3Gen}	Yes
AllGen	Yes	a_{AllGen}	Yes
GSNP	Yes	g_{SNP}	Yes
GSEG1	Yes	g_{SEG1}	Yes
GSEG4	Yes	g_{SEG4}	Yes
Random	All possible combinations of 9,841 females and 50 bulls		

[1] a_{3Gen} = pedigree relationships using three generations of ancestors, a_{AllGen} = 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.

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 and others, 2020). The mating R script was provided by Bérodier *et al*. (2021). The R script set constraints that were considered in the linear programming optimisation. The constraints used in Paper II were one mating per female and a threshold percentage for the maximum number of females per bull and herd, for which two different levels were evaluated (5% and 10%), following Bérodier *et al.* (2021).

Table 3. Descriptive statistics on the Nordic Red Dairy Cattle (RDC) females and bulls selected for mating allocations in Paper II

Trait	Females	Dataset BullVG[1]
Number of animals	9841	50
Average Nordic Total Merit (NTM)	10.7	28.4
Carriers of defect BTA12 (%)	14.7	12.0
Carriers of defect BTA23 (%)	1.1	0.0
Carriers of defect BH2 (%)	0.3	0.0
Carriers of defect PIRM/AH1 (%)	1.6	0.0
Carriers of defect AH2 (%)	1.2	0.0
Carriers of defect SMA (%)	0.30	0.0

[1]Fifty genotyped RDC bulls from the Nordic breeding cooperative VikingGenetics.

4.2.2 Results and comments

The mean value of the relationship coefficients between all possible combinations of females and males ranged from 0.009 to 0.188, and the standard deviation ranged from 0.042 to 0.047 (Table 4). The correlations between the genetic relationship coefficients were all 0.83 or higher. The strongest correlation was between a_{AllGen} and a_{3Gen} (r=0.99), and the second strongest was between g_{SEG1} and g_{SEG4} (r=0.98) (Table 5). There is a long tradition of pedigree recording in the Nordic countries, and the strong correlation found between pedigree and genomic relationships confirms the good documentation of dairy pedigrees in DFS.

Table 4. Descriptive statistics on relationships (mean, standard deviation (SD), minimum value (Min) and maximum value (Max)) between all possible combinations of 9,841 females and 50 bulls of Nordic Red Dairy Cattle analysed in Paper II

Relationship[1]	Mean	SD	Min	Max
a_{3Gen}	0.028	0.042	0	0.648
a_{AllGen}	0.066	0.042	0.003	0.667
g_{SNP}	0.009	0.047	-0.095	0.673
g_{SEG1}	0.188	0.046	0.038	0.789
g_{SEG4}	0.115	0.045	0.005	0.727

[1]a_{3Gen} = pedigree relationships using three generations of ancestors, a_{AllGen} = 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 5. Correlations between the different relationship coefficients for all possible combinations of 9,841 females and 50 bulls of Nordic Red Dairy Cattle analysed in Paper II

Relationship[1]	a_{3Gen}	a_{AllGen}	g_{SNP}	g_{SEG1}	g_{SEG4}
a_{3Gen}	1	0.99	0.88	0.83	0.87
a_{AllGen}		1	0.88	0.85	0.88
g_{SNP}			1	0.9	0.93
g_{SEG1}				1	0.98

[1] a_{3Gen} = pedigree relationships using three generations of ancestors, a_{AllGen} = 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.

It was possible to maximise economic score with limited impact on the average NTM level (Table 6 and 7). Including the cost of the known recessive genetic defect (at BTA12) when optimising mating strategies eliminated the risk of expression of that genetic defect, regardless of the genetic relationship used. In scenario 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 expression of genetic defects.

Table 6. Comparison of outcome of planned matings of 9,841 females for seven mating scenarios in Nordic Red Dairy Cattle using various comparison criteria. Fifty marketed bulls from VikingGenetics were available for matings (BullVG dataset). Maximum percent of females per bull and herd set to 5%

Comparison criterion[1]	Random	Max NTM	3Gen	AllGen	GSNP	GSEG1	GSEG4
Average Nordic Total Merit (NTM)	19.5	20.8	20.8	20.8	20.8	20.7	20.8
Average a_{3Gen} between parents	0.028	0.033	0.007	0.009	0.014	0.014	0.014
Average a_{AllGen} between parents	0.066	0.070	0.046	0.043	0.050	0.050	0.050
Average g_{SNP} between parents	0.009	0.014	-0.012	-0.016	-0.038	-0.034	-0.033
Average g_{SEG1} between parents	0.188	0.191	0.167	0.163	0.148	0.143	0.146
Average g_{SEG4} between parents	0.115	0.119	0.094	0.091	0.078	0.075	0.074
Expression of genetic defect (%)	0.4	0.4	0	0	0	0	0
Semen cost (€)	42.0	43.6	43.6	43.6	43.5	43.5	43.5
Predicted BTA12 carrier frequency (%)	13.4	12.9	13.2	12.9	12.9	12.0	12.9

[1]Average NTM level, five different genetic relationships, the probability of expression of genetic defect (at BTA12), the average cost of semen for a pregnancy, and predicted genetic defect at BTA12 carrier frequency in the next generation.

Table 7. Comparison of outcome of planned matings of 9,841 females for seven mating scenarios in Nordic Red Dairy Cattle using various comparison criteria. Fifty marketed bulls from VikingGenetics were available for matings (BullVG dataset). Maximum percent of females per bull and herd set to 10%

Comparison criterion[1]	Random	Max NTM	3Gen	AllGen	GSNP	GSEG1	GSEG4
Average Nordic Total Merit (NTM)	19.5	22.1	22.1	22.1	22.0	22.0	22.0
Average a_{3Gen} between parents	0.028	0.037	0.009	0.010	0.016	0.016	0.015
Average a_{AllGen} between parents	0.066	0.074	0.046	0.044	0.051	0.051	0.051
Average g_{SNP} between parents	0.009	0.018	-0.010	-0.013	-0.032	-0.027	-0.027
Average g_{SEG1} between parents	0.188	0.195	0.166	0.163	0.150	0.145	0.148
Average g_{SEG4} between parents	0.115	0.037	0.009	0.010	0.016	0.016	0.015
Expression of genetic defect (%)	0.4	0.04	0	0	0	0	0
Semen cost (€)	42.0	45.8	46.2	46.3	45.8	45.7	45.9
Predicted BTA12 carrier frequency (%)	13.4	7.9	7.6	7.6	8.2	7.8	7.8

[1]Average NTM level, five different genetic relationships, the probability of expression of genetic defect (at BTA12), the average semen for a pregnancy, and predicted genetic defect at BTA12 carrier frequency in the next generation.

Including a genomic relationship in the economic score kept the other genomic relationship averages at a low level. For example, with the constraint 5% females per bull and herd, including g_{SNP} in the objective function (scenario GSNP) resulted in g_{SEG1} of 0.148, compared with 0.143 in scenario GSEG1 (Table 6). Using the pedigree relationships also reduced the genomic relationships compared with scenarios Random and MaxNTM, but not as much as using genomic relationships in the objective function. Considering the example with the constraint 5% females per bull and herd, and including g_{SNP} in the objective function (scenario GSNP), the pedigree relationship scenarios resulted in g_{SEG1} of 0.167 for scenario 3Gen and 0.163 for AllGen. There were only minor differences between the scenarios with genomic relationships in their ability to reduce pedigree relationships. Including pedigree relationships in the economic scores consistently reduced pedigree relationships more than genomic relationships. For example, all scenarios optimising genomic relationships resulted in a_{AllGen} of 0.050, AllGen resulted in a_{AllGen} of 0.043, and 3Gen resulted in a_{AllGen} of 0.046 (Table 6).

The results obtained in Paper II also demonstrated the efficiency of linear programming as a method for optimising mating plans, as it maximised the economic score for all herds studied within seconds.

4.3 Mating allocations in Holstein combining genomic information and linear programming optimisation at herd level (Paper III)

4.3.1 Material and methods

Paper III explored mating allocation in Holstein using genomic information for 24,333 Holstein females born in DFS. Linear programming was used to optimise economic scores within each herd in a similar way as in Paper II, considering genetic level, genetic relationship, semen cost and the economic impact of genetic defects and, for Holstein, also polledness and beta-casein. Two datasets of bulls were used: the top 50 genotyped bulls on the Nordic Total Merit scale, and the top 25 polled genotyped bulls on the Nordic Total Merit scale.

The five different genetic relationships were calculated in the same way as in Paper II. The carrier frequencies considered for genetic defects are shown in Table 8 (for details, see Table A1 in Appendix 1).

Table 8. Descriptive statistics on the Holstein females and bulls selected for mating allocations in Paper III

	Females		
	289	Dataset	Dataset
Trait	herds	Bull50	Bull25Polled
Number of animals	24,333	50	25
Average Nordic Total Merit (NTM)	12.10	33.93	27.17
Carriers of defect HH1 (%)	3.45	2.00	16.00
Carriers of defect HH3 (%)	3.62	4.00	0.00
Carriers of defect HH4 (%)	1.31	0.00	0.00
Carriers of defect HH6 (%)	0.30	0.00	0.00
Carriers of defect HH7 (%)	0.29	0.00	0.00
Carriers of defect Blad (%)	0.27	0.00	0.00
Carriers of defect RP1 (%)	0.63	0.00	0.00
Heterozygous polled (Pp) (%)	3.74	14.00	84.00
Homozygous polled (PP) (%)	0.10	0.00	16.00
Heterozygous Beta Casein (A1A2) (%)	37.11	30.00	44.00
Homozygous Beta Casein (A2A2) (%)	57.12	66.00	48.00

For each potential mating between female *i* and bull *j*, an economic score was calculated as:

$$Score_{ij} = \left(\frac{NTM_i + NTM_j}{2} + \lambda F_{ij} + p(\text{BetaC}) \times v_{BetaC}\right) \times \text{prob}(♀) - \sum_{r=1}^{n_r} p(aa)_r \times v_r + p(P) \times v_P - \text{semen cost} \quad (4)$$

This is similar to the equation used in Paper II (eq. 3) except for p(BetaC), which is the probability of a homozygous offspring for beta-casein (A2A2), v_P is the value of a homozygous offspring for beta-casein (A2A2), p(*P*) is the probability of a polled offspring and v_P is the value of a polled offspring.

An index unit of NTM is reported to be worth €25.4 over the lifetime of a Holstein female in DFS (Fikse & Kargo, 2020). In Paper III, the economic consequence of a 1% increase in inbreeding was set to €25.4. The cost of an early abortion (caused by defects: HH1, HH3, HH4, HH6, HH7) was

assumed to be €80, based on the resulting longer calving interval (€30-40/month) and the cost of extra insemination(s) (€30) (Oskarsson & Engelbrekts, 2015; Sørensen *et al.*, 2018). Bulls carrying genetic defects BLAD and RP1 are not allowed in the breeding programme at VikingGenetics, so no cost was estimated for these defects. Different economic values (€0, €10, €50 and €100) for polledness and beta-casein (A2A2) were tested (Table 9).

Table 9. Description of the mating scenarios considered for Holstein cattle in Paper III

	Economic score includes:				
Scenario	NTM	Relation-ship	Genetic defect value	Polled value (€)	Beta-casein value (€)
MaxNTM	Yes	No	No	0	0
3Gen	Yes	a_{3Gen}	Yes	0	0
AllGen	Yes	a_{AllGen}	Yes	0	0
GSNP	Yes	g_{SNP}	Yes	0	0
GSEG1	Yes	g_{SEG1}	Yes	0	0
GSEG4	Yes	g_{SEG4}	Yes	0	0
GSNPPolled10	Yes	g_{SNP}	Yes	10	0
GSNPPolled50	Yes	g_{SNP}	Yes	50	0
GSNPPolled100	Yes	g_{SNP}	Yes	100	0
GSNPBetaC10	Yes	g_{SNP}	Yes	0	10
GSNPBetaC50	Yes	g_{SNP}	Yes	0	50
GSNPBetaC100	Yes	g_{SNP}	Yes	0	100
GSNPPolledBetaC10	Yes	g_{SNP}	Yes	10	10
GSNPPolledBetaC50	Yes	g_{SNP}	Yes	50	50
GSNPPolledBetaC100	Yes	g_{SNP}	Yes	100	100
Random	All possible combinations of females and bulls.				

Prices for sexed semen set by VikingGenetics in 2021 were used. Semen price depends on the bull's NTM and polledness status. In 2021, a dose of semen from a horned bull with NTM >35, 33-34, 30-32 and <30 cost €26, €23, €20 and €17, respectively. Semen of polled bulls (homozygous or heterozygous for the polled allele) costs €3 more than semen of horned bulls with the same NTM (Hanna Driscoll, Product manager Holstein, VikingGenetics, personal communication January 19, 2022).

SAS version 9.4 (SAS Institute Inc., Cary, NC) and R version 3.6.3 (R Core Team, 2020) were used for statistical analysis. A chi-square test was conducted in SAS to test for associations between polledness genotype and HH1, HH3 or beta-casein genotype.

4.3.2 Results and comments

The results presented are for matings between the 24,333 females selected for matings and the dataset Bull50, unless otherwise specified.

For all possible combinations of females and males, the mean value of the relationship coefficient ranged from 0.010 to 0.269, and the standard deviation ranged from 0.031 to 0.042 (Table 10). For all correlations between different genetic relationship coefficients, the value of the correlation coefficient was ≥0.69. The strongest correlation was between g_{SEG1} and g_{SEG4} (r=0.97). Further, all correlations between a_{AllGen} and genomic relationships were of similar strength (0.75-0.76), while those between a_{3gen} and the genomic relationships showed a wider range of values (0.69-0.75) (Table 11). Hence, the correlations between genomic and pedigree relationship were in general weaker that those obtained for RDC in Paper II. Pedigree depth is similar in both breeds, so the difference is most likely linked to some other factor/s. One possibility is that pedigree correctness is greater in RDC than in Holstein, due to the less common exchange of RDC bulls and their pedigrees worldwide, as most RDC animals are found within the Nordic countries.

Table 10. Descriptive statistics on relationships (mean, standard deviation (SD), minimum value (Min) and maximum value (Max)) between all possible combinations of 24,333 females and 50 bulls of the Holstein breed analysed in Paper III

Relationship coefficient[1]	Mean	SD	Min	Max
a_{3Gen}	0.015	0.031	0	0.545
a_{AllGen}	0.132	0.031	0.035	0.647
g_{SNP}	0.010	0.040	-0.106	0.576
g_{SEG1}	0.269	0.042	0.089	0.853
g_{SEG4}	0.181	0.041	0.039	0.763

[1] a_{3Gen} = pedigree relationships using three generations of ancestors, a_{AllGen} = 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 11. Correlations between the different relationship coefficients for all possible combinations of 24,333 females and 50 bulls of the Holstein breed analysed in Paper III

	Relationship				
Relationship[1]	a_{3Gen}	a_{AllGen}	g_{SNP}	g_{SEG1}	g_{SEG4}
a_{3Gen}	1	0.95	0.75	0.69	0.70
a_{AllGen}		1	0.76	0.75	0.76
g_{SNP}			1	0.88	0.87
g_{SEG1}				1	0.97

[1]a_{3Gen} = pedigree relationships using three generations of ancestors, a_{AllGen} = 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.

As found for RDC in Paper II, it was possible to reduce genetic relationships in Holstein cattle and eliminate expression of genetic defects with minimal effect on the genetic level. The results also showed that it was possible to increase the percentage of polled offspring substantially in one generation when competitive bulls were available, without any significant negative impact on other comparison criteria (Table 12 and 13). It was also possible to increase the number of homozygous beta-casein (A2A2) offspring without any negative impact on other comparison criteria (Table 14).

Bulls carrying the polled allele were less likely to be homozygous for beta-casein (A2A2) and more likely to be carriers of the genetic defect HH1. Hence, adding economic value to a monogenic trait in the economic score used for mating allocations sometimes negatively impacted another monogenic trait. Among the 24,333 mated Holstein females, polled females (Pp and PP) were less likely to be homozygous for beta-casein (A2A2) (or A2A2 females were less likely to carry the polled allele). For example, 58% of the horned females, but only 44% of the heterozygous polled (Pp) females, were homozygous for beta-casein (A2A2). The chi-square test showed a strongly significant negative association between polled and beta-casein genotype ($p<0.0001$) in the data. Adding economic value to both the polledness trait and beta-casein (A2A2) in the economic score used for mating allocations increased the expected number of polled offspring and offspring homozygous for beta-casein (A2A2) compared with GSNPPolled€0 (Table 15).

Table 12. Results of four mating scenarios investigating extra economic value for the polledness trait, based on 24,333 Holstein females. Available bulls were 50 Holstein bulls marketed by VikingGenetics (Bull50). Maximum 5% females per bull and herd

Comparison criterion	GSNP			
	Polled €0	Polled €10	Polled €50	Polled €100
Average Nordic Total Merit (NTM)	24.2	24.2	24.2	24.1
Average a_{AllGen} between parents	0.121	0.121	0.121	0.122
Average g_{SNP} between parents	-0.040	-0.040	-0.040	-0.039
At-risk matings (%)	0.00	0.00	0.00	0.00
Percentage of polled offspring	9.7	10.3	13.2	17.0
Percentage of homozygous A2A2 offspring	61.0	60.9	60.8	60.8

Table 13. Results of four mating scenarios investigating extra economic value for the polledness trait, based on 24,333 Holstein females. Available bulls were 50 Holstein bulls marketed by VikingGenetics (Bull50). Maximum 10% females per bull and herd

Comparison criterion	GSNP			
	Polled €0	Polled €10	Polled €50	Polled €100
Average Nordic Total Merit (NTM)	25.2	25.2	25.2	25.0
Average a_{AllGen} between parents	0.121	0.121	0.121	0.122
Average g_{SNP} between parents	-0.034	-0.034	-0.034	-0.034
At-risk matings (%)	0.00	0.00	0.00	0.00
Percentage of polled offspring	11.7	12.6	16.4	22.5
Percentage of homozygous A2A2 offspring	66.4	65.9	64.0	62.2

Table 14. Results of four mating scenarios investigating extra economic value for beta-casein (A2A2), based on 24,333 Holstein females. Available bulls were 50 Holstein bulls marketed by VikingGenetics (Bull50). Maximum 10% females per bull and herd

Comparison criterion	GSNP			
	BetaC €0	BetaC €10	BetaC €50	BetaC €100
Average Nordic Total Merit (NTM)	25.2	25.2	25.2	25.1
Average a_{AllGen} between parents	0.121	0.121	0.121	0.121
Average g_{SNP} between parents	-0.034	-0.034	-0.034	-0.033
At-risk matings (%)	0.00	0.00	0.00	0.00
Percentage of polled offspring	11.7	11.2	9.3	8.0
Percentage of homozygous A2A2 offspring	66.4	68.2	72.8	75.0

Table 15. Results of four mating scenarios investigating extra economic value for the polledness trait and beta-casein (A2A2), based on 24,333 Holstein females. Available bulls were 50 Holstein bulls marketed by VikingGenetics (Bull50). Maximum 10% females per bull and herd

Comparison criterion	GSNP			
	Polled BetaC €0	Polled BetaC €10	Polled BetaC €50	Polled BetaC €100
Average Nordic Total Merit (NTM)	25.2	25.2	25.2	25.1
Average a_{AllGen} between parents	0.121	0.121	0.121	0.122
Average g_{SNP} between parents	-0.034	-0.034	-0.033	-0.032
At-risk matings (%)	0.00	0.00	0.00	0.00
Percentage of polled offspring	11.7	12.2	14.1	18.5
Percentage of homozygous A2A2 offspring	66.4	67.7	70.4	71.8

When 25 polled bulls (21 Pp bulls, four PP bulls) were available for mating allocations, it was possible to further increase the expected percentage of polled offspring (Table 16). For example, when using BullPolled25 and a constraint of 5% females per bull and herd, the expected percentage of polled offspring was 60.1% in GSNPPolled100€, compared with 17.0% using Bull50. Considering the same example, the average NTM level was 20.2 using BullPolled25, compared with 24.1 using Bull50 (Table 12). The average genetic relationships using BullPolled25 were slightly higher than

using Bull50 with the same constraints and economic scores. The expected percentage of offspring homozygous for beta-casein (A2A2) was also lower. Hence, using only semen from polled bulls, which might be necessary if dehorning is banned, would be relatively costly.

Table 16. Results of four mating scenarios investigating extra economic value for the polledness trait, based on 24,333 Holstein females. Available bulls were 25 polled Holstein bulls marketed by VikingGenetics (BullPolled25). Maximum 5% females per bull and herd

Comparison criterion	GSNP			
	Polled €0	Polled €10	Polled €50	Polled €100
Average Nordic Total Merit (NTM)	20.2	20.2	20.2	20.2
Average a_{AllGen} between parents	0.121	0.121	0.121	0.121
Average g_{SNP} between parents	-0.034	-0.034	-0.034	-0.034
At-risk matings (%)	0.00	0.00	0.00	0.00
Percentage of polled offspring	59.5	59.7	60.0	60.1
Percentage of homozygous A2A2 offspring	54.3	54.3	54.3	54.3

4.4 Simulation of long-term impact of dairy cattle mating programmes using genomic information (Paper IV)

4.4.1 Material and methods

Paper IV examined the long-term impact of genomic mating allocations with stochastic simulation, where the matings followed a similar approach as used in Papers II and III. The economic scores included genetic level, a favourable monogenic trait (polledness), a recessive genetic defect and parent relationships. One unknown recessive genetic defect was also monitored. The AlphaSimR package version 1.3.4 (Gaynor *et al*., 2021) in R version 4.1.3 (R Core Team, 2020) was used to simulate a closed population under selection with discrete generations. The genomes of the founder population were created with the MaCS coalescent simulator, which was run within the AlphaSimR package, using the "CATTLE" population history (MacLeod *et al*., 2013). The founder population was generated once and was the same for all scenarios and replicates. A total of 29 chromosomes with 6,000

segregating sites per chromosome were simulated. The breeding goal trait was constructed by adding an additive trait in AlphaSimR, which was assigned 4,000 QTLs per chromosome. The breeding goal had a mean of 0 and a genetic standard deviation of 10. In addition, a SNP chip with 1,600 SNP per chromosome was constructed in AlphasimR.

Breeding animals were selected based on a breeding goal with accuracy of approximately 0.7. A population of 8,000 females across 40 herds was simulated, with 200 females in every generation selected for multiple ovulation and embryo transfer (MOET). Each donor produced 20 offspring (50:50 sex ratio). Of the 2,000 males produced via MOET, the 100 best were selected as sires of the next generation. Following MOET, the donors and the rest of the females were inseminated with sexed semen, and all offspring were assumed to be females. Thus, approximately 10,000 females were available for the next generation (sometimes less if one of the lethal genetic defects was expressed), with the top 8,000 females selected for breeding.

The simulation spanned 30 generations, and each scenario was replicated 30 times. For the first 20 generations, random matings (with the randCross function in AlphasimR) were performed among all selected animals. In the last 10 generations, matings with sexed semen were assigned based on an economic score that defined the scenario (see Table 17) and matings of the donors were still assigned at random.

In generation 20, three SNP markers were selected to represent three monogenic traits: one known (assumed) lethal recessive genetic defect with an allele frequency of approximately 0.05 (range 0.04-0.06); one unknown lethal recessive genetic defect with an allele frequency of approximately 0.09 (range 0.08-0.10), which served as a reference for risk management; and one dominant trait with an allele frequency of 0.12 (range 0.11-0.13), which represented polledness. It was assumed that conceptus/offspring homozygous for the recessive genetic defects died and thus they were excluded from breeding.

The scenarios considered were similar to those in Papers II and III (Table 17). Linear programming was used as in Papers II and III and the same constraints were applied, *i.e.* one mating per female and a threshold percentage for the maximum number of females per bull and herd of 5% or 10%.

A difference from Papers II and III was that when calculating g_{SNP}, the founder population's allele frequency was used instead of the allele frequency of all genotyped animals.

Table 17. Description of the mating scenarios considered in the simulation study in Paper IV

	Economic score includes:			
Mating scenario	TMI[1]	Relation-ship	Genetic defect value	Polled value (€)
MaxTMI	Yes	No	No	0
Ped	Yes	a_{Ped}	Yes	0
GSNP	Yes	g_{SNP}	Yes	0
PedPolled10	Yes	a_{Ped}	Yes	10
PedPolled50	Yes	a_{Ped}	Yes	50
PedPolled100	Yes	a_{Ped}	Yes	100
GSNPPolled10	Yes	g_{SNP}	Yes	10
GSNPPolled50	Yes	g_{SNP}	Yes	50
GSNPPolled100	Yes	g_{SNP}	Yes	100
Random	Matings were randomly assigned with equal number of offspring (females 1 offspring and bulls 80 offspring)			

[1]Total merit index.

In the last 10 generations of the simulation, the different scenarios were compared by (i) genetic gain in total merit index (TMI) per generation; (ii) rate of pedigree inbreeding; (iii) rate of genomic inbreeding per generation from the diagonal of the VanRaden relationship matrix (excess homozygosity relative to the base population); (iv) change in carrier frequency per generation of the known and unknown recessive alleles; (v) change in number/frequency of polled offspring per generation; (vi) number of affected conceptuses in the last 10 generations of the known and unknown genetic defects; (vii) number of bulls used per generation; and (viii) number of bulls used per generation up to the maximum number of doses allowed for the threshold level (5% and 10%) of females per bull and herd.

4.4.2 Results and comments

Inclusion of a genomic relationship in the economic score significantly reduced the rate of increase for both pedigree and genomic inbreeding compared with only maximising genetic level (MaxTMI), with a few

exceptions (Table 18-Table 21). Incorporating a pedigree relationship into the economic score slowed the rate of increase for both pedigree and genomic inbreeding compared with MaxTMI, although there were more exceptions regarding the level of significance. The scenario Random had lower rates of inbreeding and genetic gain than most other scenarios. The rates of pedigree inbreeding were similar when either genomic or pedigree relationship was included in the economic score, but using the genomic relationship led to a lower rate of genomic inbreeding (Table 21). The frequency of polled offspring increased on average per generation when the value of polledness was €50 or higher, as also found in Paper III, while it remained constant when the value was lower. The largest change in frequency of polled offspring per generation (0.037) occurred in one of the scenarios where the value of polled was €100 (GSNPPolled100) using 10% females per bull and herd (Table 21). The frequency of polled offspring increased faster over generations with the 10% females per bull and herd constraint compared with the 5% constraint. For example, in GSNPPolled100, the increase in frequency of polled offspring was 0.037 when allowing up to 10% females per bull and herd, compared with 0.028 when using the limitation of 5% females per bull and herd (Table 21). Hence most of the results in Paper IV were in line with those in Paper III.

There was great variation in the different monogenic traits monitored across replicates and scenarios. In general, using fewer bulls resulted in greater variation across replicates regarding the monogenetic traits. The scenarios using 10% females per bull and herd showed the highest variation, and the scenario Random the lowest (Figure 5). For example, the variation was higher for 10% females per bull and herd (Figure 5), even though the mean frequency was similar in both sets (Table 19-21). The scenario with a limit of 5% females per bull was more similar to Random in terms of variation than the scenario with a limit of 10% females per bull.

Table 18. Results of six mating scenarios in the last 10 generations (gen) of the simulations in Paper IV. Maximum 5% females per bull and herd

Comparison criterion	Random	MaxTMI	Ped	Ped		
				Polled €10	Polled €50	Polled €100
Genetic gain (per gen)	7.59*	7.93 (7.61-8.27)	7.91	7.87	7.89	7.87
Pedigree inbreeding rate (per gen)	0.0025*	0.0041 (0.0032-0.0052)	0.0032*	0.0033	0.0031*	0.0033
Genomic inbreeding rate (per gen)	0.0073*	0.0092 (0.0081-0.010)	0.0085	0.0085	0.0084	0.0084
Carrier frequency change of the known defect (per gen)	-0.005	-0.002 (-0.0092-0.0093)	-0.004	-0.004	-0.004	-0.003
Carrier frequency change of the unknown defect (per gen)	-0.007	-0.009 (-0.021- -0.001)	-0.008	-0.007	-0.008	-0.006
Polled offspring frequency change (per gen)	-0.005	-0.007 (-0.02-0.014)	-0.003	0.000	0.017*	0.032*
Average frequency of conceptuses affected by the known defect	0.0013	0.002 (0.0001-0.006)	0.0*	0.0*	0.0*	0.0*
Average frequency of conceptuses affected by the unknown defect	0.0049	0.0040 (0.001-0.0088)	0.0033	0.0037	0.0047	0.0048
Average (per gen) number of bulls used	NA	21	24	24	24	25
Average (per gen) number of bulls used to a maximum	NA	20	17	17	17	17

Table 19. Results of five mating scenarios in the last 10 generations (gen) of the simulations in Paper IV. Maximum 5% females per bull and herd

Comparison criterion	MaxTMI	GSNP	GSNP		
			Polled €10	Polled €50	Polled €100
Genetic gain (per gen)	7.93 (7.61-8.27)	7.86	7.82	7.87	7.83
Pedigree inbreeding rate (per gen)	0.0041 (0.0032-0.0052)	0.0031*	0.0032*	0.0031*	0.0031*
Genomic inbreeding rate (per gen)	0.0092 (0.0081-0.010)	0.0080*	0.0080*	0.0080*	0.0080*
Carrier frequency change of the known defect (per gen)	-0.002 (-0.0092-0.0093)	-0.003	-0.004	-0.004	-0.003
Carrier frequency change of the unknown defect (per gen)	-0.009 (-0.021- -0.001)	-0.006	-0.005	-0.003	-0.007
Polled offspring frequency change (per gen)	-0.007 (-0.02-0.014)	-0.004	0.00	0.015*	0.028*
Average frequency of conceptuses affected by the known defect	0.002 (0.0001-0.006)	0.00015	0.00015	0.00010	0.00017
Average frequency of conceptuses affected by the unknown defect	0.0040 (0.001-0.0088)	0.0032	0.0048	0.0040	0.0032
Average (per gen) number of bulls used	21	47	47	47	45
Average (per gen) number of bulls used to a maximum	20	13	13	13	13

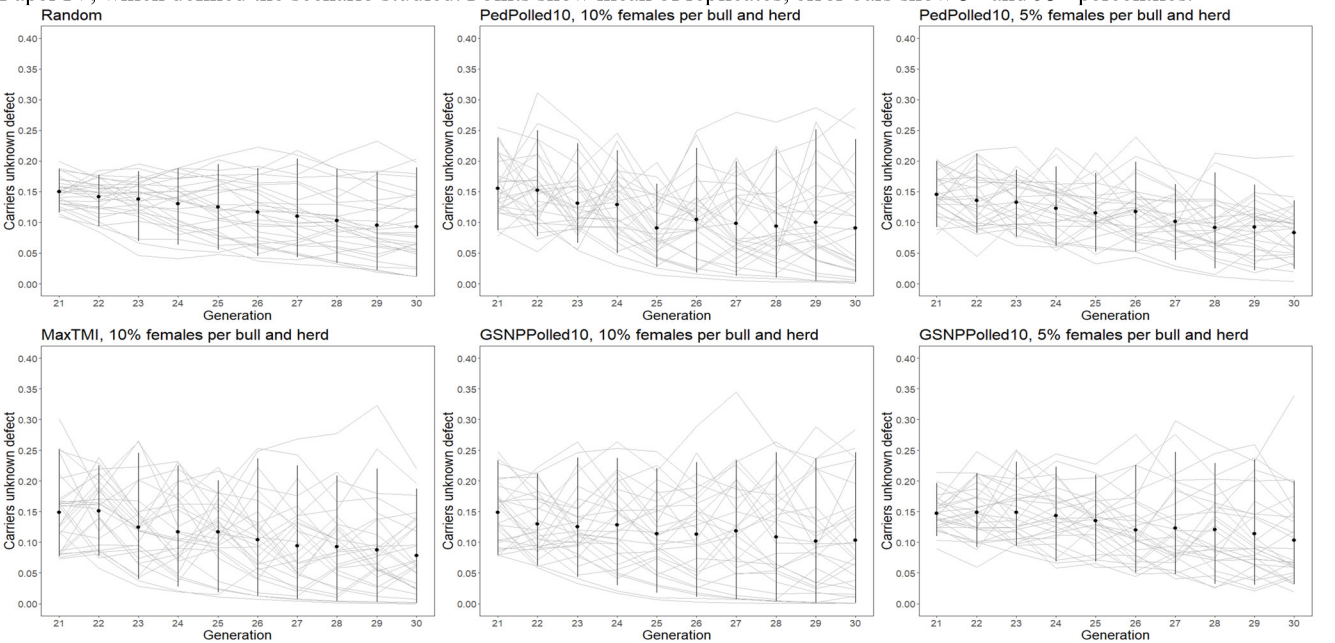
Table 20. Results of six mating scenarios in the last 10 generations (gen) of the simulations in Paper IV. Maximum 10% females per bull and herd

Comparison criterion	Random	MaxTMI	Ped	Ped		
				Polled €10	Polled €50	Polled €100
Genetic gain (per gen)	7.59*	8.00 (7.67-8.37)	7.97	7.93	7.97	7.98
Pedigree inbreeding rate (per gen)	0.0025*	0.0053 (0.0041-0.0068)	0.0039*	0.0039*	0.0039*	0.0040*
Genomic inbreeding rate (per gen)	0.0073*	0.011 (0.0091-0.0122)	0.0093	0.0091*	0.0092	0.0093
Carrier frequency change of the known defect (per gen)	-0.005	-0.005 (-0.015-0.009)	-0.007	-0.004	-0.005	-0.005
Carrier frequency change of the unknown defect (per gen)	-0.007	-0.008 (-0.021-0.011)	-0.0068	-0.0073	-0.0060	-0.007
Polled offspring frequency change (per gen)	-0.005	-0.0037 (-0.022-0.0025)	-0.0071	0.002	0.020	0.034*
Average frequency of conceptuses affected by the known defect	0.0013	0.0012 (0.00013-0.0041)	0	0.000001*	0.000001*	0.00001*
Average frequency of conceptuses affected by the unknown defect	0.0049	0.0044 (0.0004-0.0089)	0	0.004	0.005	0.005
Average (per gen) number of bulls used	NA	10	14	14	14	15
Average (per gen) number of bulls used to a maximum	NA	10	8	8	8	8

Table 21. Results of five mating scenarios in the last 10 generations (gen) of the simulations in Paper IV. Maximum 10% females per bull and herd

Comparison criterion	MaxTMI	GSNP	GSNP		
			Polled €10	Polled €50	Polled €100
Genetic gain (per gen)	8.00 (7.67-8.37)	7.98	8.02	7.94	7.88
Pedigree inbreeding rate (per gen)	0.0053 (0.0041-0.0068)	0.0039*	0.0037*	0.0038*	0.0038*
Genomic inbreeding rate (per gen)	0.011 (0.0091-0.0122)	0.0086*	0.0087*	0.0086*	0.0086*
Carrier frequency change of the known defect (per gen)	-0.005 (-0.015-0.009)	-0.0026	-0.0026	-0.0022	-0.0033
Carrier frequency change of the unknown defect (per gen)	-0.008 (-0.021-0.011)	-0.0073	-0.0043	-0.0050	-0.0071
Polled offspring frequency change (per gen)	-0.0037 (-0.022-0.0025)	-0.004	-0.0019	0.020	0.037*
Average frequency of conceptuses affected by the known defect	0.0012 (0.00013-0.0041)	0.00002	0.0003	0.0002	0.0002
Average frequency of conceptuses affected by the unknown defect	0.0044 (0.0004-0.0089)	0.004	0.004	0.004	0.005
Average (per gen) number of bulls used	10	30	30	30	30
Average (per gen) number of bulls used to a maximum	10	6	6	6	6

Figure 5. Frequency of carriers of the (assumed) unknown recessive genetic defect in the last 10 generations of the simulations in Paper IV, which defined the scenario studied. Points show mean of replicates, error bars show 5th and 95th percentiles.



The carrier frequency of both known and unknown recessive genetic defects decreased on average over generations in all scenarios analysed in Paper IV. The number of conceptuses affected by the known recessive genetic defect decreased when the cost of the genetic defect was included in the economic scores. Using pedigree relationship with a cost of the recessive genetic defect avoided almost all affected conceptuses in most scenarios and replicates. The risk of mating two carriers was slightly higher when genomic relationship was used instead of pedigree relationship with the cost of the recessive genetic defect (Table 19-Table 21). Mating of two carriers of the same genetic defect did not occur in Papers II and III.

5. General discussion

Genomic selection has brought about a revolution in dairy cattle breeding in the past 15 years. Estimating the overall value of genotyping of females at herd level is challenging, because it is influenced by multiple factors that may vary depending on country, region and herd. Several studies have tried to quantify the net benefit of genotyping candidate females for replacement using different methods and assumptions (Calus *et al*., 2015; Weigel *et al*., 2015; Newton & Berry, 2020). For DFS, Hjortø *et al*. (2015) showed that genomic testing was profitable in most cases with a genotyping price of around €30 when combined with sexed semen and beef semen. However, even though the price is lower now (around €20-25), far from all female calves are genotyped in DFS. In Paper I, the aim was to increase confidence in genomic breeding values by showing that they can predict future cow phenotypes better than parent average breeding values. An advantage of using phenotypes for validation of breeding values is that it makes it easy for farmers to understand how their animals' breeding values work in practice, when validation is against their own farm records. The correlations with adjusted phenotypes were 38-136% higher for GEBV than for PA in Red Dairy Cattle, 42-194% higher for GEBV than for PA in Holstein and 11-78% higher for GEBV than for PA in Jersey. Hence, the conclusion from the work in Paper I was that farmers in DFS can have confidence in using genomic technology on their herds. A growing reference population enhances the accuracy of GEBV, thus most likely increasing the difference in accuracy between GEBV and PA since the study. Furthermore, the research field of genomic evaluation is constantly working to improve the accuracy of GEBV (Misztal *et al*., 2020).

5.1 Economic scores and linear programming

Papers II-IV focused on how best to use genomic information when planning matings. A decision was made to use economic scoring systems to rank each potential mating, based on findings in previous studies (Carthy *et al*., 2019; Bérodier *et al*., 2021). The economic score is flexible and can be adjusted to match different economic conditions. It is also possible to include new information as it becomes available, *e.g*. on new monogenetic traits. For example, kappa-casein has been included in the SNP array since the analyses in Paper III. Specific breeding values could also be included in the economic scores, following *e.g*. the approach by Carthy *et al*. (2019), but this was not done in this thesis. An advantage of using an economic score over other methods, such as maximising genetic level with constraints on *e.g*. genetic relationships as evaluated by Bérodier *et al*. (2021), is that it avoids the risk of some females not being mated if the constraints are too stringent, or of obtaining a suboptimal solution concerning genetic relationships if the constraints are too relaxed.

It was decided to use linear programming since it has been shown to outperform sequential mating methods, because it uses simultaneous rather than sequential solving to find the economically optimal matings for each herd (Sun *et al.*, 2013; Carthy *et al*., 2019; Bérodier *et al*., 2021). It has also been shown to be faster than other mating methods such as sequential solving (Sun *et al*., 2013; Carthy *et al*., 2019). Therefore, the method is suitable for implementation in software that can assist farmers or advisors in mating decisions in real time, which was considered important in the work in this thesis.

5.2 Including recessive genetic defects in the economic score

The relatively high carrier frequency of the genetic defect at BTA12 in Red Dairy Cattle served as a good example of how the method would handle a recessive genetic defect with high carrier frequency (14.7% in females and 12% in available bulls). An economic score including a penalty for mating two carriers effectively eliminated expression of genetic defects. It was more profitable to use the carrier bull on a non-carrier female than on a carrier female. Hence, the conclusion from the work in Paper II was that linear programming can help avoid expression of genetic defects unless the

possible matings are restricted, *e.g.* if only a few non-carrier bulls are available and therefore a carrier bull has to be mated with a carrier female. The same conclusion was drawn from the results for Holstein cattle obtained in Paper III. However, in Paper IV it was found that the risk of expression of the known recessive genetic defect increased slightly when the economic score included a genomic relationship instead of a pedigree relationship. This was mainly due to low genomic relationships that made it worthwhile to mate two carriers, which was not observed in the studies in Papers II and III. One possible explanation for the latter is that too few different situations were encountered in those studies, as the analyses in both only considered one or two different bull sets and only looked one generation ahead (with no replicates because those studies were based on real data). Most recessive genetic defects examined in this thesis caused early embryonic loss, which has lower economic consequences than other defects resulting in late-term abortions or defective or dead calves. More severe defects should be assigned a higher economic cost in the economic score. This would likely decrease the probability of expression of the defect even more, as mating two carriers would be more costly. The cost of €80 assumed in this thesis reduced the frequency of mating between two carriers to almost zero, so a slightly higher cost would most likely eliminate such matings.

In the study by Bérodier *et al.* (2021), expression of recessive genetic defects could not be completely avoided when using linear programming and similar constraints as in this thesis, most likely due to more restricted bull usage in their study. For example, only eight bulls could be mated to heifers due to restriction of calving ease, while restrictions on availability of semen were also considered (Bérodier *et al.*, 2021). An earlier study by Cole (2015) used sequential solving rather than linear programming and found that more conceptuses were affected by recessive genetic defects when using an economic score (including pedigree relationship and penalty for genetic defects) compared with random mating. That study also revealed a downside of sequential solving which, unlike linear programming, cannot account for the fact that the value of one mating is affected by other matings, which is the case with a limited amount of permitted matings per bull and herd. For example, linear programming accounts for the fact that a bull carrying a recessive genetic defect brings the most value (for most cases) when it is mated to a non-carrier female if there is a maximum number of

inseminations. Hence, while linear programming can help avoid expression of genetic defects, this is probably not possible with sequential solving.

5.3 Breeding for polledness and beta-casein

Papers III and IV focused on polledness, a trait which is high on the agenda within Europe because organic farms in the EU now have to apply for a permit if they want to dehorn cattle (EU Commission Regulation No. 889/2008). In the future, dehorning might be regulated further or even banned, and not only in organic herds. The new EU regulation has increased the demand for semen from bulls carrying the polled allele, and several breeding companies have started marketing polled bulls more heavily (Hanna Driscoll, Product manager Holstein, VikingGenetics, personal communication May 24, 2023). The results in Papers III and IV showed that farmers can achieve relatively rapid changes in the polled allele frequencies with limited effect on genetic gain and rate of inbreeding. However, Paper III also showed that in Holstein, using only polled bulls would be relatively costly in terms of lower genetic level and higher inbreeding. Very few of the top ranked bulls in Holstein are homozygous for the polled allele. Moreover, it was found in Paper IV that the initial rapid increase in the frequency of polled animals observed on introducing an economic score with extra value for polledness was not always maintained in the following generations. One possible explanation is that polled animals are more related and that mating optimisation will prefer horned bulls in the next generations. Hence, breeding for a monogenic trait, such as polledness, requires a long-term perspective (as does animal breeding in general). This means that farmers and breeding companies face a difficult task of adjusting to new rules that vary in a relatively short time, whereas the time horizon of animal breeding spans multiple generations, which often equates to decades in cattle breeding.

In Paper III, beta-casein genotype of the animals was also studied, since despite the lack of confirmed benefits, some countries are seeking to increase consumption of A2 milk and some dairies pay extra for A2 milk (Bisutti *et al.*, 2022). It is still uncommon for farmers in DFS to breed to increase the percentage of A2A2 offspring, and even more uncommon to breed for beta-casein and polledness simultaneously. The results obtained in Paper III for beta-casein and polledness illustrate the interactions that can occur when

breeding for two favourable monogenic traits, since animals carrying the polled allele were found to be less likely to be homozygous for beta-casein (A2A2). It is likely that breeding for multiple monogenic traits will become more common as knowledge about the cattle genome increases, and Paper III provides a useful example of potential challenges that may arise. In general, it is advisable to monitor available monogenic traits at herd and population level to keep track of changes in allele frequencies.

5.4 Genetic relationships

The studies performed in this thesis also addressed the correlations between the different measures of genetic relationships. The correlation between the pedigree relationship and genomic relationship estimates was high in Red Dairy Cattle, 0.83-0.88 for a_{3Gen} and 0.85-0.88 for a_{AllGen} (Paper II). However, Carthy *et al.* (2019) reported a correlation of 0.57 between pedigree relationships and genomic relationship, which is lower than the values in Paper II and other studies (0.67-0.88) (VanRaden *et al*., 2011; Pryce *et al*., 2012). Pryce *et al*. (2012) found a correlation of 0.67, 0.73, 0.84 and 0.87 when the number of generations of recorded ancestry was 2, 4, 6 and 8, respectively, and concluded that pedigree depth plays a major role in the strength of correlation between pedigree relationships and genomic relationships. Lower correlations between pedigree relationship and genomic relationship estimates were found for RDC in Paper III (0.69-0.75 for a_{3Gen}, 0.75-0.76 for a_{AllGen}). Pedigree depth in the study by Pryce *et al*. (2012) was similar to that in Paper II, so the difference is most likely attributable to some other factor(s). One possibility is that the pedigree correctness is better in RDC than in Holstein, due to the less common exchange of RDC bulls and their pedigrees worldwide, as RDC are generally found within the Nordic countries. It is also important to highlight that for some populations, the effectiveness of pedigree relationships may have been overestimated in Paper IV because a perfect pedigree in terms of completeness and correctness was available. With lower pedigree correctness, the relative benefits of using genomic relationships would likely be greater. However, the results may still be applicable for RDC, since in Paper II there were equally strong correlations between pedigree and genomic relationships as in the simulations in Paper IV.

There are several arguments for using genomic estimates of relationship and inbreeding instead of pedigree. First, genomic estimates do not rely on pedigree data, which can be incorrect or have limited depth (Carthy *et al*., 2019; Makanjuola *et al*., 2020). The data used in this thesis were corrected for possible mismatches when received from the Nordic Cattle Genetic Evaluation. Hence, the benefit that genomic information brings in the form of assigning the right parents to an animal was not explored. Approximately 5% of genotyped animals in Sweden have at least one parent incorrectly assigned (Lina Baudin, expert in breeding routines, Växa personal communication, March 5, 2021). Second, even if the pedigree data are correct and complete, genomic relationships are still more accurate because they consider the fact that the genome is transmitted in chromosomes, and not as infinite unlinked loci (Hill & Weir, 2011). Third, the assumption of 50% probability of an allele being selected is not true in a population under selection (Forutan *et al*., 2018). Therefore use of genomic estimates of relationship is recommended in modern mating programmes.

In general, genomic relationships were good at keeping each other low when included in an economic score used for mating allocations, and the best approach would be to implement one of these instead of pedigree relationships. A segment-based relationship was used in this thesis, with the aim of reducing the number of runs of homozygosity (ROH) in the potential offspring. In principle, ROH are enriched for deleterious alleles that mainly cause inbreeding depression (Charlesworth & Willis, 2009). Long ROH reflect new inbreeding and are expected to contain more deleterious alleles than short ROH, due to purging and recombination through the generations (Stoffel *et al*., 2021). According to Pryce *et al*. (2014), long regions (>3 Mb) are associated with inbreeding depression for milk yield in Holstein and Jersey cattle. However, Zhang *et al*. (2015) found significantly higher enrichment of deleterious variants in short (<0.1-3 Mb) compared with long (>3 Mb) regions in the Holstein, RDC and Jersey cattle. Hence, the optimal segment length for use in segment-based relationships remains to be determined. However, this thesis showed that g_{SEG1} and g_{SEG4} keep each other low when included in an economic score, so the difference is most likely marginal for the outcome of the mating allocations. One could also speculate that in the future, inbreeding may be targeted at specific regions and chromosomes, beyond genetic defects. This could then be included in an

economic score to avoid inbreeding in a more precise way if the effects of inbreeding are not equal along all chromosomes and regions.

Papers III and IV focused on SNP-by-SNP genomic relationship coefficient (g_{SNP}), which could be argued to be closest to implementation since it is often used in genomic evaluations and therefore is already calculated and ready for use. It also is the fastest genomic relationship to calculate and, as observed in this thesis, it is relatively good at keeping the segment-based relationships low, making it an efficient implementation alternative. However, computation time aside, a segment-based relationship based on current research should be considered, because it is most likely better in prediction of inbreeding depression (Doekes *et al*., 2021). In general, more studies on this topic are needed, particularly from a Nordic perspective, in order to identify the relationship that best predicts inbreeding depression in the three major dairy cattle breeds.

5.5 Short- and long-term effects of mating programmes

Many of the results in Papers II and III were maintained over several generations in Paper IV. For example, compared with only maximising genetic level, including any genetic relationship in the economic score lowered the rate of increase in pedigree and genomic inbreeding, with minimal effect on genetic gain. In addition, including the cost of a recessive genetic defect in the score helped reduce the risk of expression. As discussed earlier, the risk of expression of the known recessive genetic defect increased slightly when the economic score included a genomic relationship instead of a pedigree relationship. However, genomic relationships resulted in more bulls being used, which was favourable for the rate of genomic inbreeding and performed equally well concerning the rate of pedigree inbreeding.

Paper IV also provided new insights into the constraint on bull usage, where using 5% instead of 10% females per bull and herd reduced the rate of inbreeding. Using more bulls, which resulted in a lower rate of inbreeding, reduced the variation in carrier frequency for the genetic defects, which lowered the probability of mating two carriers of an unknown genetic defect in some generations when carriers were widely used in previous generations. Hence, to minimise the risk of unknown recessive genetic defects, the constraint of 5% females per bull and herd is the best option. Exceeding 10% females per bull and herd cannot be recommended, as that constraint gave

higher variation in carrier frequencies than using 5%. However, if farmers want to achieve rapid changes in the frequency of polled offspring, this could be slightly more likely with 10% females per bull and herd compared with 5%, at the risk of also increasing the frequency of (unknown) genetic defects.

5.6 Final remarks

The results presented in this thesis are largely generalisable, i.e. breeders of other livestock species could adopt the mating scenarios analysed here, but they would need to be adapted to each specific situation. Further, including genomic relationships and information about genetic defects, as done in the mating studies in this thesis, requires data on both female and male genotypes. An economic score could be developed for crossbred animals where the focus is to maximise heterosis instead of minimising parent relatedness. Ungenotyped animals were not considered in this thesis, but one option for these could be to impute their genotype, as done by Carthy et al. (2019) using the method described by Gengler et al. (2007). Sun et al. (2013) suggest use of the H matrix in single-step genomic evaluation.

It is important to mention that this thesis did not fully examine all benefits of genotyping and it is important to bear in mind that the value of genotyping is likely to increase as new applications emerge, such as product or animal traceability through the food chain (Newton & Berry, 2020). Genomic selection also makes it possible to select for novel traits with high accuracy, without the farmer having to measure the trait in their herd (Henryon et al., 2014).

In conclusion, the results presented in this thesis can increase farmers' confidence in GEBV. Genomic mating plans were also assessed, both in a long-term and short-term perspective. It was shown that optimising economic score with linear programming can help avoid expression of genetic defects and increase the level of favourable traits like polledness and beta-casein (genotype A2A2). A recent study in which I participated showed that farmers in Sweden are positive to the use of modern breeding tools like genotyping, sexed semen and beef semen (Clasen et al., 2021). This thesis highlighted the benefits of genotyping for planning matings and will hopefully encourage more farmers to genotype their herds. The mating method used is relatively easy to implement, flexible and ready for future applications when more is known about the cattle genome.

6. Final conclusions

- Virgin heifer genomically enhanced breeding values predicted cow performance significantly better than did parent average breeding values for the vast majority of traits analysed in Holstein, Jersey and Red Dairy Cattle.

- Linear programming maximised the economic score for all herds studied within seconds, provided that data on genetic relationships were available, which means that it is a suitable method for implementation in mating software to be used by advisors and farmers.

- In Red Dairy Cattle and Holstein, it was possible to reduce genetic relatedness between parents with minimal effect on genetic level. Including the cost of known recessive genetic defects entirely eliminated the risk of expression of these defects.

- There were strong correlations between measures of genomic and pedigree relationships for dairy cattle in Denmark, Finland and Sweden.

- It was possible to reduce genomic relationships between parents by including pedigree measures in the economic score, but it was best done by including genomic measures.

- Genomic relationships studied were good at keeping each other low when included in an economic score used for mating allocations.

- In Holstein, it was possible to increase the percentage of polled and beta-casein homozygous (A2A2) offspring substantially in one generation when competitive bulls were available, without any significant negative effect on other mating criteria.

- Using only semen from polled bulls, which might be necessary if dehorning is banned, considerably affected the genetic level in Holstein.

- In Holstein, animals carrying the polled allele were less likely to be homozygous for beta-casein (A2A2) and more likely to be carriers of the genetic defect HH1.

- Long-term simulations revealed that, compared with only maximising genetic level, including genomic or pedigree relationship in the economic score lowered the rate of pedigree and genomic inbreeding, with minimal effect on genetic gain.

- Using genomic relationships in the economic score resulted in more bulls being used, which was favourable for lowering the rate of genomic inbreeding and performed equally well to using pedigree relationships in terms of the rate of pedigree inbreeding.

- A 5% females per bull and herd constraint lowered the variation in carrier frequency for genetic defects, which minimised the risk of mating two carriers of an unknown genetic defect in future generations after widespread use of carriers in previous generations. However, allowing 10% females per bull could accelerate the increase in frequency of the polled allele.

7. Practical recommendations

The current cost of genotyping (around €20-25 per genotype in the Nordic countries) makes it economically justifiable to genotype females for management of breeding stock. Genotyping can be combined with sexed and beef semen and/or selling or culling surplus heifers, as suggested by the literature (Hjortø *et al.*, 2015; Newton & Berry, 2020). Paper I showed that dairy farmers in the Nordic countries can make better selection decisions based on genomically enhanced breeding values than on parent average breeding values.

An economic score that is easy to customise should be developed. As new monogenic traits are identified, the economic score should incorporate these. Moreover, the economic score should reflect the specific needs and preferences of each herd. For instance, as demonstrated in this thesis, the polled trait may have a higher value for organic herds. Therefore, the first step in applying the method presented here is to define the economic score for the target herd and then set the constraints, *e.g*. number of females per bull.

SNP-by-SNP genomic relationship coefficient (g_{SNP}) was the fastest genomic relationship to calculate and was relatively powerful, keeping the segment-based relationships low, making it an efficient implementation option. However, segment-based relationships might be preferable if they can better predict inbreeding depression in future studies. This would entail a trade-off with the increased computation time required for segment-based methods. Papers II and III demonstrated that using a genomic relationship can reduce the expected genomic inbreeding in the next generation by ~1% compared with using a pedigree relationship, which would correspond to a value of approximately €25 per mating and potentially cover the cost of genomic testing. However, further studies are needed to investigate the

different types of genomic relationships and their implications for the economic performance of dairy farms. Use of a genomic relationship can be recommended if both female and bull are genotyped, as current research suggests that it is a better predictor of inbreeding depression than a pedigree relationship. In this thesis, the largest difference in inbreeding was observed when switching from a pedigree to any genomic relationship, rather than fine-tuning the genomic relationship. Therefore, using any genomic relationship for mating decisions and updating it as new research becomes available can be recommended.

When breeding for multiple monogenic traits, it is important that different outcomes of mating allocations are easily comparable. It was shown in Paper III that animals carrying the polled allele had a lower probability of being homozygous for beta-casein (A2A2 genotype) and a higher probability of being carriers of the genetic defect HH1. Therefore, assigning an economic value to a monogenic trait in the economic score used for mating allocations could have a negative impact on another monogenic trait. Hence, it is advisable to monitor the comparison criteria used in this thesis. The carrier frequency in female candidates could provide valuable information for farmers and advisors before making mating decisions in practice, as it could indicate how different defects should be weighted in a specific herd. As demonstrated in Paper II, the overall frequency of most defects could be low among all females, but the carrier frequencies could vary considerably in a specific herd.

It is important to note that the use of genomic relationships slightly increased the risk of conceptuses affected by the known recessive defect examined in Paper IV compared with pedigree relationships. Such matings are not advisable in practice and may be against the law. This could also be true for specific bull traits, *e.g.* some bulls with traits that give difficult calvings should not be used for heifers. This is important to consider in practical implementations and should be relatively easy to include in linear programming optimisation.

Another recommendation is to have a constraint on bull usage, as done in this thesis. To minimise the risk of unknown recessive genetic defects, using the 5% females per herd and bull option would be the best choice. Exceeding 10% females per bull and herd is not recommended, to avoid higher variation in carrier frequencies. However, if farmers want to achieve rapid changes in the frequency of polled offspring, there could be some benefits of using 10%

females per bull and herd. In that case, it is important to be aware of the risk of increasing the frequency of (unknown) genetic defects.

Most farms in the Nordic countries use a special ear tag that collects a tissue sample at the same time as the animal's identification is inserted into the ear. This method requires minimal extra work and allows for early sampling and decision making. The farmers also benefit from the pedigree verification, which provides a more accurate pedigree and helps to avoid unnecessary inbreeding when planning matings. However, this aspect was not addressed in Papers II-IV, which might have underestimated the advantage of genomic data over pedigree data, as the pedigree errors on genotyped animals were already corrected in the data.

This thesis demonstrated the potential of female genotyping at herd level for enhancing dairy cattle breeding. Genotyping can offer valuable information for estimating genomic breeding values, optimising mating programmes and selecting for novel traits. The cost of genotyping (currently €20-25 per genotype) may have reached a plateau in terms of possible reductions. However, the value of genotyping is likely to increase as new applications emerge, such as product or animal traceability through the food chain (Newton & Berry, 2020). Therefore, based on the results in this thesis and other literature, genotyping heifers can be profitable if farmers use the genomic results to improve selection and mating strategies.

8. Future research

A possible direction for future research is to develop an economic score for crossbred animals that accounts for heterosis effects. Currently, around 10% of cows in the Nordic countries are crossbred and this percentage may increase in the future. However, there is a lack of research on how to incorporate heterosis into economic scores for crossbred animals. Another area of interest is economic score development for beef on dairy matings.

The optimal segment length when considering genomic relationships is unclear. In general, more studies are needed on the economic consequences of inbreeding and monogenic traits in Nordic production systems. Future studies should also seek to identify chromosomal regions linked to inbreeding depression in dairy cattle and incorporate these into the economic score in order to prevent inbreeding.

It is important to develop tools that can support farmers in their decision to start genotyping based on their specific farm characteristics. Farmers may have different motives and constraints for adopting genotyping technology. Future research should develop tools that help farmers evaluate the potential benefits and costs of genotyping based on their specific farm characteristics.

Economic scores could be further extended by incorporating more monogenic traits, such as kappa-casein, which was added to the SNP array following the work in this thesis. The effects of including specific breeding values in the economic score, *e.g.* following the approach of Carthy *et al*. (2019), should also be analysed.

More surveys are needed to identify the factors influencing farmers' adoption of genotyping technology. Farmers may face various barriers or challenges when adopting genotyping technology, such as lack of awareness, knowledge, skills, trust or resources.

Another possible direction for future research is to explore economic scores for mating programmes targeting breeding animals (*e.g*. MOET mating) and how it interacts with OCS. This could help breeding companies optimise their selection decisions and allocate resources more efficiently. Future research should also investigate the use of MOET in commercial herds without the constraint of one mating per female applied in this thesis.

Finally, there are many more scenarios that are yet to be studied in a follow-up to Paper IV. These include more or different monogenic traits, with higher or lower carrier frequencies. Further, associations between monogenetic traits, *e.g.* as found between polled and beta-casein in Paper III, or associations with the breeding goal should be included. In addition, future studies should assess more active selection for monogenic traits in the MOET programme.

9. References

Bengtsson, C., Thomasen, J.R., Kargo, M., Bouquet, A., Slagboom, M. (2022). Emphasis on resilience in dairy cattle breeding: Possibilities and consequences. *J. Dairy Sci.,* 105, 7588–7599. https://doi.org/10.3168/jds.2021-21049

Berkelaar and others, (2020). lpSolve: interface to "Lp_solve" v. 5.5 to solve linear/integer programs. https://cran.rproject.org/web/packages/lpSolve/index.html.

Bérodier, M., Berg, P., Meuwissen, T., Boichard, D., Brochard, M., Ducrocq, V. (2021). Improved dairy cattle mating plans at herd level using genomic information. *Animal,* 15, 100016. https://doi.org/10.1016/j.animal.2020.100016

Bisutti, V., Pegolo, S., Giannuzzi, D., Mota, L.F.M., Vanzin, A., Toscano, A., Trevisi, E., Marsan, P.A., Brasca, M., Cecchinato, A. (2022). The β-casein (CSN2) A2 allelic variant alters milk protein profile and slightly worsens coagulation properties in Holstein cows. *J. Dairy Sci.,* 105, 3794–3809. https://doi.org/10.3168/jds.2021-21537

Browning, B.L., Browning, S.R. (2013). Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics,* 194, 459–471. https://doi.org/10.1534/genetics.113.150029

Browning, S.R., Browning, B.L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 81, 1084–1097. https://doi.org/10.1086/521987

Burnell, M. (2019). The use of sexed semen in dairy herds. *Livestock,* 24, 282–286. https://doi.org/10.12968/live.2019.24.6.282

Calus, M.P.L., Bijma, P., Veerkamp, R.F. (2015). Evaluation of genomic selection for replacement strategies using selection index theory. *J. Dairy Sci.,* 98, 6499–6509. https://doi.org/10.3168/jds.2014-9192

Carthy, T.R., McCarthy, J., Berry, D.P. (2019). A mating advice system in dairy cattle incorporating genomic information. *J. Dairy Sci.,* 102, 8210–8220. https://doi.org/10.3168/jds.2019-16283

Charlesworth, D., Willis, J.H. (2009). The genetics of inbreeding depression. *Nat. Rev. Genet.,* 10, 783–796. https://doi.org/10.1038/nrg2664

Christensen, O.F., Lund, M.S. (2010). Genomic prediction when some animals are not genotyped. *Genet. Sel. Evol.,* 42, 2. https://doi.org/10.1186/1297-9686-42-2

Clasen, J.B., Bengtsson, C., Källström, H.N., Strandberg, E., Fikse, W.F., Rydhmer, L. (2021). Dairy cattle farmers' preferences for different breeding tools. *Animal*, 15, 100409. https://doi.org/10.1016/j.animal.2021.100409

Cole, J.B. (2015). A simple strategy for managing many recessive disorders in a dairy cattle breeding program. *Genet. Sel. Evol.,* 47, 94. https://doi.org/10.1186/s12711-015-0174-9

de Cara, M.Á.R., Villanueva, B., Toro, M.Á., Fernández, J. (2013). Using genomic tools to maintain diversity and fitness in conservation programmes. *Mol. Ecol.,* 22, 6091–6099. https://doi.org/10.1111/mec.12560

de Roos, A.P.W., Hayes, B.J., Goddard, M.E. (2009). Reliability of genomic predictions across multiple populations. *Genetics*, 183, 1545–1553. https://doi.org/10.1534/genetics.109.104935

Doekes, H.P., Bijma, P., Windig, J.J. (2021). How Depressing Is Inbreeding? A Meta-Analysis of 30 Years of Research on the Effects of Inbreeding in Livestock. *Genes,* 12, 926. https://doi.org/10.3390/genes12060926

EFFAB, 2020. Code of good practice. Available at: https://www.responsiblebreeding.eu/uploads/2/3/1/3/23133976/01_general_document_2020_final-code_efabar.pdf (accessed 2.2.22).

Fikse, F., Kargo, M. (2020). The economic value of NTM, Available at: https://www.nordicebv.info/wp-content/uploads/2020/11/The-economic-value-of-NTM_final.pdf (accessed 2.2.22).

Forutan, M., Ansari Mahyari, S., Baes, C., Melzer, N., Schenkel, F.S., Sargolzaei, M. (2018). Inbreeding and runs of homozygosity before and after genomic selection in North American Holstein cattle. *BMC Genomics,* 19. https://doi.org/10.1186/s12864-018-4453-z

Gautier, M., Faraut, T., Moazami-Goudarzi, K., Navratil, V., Foglio, M., Grohs, C., Boland, A., Garnier, J.-G., Boichard, D., Lathrop, G.M., Gut, I.G., Eggen, A. (2007). Genetic and haplotypic structure in 14

european and african cattle breeds. *Genetics,* 177, 1059–1070. https://doi.org/10.1534/genetics.107.075804

Gaynor, R.C., Gorjanc, G., Hickey, J.M. (2021). AlphaSimR: an R package for breeding program simulations. *G3 GenesGenomesGenetics*, 11, jkaa017. https://doi.org/10.1093/g3journal/jkaa017

Hayes, B.J., Bowman, P.J., Chamberlain, A.J., Goddard, M.E. (2009). Invited review: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci.,* 92, 433–443. https://doi.org/10.3168/jds.2008-1646

Henryon, M., Berg, P., Sørensen, A.C. (2014). Animal-breeding schemes using genomic information need breeding plans designed to maximise long-term genetic gains. *Livest. Sci.,* 166, 38–47. https://doi.org/10.1016/j.livsci.2014.06.016

Hjortø, L., Ettema, J.F., Kargo, M., Sørensen, A.C. (2015). Genomic testing interacts with reproductive surplus in reducing genetic lag and increasing economic net return. *J. Dairy Sci.,* 98, 646–658. https://doi.org/10.3168/jds.2014-8401

Indyk, H.E., Gill, B.D., Wood, J.E., Chetikam, S., Kobayashi, T. (2021). The determination of intact β-casein in milk products by biosensor immunoassay. *J. Food Compos. Anal.*, 101, 103946. https://doi.org/10.1016/j.jfca.2021.103946

Jansen, G.B., Wilton, J.W. (1985). Selecting Mating Pairs with Linear Programming Techniques. *J. Dairy Sci.,* 68, 1302–1305. https://doi.org/10.3168/jds.S0022-0302(85)80961-9

Kolmodin, R., Strandberg, E., Madsen, P., Jensen, J., Jorjani, H. (2002). Genotype by Environment Interaction in Nordic Dairy Cattle Studied Using Reaction Norms. *Acta Agric. Scand. Sect. - Anim. Sci.,* 52, 11–24. https://doi.org/10.1080/09064700252806380

Legarra, A., Aguilar, I., Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.,* 92, 4656–4663. https://doi.org/10.3168/jds.2009-2061

Liu, H., Henryon, M., Sørensen, A.C. (2017). Mating strategies with genomic information reduce rates of inbreeding in animal breeding schemes without compromising genetic gain. *Animal,* 11, 547–555. https://doi.org/10.1017/S1751731116001786

Løvendahl, P., Difford, G.F., Li, B., Chagunda, M.G.G., Huhtanen, P., Lidauer, M.H., Lassen, J., Lund, P. (2018). Review: Selecting for improved feed efficiency and reduced methane emissions in dairy cattle. *Animal,* 12, s336–s349. https://doi.org/10.1017/S1751731118002276

MacLeod, I.M., Larkin, D.M., Lewin, H.A., Hayes, B.J., Goddard, M.E. (2013). Inferring Demography from Runs of Homozygosity in Whole-Genome Sequence, with Correction for Sequence Errors. *Mol. Biol. Evol.,* 30, 2209–2223. https://doi.org/10.1093/molbev/mst125

Makanjuola, B.O., Miglior, F., Abdalla, E.A., Maltecca, C., Schenkel, F.S., Baes, C.F. (2020). Effect of genomic selection on rate of inbreeding and coancestry and effective population size of Holstein and Jersey cattle populations. *J. Dairy Sci.,* 103, 5183–5199. https://doi.org/10.3168/jds.2019-18013

Mäntysaari, E.A., Strandén, I. (2010). Use of bivariate EBV-DGV model to combine genomic and conventional breeding value evaluations. " Proceedings of the 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany. Proc. 9th World Congr. *Genet. Appl. Livest. Prod. Leipz.,* Ger. 4. Available at: http://www.nordicebv.info/wp-content/uploads/2015/04/Use-of-bivariate-EBV-DGV-model-to-combine-genomic.pdf

Meuwissen, T.H., Hayes, B.J., Goddard, M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157, 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Misztal, I., Lourenco, D., Legarra, A. (2020). Current status of genomic evaluation. *J. Anim. Sci.,* 98, skaa101. https://doi.org/10.1093/jas/skaa101

NAV, (2019). NAV routine genetic evaluation of Dairy Cattle – data and genetic models. Available at: https://www.nordicebv.info/wp-content/uploads/2017/03/NAV-routine-genetic-evaluation-122016_FINAL.pdf https://www.nordicebv.info/wp-content/uploads/2019/06/NAV-routine-genetic-evaluation-04062019.pdf (accessed 8.22.19).

NAV, (2023). Available at: https://nordic.mloy.fi/NAVTrends (accessed 10.25.23).

Newton, J.E., Berry, D.P. (2020). On-farm net benefit of genotyping candidate female replacement cattle and sheep. *Animal,* 14, 1565–1575. https://doi.org/10.1017/S1751731120000208

Oltenacu, P., Broom, D. (2010). The impact of genetic selection for increased milk yield on the welfare of dairy cows. *Anim. Welf.,* 19, 39–49. https://doi.org/10.1017/S0962728600002220

Pryce, J.E., Hayes, B.J., Goddard, M.E. (2012). Novel strategies to minimize progeny inbreeding while maximizing genetic gain using genomic information. *J. Dairy Sci.,* 95, 377–388. https://doi.org/10.3168/jds.2011-4254

Purfield, D.C., Berry, D.P., McParland, S., Bradley, D.G. (2012). Runs of homozygosity and population history in cattle. *BMC Genet.* 13, 70. https://doi.org/10.1186/1471-2156-13-70

R Core Team, 2020. A language and environment for statistical computing.

Rajora, O.P. (2019). Population Genomics: Concepts, Approaches and Applications, *Springer International Publishing, Cham.* https://doi.org/10.1007/978-3-030-04589-0

Sargolzaei, M. (2014). SNP1101 user's guide. Version 1.0.

Schaeffer, L. r., 2006. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.,* 123, 218–223. https://doi.org/10.1111/j.1439-0388.2006.00595.x

Stock, M.L., Baldridge, S.L., Griffin, D., Coetzee, J.F. (2013). Bovine Dehorning. *Vet. Clin. North Am. Food Anim. Pract.,* 29, 103–133. https://doi.org/10.1016/j.cvfa.2012.11.001

Stoffel, M.A., Johnston, S.E., Pilkington, J.G., Pemberton, J.M. (2021). Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. *Nat. Commun.,* 12, 2972. https://doi.org/10.1038/s41467-021-23222-9

Su, G., Guldbrandtsen, B., Gregersen, V.R., Lund, M.S. (2010). Preliminary investigation on reliability of genomic estimated breeding values in the Danish Holstein population. *J. Dairy Sci.,* 93, 1175–1183. https://doi.org/10.3168/jds.2009-2192

Summer, A., Frangia, F.D., Marsan, P.A., Noni, I.D., Malacarne, M. (2020). Occurrence, biological properties and potential effects on human health of β-casomorphin 7: Current knowledge and concerns. *Crit. Rev. Food Sci. Nutr.,* 2020, 3705–3723. https://doi.org/10.1080/10408398.2019.1707157

Sun, C., VanRaden, P.M., O'Connell, J.R., Weigel, K.A., Gianola, D., (2013). Mating programs including genomic relationships and dominance effects. J. Dairy Sci. 96, 8014–8023. https://doi.org/10.3168/jds.2013-6969

Taskinen, M., Mäntysaari, E.A., Lidauer, M.H., Knürr, T., Pösö, J., Su, G., Aamand, G.P., Strandén, I. (2013). Comparison of Model Reliabilities from Single-Step and Bivariate Blending Methods. *INTERBULL Bull. 6.* https://jukuri.luke.fi/bitstream/handle/10024/481814/1331-2230-1-PB.pdf?sequence=1

Thomasen, J.R., Willam, A., Egger-Danner, C., Sørensen, A.C. (2016). Reproductive technologies combine well with genomic selection in dairy breeding programs. *J. Dairy Sci.,* 99, 1331–1340. https://doi.org/10.3168/jds.2015-9437

VanRaden, P.M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.,* 91, 4414–4423. https://doi.org/10.3168/jds.2007-0980

Wang, H., Misztal, I., Legarra, A. (2014). Differences between genomic-based and pedigree-based relationships in a chicken population, as a function of quality control and pedigree links among individuals. *J. Anim. Breed. Genet.,* 131, 445–451. https://doi.org/10.1111/jbg.12109

Weigel, K.A., Lin, S.W. (2000). Use of computerized mate selection programs to control inbreeding of Holstein and Jersey Cattle in the next generation. *J. Dairy Sci.,* 83, 822–828. https://doi.org/10.3168/jds.S0022-0302(00)74945-9

Weigel, K.A., Mikshowsky, A.A., Cabrera, V.E. (2015). Effective use of genomics in sire selection and replacement heifer management. In Proc. Western Dairy Management Conference, Reno, NV. 14. http://www.wdmc.org/2015/Weigel.pdf

Yao, C., Armentano, L.E., VandeHaar, M.J., Weigel, K.A. (2015). Short communication: Use of single nucleotide polymorphism genotypes and health history to predict future phenotypes for milk production, dry matter intake, body weight, and residual feed intake in dairy cattle. *J. Dairy Sci.,* 98, 2027–2032. https://doi.org/10.3168/jds.2014-8707

Zhang, Q., Guldbrandtsen, B., Bosse, M., Lund, M.S., Sahana, G. (2015). Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC Genomics,* 16, 542. https://doi.org/10.1186/s12864-015-1715-x

Popular science summary

Breeding work in dairy cattle is important for improving the productivity and profitability of the dairy sector. Traditionally, the main goal of breeding dairy cattle was to improve milk yield, but other desirable traits such as health and fertility have now become important breeding objectives. Recent research has also sought to include traits related to sustainability and climate impact, *e.g.* methane emissions.

In 2008, a new breeding technique called genomic selection became available for use in dairy cattle breeding. It applies information from the DNA of cows to predict how good they are in terms of different traits, such as milk yield, health, fertility and sustainability. Genomic selection is useful because it can better tell how good a cow or bull is before they have their own offspring, rather than only looking at information from their parents.

In the early days, DNA testing was expensive and was only used for candidate bulls for artificial insemination. Over time, however, the cost of genotyping has decreased significantly and more and more farmers are now genotyping their heifers. The female test results can be used to identify the best females for replacement and those to inseminate with sexed semen in order to obtain female calves with the highest breeding value. The genomic test results can also be used to decide which female to mate with which bull. For instance, DNA test results can be used to estimate genomic relationships between females and bulls. Genomic relationships are suggested to be more accurate than those relying on pedigree information because they are based on actual DNA and do not rely on pedigree data, which can be incorrectly recorded or incomplete. In addition, DNA testing gives insights into single-gene traits that can be considered in mating choices. Examples of single-gene traits are cows with no horns (polled) or special milk quality traits like caseins.

This doctoral thesis focused on how genomic information can be used at herd level, mainly considering better mating plans. The first study compared genomic breeding values and parent average breeding values for young females in terms of their ability to predict cow performance later in life. Twelve different traits in first parity were analysed, including production, conformation, fertility and other functional traits. The results for all traits showed that genomic breeding values can predict future cow performance significantly more accurately than parent average breeding values.

A second study investigated mating plan optimisation for Red Dairy Cattle at herd level based on economic score for different mating options within each herd, considering genetic level, semen cost, the economic impact of recessive genetic defects, and genomic and pedigree relationships. The mating results showed that it was possible to reduce genetic relatedness 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 measures, but it was best done with genomic measures.

A third study analysed mating plan optimisation in the Holstein dairy breed and also included two favourable single-gene traits, polledness (absence of horn) and beta-casein. Beta-casein is one of the main proteins in milk and the genetic trait mainly has two variants in cattle (A1 and A2). Cows that carry two copies of the A2 variant produce A2 milk, which is often advertised as a healthier alternative than regular cow milk. However, the human health advantages of drinking A2 milk are still under debate. The results obtained for Holstein cattle were similar to those obtained for Red Dairy Cattle as regards genetic relationships and defects. In addition, it was possible to increase the frequency of polled and beta-casein (A2A2) offspring without negatively impacting other criteria.

A final simulation study investigated the long-term impact of genomic mating allocations. The matings followed a similar approach as in the previous studies and were optimised on genetic level, a favourable single-gene trait (polledness), a recessive genetic defect and parent relationships. An (assumed) unknown recessive genetic defect was also monitored. Compared with only maximising genetic level, including any genetic relationship in the economic score lowered the rate of increase in pedigree and genomic inbreeding, with minimal effect on genetic gain. Including the

cost of a recessive genetic defect in the score helped reduce the risk of expression of that defect. Furthermore, including an economic value for polledness in the economic score increased the frequency of the polled allele in the population, without negatively impacting other comparison criteria. Using more bulls, which helped lower the rate of inbreeding, was favourable regarding the number of animals that were carriers of genetic defects, which reduced the risk of expression in future generations.

One possible direction for future research is to develop economic scores for crossbred dairy cows that take into account the benefits of heterosis. Currently, around 10% of dairy cows in Denmark, Sweden and Finland are crossbred, and this percentage may increase in the future. In addition, future studies should investigate how to calculate genomic relationships in the best way possible. There are many methods available, but it is unclear which is best for different situations and breeds. This is especially important for the Jersey breed and for Nordic Red Dairy Cattle, for which research is limited to the Nordic countries.

Populärvetenskaplig sammanfattning

Mjölkkoavel är viktigt för att förbättra produktivitet och lönsamhet inom mjölksektorn. Historiskt var huvudmålet med mjölkkoavel att förbättra mjölkavkastningen, men de senaste decennierna har andra önskvärda egenskaper såsom fruktsamhet och hälsa fått allt större fokus. På senare tid fokuserar forskningen på att inkludera nya egenskaper som relaterar till miljö- och klimatpåverkan.

År 2008 började en ny avelsteknik vid namn genomisk selektion att implementeras av avelsföretagen. Den använder information från kornas DNA för att förutsäga hur bra de är för olika egenskaper, såsom mjölkavkastning, hälsa, fruktsamhet och hållbarhet. Genomisk selektion är användbart inom mjölkkoavel eftersom den kan ge en bättre bild av hur bra en kviga kommer att bli, eller hur bra en tjur är innan den får egna avkommor jämfört med att bara titta på djurens stamtavla.

När tekniken implementerades var det relativt dyrt att DNA-testa djur och det var främst tjurar som testades som var kandidater till att bli semintjurar. Sedan dess har kostnaden sjunkit betydligt och idag (år 2023) kostar det cirka 250 kr att ta ett DNA-test. Detta gör det intressant att testa även kvigor och allt fler lantbrukare DNA-testar hela sin besättning. Resultaten kan användas till att välja ut vilka kvigor som ska insemineras med könssorterad sperma så att man är säker på att få en kvigkalv av sina bästa djur. Dessutom kan DNA-testresultaten användas till att beräkna släktskap mellan hondjur och tjur på ett bättre sätt och därmed undvika onödig inavel. Genomiska släktskap anses vara säkrare än de som baseras på stamtavlan eftersom de tittar på faktiskt DNA och inte förlitar sig på stamtavlan som kan innehålla fel eller vara ofullständig. Dessutom kan DNA-testen ge insikter om egenskaper som styrs av enskilda gener som kan beaktas vid valet av vilken tjur som ska användas till vilket hondjur (parningsplanering). Exempel på sådana egenskaper är

vissa genetiska defekter, djur som saknar horn (kulliga) eller specifika mjölkkvalitetsegenskaper såsom kaseiner.

Denna doktorsavhandling undersöker hur DNA-information kan användas på besättningsnivå, främst med hänsyn till bättre parningsplanering. I den första studien jämförde vi kvigors genomiska avelsvärden med avelsvärden baserat på enbart stamtavla vad gäller förmåga att förutsäga kornas egenskaper senare i livet. Syftet med denna studie var att öka förtroendet för genomiska avelsvärden bland lantbrukare. Vi analyserade 12 olika egenskaper, inklusive mjölkavkastning, exteriör, fruktsamhet och hälsoegenskaper. Genomiska avelsvärden förutsåg framtida egenskaper betydligt bättre jämfört med avelsvärden baserade på stamtavla.

I den andra studien optimerade vi parningar för nordiska röda raser med hjälp av ekonomiska poängsummor där vi tog hänsyn till genetisk nivå, släktskap, och sannolikhet för att genetiska defekter kommer till uttryck. Resultaten visade att vi kunde minska släktskapet mellan föräldrar med minimal påverkan på genetisk nivå. När kostnaden för recessiva genetiska defekter ingick i optimeringen försvann risken för bärarparningar. Med andra ord var det aldrig ekonomiskt fördelaktigt att inseminera ett hondjur som var bärare av en recessiv genetisk defekt med en tjur som var bärare av samma genetiska defekt. Ett längre kalvningsintervall och en extra inseminering kostar mer än vad en optimal parning för övriga faktorer kan kompensera för.

I den tredje artikeln undersökte vi parningplanering för en annan ras (holstein) och inkluderade även positiva egenskaper såsom kullighet och Beta-kasein. Resultaten överstämde väl med den andra artikeln vad det gäller släktskap och genetiska defekter. Dessutom var det möjligt att öka andelen kulliga och dubbelbärare av varianten av Beta-kasein (A2A2) utan att nämnvärt påverka de övriga jämförelsekriterierna.

Slutligen undersökte vi vilka effekter genomisk parningsplanering har på lång sikt. Denna artikel byggde på de två tidigare parningsartiklarna. Vi inkluderade genetisk nivå, kullighet, en recessiv genetisk defekt och släktskap på liknande sätt, dessutom följde vi en okänd genetisk defekt för att undersöka risker beroende på vilka parningsbeslut som togs. Bland annat så visade resultaten att jämfört med att bara maximera genetisk nivå så kunde vi genom att inkludera släktskap i den ekonomiska poängsumman minska inavelstakten och minska risken att kända genetiska defekter kommer till uttryck. Vi kunde också öka andelen kulliga kor utan större påverkan på

övriga jämförelsekriterier. Att använda fler tjurar var fördelaktigt i förhållande till inavelsökning och minskad risk för okända genetiska defeker skulle bli vanligt förekommande eller komma till uttryck.

Framtida studier skulle kunna undersöka möjligheten för ekonomiska poänsummor för korsningsdjur. I dagsläget (år 2023) är cirka 10% av korna korsningar i Danmark, Finland och Sverige och det är möjligt att det kommer öka i framtiden. Det krävs också mer forskning om de olika sätten att beräkna genomiska släktskap och hur bra de är på att förutsäga inavel och dess effekter. Detta är särskild viktigt för rasen jersey och de nordiska röda raserna för vilka forskningen är begränsad till nästan enbart Norden.

Acknowledgements

The work in this thesis was carried out at the Department of Animal Breeding and Genetics (HGEN), Swedish University of Agricultural Sciences (SLU), and at VikingGenetics in Skara and Falkenberg. It has been five years filled with a lot of studies but also a pandemic, and I have become a dad of two wonderful kids. Despite many obstacles, I have learned a lot and had a lot of fun, and I am for sure a better version of myself than when I started this journey.

I would like to acknowledge the following for supporting me in one way or another during my PhD studies:

My supervisors: **Erling Strandberg**, thank you for sharing your excellent expertise in animal breeding, you were a great inspiration as an expert in the field and on being professional while still retaining a great sense of humour. **Hans Stålhammar**, the most important person in my career, you first brought me into VikingGenetics as a trainee in 2015 and helped me become a PhD student. Your kindness and support are beyond what to expect from any human being. **Freddy Fikse,** thanks for all the tricky questions, of course, but also for sharing your knowledge with me the last 10+ years. You were a major inspiration to me. **Susanne Eriksson**, thanks for all your support and positive comments. I have never left a meeting with you without a better feeling than before. You have taught me a lot about animal breeding, writing and project leadership. **Jørn Rind Thomasen,** thank you for all your kindness, support and great fun during meetings and travel. **Martin Johnson**, thank you for your excellent expertise in the final paper and for bringing my research to Scotland, where you made me hold a seminar.

My colleagues at VikingGenetics: especially **Mia**, **Charlotte**, **Anna, Stina**, **Annika**, **Annicka**, **Sven**, **Sofia**, **Elin**, **Hanna**, **Lars**, **LP**, **Kenneth**, **Anders**, **Kasia**, **Johanna**, **Jan**, **Henrik**, and **Saija**. Thank you for your

support and for answering my strange questions for my PhD. Thanks also for sharing my studies internally and externally and for bringing the research to implementation. A special thank you to **Søren** for making my PhD studies possible and for all the support during this journey.

My colleagues at SLU: especially **Juan**, **Tomas K**, **Julie**, **Patricia**, **Paulina**, **DJ**, **Sandra**, **Stanley**, **Martin**, **Ida**, **Sallam**, **Anahit**, **Elena**, **Jussi,** **Fernando**, **Lotta**, **Cano**, **Lars**, **Christos,** and **Elisenda**. I also thank my colleagues at the other Nordic universities for the great fun and discussions during conferences and courses.

Most importantly, my warmest thanks to my beautiful family: **Camilla,** **Lykke** and **Celina** for always being there and reminding me that there are other important things in life than just PhD studies. **Mom** and **Dad**, for your support and advice from "real farmers". My sister, **Johanna**, for help and answers to questions relating cow management. My extra sister and office mate **Julie,** for all laughs, conferences, great parties, crazy ideas and support during the journey.

Finally, thanks to **LivsID** for funding externally employed doctoral students within food-related research; to all the people connected to the programme; and to **Faba**, **Nordic Cattle Genetic Evaluation**, **SEGES** and **Växa** for providing data and feedback during the project.

Appendix 1.

Table A1. Monogenic traits available for mating programmes in Denmark, Finland and Sweden

Monogenic trait	OMIA CODE	Phenotype
Bovine Leukocyte Adhesion Deficiency (BLAD)	000595-9913	Extreme susceptibility to infection and early mortality in homozygous offspring
Complex Vertebral Malformation (CVM)	001340-9913	Stillborn calf
Holstein Haplotype 1 (HH1)	000001-9913	Early abortion of homozygous conceptus
Holstein Haplotype 2 (HH2)	001824-9913	Early abortion of homozygous conceptus
Holstein Haplotype 4 (HH4)	001826-9913	Early abortion of homozygous conceptus
Holstein Haplotype 6 (HH6)	002194-9913	Early abortion of homozygous conceptus
Holstein Haplotype 7 (HH7)	001830-9913	Early abortion of homozygous conceptus
Spinal muscular atrophy (SMA)	000939-9913	Calves become weak and have problems standing, progressively worsen until they die
Arthrogryposis multiplex congenita (AMC)	002022-9913	Stillborn calf or calf death shortly after birth
Ptosis Intellectual disability, Retarded growth, and Mortality (PIRM/AH1)	001934-9913	Early abortion, PIRM/AH1 are located very closely and are Expected to be the same disease

Monogenic trait	OMIA CODE	Phenotype
Ayrshire Haplotype 2 (AH2)	002134-9913	Early abortion of homozygous conceptus
Brown Swiss Haplotype 2 (BH2)	001934-9913	Stillborn calf or calf death shortly after birth
Bos Taurus Autosome 12 (BTA12)	001901-9913	Early abortion of homozygous conceptus
Bos Taurus Autosome 23 (BTA23)	001991-9913	Stillborn calf
Jersey Haplotype 1 (JH1)	001697-9913	Embryonic death
Progressive retinal degeneration (RP1)	000866-9913	Progressive blindness in homozygous offspring
Polledness	000483-9913	Absence of horns in offspring carrying at least one copy of the polled allele (Celtic and Friesian allele considered)
Beta-casein	002033-9913	A cow produces so-called ”A2 milk” if she has two copies of the A2 allele
Kappa-casein	002400-9913	Milk protein that influences the amount of clotting that occurs, six possible genotypes (AA, AB, AE, BB, BE, EE). Milk from cows with BB genotype clots more quickly and produces the highest cheese yield, cows with EE genotype produce milk that does not clot

I

J. Dairy Sci. 103:6383–6391
https://doi.org/10.3168/jds.2019-17963

Association of genomically enhanced and parent average breeding values with cow performance in Nordic dairy cattle

C. Bengtsson,[1,2]* H. Stålhammar,[1] E. Strandberg,[2] S. Eriksson,[2] and W. F. Fikse[3]
[1]VikingGenetics Sweden AB, 53294 Skara, Sweden
[2]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden
[3]Växa Sverige, 75105 Uppsala, Sweden

ABSTRACT

This study compared the abilities of virgin heifer genomically enhanced breeding values (GEBV) and parent average breeding values (PA) to predict future cow performance. To increase confidence in genomic technology among farmers, a clear demonstration of the relationship between genomic predictions and future phenotypes is needed. We analyzed 12 different traits in first parity, including production, conformation, fertility, and other functional traits. Phenotype data were obtained from national milk recording schemes and breeding values from the Nordic Cattle Genetic Evaluation. Direct genomic breeding values were calculated using genomic BLUP and combined with traditional breeding values, using bivariate blending. The data covered 14,862 Red Dairy Cattle, 17,145 Holstein, and 7,330 Jersey genotyped virgin heifers born between 2013 and 2015 in Denmark, Finland, and Sweden. Phenotypes adjusted for systematic environmental effects were used as measures of cow performance. Two methods were used to compared virgin heifer GEBV and PA regarding their ability to predict future cow performance: (1) correlations between breeding values and adjusted phenotypes, (2) ranking cows into 4 quartiles for their virgin heifer GEBV or PA, and calculating actual cow performance for each quartile. We showed that virgin heifer GEBV predicted cow performance significantly better than PA for the vast majority of analyzed traits. The correlations with adjusted phenotypes were 38 to 136% higher for GEBV than for PA in Red Dairy Cattle, 42 to 194% higher for GEBV in Holstein, and 11 to 78% higher for GEBV in Jersey. The relative change between GEBV bottom and top quartiles compared with that between PA bottom and top quartiles ranged from 9 to 261% for RDC, 42 to 138% for Holstein, and 4 to 90% for Jersey. Hence, farmers in Denmark, Finland, and Sweden can have confidence in using genomic technology on their herds.
Key words: genomic breeding value, genotyping, dairy cow, validation

INTRODUCTION

To increase confidence in genomic technology among farmers, a clear illustration of the relationship between genomic predictions and future phenotypes is needed (Pryce and Hayes, 2012). In the early years of genomic selection, mainly bulls were tested, but genotyping of virgin heifers has become more interesting as the costs decrease (Calus et al., 2015; Hjortø et al., 2015; Ettema et al., 2017). At herd level, genomic test results can be used to (1) find the best females for breeding and replacement, (2) identify females for embryo transfer or in vitro fertilization, (3) correct parentage assignment, (4) control monogenic traits, and (5) avoid inbreeding through genomic-assisted mating plans (Pryce et al., 2012).

Genomically enhanced breeding values (**GEBV**) can be validated in different ways. Cross-validation includes dividing the available data set into validation and training sets. By masking observations of all individuals in the validation set and predicting the observations or EBV with a model based on individuals in the training set only, the correlation between masked phenotypes or EBV and predicted values for the validation individuals can be estimated. This correlation then reflects the accuracy of prediction (de Roos et al., 2009). A disadvantage of validating GEBV against conventional EBV is that training and validation sets are rarely strictly independent (Su et al., 2010). Yao et al. (2015) used genotypes and health data to predict future phenotypes, taking correlations between predicted values and phenotypes as measurements of accuracy. To illustrate the accuracy of GEBV compared with parent average breeding values (**PA**), Weigel et al. (2015) divided cows into quartiles based on their virgin heifer GEBV and sire PTA, and thereafter calculated actual cow performances for each quartile.

Received November 26, 2019.
Accepted February 27, 2020.
*Corresponding author: chben@vikinggenetics.com

Establishment of the Nordic Cattle Genetic Evaluation in 2002 has led to intensified cooperation between AI organizations in Denmark, Finland, and Sweden (**DFS**). Because the differences across the Nordic countries were small, according to a study on genotype × environment interactions by Kolmodin et al. (2002), a joint breeding program was established. The current breeding goal combines breeding values for 60 traits into 14 main breeding values, including health, reproduction, production, and conformation.

Genotyping of cows and virgin heifers in DFS started on a large scale in 2012 with the VikingGenetics genotyping project. Three breeds in the DFS countries have genomic breeding schemes: Red Dairy Cattle (**RDC**), Holstein, and Jersey. Initially, the main purpose was to include genotyped females in the reference population and thereby increase the accuracy of GEBV. This was especially important for RDC and Jersey, which had more limited reference populations based on bulls than did the Holstein breed. In 2018, close to 12% of females born in DFS were genomically tested, compared with approximately 2% in 2012, and growth potential for genomic testing remains. To date, over 250,000 females have been genotyped, and phenotypic information from over 100,000 of these animals has been recorded (Nielsen et al., 2019).

Approximately 85% of farms in DFS are enrolled in the national milk recording schemes. This enables validation of GEBV with phenotype data on a large scale, with a design having the desirable property that the validation population is strictly independent of the training population. The purpose of this study was to compare the abilities of virgin heifer GEBV and PA to predict future cow performance. To our knowledge, this has not previously been done on a large scale in 3 breeds across countries. This could be an important step to convince farmers that genomic breeding values are valuable for use on their herds for selection decisions.

MATERIALS AND METHODS

Data

Phenotype data were collected from the DFS milk recording schemes for the 3 breeds (RDC, Holstein, and Jersey). Observations from the first lactation of animals born from 2013 to 2015 were used in the analysis. To be included in the study, all animals were required to have a 305-d milk yield record. The total numbers of genotyped females in the study period with a 305-d milk yield record were 20,274 RDC, 23,910 Holstein, and 9,312 Jersey. We analyzed 12 traits in first parity: 3 milk production traits (milk yield, fat yield, and protein yield), 2 udder health traits (SCS and occurrence of clinical mastitis), 1 fertility trait (interval, in days, from first to last service, **IFL**), 2 conformation traits (udder, and feet and legs), 1 calving trait (calving ease, **CE**), 1 survival trait (survival to second calving, **survival 1–2**), 1 claw health (**CH**) trait, and 1 general health (**GH**) trait. Detailed trait definitions can be found in Table 1. For Jersey, it was not possible to analyze CH or GH, because for those traits the genomic evaluation was still under development during the study period.

Female GEBV and PA were obtained from the Nordic Cattle Genetic Evaluation (NAV, 2019). Detailed descriptions of all breeding values can be found in Table 2. Heritability in first lactation of traits used in the Nordic Cattle Genetic Evaluation can be found in Table 3, and in Table 4 average model reliabilities

Table 1. Detailed definitions of the traits studied

Trait	Phenotype definition (first lactation)
Milk yield	305-d kg of milk yield
Fat yield	305-d kg of fat yield
Protein yield	305-d kg of protein yield
SCS	SCC transformed to logarithmic scale
Clinical mastitis[1]	Clinical mastitis up to 300 d
IFL	Interval in days from first to last service
Udder	Total udder conformation score
Feet and legs	Total feet and legs conformation score
Calving ease (maternal)	First calving, recorded in 4 categories: (1) easy calving without help, (2) easy calving with help, (3) difficult calving without veterinarian help, and (4) difficult calving with veterinarian help
Claw health[1]	Records from first to second calving or up to 430 d after calving in first lactation. Claw disorders included were sole ulcer, sole hemorrhage, heel horn erosion, digital dermatitis, interdigital dermatitis, verrucose dermatitis, interdigital hyperplasia, double sole, white line separation, and corkscrew claw
General health[1]	Includes retained placenta, hormonal reproductive disorders, infective reproductive disorders, ketosis, milk fever, other metabolic diseases, other feed-related disorders, other diseases, and feet and leg problems
Survival 1–2[2]	Survival from first to second calving

[1]Defined as 1 if the animal had at least one treatment, 0 otherwise.
[2]Defined as 1 if the animal survived, 0 otherwise.

Table 2. Detailed definitions of breeding values (NAV, 2019)

Breeding value	Breeding value definition
Milk	Milk production in the first 3 lactations
Fat	Fat production in the first 3 lactations
Protein	Protein production in the first 3 lactations
Udder health	Based on records of clinical mastitis and SCC in the first 3 lactations and udder depth from first lactation. SCC and udder conformation are used as indicator traits
Fertility	Based on number of services, interval from calving to first service, interval from first to last service, non-return rate, heat strength, and conception rate. Includes records as virgin heifer to the third lactation
Udder	Linear traits combined into a group describing udder conformation. Based on the linear traits udder attachment, rear udder height, rear udder width, udder cleft/support, udder depth, teat length, teat thickness, teat placement (front), teat placement (back), and udder balance. Based on data from the first 3 lactations
Feet and legs	Linear traits combined into a group describing feet and leg conformation. Includes the linear conformation traits rear legs (side view), rear legs (rear view), hock quality, bone quality, and foot angle. Based on data from the first 3 lactations
Calving (maternal)	Including calving ease and calf survival in the first 24 h. Calving is recorded in 4 categories, as for the phenotype trait (Table 1). Calf survival is defined as 1 if the calf survived, 0 otherwise. Includes records from first to fifth calving
Claw health	Includes records from the first 3 lactations. Claw disorders included were as defined in the phenotype definition (Table 1)
General health	Genetic resistance to reproductive, digestive, and feet and leg problems. Includes the same records as the phenotype definition (Table 1). Based on data from the first 3 lactations
Longevity	Describes the genetic ability to survive. Including days from first to the fifth lactation, with a maximum of 365 d per lactation

for genotyped animals can be found. Breeding values from 36 evaluations performed between August 2014 and February 2017 were used in this study. The GEBV and PA used were based on the breeding values estimated closest in time to when the animal reached 1 yr of age. Eleven different GEBV and PA were used: milk, fat, protein, udder health, fertility, udder, feet and leg, calving (maternal), claw health, general health, and longevity. These breeding values correspond to the phenotypes listed in Table 1 but are not defined in exactly the same way. Breeding values were based on multiple lactations, whereas phenotypes were from the first lactation only. The breeding values studied are also combinations of several underlying component traits (e.g., the fertility breeding value also includes information on the interval in days between calving and first service). We calculated GEBV using bivariate blending of direct genomic values and traditional EBV (Mäntysaari and Strandén, 2010; Taskinen et al., 2013). In September 2015, the calculation of direct genomic breeding values changed from GBLUP to SNPBLUP (Nielsen et al., 2016), which was shown to give comparable results (Koivula et al., 2012). Detailed breeding value calculations can be found in NAV (2019).

To prevent virgin heifer reproductive performance from influencing the fertility breeding values used in this study, breeding values estimated after 14 mo of age were not included. For the same reason, animals genotyped after 14 mo were excluded. At the beginning of the VikingGenetics genotyping project, it was common to genotype animals up to the second lactation.

Table 3. Heritability in first lactation of traits used in the Nordic cattle genetic evaluation (NAV, 2019)[1]

Trait	RDC	Holstein	Jersey
Milk	0.41	0.43	0.44
Fat	0.35	0.36	0.35
Protein	0.41	0.35	0.38
SCC	0.12	0.13	0.11
Clinical mastitis	0.04	0.05	0.04
IFL	0.03	0.03	0.03
Udder	0.25	0.25	0.25
Feet and legs	0.20	0.20	0.20
Calving ease	0.04	0.06	0.02
Claw health[2]	0.001–0.040	0.004–0.070	0.000–0.070
General health[2]	0.003–0.01	0.004–0.034	0.004–0.013
Survival 1–2[3]	0.04	0.05	0.05

[1]RDC = Red Dairy Cattle; IFL = interval from first to last service, in days.
[2]The interval represents the range of heritability for the included subtraits (Table 1).
[3]Heritability for survival from first to second calving.

Table 4. Average model reliabilities (%) published for genotyped animals, 1 to 2 yr old, born in 2017 [Gert Pedersen Aamand, Executive Director, Nordic Cattle Genetic Evaluation (NAV, Aarhus, Denmark), personal communication, June 26, 2019]; RDC = Red Dairy Cattle

Breeding value	RDC	Holstein	Jersey
Yield	74	77	71
Udder health	66	74	63
Fertility	59	74	55
Udder	66	73	64
Feet and legs	66	66	57
Calving	54	68	43
Claw health	51	59	46
General health	50	58	45
Longevity	49	66	44

Consequently, most of the animals removed were born in 2013, which was the first year analyzed in this study. However, in the last 2 years studied, 2014 and 2015, most animals were genotyped as virgin heifers and were therefore included in the study. The number of genotyped animals also increased over the study period. Hence, most of the animals studied were born in 2014 and 2015. The numbers of genotyped animals in the birth year cohort studied (2013 to 2015), genotyped before 14 mo of age and with a 305-d milk yield record, were 14,862 RDC from 900 herds, 17,145 Holstein from 1,960 herds, and 7,330 Jersey from 235 herds.

Statistical Analysis

To obtain adjusted phenotypes for use in analysis of the predictive ability of breeding values, a larger phenotype data set was analyzed using Statistical Analysis Software (SAS) version 9.4 (SAS Institute Inc., Cary, NC). This analysis included all animals in the milk recording scheme born from 2008 to 2016, which, in total, comprised 997,797 RDC, 2,322,514 Holstein, and 240,946 Jersey. The adjusted phenotypes—that is, residual effects estimated using PROC HP MIXED with the linear model [1] described below—were named according to the respective trait; for example, adjusted milk yield was named $Milk_{Adj}$. A separate analysis was performed for each breed.

The following linear model was used for all traits:

$$y_{ijklmn} = \mu + HY_{ij} + YMC_{jkl} + CCA_{lm} + e_{ijklmn}, \quad [1]$$

where y_{ijklmn} is the observed phenotypic value in first lactation; μ is mean of the population; HY_{ij} is the fixed class effect of herd i and calving year j (2008 through 2018); YMC_{jkl} is the fixed class effect of calving year j, month k (1 through 12), and country l (Denmark, Finland, or Sweden); CCA_{lm} is the fixed class effect of country l and calving age in months as heifer m (18 to 36); and e_{ijklmn} is the random residual, $\sim ND\left(0, \sigma_e^2\right)$.

We used $HY_{ij} + YMC_{jkl}$ as contemporary groups due to small average herd size, making it difficult to use herd-year-month or herd-year-season. Country was not included in the model for Jersey, because all Jersey cows studied were located in Denmark. For further analyses, the PROC MEANS and PROC FREQ procedures in SAS were used for descriptive statistics.

Because breeding values were obtained from several routine evaluations separated in time, they were not directly comparable due to rolling base population. Linear regression analysis was used to adjust for genetic trends over time. In PROC REG, the regression coefficient was estimated between breeding values in a given evaluation and the corresponding breeding values in the last evaluation (February 2017). The linear regression model used was

$$y_i = b_0 + b_1 X + e_{ij}, \quad [2]$$

where y_i is a breeding value in the last evaluation (February 2017); b_0 is the intercept; b_1 is the regression coefficient on the corresponding breeding value (X) in a breeding evaluation performed from August 2014 to January 2017; and e_{ij} is the random residual, $\sim ND\left(0, \sigma_e^2\right)$. Breeding values were then expressed on the scale of the last evaluation, using the estimated regression parameters from Model [2] using PROC SCORE.

PROC CORR was used to calculate the correlation between breeding values (PA or GEBV) and adjusted phenotypes for each of the breeds. A 95% confidence interval using Fisher's Z transformation was used to assess the significance of the difference between correlations. The PROC RANKS procedure was used to rank cows into 4 quartiles across herds for GEBV or PA.

RESULTS

For RDC and Holstein, all correlations between breeding values and adjusted phenotypes were significantly stronger for GEBV than for PA (Table 5). For Jersey, GEBV correlations were significantly stronger for all traits except clinical mastitis, CE, and survival 1–2. The correlations with adjusted phenotypes were 42 to 194% higher for GEBV than for PA in RDC, 38 to 136% higher for GEBV in Holstein, and 11 to 78% higher for GEBV in Jersey for the different traits (Table 5). All correlations between PA and adjusted phenotypes were significantly different from zero. The highest correlation found in this study was between milk GEBV and $Milk_{Adj}$ for Jersey (0.51). One of the traits for which the correlations increased the most, IFL_{Adj} increased by over 64% for all 3 breeds when genomic information was included in the breeding values. For Jersey, the correlation between breeding value and IFL_{Adj} increased by 78% when genomic information was included.

The relative change between the GEBV bottom and top quartile (**ΔGEBV**) compared with that between the PA bottom and top quartile (**ΔPA**), ranged from 9 to 261% for RDC, 42 to 138% for Holstein, and 4 to 90% for Jersey (Table 6). However, it should be noted that the large relative percentage change between PA and GEBV for both quartiles and correlations was, in many cases, from initially low levels.

Table 5. Estimated correlations and relative change, in percent, between cow adjusted (Adj) phenotypes and respective virgin heifer parent average breeding values (PA) and genomically enhanced breeding values (GEBV) for Red Dairy Cattle (RDC), Holstein, and Jersey; N = number of cows with a phenotype; n/a = not applicable

	RDC				Holstein				Jersey			
		Correlation				Correlation				Correlation		
Trait (units)	N	PA[1]	GEBV[1]	Relative change[2] (%)	N	PA[1]	GEBV[1]	Relative change[2] (%)	N	PA[1]	GEBV[1]	Relative change[2] (%)
$Milk_{Adj}$	14,710	0.25	0.40*	63	17,039	0.31	0.45*	49	7,069	0.32	0.51*	59
Fat_{Adj}	14,571	0.21	0.31*	46	16,801	0.25	0.36*	46	7,048	0.24	0.40*	60
$Protein_{Adj}$	14,583	0.23	0.33*	52	16,902	0.26	0.38*	44	7,060	0.23	0.37*	63
SCS_{Adj}	12,834	−0.10	−0.16*	63	16,667	−0.13	−0.21*	65	6,734	−0.10	−0.18*	67
$Clinical\ mastitis_{Adj}$	12,834	−0.04	−0.06*	62	14,463	−0.06	−0.09*	38	6,447	−0.09	−0.10	13
IFL_{Adj}	14,549	−0.05	−0.09*	68	16,833	−0.06	−0.11*	64	6,451	−0.04	−0.08*	78
$Udder_{Adj}$	11,917	0.20	0.29*	42	13,412	0.23	0.37*	61	6,192	0.15	0.26*	71
$Feet\ and\ legs_{Adj}$	11,917	0.16	0.29*	73	13,413	0.14	0.20*	56	6,192	0.20	0.26*	32
$Calving\ ease_{Adj}$	11,521	−0.04	−0.07*	77	16,891	−0.05	−0.07*	88	6,691	−0.05	−0.07	37
$Claw\ health_{Adj}$	4,129	−0.06	−0.10*	74	4,829	−0.06	−0.12*	91	n/a	n/a	n/a	n/a
$General\ health_{Adj}$	13,885	−0.02	−0.05*	194	15,748	−0.03	−0.06*	94	n/a	n/a	n/a	n/a
$Survival\ 1–2_{Adj}$	14,694	0.03	0.05*	71	17,029	0.03	0.06*	136	7,053	0.03	0.03	11

[1]Negative correlations are desirable for SCS, mastitis, interval in days from first to last service (IFL), calving ease, claw health, and general health.

[2]Relative change = $\frac{\text{Correlation with GEBV} - \text{Correlation with PA}}{\text{Correlation with PA}}$, relative change in percent between GEBV correlation and PA correlation.

*Significant difference between GEBV correlation and PA correlation ($P < 0.05$).

DISCUSSION

An advantage of using phenotypes for validation of breeding values is that it makes it easy for farmers to understand how their animals' breeding values work in practice, when validation is against their own farm records. The use of phenotypes in this study was facilitated by the high rate of participation in the national milk recording schemes in the DFS countries. The extracted virgin heifer GEBV and PA, estimated before on-farm information was recorded, reflected information available to farmers at the time of selection. The maximum age at which a breeding value for a heifer was taken was 14 mo, to reflect the breeding values at first insemination for virgin heifers. For example, at that age the farmer can combine genomic selection with decisions about sexed and beef semen, as suggested in other studies (Hjortø et al., 2015; Ettema et al., 2017).

Many of the breeding values used in the present study are combinations of several underlying component traits. For example, the udder health breeding value includes data on clinical mastitis, SCC, udder attachment, and udder depth. Furthermore, most of the breeding values are based on the first 3 lactations and not only the first lactation, whereas the phenotypes studied were only from the first lactation. These 2 factors most likely resulted in somewhat weaker relationships between breeding values and phenotypes than if sub-trait breeding values for the first lactation had been used. However, those breeding values were not available for this study nor for the farmers in the studied period. Nevertheless, these factors probably had a limited influence on the relative change between GEBV and PA, which was the focus in this study.

We chose to use linear models for all traits, to represent current practice in the Nordic genetic evaluation. However, some traits could be claimed to be theoretically less well suited for a linear model, such as clinical mastitis, survival 1–2, and CE. Therefore, we tried different models to fit the data (binary distribution, Poisson distribution) in preliminary analyses for the Jersey breed, but the results were similar to those obtained using linear models.

The highest correlations obtained in this study were between production traits and breeding values. This could be expected, as production traits have the highest heritability and reliability of the traits studied (Tables 3 and 4). Mathematically, the correlation between the true breeding value and phenotype is equal to h, and the proportion of variance explained by the breeding value is h^2. However, we did not have the true breeding values in this study, and therefore the expected (squared) correlation equals the product of heritability and reliability.

Table 6. Differences in averages of adjusted (Adj) phenotypes between cows in bottom and top quartiles for virgin heifer parent average breeding values (ΔPA), genomically enhanced breeding values (ΔGEBV), and relative change in percent, respectively, for Red Dairy Cattle (RDC), Holstein, and Jersey; for trait definitions, see Table 1; n/a = not applicable

	RDC			Holstein			Jersey		
Trait[1]	ΔPA	ΔGEBV	Relative change[2] (%)	ΔPA	ΔGEBV	Relative change[2] (%)	ΔPA	ΔGEBV	Relative change[2] (%)
$Milk_{Adj}$ (kg)	708	1,171	65	1,061	1,512	42	738	1,156	57
Fat_{Adj} (kg)	28	41	45	33	48	47	28	42	54
$Protein_{Adj}$ (kg)	20	31	57	27	38	44	22	34	55
SCS_{Adj}	0.1	0.15	55	0.11	0.19	71	0.09	0.16	71
$Mastitis_{Adj}$ (score 0 or 1)	0.02	0.03	19	0.02	0.04	52	0.08	0.08	4
IFL_{Adj} (d)	7.7	12.3	60	9.1	16.1	76	7.3	12.2	68
$Udder_{Adj}$ (points)	2.5	3.5	39	2.7	4.4	62	1.7	3.3	90
Feet and $legs_{Adj}$ (points)	2	3.6	78	1.5	2.5	66	2.3	3.1	35
Calving $ease_{Adj}$ (score 1–4)	0.04	0.08	97	0.07	0.12	71	0.04	0.06	51
Claw $health_{Adj}$ (0 or 1)	0.05	0.11	111	0.07	0.14	99	n/a	n/a	n/a
General $health_{Adj}$ (0 or 1)	0.01	0.03	261	0.03	0.05	76	n/a	n/a	n/a
Survival $1–2_{Adj}$ (0 or 1)	0.03	0.03	9	0.02	0.05	138	0.04	0.04	17

[1]IFL = interval from first to last service, in days.

[2] Relative change = $\frac{\Delta GEBV - \Delta PA}{\Delta PA}$, relative change in percent between GEBV bottom and top quartile compared with PA bottom and top quartile.

The correlations with adjusted phenotypes were over 40% stronger for all production traits and breeds when genomic information was used compared with PA. The highest correlation found in this study was between milk GEBV and $Milk_{Adj}$ for Jersey. The reliability of the yield breeding values differed least from each other (Table 4), where the heritability for RDC was slightly lower than for Holstein and Jersey (Table 3). The Nordic RDC is the most genetically diverse of the 3 breeds studied, as it is a mixture of Swedish Red, Danish Red, and Finnish Ayrshire and also includes genes from Norwegian Red, Canadian Ayrshire, American Brown Swiss, and Red Holstein Friesian (NAV, 2019). Hence, less linkage disequilibrium between markers and quantitative trait loci could explain the lower correlations for RDC. The difference between top and bottom quartiles in adjusted phenotypes when using GEBV instead of PA (ΔGEBV − ΔPA; Table 6) for $Milk_{Adj}$ was lower than that reported by Weigel et al. (2015). For Holstein, the difference in our study was +450 kg with genomic information, compared with +1,104 kg in Weigel et al. (2015). However, only sire PTA (rather than PA) values were used in their study, and the results were only from 411 cows. Additionally, differences in production level and phenotypic variance most likely occurred between our study and that of Weigel et al. (2015).

In general, traits with low heritability in the present study, such as IFL, clinical mastitis, CE, CH, and GH, gained relatively more in accuracy from using genomic information than did highly heritable traits such as production. The same pattern has been reported by other studies (García-Ruiz et al., 2016; Wiggans et al., 2017).

In our study, IFL_{Adj} was one of the traits for which correlations increased the most when genomic information was included in the breeding values (over 60% for all 3 breeds). It has been established that IFL has the strongest correlation with fertility breeding value (NAV, 2019). For Jersey, the correlation between breeding value and IFL_{Adj} increased by 78% when genomic information was included. Looking at the quartiles for IFL, the difference between ΔGEBV and ΔPA was 4.6 to 7.0 d in favor of GEBV (Table 6). Consequently, virgin heifer GEBV was more effective than PA in identifying cows with poor and good fertility.

Our results also confirmed that GEBV can help in choosing animals with better udder health. The correlation between SCS_{Adj} and GEBV increased by over 55% compared with SCS_{Adj} and PA for all 3 breeds. The udder health trait with the highest heritability is SCS (Table 3), and one could expect a stronger correlation compared with clinical mastitis. Weigel et al. (2015) found that SCS showed almost no difference between quartiles for PA, even though their study had greater differences between quartiles for genomic values. In the present study, we also found significant differences for correlations between Clinical $Mastitis_{Adj}$ and udder health breeding values for Holstein and RDC but not for Jersey. However, the Jersey correlation between PA udder health and $Mastitis_{Adj}$ was relatively strong, which indicates that the conventional evaluation works well for this trait, possibly owing to higher clinical mastitis frequency among Jersey cows than among

Holstein and RDC cows (Appendix Tables A1 and A2). In the genotyped data set, the clinical mastitis frequency for Jersey was 17%, compared with 7% for Holstein and 6% for RDC. This was also reflected in the differences between quartiles, where both ΔPA and ΔGEBV differences were larger for Jersey. However, we found no differences between PA and GEBV in their ability to predict the future adjusted phenotype for Jersey (Table 6). The low clinical mastitis frequencies for RDC and Holstein made it more difficult to detect differences between GEBV and PA. The correlations between breeding values for udder health and clinical mastitis are stronger in the second and third lactations (NAV, 2019). It would have been interesting to study the second and third lactations, but for most animals these had not been completed at the time of this study.

For RDC and Holstein, we discovered significantly stronger correlations between calving GEBV and CE_{Adj} than between calving PA and CE_{Adj}. However, with fewer genotyped heifers and some animals lacking CE information, it was not possible to draw a similar conclusion for Jersey. Analysis also revealed fewer calving problems for Jersey, for which the score was on average 1.06 in the genotyped group, compared with 1.24 for RDC and 1.23 for Holstein (Appendix Tables A1 and A2). Further, heritability and GEBV reliability were also lower for the calving trait in Jersey compared with Holstein and RDC (Tables 3 and 4), which might explain why it was not possible to detect significant differences between GEBV correlation and PA correlation. The difference between quartiles was also smaller for Jersey (+0.02) compared with RDC (+0.04) and Holstein (+0.05) when genomic information was included in the breeding value (Table 6).

Phenotypes for both conformation traits were significantly more strongly correlated with conformation GEBV than with conformation PA. On examining the difference between conformation quartiles (Table 6), it was also possible to see that the prediction improved when genomic information was included. The difference between the top and bottom 25% (ΔGEBV − ΔPA) increased by between 0.8 and 1.6 scoring points for feet and leg conformation, and between 1.0 and 1.7 scoring points for udder conformation, when genomic information was used (Table 6). Thus, GEBV can be more effective than PA in predicting future conformation.

We also disovered a lack of phenotypes for the CH trait (Table 5), which might have affected the results for that trait. Nevertheless, for both RDC and Holstein, the correlations between CH GEBV and CH_{Adj} were significantly stronger than the correlations between CH PA and CH_{Adj}. For Jersey, it was not possible to compare GEBV and PA regarding their ability to predict future cow claw health, because the genomic evaluation for CH in that breed was only established in 2018.

For both Holstein and RDC, the correlations between general health GEBV and GH_{Adj} were significantly stronger than the correlations between general health PA and GH_{Adj}. The correlations between GH breeding value and GH_{Adj} increased by 99% for Holstein and 194% for RDC when using genomic selection. From the quartile differences (+0.02) for both Holstein and RDC, the benefit of using genomic selection was not equally clear. The heritability of the GH trait is low, and the trait is influenced by the environment to a large extent (Table 3). For GH, the genomic evaluation for Jersey was under development at the time of the study.

For RDC and Holstein, significantly stronger correlations occurred between longevity GEBV and Survival $1–2_{Adj}$ than between longevity PA and Survival $1–2_{Adj}$. For Jersey, we found no differences between the abilities of GEBV and PA to predict future survival performance. Further, looking at differences between the quartiles (ΔGEBV − ΔPA) for Holstein (+0.03), Jersey (±0) and RDC (±0), it was not possible to see the benefit of genomic selection for RDC and Jersey (Table 6). The reliability of survival GEBV was also lower for Jersey than for Holstein and RDC (Table 4). Survival is strongly affected by farmer decisions, the environment, and other functional and health traits (Kargo et al., 2014), which also could explain the results. It would have been interesting to study survival in later lactations. The longevity breeding value includes data to the end of the fifth lactation, and the correlations between longevity breeding value and survival are stronger in later lactations (NAV, 2019).

CONCLUSIONS

We showed that virgin heifer GEBV predicted cow performance significantly better than did PA for the vast majority of analyzed traits in Red Dairy Cattle, Jersey, and Holstein. Thus, farmers in Denmark, Finland, and Sweden can have confidence in using genomic technology on their herds for selection decisions. Traits with low heritability, such as interval from first to last insemination, clinical mastitis, calving ease, claw health, and general health, gained relatively more from inclusion of genomic information than did highly heritable traits such as production.

ACKNOWLEDGMENTS

This study was funded by the Swedish project LivsID—Externally employed doctoral students within food-related research. We offer special thanks to Faba

(Hollola, Finland), Nordic Cattle Genetic Evaluation (Aarhus, Denmark), SEGES (Aarhus, Denmark), and Växa Sverige (Uppsala, Sweden) for providing us with data. The authors have not stated any conflicts of interest.

REFERENCES

Calus, M. P. L., P. Bijma, and R. F. Veerkamp. 2015. Evaluation of genomic selection for replacement strategies using selection index theory. J. Dairy Sci. 98:6499–6509. https://doi.org/10.3168/jds.2014-9192.

de Roos, A. P. W., B. J. Hayes, and M. E. Goddard. 2009. Reliability of genomic predictions across multiple populations. Genetics 183:1545–1553. https://doi.org/10.1534/genetics.109.104935.

Ettema, J. F., J. R. Thomasen, L. Hjortø, M. Kargo, S. Østergaard, and A. C. Sørensen. 2017. Economic opportunities for using sexed semen and semen of beef bulls in dairy herds. J. Dairy Sci. 100:4161–4171. https://doi.org/10.3168/jds.2016-11333.

García-Ruiz, A., J. B. Cole, P. M. VanRaden, G. R. Wiggans, F. J. Ruiz-López, and C. P. Van Tassell. 2016. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. Proc. Natl. Acad. Sci. 113:E3995–E4004. https://doi.org/10.1073/pnas.1519061113

Hjortø, L., J. F. Ettema, M. Kargo, and A. C. Sørensen. 2015. Genomic testing interacts with reproductive surplus in reducing genetic lag and increasing economic net return. J. Dairy Sci. 98:646–658. https://doi.org/10.3168/jds.2014-8401.

Kargo, M., L. Hjortø, M. Toivonen, J. A. Eriksson, G. P. Aamand, and J. Pedersen. 2014. Economic basis for the Nordic Total Merit Index. J. Dairy Sci. 97:7879–7888. https://doi.org/10.3168/jds.2013-7694.

Koivula, M., I. Strandén, G. Su, and E. A. Mäntysaari. 2012. Different methods to calculate genomic predictions—Comparisons of BLUP at the single nucleotide polymorphism level (SNP-BLUP), BLUP at the individual level (G-BLUP), and the one-step approach (H-BLUP). J. Dairy Sci. 95:4065–4073. https://doi.org/10.3168/jds.2011-4874.

Kolmodin, R., E. Strandberg, P. Madsen, J. Jensen, and H. Jorjani. 2002. Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. Acta Agric. Scand. Sect. A—Anim. Sci. 52:11–24. https://doi.org/10.1080/09064700252806380.

Mäntysaari, E. A., and I. Strandén. 2010. Use of bivariate EBV-DGV model to combine genomic and conventional breeding value evaluations. Proc. 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Germany. http://www.nordicebv.info/wp-content/uploads/2015/04/Use-of-bivariate-EBV-DGV-model-to-combine-genomic.pdf

NAV (Nordic Cattle Genetic Evaluation). 2019. NAV routine evaluation 5. Accessed Jan. 14, 2020. https://www.nordicebv.info/wp-content/uploads/2019/06/NAV-routine-genetic-evaluation-04062019.pdf.

Nielsen, U. S., G. P. Pedersen, J. Pösö, and L. P. Sørensen. 2016. Status and plans—Genomic prediction and traditional evaluation. NAV Workshop 2016. Accessed Jan. 14, 2020. http://nordicebv.info/wp-content/uploads/2016/01/Workshop_2016-Genomic-prediction-and-traditional-evaluation-USN-GAP.pdf.

Nielsen, U. S., G. P. Pedersen, and L. P. Sørensen. 2019. Status and plans—Genomic prediction. NAV Workshop 2019. Accessed Jan. 14, 2020. https://www.nordicebv.info/wp-content/uploads/2019/01/Genomic-prediction-USN-NAV-workshop-jan-2019.pdf.

Pryce, J., and B. Hayes. 2012. A review of how dairy farmers can use and profit from genomic technologies. Anim. Prod. Sci. 52:180. https://doi.org/10.1071/AN11172.

Pryce, J. E., B. J. Hayes, and M. E. Goddard. 2012. Genotyping dairy females can improve the reliability of genomic selection for young bulls and heifers and provide farmers with new management tools. In Proc. 38th International Committee for Animal Recording (ICAR) Session, May 28–June 1, 2012, Cork, Ireland. http://www.icar.org/Cork_2012/Manuscripts/Published/Pryce%202.pdf.

Su, G., B. Guldbrandtsen, V. R. Gregersen, and M. S. Lund. 2010. Preliminary investigation on reliability of genomic estimated breeding values in the Danish Holstein population. J. Dairy Sci. 93:1175–1183. https://doi.org/10.3168/jds.2009-2192.

Taskinen, M., E. A. Mäntysaari, M. H. Lidauer, T. Knürr, J. Pösö, G. Su, G. P. Aamand, and I. Strandén. 2013. Comparison of model reliabilities from single-step and bivariate blending methods. Interbull Bull. 6:246–251. https://jukuri.luke.fi/bitstream/handle/10024/481814/1331-2230-1-PB.pdf?sequence=1.

Weigel, K. A., A. A. Mikshowsky, and V. E. Cabrera. 2015. Effective use of genomics in sire selection and replacement heifer management. In Proc. Western Dairy Management Conference, Reno, NV. http://www.wdmc.org/2015/Weigel.pdf.

Wiggans, G. R., J. B. Cole, S. M. Hubbard, and T. S. Sonstegard. 2017. Genomic selection in dairy cattle: The USDA experience. Annu. Rev. Anim. Biosci. 5:309–327. https://doi.org/10.1146/annurev-animal-021815-111422.

Yao, C., L. E. Armentano, M. J. VandeHaar, and K. A. Weigel. 2015. Short communication: Use of single nucleotide polymorphism genotypes and health history to predict future phenotypes for milk production, dry matter intake, body weight, and residual feed intake in dairy cattle. J. Dairy Sci. 98:2027–2032. https://doi.org/10.3168/jds.2014-8707.

ORCIDS

C. Bengtsson https://orcid.org/0000-0001-5784-2557
H. Stålhammar https://orcid.org/0000-0002-7354-651X
E. Strandberg https://orcid.org/0000-0001-5154-8146
S. Eriksson https://orcid.org/0000-0003-3357-5065
W. F. Fikse https://orcid.org/0000-0003-0578-0184

APPENDIX

Table A1. First-lactation descriptive statistics [average, SD, and number of animals (N)] for all animals born in 2013, 2014, and 2015 in Denmark, Finland, and Sweden

	Red Dairy Cattle			Holstein			Jersey		
Trait[1]	Mean	SD	N	Mean	SD	N	Mean	SD	N
Milk yield (kg)	8,022	1,432	217,245	8,984	1,660	601,353	6,263	1,126	61,105
Fat yield (kg)	355	63	210,289	363	63	592,887	371	61	61,100
Protein yield (kg)	288	50	210,295	308	53	592,914	259	43	61,105
SCS	0.68	0.42	200,178	0.67	0.37	553,293	0.8	0.37	57,807
Clinical mastitis (score 0 or 1)	0.06	0.24	182,038	0.09	0.29	495,373	0.16	0.37	52,458
IFL (d)	45.86	61.72	177,969	44.36	61.7	491,602	43.17	61.42	53,503
Udder (points)	79.79	5.34	96,273	80.06	5.28	244,125	80	5.53	38,496
Feet and legs (points)	80.02	5.34	96,288	80.18	4.99	244,144	80.07	5.5	38,496
Calving ease (maternal; score 1–4)	1.24	0.52	139,956	1.23	0.5	471,599	1.06	0.32	55,967
Claw health (0 or 1)	0.51	0.5	45,466	0.64	0.48	144,002	0.51	0.5	13,225
General health (0 or 1)	0.12	0.32	185,577	0.16	0.36	519,512	0.15	0.36	54,258
Survival 1–2 (0 or 1)	0.66	0.47	212,226	0.69	0.46	588,990	0.73	0.44	60,921

[1]IFL = interval from first to last service.

Table A2. First-lactation descriptive statistics [average, SD, and number of animals (N)] for all animals genotyped and qualified for analysis born in 2013, 2014, and 2015 in Denmark, Finland, and Sweden

	Red Dairy Cattle			Holstein			Jersey		
Trait[1]	Mean	SD	N	Mean	SD	N	Mean	SD	N
Milk yield (kg)	8,473	1,312	14,710	9,452	1,579	17,039	6,451	1,024	7,069
Fat yield (kg)	374	56	14,571	383	58	16,801	384	55	7,048
Protein yield (kg)	306	45	14,583	329	51	16,902	269	40	7,060
SCS	0.64	0.40	12,834	0.61	0.35	16,667	0.76	0.36	6,734
Clinical mastitis (score 0 or 1)	0.06	0.24	12,834	0.07	0.27	14,463	0.17	0.36	6,447
IFL (d)	42.03	58.53	14,549	39.87	57.92	16,833	40.36	59.85	6,451
Udder (points)	80.33	5.20	11,917	81.46	4.85	13,412	80.72	5.25	6,192
Feet and legs (points)	80.12	5.26	11,917	81.06	4.71	13,413	80.80	5.21	6,192
Calving ease (maternal; score 1–4)	1.22	0.5	11,521	1.23	0.50	16,891	1.07	0.33	6,691
Claw health (0 or 1)	0.54	0.50	4,129	0.59	0.49	4,829	0.55	0.49	1,096
General health (0 or 1)	0.15	0.35	13,885	0.14	0.35	15,748	0.15	0.35	6,626
Survival 1–2 (0 or 1)	0.69	0.46	14,694	0.71	0.45	17,029	0.75	0.43	7,053

[1]IFL = interval from first to last service.

II

J. Dairy Sci. 105:1281–1297
https://doi.org/10.3168/jds.2021-20849

Mating allocations in Nordic Red Dairy Cattle using genomic information

C. Bengtsson,[1,2]* H. Stålhammar,[1] J. R. Thomasen,[1] S. Eriksson,[2] W. F. Fikse,[3] and E. Strandberg[2]
[1]VikingGenetics, VikingGenetics Sweden AB, 53294 Skara, Sweden
[2]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Box 7023, 75007 Uppsala, Sweden
[3]Växa Sverige, Växa Sverige, Box 288, 75105 Uppsala, Sweden

ABSTRACT

In this study, we compared mating allocations in Nordic Red Dairy Cattle using genomic information. We used linear programming to optimize different economic scores within each herd, considering genetic level, semen cost, the economic impact of recessive genetic defects, and genetic relationships. We selected 9,841 genotyped females born in Denmark, Finland, or Sweden in 2019 for mating allocations. We used 2 different pedigree relationship coefficients, the first tracing the pedigree 3 generations back from the parents of the potential mating and the second based on all available pedigree information. We used 3 different genomic relationship coefficients, 1 SNP-by-SNP genomic relationship and 2 based on shared genomic segments. We found high correlations ($\geq$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 measures, but it was best done with genomic measures. Linear programming maximized 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.
Key words: genomic relationships, pedigree relationships, mating program, linear programming

INTRODUCTION

Mating programs are an important support tool for livestock breeders, helping them to identify the best parental matings to maximize 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 markers can give information about major genes and genetic defects. Minimizing the probability of obtaining offspring homozygous for a lethal recessive genetic defect is of economic importance for farmers (Pryce et al., 2012). Further, the EFFAB (European Forum of Farm Animal Breeders, Brussels, Belgium) code of good practice states that breeding organizations should improve health and welfare by reducing the incidence of genetic defects (EFFAB, 2020).

SNP markers also offer the possibility to reduce genomic relationships between parents when making mating plans. Various methods have been proposed for calculation of genomic relationships, including SNP-by-SNP relationships as described by, for example, VanRaden (2008). Further, methods using shared genomic segments, as described by, for example, de Cara et al. (2013), aim to reduce the number of runs of homozygosity (**ROH**) in the offspring. Genomic estimates of relationships are suggested to be more accurate than pedigree information because they do not rely on pedigree completeness or correctness (Pryce et al., 2012; Sun et al., 2013; Carthy et al., 2019) and also because pedigree relationships incorrectly assume infinite, unlinked loci (Hill and Weir, 2011). Furthermore, genomic estimates of relationships can differentiate between animals with the same pedigree relationship that have inherited partly different genetic variants from their parents.

At population level, various genomic relationships have been compared previously with pedigree measures using optimum contribution selection (**OCS**; Sonesson et al., 2012; Henryon et al., 2019; Meuwissen et al., 2020). Sonesson et al. (2012) concluded that genomic selection needs genomic control of inbreeding. In contrast, using pedigree relationships in OCS, rather than genomic relationships, has been shown to achieve more true genetic gain in the long term (Henryon et al., 2019). Further, Meuwissen et al. (2020) illustrated

Received June 10, 2021.
Accepted September 8, 2021.
*Corresponding author: christian.bengtsson@slu.se or chben@vikinggenetics.com

that different relationship matrices are preferred when aiming for maintain heterozygosity or when controlling genetic drift, where the latter prevents genetic defects from drifting to high frequencies and random drift of functional traits.

Several known recessive genetic defects in Nordic Red Dairy Cattle (**RDC**) are included in the SNP chip currently used for genotyping, and additional genetic defects are included as they are detected. At the beginning of 2020, the carrier status of 6 genetic defects in RDC was automatically provided with the genomic test. Besides reducing genetic relationships, other relevant information (e.g., genetic level, semen cost, the economic impact of recessive genetic defects) has to be considered when making mating plans. An economic score for each potential mating, which combines and weighs all economically relevant information, has been proposed (Pryce et al., 2012; Carthy et al., 2019; Bérodier et al., 2021). Using linear programming to maximize every herd's mean economic score, subject to necessary constraints, is a fast and effective method (Carthy et al., 2019; Bérodier et al., 2021). Further, linear programming has been shown to outperform other mating methods such as sequential mate allocation (Sun et al., 2013; Carthy et al., 2019; Bérodier et al., 2021).

There are several mating programs available in the Nordic countries, but to our knowledge none takes into account genomic relationships to plan matings. In total numbers, RDC is the second most common dairy breed in the Nordic countries Sweden, Finland, and Denmark, with approximately 200,000 cows in the milk recording scheme. Nordic Red Dairy Cattle are a mixture of Swedish Red, Danish Red, and Finnish Ayrshire, and historically also contain genes from Norwegian Red, Canadian Ayrshire, American Brown Swiss, and Red Holstein-Friesian (NAV, 2019). Genotyping of RDC started on a large scale in 2012, with the VikingGenetics genotyping project. From 2012 to 2020, more than 100,000 RDC females and 20,000 RDC males were genotyped. Approximately 20% of the RDC females born in 2019 were genotyped.

Our objective in this study was to investigate the ability of different approaches for mating allocation to maximize expected genetic level, limiting parent relationship and minimizing the probability of expression of genetic defects, in the next generation. We investigated all scenarios at herd level with real data. We used linear programming to optimize different economic scores within each herd, considering genetic level, semen cost, the economic impact of recessive genetic defects, and 5 different measures of relationships (2 pedigree based and 3 genomic based).

MATERIALS AND METHODS

Breeding values, pedigree data, SNP data, and data on the carrier status of genetic defects were obtained from the Nordic Cattle Genetic Evaluation (NAV, 2019).

Genotype Data

The SNP information for all genotyped RDC animals born between 2011 and 2020 in Denmark, Finland, and Sweden was available. Nordic Cattle Genetic Evaluation uses the Illumina 50k chip (Illumina Inc.) as the standard for genomic prediction, and genotypes from lower-density chips were imputed by NAV to 50k with FImpute software (Sargolzaei et al., 2014). From late 2018 onward, most of the animals were genotyped with a EuroG MD beadchip (Borchersen, 2019). In total, the data included genotypes from 149,943 animals (28,337 males and 121,606 females).

In RDC, several known recessive genetic defects are segregating (Wu et al., 2020). Genotype information for

Table 1. Known recessive genetic defects, and their effect if homozygous, available with a genomic test for Nordic Red Dairy Cattle

Recessive genetic defect	Effect if homozygous
BTA12 OMIA 001901–9913	Early abortion, between the first and fifth month of gestation (Kadri et al., 2014)
BTA23 OMIA 001991–9913	Stillborn calf (Sahana et al., 2016)
Brown Swiss haplotype 2 (BH2) OMIA 001939–9913	Stillborn calf or calf death shortly after birth (Schwarzenbacher et al., 2016)
Ptosis intellectual disability, retarded growth, and mortality (PIRM/AH1) OMIA 001934–9913	Early abortion within 100 d of gestation. Inhibited growth if calves are born. PIRM/AH1 are located very close together and are expected to be the same disease (Guarini et al., 2019).
Ayrshire haplotype 2 (AH2) OMIA 002134–9913	Early abortion within 56 d of gestation (Guarini et al., 2019)
Spinal muscular atrophy (SMA) OMIA 000939–9913	Calves become weak and have problems standing, progressively worsen until they die; seen in wk 1–12 (Krebs et al., 2007)

Table 2. Descriptive statistics on the Nordic Red Dairy Cattle females and bulls selected for mating allocations[1]

Trait	Females	Data set BullVG	Data set BullAll
Number of animals	9,841	50	50
Average Nordic total merit (NTM)	10.7	28.4	25.2
Carriers of defect BTA12 (%)	14.7	12.0	14.0
Carriers of defect BTA23 (%)	1.1	0.0	2.0
Carriers of defect BH2 (%)	0.3	0.0	0.0
Carriers of defect PIRM/AH1 (%)	1.6	0.0	0.0
Carriers of defect AH2 (%)	1.2	0.0	0.0
Carriers of defect SMA (%)	0.30	0.0	0.0

[1]BullVG = 50 genotyped RDC bulls from the Nordic breeding cooperative VikingGenetics; BullAll = 50 genotyped RDC bulls born between January 2017 and August 2019.

a total of 6 genetic defects (Table 1) has been derived by SEGES (Skejby, Denmark) for NAV from SNPs in the EuroG MD beadchip.

Breeding Values

Genomic breeding values from the NAV evaluation performed in May 2020 were used in this study. The total merit index used was Nordic total merit (**NTM**), which at the time of this study was composed of 15 subindices, including yield index, youngstock survival, longevity, growth, udder health, udder, feet and legs, frame, hoof health, milkability, daughter fertility, general health, temperament, calving maternal, and calving direct. Nordic total merit is expressed in standardized units with a mean of 0 and a genetic standard deviation of 10 (NAV, 2019).

Data Selection

Females. We selected 9,841 genotyped females born in Denmark, Finland, or Sweden in 2019 for mating allocations (Table 2). In late 2018, a new SNP array for genotyping was introduced in these countries, which included the 6 known genetic defects listed in Table 1. Hence, 2019 was the first year with complete information about the 6 genetic defects we considered in our mating allocations. All females included belonged to herds with 20 or more genotyped females in 2019. In total, 234 herds were represented, with an average of 42 genotyped females per herd (the smallest number of genotyped females in a herd was 20 and the largest was 244). Descriptive herd statistics on the carrier frequency of the different genetic defects can be found in Table 3.

Bulls. We used 2 data sets on bulls (Table 2), which were potential mates of the 9,841 selected females. The first bull data set (**BullVG**) included 50 genotyped RDC bulls from the Nordic breeding cooperative VikingGenetics. These bulls were born between January 2017 and August 2019. Since it became possible, RDC bulls have been subjected to additional tests for the 6 genetic defects considered here, enabling us to use older bulls than females in our mating allocations. At VikingGenetics, the program EVA (Berg et al., 2006) is used for OCS using pedigree relationships (Jakob Lykke Voergaard, product manager, VikingRed, VikingGenetics, personal communication, January 11, 2021). The bulls we chose as potential mates in this study were the top available RDC bulls based on the NTM scale for which semen was marketed. There were 32 sires of the bulls in BullVG. In total, 6 of the 50 bulls were carriers of the recessive genetic defect at BTA12. None of the other genetic defects in Table 1 was present in BullVG. The highest-ranked carrier bull of the genetic defect at BTA12 was number 13 on the NTM scale.

The second bull data set (**BullAll**) also consisted of 50 genotyped RDC bulls born between January 2017 and August 2019. We removed the requirement to use only marketed semen, to eliminate any pre-selection

Table 3. Herd descriptive statistics (n = 234) of the carrier frequency (proportion of heterozygotes) of the 6 known genetic defects in Nordic Red Dairy Cattle[1]

Heading	BTA12	BTA23	BH2	PIRM/AH1	AH2	SMA
Mean (%)	15.0	1.3	0.3	1.8	1.4	0.3
Min (%)	0.0	0.0	0.0	0.0	0.0	0.0
Max (%)	36.0	9.5	9.5	17.4	21.0	6.2
First quartile (%)	10.2	0.0	0.0	3.0	0.0	0.0
Third quartile (%)	19.2	2.2	0.0	0.0	2.0	0.0

[1]Mean = mean of all herds carrier frequency; Min = minimum percent of carriers in any herds; Max = maximum percent of carriers in any herds.

for the breeding program based on bull carrier status. Further, we selected 50 bulls in a row on the NTM ranking so that a carrier of genetic defect at BTA12 would be ranked number 3 and that the bull data set in total would contain a higher carrier frequency. There were 33 sires of the 50 bulls in BullAll. In BullAll, 7 bulls were carriers of genetic defect at BTA12 and one was a carrier of genetic defect at BTA23 (Table 2). The carrier of genetic defect at BTA23 was number 19 on the NTM ranking.

Relationship Measures

Pedigree Relationships. We used 2 different pedigree relationship coefficients. To reflect the current Nordic mating programs, which use limited number of generations when calculating relationships, the first relationship coefficient traced the pedigree 3 generations back from the parents of the potential mating ($\mathbf{a_{3Gen}}$). The second pedigree relationship coefficient was based on all available pedigree information ($\mathbf{a_{AllGen}}$). The discrete generation equivalent (Woolliams and Mäntysaari, 1995) for the mated animals was 18.0 and the equivalent complete generations (Maignel et al., 1996) was 12.6. The 5-generation pedigree completeness for genotyped animals was 99.4%.

The pedigree file contained 48,434,951 animals. For most cases, the pedigree for genotyped animals was already corrected for mismatches by NAV. We found only 7 genotyped animals with mismatching parents, and they were excluded from further analyses. The pedigree relationship coefficients were estimated in RelaX2 software (Strandén and Vuori, 2006), which uses an algorithm modified from Meuwissen and Luo (1992).

Genomic Relationships. We used 3 different genomic relationship coefficients, one SNP-by-SNP genomic relationship and 2 based on shared genomic segments. The SNP-by-SNP genomic relationship coefficient ($\mathbf{g_{SNP}}$) between animals i and j was calculated according to VanRaden (2008):

$$g_{SNP_{ij}} = \frac{\sum_m \left(x_{im} - 2p_m\right) \times \left(x_{jm} - 2p_m\right)}{2\sum_m p_m \ (1 - p_m)},$$

where x_{im} and x_{jm} are the genotype scores of animal i and animal j at marker m, coded: 0 = homozygote, 1 = heterozygote, and 2 = alternative homozygote, and p_m is the frequency of the alternative allele of marker m in the founder population. Because we did not know the founder population frequency, we instead used the allele frequency of all 149,943 genotyped RDC animals available for this study, as is common practice for genomic evaluation (Wang et al., 2014). We used the software SNP1101 to calculate the SNP-by-SNP genomic relationship coefficients (Sargolzaei, 2014).

The 2 genomic relationship coefficients based on shared genomic segments were calculated following de Cara et al. (2013):

$$g_{SEG_{ij}} = \frac{\sum_k \sum_{ai=1}^{2} \sum_{bj=1}^{2} \left[L_{SEGk}\left(a_i b_j\right)\right]}{2L_{AUTO}},$$

where L_{SEGk} is the length (in base pairs) of the kth shared segment measured over homolog a of animal i and homolog b of animal j, and L_{AUTO} is the total length of the autosomes covered by the SNP in base pairs.

The 2 segment-based genomic relationship coefficients were based on different minimum lengths of segments: 1 cM ($\mathbf{g_{SEG1}}$) and 4 cM ($\mathbf{g_{SEG4}}$), assuming 1 cM = 1,000,000 bp (Gautier et al., 2007). The lengths of segments were chosen to represent short and long segments, similarly to other studies (Zhang et al., 2015; Forutan et al., 2018; Makanjuola et al., 2020; Martikainen et al., 2020). Phasing of genotypes was done in Beagle 4.1 with default settings (Browning and Browning, 2007), and segments of minimum desired length were extracted in RefineIBD with the default setting except for the logarithm of the odds (**LOD**) score (base 10 log of the likelihood ratio), where we used LOD = 0.1 (Browning and Browning, 2013). The LOD score is used to prune out shared segments that are not common in the population. Hence, default LOD = 3.0 in RefineIBD was considered too high for our purposes, as in a recent study (Olsen et al., 2020).

Economic Score

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

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

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 a 1% increase in inbreeding, F_{ij} is the pedigree or genomic based co-ancestry (relationship/2), prob(Fem) 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, and semen cost is the average amount (€) spent on semen for a pregnancy.

The value of 1 index unit of NTM was approximated to be €24.8, based on the value per NTM unit and year (€9.2) the average and production lifetime (2.7 yr; Fikse and Kargo, 2020). We only considered sexed semen and assumed a 0.9 probability of producing a female conceptus, which is the minimum expected sexing rate for most sexing technologies (Burnell, 2019). Sexed semen is gaining popularity in the Nordic countries and is combined with the use of beef semen to get the number of heifers needed for the next generation. It is expected that most of the semen sold by VikingGenetics in future will be sexed dairy semen and beef semen (Jakob Lykke Voergaard, product manager, VikingRed, VikingGenetics, personal communication, January 11, 2021).

The economic consequence of a 1% increase in inbreeding was set to €24.8. The current version of the Swedish mating program "Genvägen" uses a penalty of 1 NTM unit per 1% increase in inbreeding, which would mean €24.8 (Lina Baudin, expert in breeding routines, Växa Sverige, personal communication, March 5, 2021). To our knowledge, no such values have been calculated specifically for the RDC breed, and therefore in a sensitivity analysis we set the economic consequence of a 1% increase in inbreeding to €10.0, €24.8, or €40.0. The analysis was performed with BullVG and scenarios maximizing economic scores, including all available information and a maximum of 5% females per bull and herd.

The costs associated with genetic defects were based on economic effects of health disorders estimated by Oskarsson and Engelbrekts (2015) and the economic assumptions behind the NTM (Sørensen et al., 2018). We assumed the cost of an early abortion (genetic defect at BTA12, PIRM/AH1, AH2) to be €80, based on the resulting longer calving interval (€30–€40/mo) and the cost of extra insemination(s) (€30). We assumed the cost of a later abortion or an early calf death to be €160 (genetic defect SMA, BH2, and at BTA23).

We used the prices for sexed semen set by VikingGenetics in 2020, where a semen dose for a bull with a NTM of 30 or more cost €26, with a NTM of 25 to 30 cost €22.5, and with a NTM of 20 to 25 cost €19 (Jakob Lykke Voergaard, product manager, VikingRed, VikingGenetics, personal communication, January 11, 2021). We multiplied the semen price by 1.8, which is the average number of inseminations needed for a pregnancy in RDC (Sørensen et al., 2018).

Mating Scenarios

In addition to the economic scores that included all available information described above, we investigated mating scenarios without the penalty for genetic defects. In addition, we investigated scenarios that only aimed to reduce the genetic relationships. Detailed information about the mating scenarios can be found in Table 4.

Mate Allocation

Mate allocation was programmed in R version 3.6.3 (R Core Team, 2020). Linear programming optimization was performed with the 'Lp_solve' package in R (Berkelaar et al., 2020). The mating R script was provided by Bérodier et al. (2021). The R script set up constraints that were considered in the linear programming optimization. We used the constraints: one mating per female and a threshold percentage for the maximum number of females per bull and herd, for which we evaluated 2 different levels, 5% and 10%, similarly to Bérodier et al. (2021). The threshold for the number of females per bull and herd was in line with current recommendations given by Swedish breeding advisors.

The planned matings achieved from each scenario were compared by (1) average NTM; (2) average genetic relationships (a_{3Gen}, a_{AllGen}, g_{SNP}, g_{SEG1}, g_{SEG4}); (3) the probability of expression of genetic defects, including genetic defect at BTA12, using bull set BullVG, and including genetic defects at BTA12 and BTA23 using bull set BullAll; (4) the average cost of semen for a pregnancy, calculated in the same way as in the economic score; (5) the total number of bulls used; (6) the number of bulls used to the maximum number of doses based on the threshold (5% and 10%) of females per bull and herd; (7) average pedigree relationship among all planned matings, calculated similarly to a_{AllGen}; and (8) predicted carrier frequency of genetic defect at BTA12 using BullVG, and predicted carrier frequency of genetic defects at BTA12 and BTA23 using BullAll, calculated as 50% of the cases when a parent was a carrier divided by the total number of matings. The predicted carrier frequency in the next generation did not include homozygotes for the genetic defects, which were included in the probability of expression of genetic defects.

Statistical Analysis

SAS software version 9.4 (SAS Institute Inc.) and R version 3.6.3 (R Core Team, 2020) were used for statistical analysis.

RESULTS

All results are presented for the selected females and bulls in BullVG, unless otherwise specified.

Genetic Relationship Coefficients

The mean value of the relationship coefficients between all possible combinations of females and males ranged from 0.009 to 0.188, and the standard deviation ranged from 0.042 to 0.047 (Table 5). The correlations between the genetic relationship coefficients were all 0.83 or higher. The strongest correlation was between a_{AllGen} and a_{3Gen} (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 6). The coefficients of regression on a_{AllGen} were close to 1, highest for a_{3Gen} and g_{SNP} and somewhat lower for g_{SEG1} and g_{SEG4} (Figure 1).

Table 4. Description of the 15 different mating scenarios considered

	Economic score includes					
Scenario[1]	Nordic total merit, NTM	Relationship[2]	Genetic defect value	Sexed semen	Semen cost	Linear programming objective[3]
MaxNTM	Yes	No	No	Yes	Yes	Max
3Gen	Yes	a_{3Gen}	Yes	Yes	Yes	Max
AllGen	Yes	a_{AllGen}	Yes	Yes	Yes	Max
GSNP	Yes	g_{SNP}	Yes	Yes	Yes	Max
GSEG1	Yes	g_{SEG1}	Yes	Yes	Yes	Max
GSEG4	Yes	g_{SEG4}	Yes	Yes	Yes	Max
3Gen_NoDefect	Yes	a_{3Gen}	No	Yes	Yes	Max
AllGen_NoDefect	Yes	a_{AllGen}	No	Yes	Yes	Max
GSNP_NoDefect	Yes	g_{SNP}	No	Yes	Yes	Max
GSEG1_NoDefect	Yes	g_{SEG1}	No	Yes	Yes	Max
GSEG4_NoDefect	Yes	g_{SEG4}	No	Yes	Yes	Max
3Gen_Min	No	a_{3Gen}	No	Yes	No	Min
AllGen_Min	No	a_{AllGen}	No	Yes	No	Min
GSNP_Min	No	g_{SNP}	No	Yes	No	Min
GSEG1_Min	No	g_{SEG1}	No	Yes	No	Min
GSEG4_Min	No	g_{SEG4}	No	Yes	No	Min
Random	All possible combinations of 9,841 females and 50 bulls					

[1]MaxNTM: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, and semen cost. 3Gen: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a pedigree relationship including 3 generations of ancestors (a_{3Gen}), and penalty for genetic defects. AllGen: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a pedigree relationship including all available ancestors (a_{AllGen}), and penalty for genetic defects. GSNP: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship calculated according to VanRaden (2008) (g_{SNP}), and penalty for genetic defects. GSEG1: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 1 cM (g_{SEG1}), and penalty for genetic defects. GSEG4: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 4 cM (g_{SEG4}), and penalty for genetic defects. 3Gen_NoDefect: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a pedigree relationship including 3 generations of ancestors (a_{3Gen}). AllGen_NoDefect: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a pedigree relationship including all available ancestors (a_{AllGen}). GSNP_NoDefect: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a genomic relationship calculated according to VanRaden (2008) (g_{SNP}). GSEG1_NoDefect: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 1 cM (g_{SEG1}). GSEG4_NoDefect: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 4 cM (g_{SEG4}). 3Gen_Min: mating scenario where mates were selected based on minimizing an economic score including a pedigree relationship including 3 generations of ancestors. AllGen_Min: mating scenario where mates were selected based on minimizing an economic score including a pedigree relationship including all available ancestors. GSNP_Min: mating scenario where mates were selected based on minimizing an economic score including a genomic relationship calculated according to VanRaden (2008). GSEG1_Min: mating scenario where mates were selected based on minimizing an economic score, including a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 1 cM. GSEG4_Min: mating scenario where mates were selected based on minimizing an economic score including a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 4 cM.

[2]a_{3Gen} = pedigree relationships using 3 generations of ancestors; a_{AllGen} = 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) cM.

[3]The objective of linear programming is to maximize (Max) or minimize (Min) the economic score.

Table 5. Descriptive statistics on relationships [mean, SD, minimum value (Min), and maximum value (Max)] between all possible combinations of 9,841 females and 50 bulls[1]

Relationship	Mean	SD	Min	Max
a_{3Gen}	0.028	0.042	0	0.648
a_{AllGen}	0.066	0.042	0.003	0.667
g_{SNP}	0.009	0.047	−0.095	0.673
g_{SEG1}	0.188	0.046	0.038	0.789
g_{SEG4}	0.115	0.045	0.005	0.727

[1]a_{3Gen} = pedigree relationships using 3 generations of ancestors; a_{AllGen} = 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) cM.

Table 6. Correlations between the different relationship coefficients for all possible combinations of 9,841 females and 50 bulls[1]

Relationship	a_{3Gen}	a_{AllGen}	g_{SNP}	g_{SEG1}	g_{SEG4}
a_{3Gen}	1	0.99	0.88	0.83	0.87
a_{AllGen}		1	0.88	0.85	0.88
g_{SNP}			1	0.9	0.93
g_{SEG1}				1	0.98

[1]a_{3Gen} = pedigree relationships using 3 generations of ancestors; a_{AllGen} = 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) cM.

Mate Allocation

Using BullVG. It was possible to maximize economic score with limited impact on the average NTM level (Table 7). Including the cost of the known recessive genetic defect (at BTA12) when optimizing mating strategies eliminated the risk of expression of the genetic defect, regardless of which genetic relationship was used. In MaxNTM (mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, and semen cost), the NTM level improved compared with Random (all possible combinations of 9,841 females and 50 bulls), but it resulted in higher average genetic relationship coefficients than Random and did not reduce the probability of expression of genetic defects.

Including a genomic relationship in the economic score also kept the other genomic relationship averages at a low level. For example, with the constraint 5% females per bull and herd, including g_{SNP} in the objective function (scenario GSNP) resulted in a g_{SEG1} of

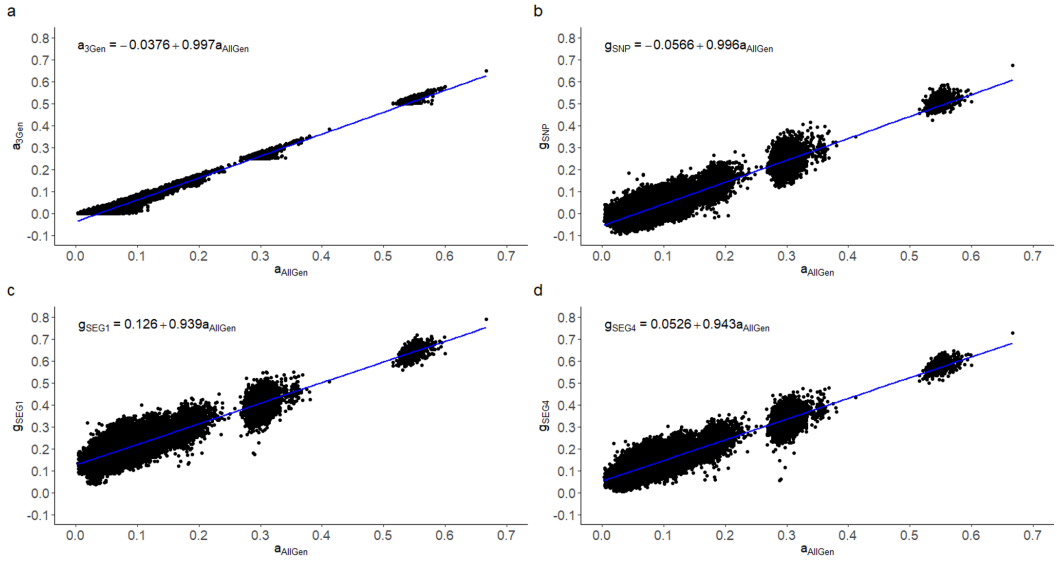


Figure 1. (a) Relationship coefficients estimated from pedigree data with 3 generations of ancestors (a_{3Gen}), (b) relationship coefficients estimated from SNP data (g_{SNP}; VanRaden, 2008), (c) relationship coefficients estimated from shared genomic segments with a minimum segment length of 1 cM (g_{SEG1}), and (d) minimum length of 4 cM (g_{SEG4}; de Cara et al., 2013), all plotted against relationship coefficients estimated from pedigree data using all available ancestors (a_{AllGen}). The diagrams include relationships for all possible combinations of 9,841 Nordic Red Dairy Cattle females and 50 bulls.

Table 7. Comparison of outcome of planned matings of 9,841 females for 13 mating scenarios in Nordic Red Dairy Cattle using various comparison criteria[1,2]

		5% females/bull scenarios[3]						10% females/bull scenarios[3]					
Comparison criterion	Random	Max NTM	3Gen	AllGen	GSNP	GSEG1	GSEG4	Max NTM	3Gen	AllGen	GSNP	GSEG1	GSEG4
Average Nordic total merit (NTM)	19.5	20.8	20.8	20.8	20.8	20.7	20.8	22.1	22.1	22.1	22.0	22.0	22.0
Average a_{3Gen} between parents	0.028	0.033	0.007	0.009	0.014	0.014	0.014	0.037	0.009	0.010	0.016	0.016	0.015
Average a_{AllGen} between parents	0.066	0.070	0.046	0.043	0.050	0.050	0.050	0.074	0.046	0.044	0.051	0.051	0.051
Average g_{SNP} between parents	0.009	0.014	−0.012	−0.016	−0.038	−0.034	−0.033	0.018	−0.010	−0.013	−0.032	−0.027	−0.027
Average g_{SEG1} between parents	0.188	0.191	0.167	0.163	0.148	0.143	0.146	0.195	0.166	0.163	0.150	0.145	0.148
Average g_{SEG4} between parents	0.115	0.119	0.094	0.091	0.078	0.075	0.074	0.125	0.094	0.091	0.080	0.078	0.076
Probability of expression of genetic defect (%)	0.4	0.4	0	0	0	0	0	0.04	0	0	0	0	0
Average cost of semen for a pregnancy (€)	42.0	43.6	43.6	43.6	43.5	43.5	43.5	45.8	46.2	46.3	45.8	45.7	45.9
Number of bulls used	NA[4]	39	46	46	50	48	48	15	18	18	28	27	26
Number of bulls used to a maximum	NA	20	18	16	10	12	13	10	8	7	3	3	4
Average a_{AllGen} between planned matings	NA	0.089	0.089	0.088	0.088	0.088	0.088	0.107	0.106	0.106	0.104	0.105	0.105
Predicted BTA12 carrier frequency in the next generation (%)	13.4	12.9	13.2	12.9	12.9	12.0	12.9	7.9	7.6	7.6	8.2	7.8	7.8

[1]Fifty marketed bulls from VikingGenetics were available for matings (BullVG data set).

[2]Average NTM level, 5 different genetic relationships, the probability of expression of genetic defect (at BTA12), the average cost of semen for a pregnancy, the number of bulls used, the number of bulls used to a maximum number of doses based on the 5% and 10% constraint of females per bull and herd, average pedigree relationship between all planned matings, and predicted genetic defect at BTA12 carrier frequency in the next generation. MaxNTM: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, and semen cost. 3Gen: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a pedigree relationship including 3 generations of ancestors (a_{3Gen}), and penalty for genetic defects. AllGen: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a pedigree relationship including all available ancestors (a_{AllGen}), and penalty for genetic defects. GSNP: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship calculated according to VanRaden (2008) (g_{SNP}), and penalty for genetic defects. GSEG1: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 1 cM (g_{SEG1}), and penalty for genetic defects. GSEG4: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 4 cM (g_{SEG4}), and penalty for genetic defects.

[3]Maximum percentage of females per bull and herd set to 5% or 10%.

[4]Not applicable.

0.148, compared with 0.143 with GSEG1 (Table 7). 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. Considering the example with the constraint 5% females per bull and herd, and including g_{SNP} in the objective function (scenario GSNP), the pedigree relationship scenarios resulted in a g_{SEG1} of 0.167 for 3Gen and 0.163 for AllGen. There were only minor differences between the scenarios with genomic relationships in their ability to reduce pedigree relationships. Including pedigree relationships in the economic scores consistently reduced pedigree relationships more than genomic relationships. For example, all scenarios optimizing genomic relationships resulted in a_{AllGen} of 0.050, AllGen resulted in a_{AllGen} of 0.043, and 3Gen resulted in a_{AllGen} of 0.046 (Table 7).

Using BullAll. For the bull set BullAll, including the costs of the known recessive genetic defects (at BTA12 and BTA23) when optimizing mating strategies entirely eliminated the risk of expression of a genetic defect, regardless of which genetic relationship was used in the objective function (Table 8).

Bull Usage. The number of bulls used in the scenarios considering genomic relationships was always higher than in the scenarios considering pedigree relationships. Furthermore, fewer bulls were used for the maximum number of permitted inseminations considering genomic relationships compared with scenarios considering pedigree relationships with the same constraints. Minor differences were observed in the average pedigree relationship between all planned matings using the same threshold for females per bull and herd.

Predicted Carrier Frequency in the Next Generation

The predicted carrier frequency in the next generation was half the carrier frequencies in Table 2 for the genetic defects not present in bull set BullVG (at BTA23, BH2, PIRM/AH1, AH2, SMA) and bull set BullAll (BH2, PIRM/AH1, AH2, SMA). Further, the predicted carrier frequencies of known genetic defects in the next generation depended on the proportion of carrier bull used. Using a maximum of 10% females per bull and herd resulted in considerably lower carrier frequencies in the next generation (Table 7). In this case, the best carrier bull was ranked number 13 on the NTM scale and that bull was rarely chosen in any of the mating allocations. However, when using a maximum constraint of 5% females per bull and herd, the predicted carrier frequency in the next generation was higher than with a maximum constraint of 10% females per bull and herd. The bull ranked number 13 and the other lower-ranked bulls on the NTM scale were then required to be used due to the constraint. When using bull set BullAll, more carrier bulls were ranked high on the NTM scale. Hence, it resulted in higher predicted carrier frequency in the next generation as a consequence of carrier bulls being selected more often (Table 8) than when using bull set BullVG (Table 7).

Alternative Scenarios

Results for scenarios excluding genetic defects from the objective function showed a probability of expression of genetic defect without the penalty for defects in the economic score (Table 9). Including g_{SEG1} resulted in the lowest probability of expression of genetic defects. There were only minor changes for the other result parameters compared with when the penalty was included.

Results for scenarios minimizing parents' genetic relationships showed a lower average NTM level than the other scenarios, because they were not optimized with respect to NTM (Table 10). Furthermore, the average a_{AllGen} between planned matings was improved (e.g., 0.083 in AllGen_Min to 0.089 in MaxNTM). In MaxNTM, the average a_{AllGen} relationship was 0.070 (Table 9). Compared with scenarios maximizing economic scores, including all information except the defect penalty (Table 9), the genetic relationships could be reduced slightly more. For example, in AllGen_NoDefect, the a_{AllGen} relationship was 0.043 (Table 9) and in AllGen_Min it was 0.040 (Table 10). Similarly, in GSNP_NoDefect g_{SNP} was −0.038 and in GSNP_Min it was −0.044. Further, in the scenarios aimed at only minimizing the parents' genetic relationship, we observed a probability of expression of a genetic defect. AllGen_Min and GSEG4_Min resulted in a 0.2% probability of expression of a genetic defect and GSEG1_Min in 0.1% probability, compared with 0.4% probability in Random and MaxNTM.

Effect of Constraints Used in Mate Allocation

Changing the maximum number of females per bull and herd from 5% to 10% resulted in a higher NTM, and the increase was greater for BullVG (1.2–1.4 NTM units) than for BullAll (0–0.2 NTM units; Table 7–8), owing to more variation in NTM level in BullVG than in BullAll. Lower variation in NTM level led to genetic relationships being more decisive in mating optimization, which in turn led to fewer bulls being used to their maximum number of inseminations based on the constraints 5% and 10% females per bull and herd. For example, in 3Gen, using the constraint 5% females per

Table 8. Comparison of outcomes of planned matings of 9,841 females for 13 mating scenarios in Nordic Red Dairy Cattle using various comparison criteria[1,2]

		5% females/bull scenarios[3]						10% females/bull scenarios[3]					
Comparison criterion	Random	Max NTM	3Gen	AllGen	GSNP	GSEG1	GSEG4	Max NTM	3Gen	AllGen	GSNP	GSEG1	GSEG4
Average Nordic total merit (NTM)	18.0	18.2	18.2	18.1	18.1	18.1	18.1	18.4	18.3	18.3	18.1	18.1	18.2
Average a_{3Gen} between parents	0.003	0.028	0.006	0.007	0.014	0.013	0.013	0.031	0.006	0.007	0.013	0.013	0.013
Average a_{AllGen} between parents	0.066	0.066	0.045	0.042	0.051	0.050	0.050	0.069	0.046	0.042	0.050	0.050	0.050
Average g_{SNP} between parents	0.010	0.010	−0.011	−0.016	−0.041	−0.033	−0.033	0.012	−0.010	−0.015	−0.042	−0.031	−0.031
Average g_{SEG1} between parents	0.185	0.183	0.164	0.160	0.144	0.135	0.137	0.187	0.164	0.159	0.145	0.131	0.134
Average g_{SEG4} between parents	0.113	0.112	0.093	0.090	0.075	0.070	0.068	0.115	0.093	0.090	0.076	0.068	0.066
Probability of expression of genetic defects (%)	0.5	0.5	0	0	0	0	0	0.6	0	0	0	0	0
Average cost of semen for a pregnancy (€)	38.0	39.8	40.0	39.7	39.1	39.4	39.4	40.5	40.4	40.0	36.3	38.1	37.9
Number of bulls used	NA[4]	39	44	45	50	50	50	15	32	38	50	48	49
Number of bulls used to a maximum	NA	20	8	4	0	3	1	10	1	0	0	0	0
Average a_{AllGen} between planned matings	NA	0.086	0.084	0.083	0.083	0.083	0.083	0.098	0.092	0.089	0.085	0.088	0.087
Predicted BTA12 carrier frequency in the next generation (%)	14.4	14.9	16.2	16.8	15.3	13.3	13.8	16.4	17.8	16.8	15.6	13.6	14.4
Predicted BTA23 carrier frequency in the next generation (%)	1.6	2.6	2.2	1.5	2.2	2.5	2.4	0.6	0.1	0.7	2.0	2.8	2.6

[1]Available bulls were 50 bulls selected with a higher carrier frequency of the genetic defects than the marketed bulls from VikingGenetics (BullAll).

[2]Average NTM level, 5 different genetic relationships, the probability of expression of genetic defects (at BTA12, BTA23), the average cost of semen for a pregnancy, the number of bulls used, the number of bulls used to a maximum number of doses based on the 5% and 10% constraint of females per bull and herd, average pedigree relationship between all planned matings, and predicted genetic defect at BTA12 and BTA23 carrier frequency in the next generation. MaxNTM: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, and semen cost. 3Gen: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a pedigree relationship including 3 generations of ancestors (a_{3Gen}), and penalty for genetic defects. AllGen: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a pedigree relationship including all available ancestors (a_{AllGen}), and penalty for genetic defects. GSNP: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship calculated according to VanRaden (2008) (g_{SNP}), and penalty for genetic defects. GSEG1: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 1 cM (g_{SEG1}), and penalty for genetic defects. GSEG4: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 4 cM (g_{SEG4}), and penalty for genetic defects.

[3]Maximum percent of females per bull and herd set to 5% or 10%.

[4]Not applicable.

Table 9. Comparison of outcome of planned matings of 9,841 females for 6 mating scenarios in Nordic Red Dairy Cattle using various comparison criteria[1,2]

		Scenarios without penalty for defects				
Comparison criterion	Max NTM	3Gen_No Defect	AllGen_No Defect	GSNP_No Defect	GSEG1_No Defect	GSEG4_No Defect
Average Nordic total merit (NTM)	20.8	20.8	20.8	20.8	20.7	20.8
Average a_{3Gen} between parents	0.033	0.007	0.009	0.014	0.014	0.014
Average a_{AllGen} between parents	0.070	0.046	0.043	0.050	0.050	0.050
Average g_{SNP} between parents	0.014	−0.012	−0.016	−0.038	−0.034	−0.033
Average g_{SEG1} between parents	0.191	0.167	0.163	0.148	0.143	0.146
Average g_{SEG4} between parents	0.119	0.094	0.091	0.078	0.075	0.074
Probability of expression of genetic defect (%)	0.4	0.4	0.4	0.4	0.2	0.3
Average cost of semen for a pregnancy (€)	43.6	43.6	43.6	43.5	43.5	43.5
Number of bulls used	39	46	45	50	48	47
Number of bulls used to a maximum	20	18	16	10	12	13
Average a_{AllGen} between planned matings	0.089	0.089	0.088	0.088	0.088	0.088
Predicted BTA12 carrier frequency in the next generation (%)	12.9	13.0	12.8	13.1	12.6	12.8

[1]Fifty marketed bulls from VikingGenetics were available for matings (BullVG). Maximum percentage of females per bull and herd set to 5%.

[2]Average NTM level, 5 different genetic relationships, the probability of expression of genetic defect (at BTA12), the average cost of semen for a pregnancy, the number of bulls used, the number of bulls used to a maximum number of doses based on the 5% of females per bull and herd, average pedigree relationship between all planned matings, and predicted genetic defect at BTA12 carrier frequency in the next generation. MaxNTM: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, and semen cost. 3Gen_NoDefect: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a pedigree relationship including 3 generations of ancestors (a_{3Gen}). AllGen_NoDefect: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a pedigree relationship including all available ancestors (a_{AllGen}). GSNP_NoDefect: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a genomic relationship calculated according to VanRaden (2008) (g_{SNP}). GSEG1_NoDefect: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 1 cM (g_{SEG1}). GSEG4_NoDefect: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 4 cM (g_{SEG4}).

Table 10. Comparison of outcome of planned matings of 9,841 females for 6 mating scenarios in Nordic Red Dairy Cattle using various comparison criteria[1,2]

		Scenarios minimizing relationships				
Comparison criterion	Max NTM	3Gen _Min	AllGen _Min	GSNP _Min	GSEG1 _Min	GSEG4 _Min
Average Nordic total merit (NTM)	20.8	19.2	19.6	19.5	19.7	19.7
Average a_{3Gen} between parents	0.033	0.004	0.007	0.013	0.013	0.013
Average a_{AllGen} between parents	0.070	0.044	0.040	0.050	0.049	0.049
Average g_{SNP} between parents	0.014	−0.015	−0.019	−0.044	−0.036	−0.036
Average g_{SEG1} between parents	0.191	0.167	0.160	0.145	0.137	0.140
Average g_{SEG4} between parents	0.119	0.094	0.088	0.075	0.071	0.068
Probability of expression of genetic defect (%)	0.4	0.4	0.2	0.4	0.1	0.2
Average cost of semen for a pregnancy (€)	43.6	41.6	41.7	41.9	42.2	42.1
Number of bulls used	39	50	49	50	50	50
Number of bulls used to a maximum	20	2	0	0	1	0
Average a_{AllGen} between all planned matings	0.089	0.084	0.083	0.083	0.084	0.083
Predicted BTA12 carrier frequency in the next generation (%)	12.9	12.9	10.1	13.5	10.4	11.0

[1]Fifty marketed bulls from VikingGenetics were available for matings (BullVG). Maximum percentage of females per bull and herd set to 5%.

[2]Average NTM level, 5 different genetic relationships, the probability of expression of genetic defect (at BTA12), the average cost of semen for a pregnancy, the number of bulls used, the number of bulls used to a maximum number of doses based on the 5% of females per bull and herd, average pedigree relationship between all planned matings, and predicted genetic defect at BTA12 carrier frequency in the next generation. MaxNTM: mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, and semen cost. 3Gen_Min: mating scenario where mates were selected based on minimizing an economic score including a pedigree relationship including 3 generations of ancestors (a_{3Gen}). AllGen_Min: mating scenario where mates were selected based on minimizing an economic score including a pedigree relationship including all available ancestors (a_{AllGen}). GSNP_Min: mating scenario where mates were selected based on minimizing an economic score including a genomic relationship calculated according to VanRaden (2008) (gSNP). GSEG1_Min: mating scenario where mates were selected based on minimizing an economic score, including a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 1 cM (g_{SEG1}). GSEG4_Min: mating scenario where mates were selected based on minimizing an economic score including a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 4 cM (g_{SEG4}).

bull and herd resulted in 18 bulls being used to the maximum when using BullVG (Table 7), and 8 bulls being used to the maximum when using BullAll (Table 8). Furthermore, changing the maximum number of females per bull and herd from 5% to 10% increased the average a_{AllGen} among planned matings. The threshold for the maximum number of females per bull and herd thus seems to be most influential for a_{AllGen} among planned matings, and we saw only minor differences between scenarios with the same threshold.

The total cost of semen for a pregnancy increased on changing the maximum number of females per bull and herd from 5% to 10%, because it was more profitable to use bulls from the highest price category more extensively. In general, there were minor differences between scenarios in total cost of semen with the same constraints. Some differences occurred with these constraints if many bulls had NTM close to the price category borders. For example, in BullAll, allowing a maximum of 10% females per bull and herd meant that many bulls had NTM close to 25, which was the price category border.

Sensitivity Analysis

Changing the economic consequence of a 1% increase in inbreeding from €10.0 to €40.0 did not change the average a_{3Gen} (0.07) or a_{AllGen} (0.043), whereas g_{SNP} changed slightly from −0.039 using €40.0 to −0.036 using €10.0, average g_{SEG1} changed from 0.141 using €40.0 to 0.145 using €10.0, and g_{SEG4} changed from 0.072 using €40.0 and 0.075 using €10.0. The average NTM level was kept between 20.6 and 20.8, and no risk of expression of a known genetic defect.

DISCUSSION

The results we present here 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 optimizing mating strategies eliminated expression of known genetic defects, regardless of the genetic relationship used. There is a long tradition of recording in the Nordic countries, and the strong correlation between pedigree and genomic relationships that we estimated confirms that dairy pedigrees are well documented in the Nordic countries. The results of the sensitivity analysis showed that the mating results are robust in the inbreeding penalty range tested. Furthermore, the genetic relationship was reduced only slightly more when using an economic score designed to only reduce the different genetic relationships than when using an economic score including all available information.

Genetic Relationships

The correlation between the pedigree relationship and genomic relationship estimates was high, ≥0.83 for a_{3Gen}, and ≥0.85 for a_{AllGen} (Table 6). Carthy et al. (2019) reported a 0.57 correlation 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, this corresponded to a correlation of 0.67, 0.73, 0.84, and 0.87, respectively. Similarly to our study, they also found that the reduction in genetic relationship was dependent on the way genetic relationships were evaluated. For example, including genomic relationships in an economic score was superior to including pedigree relationships when the goal was to reduce a genomic relationship (Pryce et al., 2012).

Compared with other common dairy cattle breeds, the estimated average genetic relationship between parents was low in the present study. The average pedigree relationship coefficient was approximately half that found by Bérodier et al. (2021) for the Montbéliarde breed, with slightly less pedigree information available (9.7–10.0 equivalent complete generations compared with 12.6 in our study). Carthy et al. (2019) found an average pedigree relationship for Holstein-Friesian in their mating replicates of 6.24%, which is higher than in all our scenarios including genetic relationships (Tables 7–9). However, in Carthy et al. (2019), the only information given was that animals were traced back at least 5 generations, where possible, but with no further information about pedigree completeness and therefore it is hard to compare their values with our study. Our average genomic relationship coefficients were also low compared with those in Makanjuola et al. (2020), who investigated genetic relationships in North American Jersey and Holstein. Using a segment length of 1,000,000 bp, similar to us, their f_{SEG} co-ancestry of 15.84% for Holstein and 23.46% for Jersey should correspond to half our g_{SEG1} value, which for all potential mating with bulls in the set BullVG was 9.44% ($g_{SEG1}/2$) (Table 5). The low genetic relationship in RDC can be explained by the different breeds included over time in the RDC breeding program, which has included a mixture of Swedish Red, Danish Red, and Finnish Ayrshire, plus some genes from Norwegian Red, Canadian Ayrshire, American Brown Swiss, and Red Holstein-Friesian (NAV, 2019). We noticed that the mating program favored bulls with a high percentage of breeds other than Swedish Red, Danish Red, and Finnish Ayrshire. All

bulls we mated qualified for the joint Nordic breeding program (VikingRed), where proportions of up to 25% of other breeds are allowed (Jakob Lykke Voergaard, product manager, VikingRed, VikingGenetics, personal communication, January 11, 2021). However, some national herdbooks require a lower percentage of other breeds [e.g., the Swedish Red herdbook (Swedish Red Cattle Association, Hörby, Sweden)]. Hence, a higher average relationship coefficient might be obtained with more strict selection of bulls with regard to breed percentages.

Using Genomic or Pedigree Relationships

An argument for using genomic estimates of inbreeding and relationships is that they do not rely on pedigree data, which can have limited depth or be incorrect (Carthy et al., 2019; Makanjuola et al., 2020; Bérodier et al., 2021). Nordic Cattle Genetic Evaluation had corrected the pedigree in most cases for possible mismatches using genomic information. Hence, we did not fully reveal the benefit that a genomic relationship brings in terms of assigning the right parents to an animal. In Sweden approximately 5% of genotyped animals have at least one parent incorrectly reported (Lina Baudin, expert in breeding routines, Växa Sverige, personal communication, March 5, 2021). Further, if a population is under selection, the assumption of 50% chance of each allele being selected is not true. In combination, this leads to pedigree inbreeding often underestimating true inbreeding (as identical by descent from a given base population) compared with ROH-based inbreeding (Forutan et al., 2018). Furthermore, even if pedigree is correct and deep, genomic relationships are more accurate because they consider correctly that genome is transmitted in chromosomes and not as infinite unlinked loci (Hill and Weir, 2011).

Our goal using segment-based relationships was to reduce the number of ROH in the potential offspring. ROH are suggested to be a good predictor of inbreeding depression in Finnish Ayrshire (Martikainen et al., 2017, 2020), and also in humans (Szpiech et al., 2013). In theory, ROH are enriched for deleterious alleles that mainly cause inbreeding depression (Charlesworth and Willis, 2009). In general, long ROH, reflecting new inbreeding, should contain more deleterious alleles than short ROH due to purging and recombination along with generations (Stoffel et al., 2021). Regions affecting milk and fertility lie between 1 and 14 Mb (Martikainen et al., 2020). In addition, Martikainen et al. (2017) found that pedigree inbreeding did not indicate inbreeding depression for fertility, but inbreeding based on ROH did. Further, longer regions of ROH (>3 Mb) in Holstein and Jersey have been found to be associated with inbreeding depression in milk (Pryce et al., 2014). However, Zhang et al. (2015) found that enrichment of deleterious variants was significantly higher in short (<0.1 to 3 Mb) compared with long (>3 Mb) regions in RDC, Holstein, and Jersey. Hence, it is not clear what segment length is optimal for use in segment-based relationships.

The scales of the different genetic relationship coefficients used differed (Table 5). In particular the means were different, but there were also some differences in the standard deviations. Hence, the relationships were difficult to compare directly. However, in general, genomic relationships were better at reducing pedigree relationships than pedigree relationships were at reducing genomic relationships (see e.g., Table 7). For example, the economic score 3Gen resulted in an average a_{3Gen} of 0.007 and the score GSEG1 resulted in an average a_{3Gen} of 0.014, compared with a_{3Gen} of 0.028 in Random. Hence, the relative difference in change [(0.028 − 0.014)/(0.028 − 0.007)] was 67%. Furthermore, using GSEG1 reduced g_{SEG1} compared with Random from 0.188 to 0.143, and 3Gen reduced g_{SEG1} to 0.167, that is, the relative difference [(0.188 − 0.167)/(0.188 − 0.143)] was 47%. Furthermore, there were only minor differences for genomic relationships in their ability to reduce pedigree relationships. Hence, using any of the genomic relationships could be an overall better and safer option than using pedigree relationships in keeping all average relationships studied low.

In our study, a_{AllGen} was better than a_{3Gen} at reducing the average genomic relationships (see e.g., Table 7), suggesting that the Nordic breeding organizations should use more generations when calculating pedigree relationships for nongenotyped animals if they want to control genomic relationships. This finding was expected since the depth of the pedigree plays a major role for the strength of correlation between pedigree relationships and genomic relationships in dairy cattle (Pryce et al., 2012) and similar results have also been reported in chicken (Wang et al., 2014). Furthermore, the use of any genomic relationship worked well to keep other genomic relationships low in this study, which was expected based on the strong correlations between the different genomic relationships (Table 6).

At the population level using OCS, Henryon et al. (2019) suggested that pedigree relationships realize more long-term true genetic gain than genomic relationships. However, Meuwissen et al. (2020) concluded that the choice of relationship matrix depends on which objective it should serve. Genomic relationships based on ROH resulted in allele frequency changes toward 0.5, which is clearly unfavorable if the focus is managing genetic defects. Furthermore, using genomic relationships based on VanRaden (2008) resulted in low drift,

but at the cost of a high rate of increase in homozygosity. A genomic relationship based on linkage analysis, which requires both pedigree and marker information, achieved the highest genetic gain per unit of inbreeding and kept the drift-based inbreeding within the target rate (Meuwissen et al., 2020). A downside with our study is that we only looked one generation ahead, instead of many generations as in OCS studies. Further, farmers are most likely to be mainly interested in their own herd's genetic level and have to rely on breeding organizations to offer bulls with different pedigrees, so that inbreeding depression and mating of carriers of yet unknown defects can be avoided. We were unable to draw any conclusions on which estimate of genetic relationship is best for mating plans with regard to producing offspring with low inbreeding depression and avoiding expression of unknown recessive genetic defects, balanced with high genetic gain. More studies are needed to identify the different types of genetic relationships and their future economic impact for farmers.

Recessive Genetic Defects

Carrier frequencies of the recessive genetic defects were lower in the mated bulls than in the females (Table 2). The strategy applied in VikingGenetics is to only select a carrier bull if it is genetically superior or has a valuable pedigree for preserving genetic diversity (Jakob Lykke Voergaard, product manager, VikingRed, VikingGenetics, personal communication, January 11, 2021). We observed higher frequencies of genetic defects at BTA12 and BTA23 when we removed the requirement to have marketed semen, and we tried to reflect this with the bull set BullAll (Table 8). An economic score including a penalty for mating 2 carriers effectively eliminated expression of genetic defects. It was more profitable to use the carrier bull on a noncarrier female than on a carrier female. Linear programming can help avoid expression of genetic defects unless the possible matings are restricted (e.g., if only a few noncarrier bulls are available and therefore a carrier bull has to be mated with a carrier female). Bérodier et al. (2021) considered known recessive genetic defects similar to this study and found that linear programming was better than random and sequential mating in reducing the number of genetic defects expressed. However, they could not completely avoid the expression of recessive genetic defects, most likely due to a more restricted bull usage compared with our study. For example, only 8 bulls could be mated to heifers due to restriction of calving ease, and they also included restrictions on availability of semen which we did not consider.

It is worth highlighting that even though the overall frequency (Table 2) was low among all females for all defects except genetic defect at BTA12, the carrier frequencies in some herds were much higher than in other herds (Table 3). The carrier frequency in female candidates could be valuable information for farmers and advisors before deciding on matings in practice, by indicating how different defects should be considered in a specific herd.

We observed higher carrier frequencies in the next generation for the genetic defect at BTA23 using BullAll (Table 8) than in the mated females (Table 2). In general, we saw no clear pattern in the economic score that performed best regarding the carrier frequency in the next generation. Further, we believe that the carrier frequency in the next generation is situation specific for the available bull sets, with regard to the NTM ranking of the carrier bulls, constraints, and genetic relationship. Note that higher carrier frequencies in the next generation could be expected if many bulls carrying defect alleles were represented at the top of the total merit ranking. In reality, this is not expected to occur with the current bull selection strategy at VikingGenetics. However, it could occur if bulls to be used in a herd were selected without consideration of their carrier status.

No Penalty for Genetic Defects

In scenario GSEG1_NoDefect, the probability of expression of genetic defect at BTA12 was less than in scenarios AllGen_NoDefect, 3Gen_NoDefect, and GSNP_NoDefect (Table 9), and slightly lower than in scenario GSEG4_NoDefect. According to Wu et al. (2020), the genetic defect at BTA12 region is approximately 2.6 Mb and would not be captured in g_{SEG4}. This might explain why we saw a slightly higher probability of expression of genetic defect at BTA12 in GSEG4_NoDefect compared with GSEG1_NoDefect. Further, in the scenarios aiming to minimize the different genetic relationships, GSEG1_Min had the lowest probability of expression of genetic defect at BTA12, but AllGen_Min and GSEG4_Min also reduced the probability of expression of genetic defect at BTA12 compared with Random and MaxNTM. Hence, it seems that minimizing some genetic relationships also helped lower, or at least did not increase, the probability of expression of genetic defect at BTA12.

Economic Assumptions

In the absence of estimates of RDC inbreeding depression, we used the penalty of €24.8 per 1% increase in inbreeding, which corresponded to the current version of the Swedish mating program penalty of 1 NTM unit per 1% increase in inbreeding. This value is in line

with that estimated for Holstein of US$25 (about €20) (Cole, 2015) or US$24 (Smith et al., 1998). Pryce et al. (2012) used a range up to AU$20 (about €13). When the penalty for 1% increase in inbreeding was increased to €40 or decreased to €10 in our sensitivity analysis, only minor changes in the different average relationships were observed. Furthermore, the average NTM level was kept at the same level, and no expression of known genetic defects was observed. Hence, the mating results seemed not to be sensitive in the inbreeding penalty range tested.

Regarding the economic assumption for the recessive genetic defects, no economic costs have been specifically calculated for the defects considered in our study. Our value of €80 for an early abortion was in line with Segelke et al. (2016), who estimated a cost of €70, and Bérodier et al. (2021) who estimated €75. We assumed the cost of a later abortion or an early calf death to be €160 (genetic defect SMA, BH2, and at BTA23). Oskarsson and Engelbrekts (2015) estimated the cost to be €100 to €150 in Sweden, and NTM calculations estimate the cost to be €200 to €340 (Sørensen et al., 2018). Cole (2015) estimated a stillbirth cost of US$150.

There are also most likely differences between farms within each country, such as costs associated with genetic defects. The economic score is a relatively simple calculation that demands little computer power, and it could be adjusted to match economic conditions on a specific farm.

Implementation Opportunities

Many studies have pointed out that linear programming outperforms sequential mating methods because it uses simultaneous rather than sequential solving to find the economically optimal matings for each herd (Sun et al., 2013; Carthy et al., 2019; Bérodier et al., 2021). Therefore, we decided to focus on linear programming and different economic scores and not compare different mating methods. Once the relationships (and NTM and genetic defects) had been calculated, linear programming on a regular laptop maximized the economic score for all herds studied within seconds. This means that the method is suitable for implementing in mating software to be used by advisors and farmers. The most time-consuming calculation of the whole procedure for mating planning was phasing genotypes and extracting the genomic segments, and today this has to be done on a more powerful computer. Genotype phasing and estimating allele frequencies also require information from more than a single farm. This should thus be done at central level and the genetic relationship coefficients should then be made available for downloading to the mating program. Here, g_{SNP} used allele frequencies in the current population, which are easy to obtain and often used in genomic evaluation. Further, g_{SNP} was the fastest genomic relationship to calculate and it was powerful at keeping both g_{SEG1} and g_{SEG4} low, making it an efficient implementation option. However, a segment-based relationships should be considered if future studies show they better predict inbreeding depression.

We mated all animals in a herd at a specific time, which would not be the case in a real situation because mating planning is usually performed more than once annually for each herd. For example, in Sweden, mating planning is typically performed 3 to 6 times/yr (Thure Bjerketorp, responsible for breeding advisors, Växa Sverige, personal communication, July 27, 2021). However, we were also only able to study animals born in 2019, because older animals were missing information about genetic defects. Hence, in reality there would be more animals from several birth years to mate, and the number of animals we considered will most likely be in line with a typical mating planning. However, mating planning on a subset of the herd, a third at a time, say, can be expected to be somewhat suboptimal.

The mating scenarios presented here could also be adopted by other breeds or other livestock species. However, we believe the detailed planning at the individual level is quite unique for dairy cattle, at least at the commercial herd level. Further, including genomic relationships and information about genetic defects, similar to this study, requires genotypes of both females and males. An economic score could also be developed for crossbred animals where the focus is to maximize heterosis instead of minimizing parent relationships. In this study, we did not consider ungenotyped animals. Other studies have proposed methods to impute ungenotyped animals [e.g., Carthy et al. (2019) used the method proposed by Gengler et al. (2007)], or one could use the combined genomic and pedigree relationship matrix **H** that is used in single-step genomic evaluations, as suggested by Sun et al. (2013).

CONCLUSIONS

We studied mating allocations in RDC 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 expression of the 6 known genetic defects. It was possible to reduce genomic relationships between parents with pedigree measures, but it was best done with genomic measures. More studies are needed to identify the different types of genetic relationships and their future economic impact for farmers. Linear programming maximized 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.

ACKNOWLEDGMENTS

This study was funded by the Swedish project LivsID, with externally employed doctoral students within food-related research. We offer our special thanks to FABA (Hollola, Finland), Nordic Cattle Genetic Evaluation (Aarhus, Denmark), SEGES (Aarhus, Denmark), and Växa Sverige (Uppsala, Sweden) for providing us with data. We also thank Bérodier et al. (2021) for providing their mating R script. The authors have not stated any conflicts of interest.

REFERENCES

Berg, P., J. Nielsen, and M. K. Sørensen. 2006. EVA: Realized and predicted optimal genetic contributions. Book of Abstracts: CD Commun. 27-09. Proc. 8th World Congr. Genet. Appl. Livest. Prod. Belo Horiz., Minas Gerais, Brazil.

Berkelaar and Associates. 2020. LpSolve: Interface to "Lp_solve" v. 5.5 to solve linear/integer programs. Accessed January 26, 2021. https://rdrr.io/cran/lpSolve/.

Bérodier, M., P. Berg, T. Meuwissen, D. Boichard, M. Brochard, and V. Ducrocq. 2021. Improved dairy cattle mating plans at herd level using genomic information. Animal 15:100016. https://doi.org/10.1016/j.animal.2020.100016.

Borchersen, S. 2019. Launch of improved EuroGenomics genotyping beadchip: EuroG MD. Accessed April 27, 2021. https://www.eurogenomics.com/actualites/eurogenomics-new-eurog-md-beadchip.html.

Browning, B. L., and S. R. Browning. 2013. Improving the accuracy and efficiency of identity-by-descent detection in population data. Genetics 194:459–471. https://doi.org/10.1534/genetics.113.150029.

Browning, S. R., and B. L. Browning. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am. J. Hum. Genet. 81:1084–1097. https://doi.org/10.1086/521987.

Burnell, M. 2019. The use of sexed semen in dairy herds. Livestock (Lond) 24:282–286. https://doi.org/10.12968/live.2019.24.6.282.

Carthy, T. R., J. McCarthy, and D. P. Berry. 2019. A mating advice system in dairy cattle incorporating genomic information. J. Dairy Sci. 102:8210–8220. https://doi.org/10.3168/jds.2019-16283.

Charlesworth, D., and J. H. Willis. 2009. The genetics of inbreeding depression. Nat. Rev. Genet. 10:783–796. https://doi.org/10.1038/nrg2664.

Cole, J. B. 2015. A simple strategy for managing many recessive disorders in a dairy cattle breeding program. Genet. Sel. Evol. 47:94. https://doi.org/10.1186/s12711-015-0174-9.

de Cara, M. Á. R., B. Villanueva, M. Á. Toro, and J. Fernández. 2013. Using genomic tools to maintain diversity and fitness in conservation programmes. Mol. Ecol. 22:6091–6099. https://doi.org/10.1111/mec.12560.

EFFAB. 2020. Code of good practice. Accessed February 2, 2021. https://www.responsiblebreeding.eu/uploads/2/3/1/3/23133976/animal_breeding_-_brochure_code-efabar.pdf.

Fikse, F., and M. Kargo. 2020. The economic value of NTM. Accessed February 2, 2021. https://www.nordicebv.info/wp-content/uploads/2020/11/The-economic-value-of-NTM_final.pdf.

Forutan, M., S. Ansari Mahyari, C. Baes, N. Melzer, F. S. Schenkel, and M. Sargolzaei. 2018. Inbreeding and runs of homozygosity before and after genomic selection in North American Holstein cattle. BMC Genomics 19:98. https://doi.org/10.1186/s12864-018-4453-z.

Gautier, M., T. Faraut, K. Moazami-Goudarzi, V. Navratil, M. Foglio, C. Grohs, A. Boland, J.-G. Garnier, D. Boichard, G. M. Lathrop, I. G. Gut, and A. Eggen. 2007. Genetic and haplotypic structure in 14 european and african cattle breeds. Genetics 177:1059–1070. https://doi.org/10.1534/genetics.107.075804.

Gengler, N., P. Mayeres, and M. Szydlowski. 2007. A simple method to approximate gene content in large pedigree populations: Application to the myostatin gene in dual-purpose Belgian Blue cattle. Animal 1:21–28. https://doi.org/10.1017/S1751731107392628.

Guarini, A. R., M. Sargolzaei, L. F. Brito, V. Kroezen, D. A. L. Lourenco, C. F. Baes, F. Miglior, J. B. Cole, and F. S. Schenkel. 2019. Estimating the effect of the deleterious recessive haplotypes AH1 and AH2 on reproduction performance of Ayrshire cattle. J. Dairy Sci. 102:5315–5322. https://doi.org/10.3168/jds.2018-15366.

Henryon, M., H. Liu, P. Berg, G. Su, H. M. Nielsen, G. T. Gebregiwergis, and A. C. Sørensen. 2019. Pedigree relationships to control inbreeding in optimum-contribution selection realise more genetic gain than genomic relationships. Genet. Sel. Evol. 51:39. https://doi.org/10.1186/s12711-019-0475-5.

Hill, W. G., and B. S. Weir. 2011. Variation in actual relationship as a consequence of Mendelian sampling and linkage. Genet. Res. (Camb.) 93:47–64. https://doi.org/10.1017/S0016672310000480.

Kadri, N. K., G. Sahana, C. Charlier, T. Iso-Touru, B. Guldbrandtsen, L. Karim, U. S. Nielsen, F. Panitz, G. P. Aamand, N. Schulman, M. Georges, J. Vilkki, M. S. Lund, and T. Druet. 2014. A 660-Kb deletion with antagonistic effects on fertility and milk production segregates at high frequency in Nordic Red Cattle: Additional evidence for the common occurrence of balancing selection in livestock. PLoS Genet. 10:e1004049. https://doi.org/10.1371/journal.pgen.1004049.

Krebs, S., I. Medugorac, S. Rother, K. Strasser, and M. Forster. 2007. A missense mutation in the 3-ketodihydrosphingosine reductase FVT1 as candidate causal mutation for bovine spinal muscular atrophy. Proc. Natl. Acad. Sci. USA 104:6746–6751. https://doi.org/10.1073/pnas.0607721104.

Maignel, L., D. Boichard, and E. Verrier. 1996. Genetic variability of French dairy breeds estimated from pedigree information. No 14 1996 Proc. Open Sess. Interbull Annu. Meet. 7.

Makanjuola, B. O., F. Miglior, E. A. Abdalla, C. Maltecca, F. S. Schenkel, and C. F. Baes. 2020. Effect of genomic selection on rate of inbreeding and coancestry and effective population size of Holstein and Jersey cattle populations. J. Dairy Sci. 103:5183–5199. https://doi.org/10.3168/jds.2019-18013.

Martikainen, K., M. Koivula, and P. Uimari. 2020. Identification of runs of homozygosity affecting female fertility and milk production traits in Finnish Ayrshire cattle. Scientific Rep. 10:3804. https://doi.org/10.1038/s41598-020-60830-9.

Martikainen, K., A. M. Tyrisevä, K. Matilainen, J. Pösö, and P. Uimari. 2017. Estimation of inbreeding depression on female fertility in the Finnish Ayrshire population. J. Anim. Breed. Genet. 134:383–392. https://doi.org/10.1111/jbg.12285.

Meuwissen, T., and Z. Luo. 1992. Computing inbreeding coefficients in large populations. Genet. Sel. Evol. 24:305–313. https://doi.org/10.1186/1297-9686-24-4-305.

Meuwissen, T. H. E., A. K. Sonesson, G. Gebregiwergis, and J. A. Woolliams. 2020. Management of genetic diversity in the era of genomics. Front. Genet. 11:880. https://doi.org/10.3389/fgene.2020.00880.

NAV. 2019. NAV routine genetic evaluation of dairy cattle – Data and genetic models. Accessed Aug. 22, 2019. https://www.nordicebv.info/wp-content/uploads/2019/06/NAV-routine-genetic-evaluation-04062019.pdf.

Olsen, H. F., S. Tenhunen, N. I. Dolvik, D. I. Våge, and G. Klemetsdal. 2020. Segment-based coancestry, additive relationship and genetic variance within and between the Norwegian and the Swedish Fjord horse populations. Acta Agric. Scand. Anim. Sci. 69:118–126. https://doi.org/10.1080/09064702.2019.1711155.

Oskarsson, M., and E. Engelbrekts. 2015. Kostnader för hälsostörningar hos mjölkkor. February 2, 2021. https://docplayer.se/28330924-Kostnader-for-halsostorningar-hos-mjolkkor.html.

Pryce, J. E., M. Haile-Mariam, M. E. Goddard, and B. J. Hayes. 2014. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. Genet. Sel. Evol. 46:71. https://doi.org/10.1186/s12711-014-0071-7.

Pryce, J. E., B. J. Hayes, and M. E. Goddard. 2012. Novel strategies to minimize progeny inbreeding while maximizing genetic gain using genomic information. 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.

Sahana, G., T. Iso-Touru, X. Wu, U. S. Nielsen, D.-J. de Koning, M. S. Lund, J. Vilkki, and B. Guldbrandtsen. 2016. A 0.5-Mbp deletion on bovine chromosome 23 is a strong candidate for stillbirth in Nordic Red cattle. Genet. Sel. Evol. 48:35. https://doi.org/10.1186/s12711-016-0215-z.

Sargolzaei, M. 2014. SNP1101 User's Guide. Version 1.0.

Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15:478. https://doi.org/10.1186/1471-2164-15-478.

Schwarzenbacher, H., J. Burgstaller, F. R. Seefried, C. Wurmser, M. Hilbe, S. Jung, C. Fuerst, N. Dinhopl, H. Weissenböck, B. Fuerst-Waltl, M. Dolezal, R. Winkler, O. Grueter, U. Bleul, T. Wittek, R. Fries, and H. Pausch. 2016. A missense mutation in TUBD1 is associated with high juvenile mortality in Braunvieh and Fleckvieh cattle. BMC Genomics 17:400. https://doi.org/10.1186/s12864-016-2742-y.

Segelke, D., H. Täubert, F. Reinhardt, and G. Thaller. 2016. Considering genetic characteristics in German Holstein breeding programs. J. Dairy Sci. 99:458–467. https://doi.org/10.3168/jds.2015-9764.

Smith, L. A., B. G. Cassell, and R. E. Pearson. 1998. The effects of inbreeding on the lifetime performance of dairy cattle. J. Dairy Sci. 81:2729–2737. https://doi.org/10.3168/jds.S0022-0302(98)75830-8.

Sonesson, A. K., J. A. Woolliams, and T. H. Meuwissen. 2012. Genomic selection requires genomic control of inbreeding. Genet. Sel. Evol. 44:27. https://doi.org/10.1186/1297-9686-44-27.

Sørensen, L. P., J. Pedersen, M. Kargo, U. S. Nielsen, F. Fikse, J.-Å. Eriksson, J. Pösö, R. S. Stephansen, and P. Aamand. 2018. Review of Nordic total merit index full report November 2018. https://www.nordicebv.info/wp-content/uploads/2018/11/2018.11.06-NTM-2018-report-Full.pdf.

Stoffel, M. A., S. E. Johnston, J. G. Pilkington, and J. M. Pemberton. 2021. Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. Nat. Commun. 12:2972. https://doi.org/10.1038/s41467-021-23222-9.

Strandén, I., and K. Vuori. 2006. RelaX2 program for pedigree analysis, user's guide for version 1.65. Pages 27–30 in Proceedings of the 8th World Congress on Genetics Applied to Livestock Production. Instituto Prociência.

Sun, C., P. M. VanRaden, J. R. O'Connell, K. A. Weigel, and D. Gianola. 2013. Mating programs including genomic relationships and dominance effects. J. Dairy Sci. 96:8014–8023. https://doi.org/10.3168/jds.2013-6969.

Szpiech, Z. A., J. Xu, T. J. Pemberton, W. Peng, S. Zöllner, N. A. Rosenberg, and J. Z. Li. 2013. Long runs of homozygosity are enriched for deleterious variation. Am. J. Hum. Genet. 93:90–102. https://doi.org/10.1016/j.ajhg.2013.05.003.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423. https://doi.org/10.3168/jds.2007-0980.

VanRaden, P. M., K. M. Olson, D. J. Null, and J. L. Hutchison. 2011. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. J. Dairy Sci. 94:6153–6161. https://doi.org/10.3168/jds.2011-4624.

Wang, H., I. Misztal, and A. Legarra. 2014. Differences between genomic-based and pedigree-based relationships in a chicken population, as a function of quality control and pedigree links among individuals. J. Anim. Breed. Genet. 131:445–451. https://doi.org/10.1111/jbg.12109.

Woolliams, J. A., and E. A. Mäntysaari. 1995. Genetic contributions of Finnish Ayrshire bulls over four generations. Anim. Sci. 61:177–187. https://doi.org/10.1017/S1357729800013709.

Wu, X., M. Mesbah-Uddin, B. Guldbrandtsen, M. S. Lund, and G. Sahana. 2020. Novel haplotypes responsible for prenatal death in Nordic Red and Danish Jersey cattle. J. Dairy Sci. 103:4570–4578. https://doi.org/10.3168/jds.2019-17831.

Zhang, Q., B. Guldbrandtsen, M. Bosse, M. S. Lund, and G. Sahana. 2015. Runs of homozygosity and distribution of functional variants in the cattle genome. BMC Genomics 16:542. https://doi.org/10.1186/s12864-015-1715-x.

ORCIDS

C. Bengtsson https://orcid.org/0000-0001-5784-2557
H. Stålhammar https://orcid.org/0000-0002-7354-651X
J. R. Thomasen https://orcid.org/0000-0001-7068-0572
S. Eriksson https://orcid.org/0000-0003-3357-5065
W. F. Fikse https://orcid.org/0000-0003-0578-0184
E. Strandberg https://orcid.org/0000-0001-5154-8146

III

J. Dairy Sci. 106:3359–3375
https://doi.org/10.3168/jds.2022-22926

Mating allocations in Holstein combining genomic information and linear programming optimization at the herd level

C. Bengtsson,[1,2]* H. Stålhammar,[1] J. R. Thomasen,[1] W. F. Fikse,[3] E. Strandberg,[2] and S. Eriksson[2]
[1]VikingGenetics, VikingGenetics Sweden AB, 53294 Skara, Sweden
[2]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Box 7023, 75007 Uppsala, Sweden
[3]Växa Sverige, Växa Sverige, Box 288, 75105 Uppsala, Sweden

ABSTRACT

In this study, we explored mating allocation in Holstein using genomic information for 24,333 Holstein females born in Denmark, Finland, and Sweden. We used 2 data sets of bulls: the top 50 genotyped bulls and the top 25 polled genotyped bulls on the Nordic total merit scale. We used linear programming to optimize economic scores within each herd, considering genetic level, genetic relationship, semen cost, the economic impact of genetic defects, polledness, and β-casein. We found that it was possible to reduce genetic relationships and eliminate expression of genetic defects with minimal effect on the genetic level in total merit index. Compared with maximizing only Nordic total merit index, the relative frequency of polled offspring increased from 13.5 to 22.5%, and that of offspring homozygous for β-casein (A2A2) from 66.7 to 75.0% in one generation, without any substantial negative impact on other comparison criteria. Using only semen from polled bulls, which might become necessary if dehorning is banned, considerably reduced the genetic level. We also found that animals carrying the polled allele were less likely to be homozygous for β-casein (A2A2) and more likely to be carriers of the genetic defect HH1. Hence, adding economic value to a monogenic trait in the economic score used for mating allocation sometimes negatively affected another monogenetic trait. We recommend that the comparison criteria used in this study be monitored in a modern genomic mating program.
Key words: mating allocation, polledness, Nordic total merit, β-casein (A2A2)

INTRODUCTION

Historically, mating programs at the herd level aim to maximize genetic value while minimizing expected inbreeding using pedigree information (Weigel and Lin, 2000). Genotyping provides breeders with new insights at the single nucleotide level that can be used in mating programs. For instance, SNP markers offer the possibility to calculate genomic relationships between potential parents. Genomic estimates of relationships are expected to be more accurate than when using pedigree information, because they do not rely on pedigree completeness or correctness. Genomic relationships can also differentiate between animals with the same pedigree that inherit partly different genetic variants from their parents (VanRaden, 2008; de Cara et al., 2013). In addition, SNP markers provide information about certain known monogenic traits such as defects, as well as some desired traits.

Holstein is the most common cattle breed in Denmark, Finland, and Sweden (**DFS**), with approximately 600,000 milk-recorded cows. Genotyping of females has attracted great interest in DFS in the past decade, and today approximately 25% of all females born are genotyped. However, current (2022) mating programs in the Nordic countries still use pedigree relationship information and ban at-risk matings for recessive genetic defects. The SNP array (Borchersen, 2019) currently used for genotyping in DFS includes 7 Holstein recessive genetic defects, polledness, and β-CN status. Minimizing the risk of obtaining offspring homozygous for recessive genetic defects has an economic value for farmers (Pryce et al., 2012) and is also important for animal health and welfare (EFFAB, 2020).

Other types of monogenic traits, such as horn status, also influence animal welfare. For decades, dehorning of cattle has been common practice. Dehorning is performed for several reasons, including reduced risk of injury to other cattle and improved safety for animal keepers. However, dehorning has been shown to cause behavioral, neuroendocrine, and physiological changes, indicating it to be a stressful and painful experience (Stock et al., 2013). Since 2022, organic farms in the European Union have to seek a permit if they want to dehorn their cattle (EU Commission Regulation No

Received October 19, 2022.
Accepted December 19, 2022.
*Corresponding author: chben@vikinggenetics.com

889/2008; EU, 2008). The cost of dehorning in DFS is estimated to be between €2.7 and €7.3 per head, considering veterinary costs, gas/electricity, and extra labor (Sørensen et al., 2018). However, this estimate does not consider the current situation, in which dehorning is strictly regulated in organic herds in the European Union.

Another example of a monogenic trait of economic importance is β-casein variant. Animals that are homozygotic for the A2 allele produce so-called A2 milk, which is often marketed as a healthier option than regular cow milk, although the human health benefits of consuming A2 milk are still being debated (Summer et al., 2020). Despite this lack of confirmed benefits, some countries are seeking to increase consumption of A2 milk and some dairies pay extra for A2 milk (Bisutti et al., 2022).

The new genetic insights and possibilities available require new methods that combine relevant information based on their economic value when setting up mating plans. Several studies have created economic scoring systems to rank each potential mating (Carthy et al., 2019; Bérodier et al., 2021; Bengtsson et al., 2022). The economic score often includes genetic level, expected inbreeding, the probability of conceiving an offspring homozygous for a genetic defect, and semen price (Bérodier et al., 2021; Bengtsson et al., 2022). The economic score is flexible and can be adjusted to match economic conditions on a specific farm, such as a price premium for A2 milk or polled animals. Using linear programming to maximize every herd's mean economic score, subject to necessary constraints, is a fast and effective method (Carthy et al., 2019; Bérodier et al., 2021). Linear programming has also been shown to outperform other mating methods, such as sequential mate allocation (Sun et al., 2013; Carthy et al., 2019; Bérodier et al., 2021).

Our objective in this study was to investigate the ability of different approaches for mating allocation in DFS Holstein, considering polledness, β-CN, and several recessive genetic defects. We also optimized the mating allocations on total merit index while limiting parent relationships. We investigated all mating allocations at the herd level with real data and used linear programming to optimize different economic scores within each herd.

MATERIALS AND METHODS

Breeding values, pedigree data, SNP data, and data on monogenic traits were obtained from the Nordic Cattle Genetic Evaluation (**NAV**) database (NAV, 2019). No ethical approval was needed for this study because no animal procedures were performed.

Genotype Data

Single nucleotide polymorphism information was available for all genotyped Holstein animals born between 2011 and 2020 in Denmark, Finland, and Sweden. The NAV database uses the Illumina 50k chip (Illumina Inc.) as standard for genomic prediction, and all lower-density chips are imputed by NAV to that format using FImpute (Sargolzaei et al., 2014). The EuroG MD beadchip (Borchersen, 2019) has been used since late 2018. In total, genotypes for 261,198 animals (225,298 females and 35,900 males) were available.

Total Merit

We used Nordic Total Merit (**NTM**) values from the NAV breeding evaluation performed in May 2020, which are expressed in standardized units with a mean of 0 and genetic standard deviation of 10. At the time of data extraction, NTM was composed of 15 sub-indices, covering yield index, longevity, growth, youngstock survival, udder health, udder, feet and legs, frame, hoof health, milkability, daughter fertility, general health, temperament, calving direct, and calving maternal (NAV, 2019).

Data Selection

Females. We selected 289 herds that had genotyped more than 40 Holstein females born in 2019. In total, 24,333 Holstein females were available for mating allocations. The EuroG MD beadchip (Borchersen, 2019), used since late 2018, includes information about all monogenic traits considered in this study (Table 1).

Bulls. We used 2 data sets of bulls, Bull50 and Bull25Polled (Table 2). The main bull data set (Bull50) included the top 50 genotyped bulls on the NTM scale, available from the Nordic breeding cooperative VikingGenetics. The data set Bull25Polled included the top 25 genotyped polled bulls on the NTM scale, also available from VikingGenetics, comprising 21 heterozygous polled (**Pp**) bulls and 4 homozygous polled bulls (**PP**). Bulls in both data sets were born between January 2017 and August 2019. At VikingGenetics, the program EVA (Berg et al., 2006) is used for optimum contribution selection to select breeding animals using pedigree relationships (Hanna Driscoll, product manager Holstein, VikingGenetics; personal communication, January 19, 2022).

Relationship Measures

Pedigree Relationships. Two pedigree relationships were calculated. The first relationship coefficient

Table 1. Description of monogenic traits considered in this study, code used in the Online Mendelian Inheritance in Animals (OMIA) database, and the effect in conceptus or offspring, available with a genomic test in Holstein

Monogenic trait	OMIA code	Description
Holstein Haplotype 1 (HH1)	000001-9913	Early abortion of homozygous conceptus[1]
Holstein Haplotype 3 (HH3)	001824-9913	Early abortion of homozygous conceptus[2]
Holstein Haplotype 4 (HH4)	001826-9913	Early abortion of homozygous conceptus[3]
Holstein Haplotype 6 (HH6)	002194-9913	Early abortion of homozygous conceptus[3]
Holstein Haplotype 7 (HH7)	001830-9913	Early abortion of homozygous conceptus[3]
Bovine leukocyte adhesion deficiency (BLAD)	000595-9913	Extreme susceptibility to infection and early mortality in homozygous offspring[4]
Progressive retinal degeneration (RP1)	000866-9913	Progressive blindness in homozygous offspring[5]
Polledness	000483-9913	Absence of horns in offspring carrying at least one copy of the polled allele (Celtic and Friesian allele considered)[6]
β-CN	002033-9913	A cow produces so-called A2 milk if she has 2 copies of the A2 allele[7]

[1]Adams et al. (2016).
[2]Daetwyler et al. (2014).
[3]Fritz et al. (2013).
[4]Schuster et al. (1992).
[5]Bradley et al. (1982).
[6]Medugorac et al. (2012).
[7]Gallinat et al. (2013).

traced the pedigree 3 generations back from the parents of the potential mating ($\mathbf{a_{3Gen}}$), reflecting the current Nordic mating programs. The second pedigree relationship coefficient was based on all available pedigree information ($\mathbf{a_{AllGen}}$).

For most cases, the pedigree for genotyped animals had already been corrected for mismatches by NAV. We found 143 genotyped animals with missing or mismatching parents, which were excluded from further analyses. The discrete generation equivalent (Woolliams and Mäntysaari, 1995) for the mated animals was 16.0, and the equivalent for complete generations (Maignel et al., 1996) was 12.7. The 5-generation pedigree completeness for mated animals was 99.4%.

Genomic Relationships. Three genomic relationship coefficients were used, one SNP-by-SNP genomic relationship and 2 based on shared genomic segments. The SNP-by-SNP genomic relationship coefficient ($\mathbf{g_{SNP}}$) was calculated according to VanRaden (2008), using the software SNP1101 (Sargolzaei, 2014), as follows:

$$g_{SNP_{ij}} = \frac{\sum_m \left(x_{im} - 2p_m\right) \times \left(x_{jm} - 2p_m\right)}{2\sum_m p_m \left(1 - p_m\right)},$$

where x_{im} and x_{jm} are the genotype scores of animal i and animal j at marker m, coded as 0 = homozygote, 1 = heterozygote, and 2 = alternative homozygote; and p_m is the frequency of the alternative allele of marker m in the founder population. Because we did not know the founder population frequency, the allele frequency

Table 2. Descriptive statistics on the Holstein females and bulls selected for mating allocations

Trait	Females 289 herds	Data set	
		Bull50	Bull25Polled
Number of animals	24,333	50	25
Average Nordic Total Merit (NTM)	12.10	33.93	27.17
Carriers of defect HH1 (%)	3.45	2.00	16.00
Carriers of defect HH3 (%)	3.62	4.00	0.00
Carriers of defect HH4 (%)	1.31	0.00	0.00
Carriers of defect HH6 (%)	0.30	0.00	0.00
Carriers of defect HH7 (%)	0.29	0.00	0.00
Carriers of defect BLAD (%)	0.27	0.00	0.00
Carriers of defect RP1 (%)	0.63	0.00	0.00
Heterozygous polled (Pp) (%)	3.74	14.00	84.00
Homozygous polled (PP) (%)	0.10	0.00	16.00
Heterozygous β-casein (A1A2) (%)	37.11	30.00	44.00
Homozygous β-casein (A2A2) (%)	57.12	66.00	48.00

of all genotyped Holstein was used. Using observed allele frequency instead of founder population frequency is an approximation often used for genomic evaluation (Wang et al., 2014).

The 2 genomic relationship coefficients based on shared genomic segments ($\mathbf{g_{SEG}}$) were calculated following de Cara et al. (2013):

$$g_{SEG_{ij}} = \frac{\sum_k \sum_{ai=1}^{2} \sum_{bj=1}^{2} \left[L_{SEGk} \left(a_i b_j \right) \right]}{2 L_{AUTO}},$$

where L_{SEGk} is the length (in bp) of the kth shared segment measured over homolog a of animal i and homolog b of animal j, and L_{AUTO} is the total length of the autosomes covered by the SNP (in bp).

The 2 segment-based genomic relationship coefficients were based on different minimum lengths of segments: 1 cM ($\mathbf{g_{SEG1}}$) and 4 cM ($\mathbf{g_{SEG4}}$), assuming 1 cM = 1,000,000 bp (Gautier et al., 2007). These segment lengths were chosen to represent short and long segments, similarly to other studies (Zhang et al., 2015; Martikainen et al., 2017; Forutan et al., 2018; Makanjuola et al., 2020). Phasing of genotypes was performed in Beagle 4.1 with default settings (Browning and Browning, 2007), and segments of minimum chosen length were extracted in RefineIBD with the default setting except for the logarithm of odds (**LOD**) score (base 10 log of the likelihood ratio), where we used LOD = 0.1 (Browning and Browning, 2013). The LOD score is used to prune out shared segments that are not common in the population. Hence, default LOD = 3.0 in RefineIBD was considered too high for our purposes, as reported in a recent study (Olsen et al., 2020).

Mate Allocation

Mate allocation was programmed in R version 3.6.3 (https://www.r-project.org/), using the "Lp_solve" package (Berkelaar, 2020). A mating linear programming problem has several integer properties. However, linear programming can be used instead of integer programming because the coefficient matrix has a structure that guarantees integer solutions if the right hand side of the equation are integers (Jansen and Wilton, 1985). Lp_solve is a mixed integer linear programming solver, and hence is suitable for the mating linear programming problem. A mating R script was provided by Bérodier et al. (2021) and modified to allow it to handle favorable monogenic traits. The R script set up constraints considered in linear programming optimization. We used the following constraints: 1 mating per female and a threshold percentage for the maximum number of females per bull and herd, for which we evaluated 2 levels, 5% and 10%, similarly to Bérodier et al. (2021). The threshold for the number of females per bull and herd was in line with current recommendations in DFS.

Economic Score

For each potential mating between female i and bull j, we calculated an economic score:

$$Score_{ij} = \left(\frac{NTM_i + NTM_j}{2} + \lambda \mathrm{F}_{ij} + p\left(\mathrm{BetaC}\right) \times \mathrm{v}_{BetaC} \right) \times \mathrm{prob}\left(\mathrm{Fem}\right) - \sum_{r=1}^{\mathrm{n}_r} p\left(\mathrm{aa}\right)_r \times \mathrm{v}_r + p\left(\mathrm{P}\right) \times \mathrm{v}_\mathrm{P} - \mathrm{semen\ cost},$$

where NTM_i and NTM_j are the values in euros (€) of the Nordic Total Merit units for female i and bull j, λ is the economic consequence of a 1% increase in inbreeding, F_{ij} is the pedigree- or genome-based co-ancestry (relationship/2), p(BetaC) is the probability of a homozygous offspring for β-CN (A2A2), v_{BetaC} is the value of a homozygous offspring for β-CN (A2A2), $\mathrm{prob}\left(\mathrm{Fem}\right)$ is the probability of producing a female conceptus, n_r is the number of recessive genetic defects considered, $p(\mathrm{aa})_r$ is the probability of expression of genetic defect r, v_r is the economic cost associated with recessive genetic defect r, p(P) is the probability of a polled offspring, v_P is the value of a polled offspring, and semen cost is the average amount (€) spent on semen for a pregnancy.

An index unit of NTM is worth €25.4 over the lifetime of a Holstein female in DFS (Fikse and Kargo, 2020). We considered sexed semen with 0.9 probability of producing a female conceptus (Burnell, 2019). The economic consequence of a 1% increase in inbreeding was set to €25.4. The Swedish mating program "Genvägen" uses a penalty of 1 NTM unit per 1% increase in inbreeding, which would correspond to €25.4 (Lina Baudin, expert in breeding routines, Växa Sverige; personal communication, March 5, 2021). This is in line with other studies citing US$25 (about €25; Cole, 2015) and US$24 (Smith et al., 1998).

We assumed the cost of an early abortion (HH1, HH3, HH4, HH6, HH7; Table 1) to be €80, based on the resulting longer calving interval (€30–€40/month) and the cost of extra insemination(s) (€30; Oskarsson and Engelbrekts, 2015; Sørensen et al., 2018). Bulls carrying BLAD and RP1 are not allowed in the breeding program at VikingGenetics, so we did not estimate any cost for them. We tested different economic val-

Table 3. Description of the mating scenarios considered[1]

Scenario	Economic score includes				
	NTM	Relationship	Genetic defect value	Polled value (€)	β-casein value (€)
MaxNTM	Yes	No	No	0	0
3Gen	Yes	a_{3Gen}	Yes	0	0
AllGen	Yes	a_{AllGen}	Yes	0	0
GSNP	Yes	g_{SNP}	Yes	0	0
GSEG1	Yes	g_{SEG1}	Yes	0	0
GSEG4	Yes	g_{SEG4}	Yes	0	0
GSNPPolled10	Yes	g_{SNP}	Yes	10	0
GSNPPolled50	Yes	g_{SNP}	Yes	50	0
GSNPPolled100	Yes	g_{SNP}	Yes	100	0
GSNPBetaC10	Yes	g_{SNP}	Yes	0	10
GSNPBetaC50	Yes	g_{SNP}	Yes	0	50
GSNPBetaC100	Yes	g_{SNP}	Yes	0	100
GSNPPolledBetaC10	Yes	g_{SNP}	Yes	10	10
GSNPPolledBetaC50	Yes	g_{SNP}	Yes	50	50
GSNPPolledBetaC100	Yes	g_{SNP}	Yes	100	100
Random	All possible combinations of females and bulls				

[1]MaxNTM = mating scenario where mates were selected based on maximizing an economic score including Nordic Total Merit (NTM), sexed semen, and semen cost; 3Gen = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a pedigree relationship including 3 generations of ancestors (a_{3Gen}), and a penalty for genetic defects; AllGen = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a pedigree relationship including all available ancestors (a_{AllGen}), and a penalty for genetic defects; GSNP = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship calculated according to VanRaden (2008) (g_{SNP}), and a penalty for genetic defects; GSEG1 = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 1 cM (g_{SEG1}), and a penalty for genetic defects; GSEG4 = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 4 cM (g_{SEG4}), and a penalty for genetic defects; Polled €0, €10, €50, €100 = economic value of a polled offspring, added to the economic score GSNP; BetaC €0, €10, €50, €100 = economic value of an offspring homozygous for β-CN (A2A2), added to the economic score GSNP; Polled BetaC €0, €10, €50, €100 = economic value of a polled offspring and offspring homozygous for β-CN (A2A2), added to the economic score GSNP.

ues (€0, €10, €50, and €100) for polledness and β-CN (A2A2).

We used the prices for sexed semen set by VikingGenetics in 2021. The semen price depends on the bull's NTM and polledness status. A dose of semen from a horned bull with NTM >35, 33 to 34, 30 to 32, and <30 costs €26, €23, €20, and €17, respectively. Semen of polled bulls (homozygous or heterozygous for the polled allele) costs €3 more than semen of horned bulls with the same NTM (Hanna Driscoll, product manager Holstein, VikingGenetics; personal communication, January 19, 2022). Detailed information about the mating scenarios can be found in Table 3. Sexed semen and semen cost were considered in all scenarios. The objective in linear programming was always to maximize the economic score.

Mating Allocation

The suggested planned matings were compared by (1) average NTM; (2) average genetic relationships (a_{3Gen}, a_{AllGen}, g_{SNP}, g_{SEG1}, g_{SEG4}); (3) at-risk matings, as a percentage of matings of 2 carriers of the same recessive genetic defects (the most common defects, HH1 and HH3); (4) average cost of semen for a pregnancy, calculated in the same way as in the economic score; (5) total number of bulls used; (6) number of bulls used to the maximum number of doses allowed on the threshold (5 and 10%) of females per bull and herd; and (7) predicted carrier frequency of HH1 and HH3 in the next generation (%), calculated from the proportion of matings with a carrier (assuming a 50% probability of the defect allele being inherited from a carrier parent); (8) predicted percentage of polled offspring; and (9) predicted percentage of offspring homozygous for β-CN (A2A2) in the next generation.

Statistical Analysis

We used SAS software version 9.4 (SAS Institute Inc.) and R version 3.6.3 (https://www.r-project.org/) for statistical analysis. A chi-squared test was conducted

Table 4. Descriptive statistics on relationships (mean, SD, minimum and maximum values) between all possible combinations of 24,333 females and 50 bulls

Relationship coefficient[1]	Mean	SD	Minimum	Maximum
a_{3Gen}	0.015	0.031	0	0.545
a_{AllGen}	0.132	0.031	0.035	0.647
g_{SNP}	0.010	0.040	−0.106	0.576
g_{SEG1}	0.269	0.042	0.089	0.853
g_{SEG4}	0.181	0.041	0.039	0.763

[1]Coefficients: a_{3Gen} = pedigree relationships using 3 generations of ancestors, a_{AllGen} = 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) cM.

Table 5. Correlation between the different relationship coefficients for all possible combinations of 24,333 females and 50 bulls[1]

	Relationship				
Relationship	a_{3Gen}	a_{AllGen}	g_{SNP}	g_{SEG1}	g_{SEG4}
a_{3Gen}	1	0.95	0.75	0.69	0.70
a_{AllGen}		1	0.76	0.75	0.76
g_{SNP}			1	0.88	0.87
g_{SEG1}				1	0.97

[1]Coefficients: a_{3Gen} = pedigree relationships using 3 generations of ancestors, a_{AllGen} = 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) cM.

in SAS to test association between polledness genotype and HH1, HH3, or β-CN genotype.

RESULTS

The presented mating results are between the 24,333 females selected for matings and data set Bull50, unless otherwise specified.

Genetic Relationship Coefficients

For all possible combinations of females and males, the mean value of the relationship coefficient ranged from 0.010 to 0.269, and the standard deviation ranged from 0.031 to 0.042 (Table 4). For all correlations between different genetic relationship coefficients, the value of correlation coefficient was ≥0.69. The strongest correlation was between g_{SEG1} and g_{SEG4} (r = 0.97). Further, all correlations between a_{AllGen} and genomic relationships were of similar strength (0.75–0.76), whereas those between a_{3Gen} and the genomic relationships showed a wider range (0.69–0.75; Table 5). The coefficients of regression from genomic relationship coefficients on a_{AllGen} were all close to 1. They were highest for g_{SEG1} and g_{SEG4}, and somewhat lower for a_{3Gen} and g_{SNP} (Figure 1).

Mate Allocation

Using Bull50. In scenario MaxNTM, the NTM level improved compared with scenario Random (Table 6), but the genetic relationship did not decrease. Including the cost of the known recessive genetic defects when optimizing mating strategies avoided at-risk matings (mating of 2 animals carrying the same recessive genetic defect). In 3Gen, Allgen, GSNP, GSEG1, and GSEG4, all genetic relationships were decreased compared with Random and MaxNTM. Including pedigree relationships in the economic score decreased genomic relationships compared with Random and MaxNTM, but they were further decreased when using a genomic relationship.

The number of bulls used in the scenarios considering genomic relationships was generally higher (49 to 50) than in the scenarios considering pedigree relationships (32 to 36) and the difference was even larger when allowing 10% females per bull. Furthermore, fewer bulls were used for the maximum number of permitted inseminations considering genomic relationships compared with scenarios considering pedigree relationships with the same constraints. We observed a lower percentage of polled offspring when more bulls were used; for example, 15.7% in scenario 3Gen compared with 7.5% in GSEG4.

Including an extra economic value for the polledness trait in the economic score used for mating allocations increased the expected percentage of polled offspring in the next generation (Table 7). For example, when using a constraint of 5% females per bull and herd, the expected percentage of polled offspring increased from 9.7% in GSNP to 17.0% in GSNPPolled€100. In general, the other mating parameters were minimally affected when adding economic value to the polledness trait, with the same constraints. However, when using a constraint of 10% females per bull, we observed a decline in the expected percentage of β-CN (A2A2) offspring: 66.4% in GSNPPolled€0 and 62.2% in GSNPPolled€100.

Including an economic value for β-CN (A2A2) in the economic score used for mating allocations increased the expected percentage of offspring homozygous for β-CN (A2A2), with a minor effect on the average NTM level and genetic relationships (Table 8). The highest percentage of offspring homozygous for β-CN (A2A2) was observed in Beta-C€100 (75.0%) with a constraint of 10% females per herd and bull. We observed a decline in the expected percentage of polled offspring

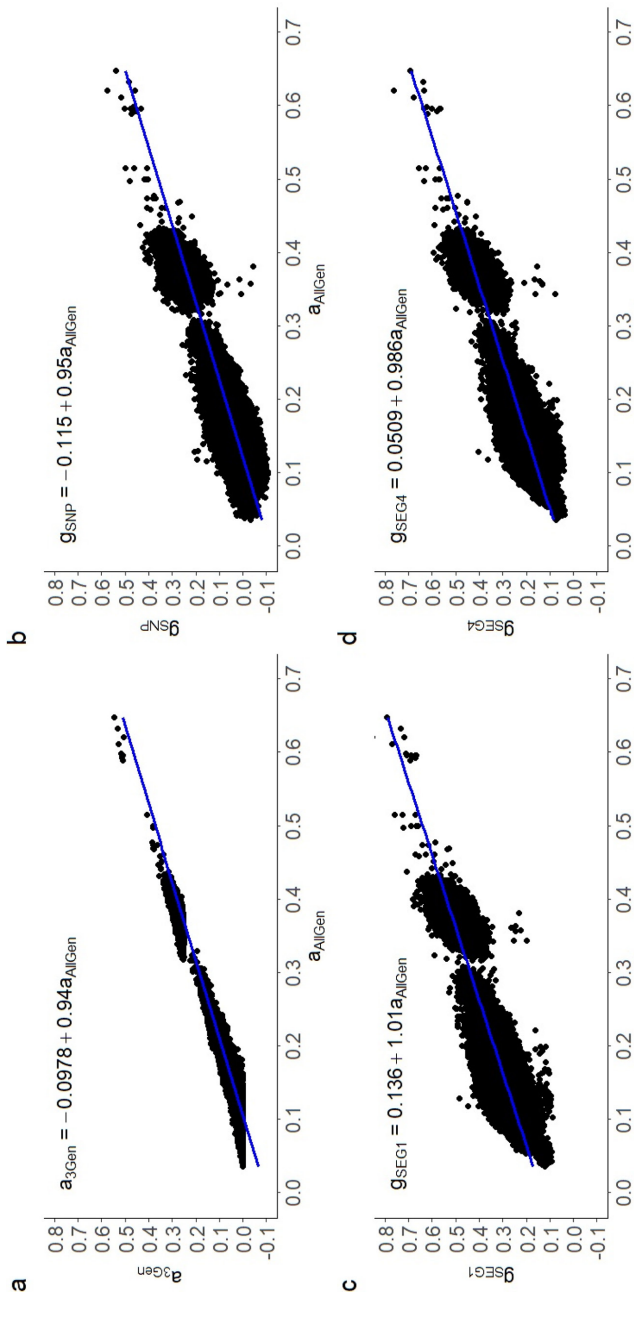


Figure 1. Relationship coefficients estimated from (a) pedigree data with 3 generations of ancestors (a_{3Gen}), (b) SNP data (g_{SNP}; VanRaden, 2008), (c) shared genomic segments with a minimum segment length of 1 cM (g_{SEG1}), and (d) shared genomic segments with a minimum length of 4 cM (g_{SEG4}; de Cara et al., 2013), all plotted against relationship coefficients estimated from pedigree data using all available ancestors (a_{AllGen}). The relationships shown are for all possible combinations of 24,333 Holstein females and 50 bulls.

Table 6. Results of 13 mating scenarios, including 24,333 Holstein females; available bulls were 50 Holstein bulls marketed by VikingGenetics (Bull50)

		5% females/bull scenarios[2]						10% females/bull scenarios[2]					
Comparison criterion[1]	Random	Max NTM	3Gen	AllGen	GSNP	GSEG1	GSEG4	Max NTM	3Gen	AllGen	GSNP	GSEG1	GSEG4
Average Nordic Total Merit (NTM)	23.0	24.3	24.3	24.3	24.2	24.2	24.2	25.4	25.4	25.3	25.2	25.1	25.1
Average a_{3Gen} between parents	0.026	0.028	0.010	0.012	0.018	0.017	0.017	0.026	0.012	0.014	0.018	0.018	0.018
Average a_{AllGen} between parents	0.132	0.132	0.115	0.113	0.121	0.120	0.120	0.129	0.116	0.114	0.121	0.120	0.120
Average g_{SNP} between parents	0.010	0.011	−0.006	−0.009	−0.040	−0.035	−0.033	0.011	−0.003	−0.004	−0.034	−0.030	−0.028
Average g_{SEG1} between parents	0.269	0.268	0.253	0.250	0.226	0.218	0.221	0.268	0.256	0.254	0.231	0.223	0.226
Average g_{SEG4} between parents	0.181	0.180	0.165	0.162	0.141	0.135	0.133	0.180	0.169	0.166	0.146	0.139	0.137
At-risk matings (%)	0.20	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average cost of semen for a pregnancy (€)	44.4	48.8	48.8	48.6	48.4	48.1	48.2	50.7	50.9	50.4	50.2	49.5	49.6
Number of bulls used	NA[3]	30	32	36	50	50	49	13	16	17	46	39	36
Number of bulls used to a maximum	NA	20	16	16	8	7	10	10	8	7	2	3	4
Predicted HH1 carrier frequency in the next generation (%)	2.7	1.7	1.7	1.7	1.8	1.8	1.8	1.7	1.7	1.7	1.7	1.7	1.7
Predicted HH3 carrier frequency in the next generation (%)	3.8	4.1	4.1	4.1	3.9	4.3	4.2	1.8	2.4	3.2	2.6	3.1	3.1
Percentage of polled offspring	8.8	10.6	10.7	9.5	9.7	8.2	8.4	13.5	15.7	11.8	11.7	7.6	7.5
Percentage of homozygous A2A2 offspring	61.2	61.8	60.6	61.1	61.0	62.5	62.1	66.7	64.1	65.3	66.4	67.1	67.2

[1]Average NTM level, 5 different genetic relationships, at-risk matings: percentage of matings with 2 carriers of the same recessive genetic defect HH1, HH3, the average cost of semen for a pregnancy, the number of bulls used, the number of bulls used to a maximum number of doses based on the 5% and 10% constraint of females per bull and herd, predicted genetic defect HH1 and HH3 carrier frequency in the next generation, predicted percentage of polled offspring in the next generation, and predicted percentage of offspring homozygous for β-CN (A2A2) in the next generation.

[2]Maximum percentage of females per bull and herd set to 5% or 10%. MaxNTM = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, and semen cost; 3Gen = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a pedigree relationship including 3 generations of ancestors (a_{3Gen}), and a penalty for genetic defects; AllGen = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a pedigree relationship including all available ancestors (a_{AllGen}), and a penalty for genetic defects; GSNP = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship calculated according to VanRaden (2008) (g_{SNP}), and a penalty for genetic defects; GSEG1 = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 1 cM (g_{SEG1}), and a penalty for genetic defects; GSEG4 = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, and a genomic relationship based on shared genomic segment calculated according to de Cara et al. (2013) with a minimum genomic segment length of 4 cM (g_{SEG4}), and a penalty for genetic defects.

[3]NA = not applicable.

Table 7. Results of 9 mating scenarios investigating extra economic value for the polledness trait, including 24,333 Holstein females; available bulls were 50 Holstein bulls marketed by VikingGenetics (Bull50)

		5% females/bull scenarios[2]				10% females/bull scenarios[2]			
		GSNP				GSNP			
Comparison criterion[1]	Random	Polled €0	Polled €10	Polled €50	Polled €100	Polled €0	Polled €10	Polled €50	Polled €100
Average Nordic Total Merit (NTM)	23.0	24.2	24.2	24.2	24.1	25.2	25.2	25.2	25.0
Average a_{AllGen} between parents	0.132	0.121	0.121	0.121	0.122	0.121	0.121	0.121	0.122
Average g_{SNP} between parents	0.010	−0.040	−0.040	−0.040	−0.039	−0.034	−0.034	−0.034	−0.034
At-risk matings (%)	0.2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average cost of semen for a pregnancy (€)	44.4	48.4	48.5	48.7	48.7	50.2	50.3	50.7	50.8
Number of bulls used	NA	50	50	49	49	46	46	45	41
Number of bulls used to a maximum	NA	8	8	7	7	2	2	3	3
Predicted HH1 carrier frequency in the next generation (%)	2.7	1.8	1.8	2.3	3.8	1.7	1.7	1.8	2.4
Predicted HH3 carrier frequency in the next generation (%)	3.8	3.9	3.8	3.7	3.4	2.6	2.5	2.4	2.2
Percentage of polled offspring	8.8	9.7	10.3	13.2	17.0	11.7	12.6	16.4	22.5
Percentage of homozygous A2A2 offspring	61.2	61.0	60.9	60.8	60.8	66.4	65.9	64.0	62.2

[1]Average NTM level, 2 different genetic relationships, at-risk matings: percentage of matings with 2 carriers of the same recessive genetic defect HH1, HH3, the average cost of semen for a pregnancy, the number of bulls used, the number of bulls used to a maximum number of doses based on the 5% and 10% constraint of females per bull and herd, predicted genetic defect HH1 and HH3 carrier frequency in the next generation, predicted percentage of polled offspring in the next generation, and predicted percentage of offspring homozygous for β-CN (A2A2) in the next generation. a_{AllGen} = a pedigree relationship including all available ancestors.

[2]Maximum percentage of females per bull and herd set to 5% or 10%. GSNP = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship calculated according to VanRaden (2008) (g_{SNP}), and penalty for genetic defects. Polled €0, €10, €50, €100 = economic value of a polled offspring, added to the economic score GSNP.

Table 8. Results of 9 mating scenarios investigating extra economic value for β-CN (A2A2), including 24,333 Holstein females; available bulls were 50 Holstein bulls marketed by VikingGenetics (Bull50)

		5% females/bull scenarios[2]				10% females/bull scenarios[2]			
		GSNP				GSNP			
Comparison criterion[1]	Random	BetaC €0	BetaC €10	BetaC €50	BetaC €100	BetaC €0	BetaC €10	BetaC €50	BetaC €100
Average Nordic Total Merit (NTM)	23.0	24.2	24.2	24.2	24.1	25.2	25.2	25.2	25.1
Average a_{AllGen} between parents	0.132	0.121	0.121	0.121	0.122	0.121	0.121	0.121	0.121
Average g_{SNP} between parents	0.010	−0.040	−0.040	−0.040	−0.039	−0.034	−0.034	−0.034	−0.033
At-risk matings (%)	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average cost of semen for a pregnancy (€)	44.4	48.4	48.4	48.2	47.7	50.2	50.2	49.9	49.7
Number of bulls used	NA	50	50	48	48	46	45	46	46
Number of bulls used to a maximum	NA	8	8	6	6	2	2	3	2
Predicted HH1 carrier frequency in the next generation (%)	2.7	1.8	1.8	1.8	1.9	1.7	1.7	1.7	1.8
Predicted HH3 carrier frequency in the next generation (%)	3.8	3.9	3.5	2.4	2.3	2.6	2.3	1.9	1.9
Percentage of polled offspring	8.8	9.7	9.5	8.8	7.9	11.7	11.2	9.3	8.0
Percentage of homozygous A2A2 offspring	61.2	61.0	62.9	69.2	72.9	66.4	68.2	72.8	75.0

[1]Average NTM level, 2 different genetic relationships, at-risk matings: percentage of matings with 2 carriers of the same recessive genetic defect HH1, HH3, the average cost of semen for a pregnancy, the number of bulls used, the number of bulls used to a maximum number of doses based on the 5% and 10% constraint of females per bull and herd, predicted genetic defect HH1 and HH3 carrier frequency in the next generation, predicted percentage of polled offspring in the next generation, and predicted percentage of offspring homozygous for β-CN (A2A2) in the next generation. a_{AllGen} = a pedigree relationship including all available ancestors.

[2]Maximum percentage of females per bull and herd set to 5% or 10%. GSNP = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship calculated according to VanRaden (2008) (g_{SNP}), and a penalty for genetic defects. BetaC €0, €10, €50, €100 = economic value of an offspring homozygous for β-CN (A2A2), added to the economic score GSNP.

when adding economic value to β-CN (A2A2) in the economic score.

Adding economic value to both the polledness trait and β-CN (A2A2) in the economic score used for mating allocations increased the expected number of polled offspring and offspring homozygous for β-CN (A2A2) compared with GSNPPolled€0 (Table 9). Using both constraints of 5% and 10% females per herd and bull, a simultaneous increase in the 2 traits occurred as the economic value increased.

Using BullPolled25. When 25 polled bulls (21 Pp bulls, 4 PP bulls) were available for mating allocations, it was possible to further increase the expected percentage of polled offspring (Table 10). For example, when using BullPolled25 and a constraint of 5% females per herd and bull, the expected percentage of polled offspring was 60.1% in GSNPPolled100€, compared with 17.0% using Bull50. Considering the same example, the average NTM level was 20.2 using BullPolled25 compared with 24.1 using Bull50. The average genetic relationships using BullPolled25 were slightly higher than those using Bull50 with the same constraints and economic scores. The expected percentage of offspring homozygous for β-CN (A2A2) was lower and the predicted HH1 carrier frequency was higher, for BullPolled25 compared with Bull50.

Association Between Monogenic Traits

Among the 24,333 mated females, polled females (Pp and PP) were less likely to be homozygous for β-CN (A2A2) (or A2A2 females were less likely to carry the polled allele; Figure 2). For example, 58% of the horned females but only 44% of the heterozygous polled (Pp) females were homozygous for β-CN (A2A2). The chi-squared test showed a significant unfavorable association between polled and β-CN genotype ($P < 0.0001$) in the data. Polled females were also more likely to be HH1 carriers (or HH1 carriers were more likely to be polled). For example, 23% of the heterozygous polled females were carriers of HH1, whereas only 3% of the horned animals were carriers (Figure 3). The chi-squared test showed a significant unfavorable association between polledness and HH1 genotype ($P < 0.0001$) in the study data. We observed no association between polledness and HH3 genotype (results not shown).

DISCUSSION

We explored mating allocations in Holstein dairy cattle, taking into account genomic information. The results showed that it was possible to reduce genetic relationships and eliminate expression of genetic defects with minimal effect on the genetic level, as we found previously in a study on Red Dairy Cattle (Bengtsson et al., 2022). The results also showed that it was possible to increase the percentage of polled offspring substantially in one generation when competitive bulls were available, without any significant negative effect on other comparison criteria. It was also possible to increase the number of homozygous β-CN (A2A2) offspring without any negative effect on other comparison criteria. Using only semen from polled bulls, which might be necessary if dehorning is banned, had a substantial impact at the genetic level. We also found that animals in this study carrying the polled allele were less likely to be homozygous for β-CN (A2A2) and more likely to be carriers of the genetic defect HH1. Hence, adding economic value to a monogenic trait in the economic score used for mating allocations sometimes negatively affected another monogenetic trait. Therefore, it may be necessary to monitor comparison criteria, as used in this study, in a modern genomic mating program.

Breeding for the Polledness Trait

Polled calves can easily be achieved by mating all females to homozygous (PP) bulls. However, no homozygous polled bulls were available in Bull50. Other authors have highlighted the absence of competitive homozygous polled bulls (Spurlock et al., 2014; Mueller et al., 2019). The reason for the difference in genetic level is not clear. Other authors have hypothesized that it could be due to lack of selection emphasis on production traits of polled bulls. Alternatively, it could be due to pleiotropic effects of chromosomal segments, or genes linked to the polled locus could contribute to a poorer genetic level for production traits (Spurlock et al., 2014). At the population level, it has been shown to take somewhere between 10 and 25 generations to get most bulls homozygous polled, from a starting allele frequency of 0.03 (Scheper et al., 2016), which is between the polled allele frequency of the mated females and bulls in this study. The large difference in number of generations required depends on many factors, including available tools such as level of genotyping and the goal of genetic gain and inbreeding. Hence, 100% homozygous (PP) bulls cannot be expected in the DFS Holstein population in the near future.

In this study, the economic value for the polledness trait had to be higher than €50 before we observed a fundamental change in the expected number of polled offspring (Table 7). We observed a lower percentage of polled offspring when more bulls were used; for example, 15.7% in 3Gen compared with 7.5% in GSEG4. This was because heterozygous polled bulls were more commonly ranked in the top half of the Bull50 data set than in the bottom half. The high ranking of hetero-

Table 9. Results of 9 mating scenarios investigating extra economic value for the polledness trait and β-CN (A2A2), including 24,333 Holstein females; available bulls were 50 Holstein bulls marketed by VikingGenetics (Bull50)

		5% females/bull scenarios[2]				10% females/bull scenarios[2]			
		GSNP				GSNP			
Comparison criterion[1]	Random	Polled BetaC €0	Polled BetaC €10	Polled BetaC €50	Polled BetaC €100	Polled BetaC €0	Polled BetaC €10	Polled BetaC €50	Polled BetaC €100
Average Nordic Total Merit (NTM)	23.0	24.2	24.2	24.2	24.1	25.2	25.2	25.2	25.1
Average a_{AllGen} between parents	0.132	0.121	0.121	0.122	0.122	0.121	0.121	0.121	0.122
Average g_{SNP} between parents	0.010	−0.040	−0.04	−0.039	−0.038	−0.034	−0.034	−0.033	−0.032
At-risk matings (%)	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average cost of semen for a pregnancy (€)	44.4	48.4	48.5	48.6	48.5	50.2	50.3	50.5	50.4
Number of bulls used	NA	50	50	48	48	46	45	44	41
Number of bulls used to a maximum	NA	8	8	7	7	2	2	2	2
Predicted HH1 carrier frequency in the next generation (%)	2.7	1.8	1.8	2.7	4.0	1.7	1.7	1.9	3.0
Predicted HH3 carrier frequency in the next generation (%)	3.8	3.9	3.5	2.2	2.0	2.6	2.3	1.9	1.8
Percentage of polled offspring	8.8	9.7	10.2	13.1	15.4	11.7	12.2	14.1	18.5
Percentage of homozygous A2A2 offspring	61.2	61.0	62.8	68.5	71.1	66.4	67.7	70.4	71.8

[1]Average NTM level, 2 different genetic relationships, at-risk matings: percentage of matings with 2 carriers of the same recessive genetic defect HH1, HH3, the average cost of semen for a pregnancy, the number of bulls used, the number of bulls used to a maximum number of doses based on the 5% and 10% constraint of females per bull and herd, predicted genetic defect HH1 and HH3 carrier frequency in the next generation, predicted percentage of polled offspring in the next generation, and predicted percentage of offspring homozygous for β-CN (A2A2) in the next generation. a_{AllGen} = a pedigree relationship including all available ancestors.

[2]Maximum percentage of females per bull and herd set to 5% or 10%. GSNP = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship calculated according to VanRaden (2008) (g_{SNP}), and a penalty for genetic defects. Polled BetaC €0, €10, €50, €100 = economic value of a polled offspring and offspring homozygous for β-CN (A2A2), added to the economic score GSNP.

Table 10. Results of 9 mating scenarios investigating extra economic value for the polledness trait, including 24,333 Holstein females; available bulls were 25 polled Holstein bulls marketed by VikingGenetics (BullPolled25)

		5% females/bull scenarios[2]				10% females/bull scenarios[2]			
		GSNP				GSNP			
Comparison criterion[1]	Random	Polled €0	Polled €10	Polled €50	Polled €100	Polled €0	Polled €10	Polled €50	Polled €100
Average Nordic Total Merit (NTM)	19.6	20.2	20.2	20.2	20.2	22.4	22.4	22.4	22.3
Average a_{AllGen} between parents	0.129	0.121	0.121	0.121	0.121	0.125	0.125	0.125	0.125
Average g_{SNP} between parents	0.010	−0.034	−0.034	−0.034	−0.034	−0.028	−0.028	−0.027	−0.027
At-risk matings (%)	0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average cost of semen for a pregnancy (€)	41.42	41.91	41.91	41.91	41.91	46.10	46.10	46.10	46.10
Number of bulls used	NA	25	25	25	25	20	20	19	19
Number of bulls used to a maximum	NA	16	17	17	17	7	6	6	5
Predicted HH1 carrier frequency in the next generation (%)	10.0	10.2	10.1	10.1	10.1	8.0	7.9	7.5	8.0
Predicted HH3 carrier frequency in the next generation (%)	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8
Percentage of polled offspring	59.0	59.5	59.7	60.0	60.1	53.6	54.3	57.3	60.9
Percentage of homozygous A2A2 offspring	53.0	54.3	54.3	54.3	54.3	57.7	57.7	57.9	56.0

[1]Average NTM level, 2 different genetic relationships, percentage of matings with 2 carriers of the same genetic defect (HH1, HH3), the average cost of semen for a pregnancy, the number of bulls used, the number of bulls used to a maximum number of doses based on the 5% and 10% constraint of females per bull and herd, predicted genetic defect HH1 and HH3 carrier frequency in the next generation, predicted percentage of polled offspring in the next generation, and predicted percentage of offspring homozygous for β-CN (A2A2) in the next generation. a_{AllGen} = a pedigree relationship including all available ancestors.

[2]Maximum percentage of females per bull and herd set to 5% or 10%. GSNP = mating scenario where mates were selected based on maximizing an economic score including NTM, sexed semen, semen cost, a genomic relationship calculated according to VanRaden (2008) (g_{SNP}), and a penalty for genetic defects. Polled €0, €10, €50, €100 = economic value of a polled offspring, added to the economic score GSNP.

zygous polled bulls on the NTM scale was surprising compared with the findings of Spurlock et al. (2014) and Mueller et al. (2019), where polled bulls were not competitive on net merit. We believe that the high ranking of heterozygous polled bulls is mainly a coincidence. However, it sets the Nordic Holstein in a good position to spread the polled allele without compromising on genetic level. When using BullPolled25, the frequency of polled offspring further increased to 60.9% in GSNP-polled€100, using a constraint of 10% females per herd and bull. However, the NTM level was lower using BullPolled25 compared with using Bull50. We also observed a negative effect on the number of offspring homozygous for β-CN (A2A2) and more carriers of the genetic defect HH1 in the next generation. Hence, the benefit of having more polled animals should be weighed carefully against the negative effect that this might have on other comparison criteria.

Some of the homozygous polled bulls in BullPolled25 were not used to their maximum allowed usage, mainly because their genetic level was too low (Table 10). For example, using a constraint of 10% females per herd and bull, the number of polled offspring did not increase compared with using the 5% females per herd bull constraint, even if the homozygous polled bulls were allowed to be used more. Hence, the highest economic value for polledness (€100) considered in this study was not enough for the homozygous polled bulls to be used to their maximum allowed usage.

Using BullPolled25, the predicted number of carriers of the genetic defect HH1 increased substantially in the next generation compared with the number of HH1 carriers among the mated females (Table 10). Hence, as we observed for the mated females (Figure 3), the top polled bulls on the NTM scale seemed more likely to be HH1 carriers. We believe it is unlikely that only polled bulls would be used for the whole Nordic Holstein population; thus, the increase in the number of HH1 carriers would be smaller in practice. Breeding companies could also limit the usage of polled carriers of HH1, for example, by stopping selling polled carriers of HH1 after fewer doses than usual. Despite the higher percentage of bulls carrying HH1, at-risk mating could be avoided. We argue that these scenarios show the importance of monitoring genetic defects at the population level because, even if at-risk matings were avoided, there could be a risk of genetic defects increasing in frequency.

Breeding for the β-Casein Trait

It was possible to increase the percentage of offspring homozygous for β-CN (A2A2) with a minor effect on the average NTM level and genetic relationships (Table 8). The A2 allele has been associated with a positive effect on milk yield traits (Freyer et al., 1999). Olenski et al. (2010) found a positive effect on milk and protein yield, but a negative effect on fat percentage. Our results confirm that the A2A2 bulls were at a competitive NTM level. The highest expected percentage of offspring homozygous for β-CN (A2A2) was observed in Beta-C€100 (75.0%), achieved with 10% females per herd and bull constraint, compared with 61.2% in the Random, β-CN, and Polledness scenarios. This difference was because homozygous animals are required to achieve the desired milk type for β-CN, whereas only one polled allele is needed to achieve the preferred phenotype for polledness. Hence, achieving 100% desired milk type for β-CN in one generation by only using β-CN (A2A2) bulls is impossible if the A1 allele is still segregating, as in the Nordic Holstein female population.

Breeding for Both Polledness and β-Casein

The expected number of polled offspring declined when adding value to β-CN (A2A2) in the economic score. Hence, bulls that were A2A2 were less likely to carry the polled allele (or polled bulls were less likely to be A2A2). This was also the case for the mated females (Figure 2). To our knowledge, no other study has investigated this. However, when giving both polledness and β-CN an economic value in the economic score, it was possible to increase the number of polled offspring and offspring homozygous for β-CN (A2A2) simultaneously, with little effect on NTM (Table 9).

Other Mating Studies

A few recent studies have used linear programming for genomic mating allocation (Carthy et al., 2019; Bérodier et al., 2021; Bengtsson et al., 2022). Carthy et al. (2019) only included genetic level and a genetic relationship in their economic score, whereas Bérodier et al. (2021) and Bengtsson et al. (2022) used an economic score similar to our scenarios 3Gen, Allgen, GSNP, GSEG1, and GSEG4. Bérodier et al. (2021) found that linear programming was better than random and sequential mating in reducing the number of recessive genetic defects expressed. However, they could not completely avoid the expression of recessive genetic defects due to restrictions in the matings. For example, only 8 bulls could be mated to heifers due to restrictions for calving ease. Bengtsson et al. (2022) found that at-risk mating could be avoided if the economic value for recessive genetic defects were included in the economic

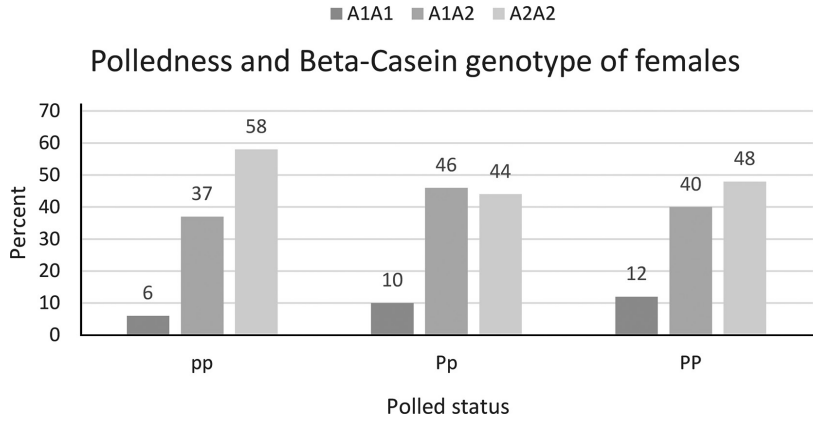


Figure 2. Polledness (pp = horned, Pp = heterozygous polled, PP = homozygous polled) and β-CN genotype (A1A1, A1A2, and A2A2) among the 24,333 mated Holstein females. A cow homozygous for the A2 allele produces so-called A2 milk.

score, which is similar to our findings in 3Gen, Allgen, GSNP, GSEG1, and GSEG4 scenarios (Table 6). There were also higher genetic defect carrier frequencies, up to 14% among females and available bulls, in Bengtsson et al. (2022). Hence, we argue that linear programming can help avoid the expression of genetic defects unless possible matings are restricted (only a few noncarrier bulls are available and a carrier female has to be mated with a carrier bull).

Genetic Relationships

We found correlations between pedigree relationship and genomic relationship estimates of $\geq$0.69 for a_{3Gen} and $\geq$0.75 for a_{AllGen}, which were within the range reported in other studies (0.57–0.88; VanRaden et al., 2011; Pryce et al., 2012; Carthy et al., 2019; Bengtsson et al., 2022). Pedigree depth is important for a strong correlation between pedigree and genomic relationships

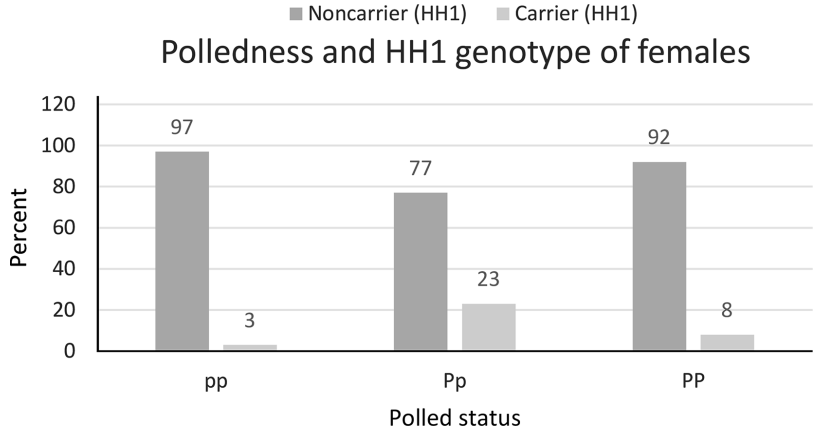


Figure 3. Polledness (pp = horned, Pp = heterozygous polled, PP = homozygous polled) and carrier status for the HH1 genotype among the 24,333 mated Holstein females. A conceptus homozygous for the HH1 allele results in an early abortion.

(Pryce et al., 2012). This was evident in our study, where we found stronger correlations between a_{AllGen} and genomic relationship than between a_{3Gen} and genomic relationship. In our previous study in Red Dairy Cattle (Bengtsson et al., 2022), we found stronger correlations ($\geq$0.83) between pedigree and genomic relationship estimates than were found for Holstein in this study. Pedigree depth was similar to that in Bengtsson et al. (2022), so the difference is most likely linked to some other factor(s). One possibility is that the pedigree correctness is greater in Red Dairy Cattle than in Holstein due to the less common exchange of bulls and their pedigrees worldwide for Red Dairy Cattle, where most animals are kept within the Nordic countries.

There are several arguments for using genomic estimates of relationship and inbreeding instead of pedigree. First, they do not rely on pedigree data, which can be incorrect or have limited depth (Carthy et al., 2019; Makanjuola et al., 2020). Our data were corrected for possible mismatches by the Nordic Cattle Genetic Evaluation, and hence we did not explore the benefit that genomic information brings in the form of assigning the right parents to an animal. Approximately 5% of genotyped animals in Sweden have at least one parent incorrectly assigned (Lina Baudin, expert in breeding routines, Växa Sverige; personal communication, March 5, 2021). Second, even if the pedigree data are correct and complete, genomic relationships are still more accurate because they consider the fact that the genome is transmitted in chromosomes and not as infinite unlinked loci (Hill and Weir, 2011). Third, the assumption of 50% probability of an allele being selected is not true in a population under selection (Forutan et al., 2018). Hence, we argue that genomic estimates should be prioritized in a modern mating program.

In general, genomic relationships were good at keeping each other low when included in an economic score used for mating allocations, and the largest benefit would be to implement one of these instead of pedigree relationships. Using a segment-based relationship, we aimed to reduce the number of runs of homozygosity (**ROH**) in the potential offspring. In a meta-analysis on the effects of inbreeding in livestock, Doekes et al. (2021) showed that genomic measures were a better indicator of inbreeding depression than pedigree measures, but found no differences between SNP-based measures and ROH. However, those authors highlighted the limited number of studies investigating ROH and inbreeding depression and scale and arbitrary definitions of ROH. In principle, ROH are enriched for deleterious alleles that mainly cause inbreeding depression (Charlesworth and Willis, 2009). Long ROH reflect new inbreeding and are expected to contain more deleterious alleles than short ROH, due to purging and recombination through the generations (Stoffel et al., 2021). Pryce et al. (2014) found that long regions (>3 Mb) were associated with inbreeding depression for milk yield in Holstein and Jersey cattle. However, Zhang et al. (2015) found that enrichment of deleterious variants was significantly higher in short (<0.1 to 3 Mb) than in long (>3 Mb) regions in the Holstein, Red Dairy Cattle, and Jersey. Hence, the optimal segment length for use in segment-based relationships remains to be determined. However, we showed that g_{SEG1} and g_{SEG4} kept each other low when included in an economic score, so the difference is most likely marginal for the outcome of the mating allocations.

The number of bulls used in the scenarios considering genomic relationships was, in general, higher than in the scenarios considering pedigree relationships (Table 6). We believe that the primary explanation for this is that genomic relationships can capture variations not detected by the pedigree, which makes some of the lower ranked bulls on the NTM scale being used more.

Economic Assumptions

The cost of dehorning in DFS is estimated to range between €2.7 and €7.3 per animal, considering veterinary costs, gas/electricity, and extra labor (Sørensen et al., 2018). Thompson et al. (2017) estimated the cost of dehorning in the United States to be between $6 and $25 per head. However, such calculations do not consider the current situation, where dehorning is strictly regulated in organic herds in the European Union. If dehorning is completely banned, farmers may be more or less forced to breed polled animals. Consequently, it is difficult to place an economic value on the polledness trait. We tackled that problem by testing a large range of economic values of the polledness trait. In addition, we used only polled bulls in BullPolled25, to represent a situation where farmers are forced to breed polled animals.

In some countries, demand for and the price of A2 milk have increased (Bisutti et al., 2022). For a farmer aiming to produce A2 milk, a female not carrying 2 copies of the A2 allele might be substantially less valuable than a female that does. The exact value for A2 milk is difficult to quantify, and most likely varies between farms. In DFS, the demand for A2 milk is still limited, to our knowledge. Hence, we believe it is uncommon for farmers in DFS to breed to increase the percentage of A2A2 offspring, and even more uncommon to breed for β-CN and more polled animals simultaneously. However, our results for β-CN and polledness illustrate the interactions that can occur when breeding for 2 favorable monogenic traits. It is also likely that new monogenic traits (e.g., κ-casein) will be added to the

SNP array (Chessa et al., 2020) or unknown monogenic traits may be discovered. The methods used in this study could also be adopted by other breeds or livestock species where other monogenic traits may be of economic importance.

The defects we considered in mating allocations all cause early abortions. Our value of €80 for an early abortion was in line with Segelke et al. (2016), who estimated a cost of €70, and Bérodier et al. (2021), who estimated a cost of €75. There are differences between countries in the cost of an insemination (Sørensen et al., 2018). The economic score could be made more farm-specific by adjusting the calculation to match the conditions on a specific farm.

We used a penalty of €25.4 per 1% increase in inbreeding, which is in line with the US$25 (about €20) used by Cole (2015) and US$24 used by Smith et al. (1998). Pryce et al. (2012) used a range up to AU$20 (about €13), whereas Bengtsson et al. (2022) tested €10 to €40 and found that mating results were not sensitive in that range. Hence, even if the cost for inbreeding in Nordic Holstein is still unknown, €25.4 appears to be a reasonable estimate.

Implementation Opportunities

We decided to use linear programming in this study because it has been shown to outperform other mating methods such as sequential solving (Sun et al., 2013; Carthy et al., 2019). When data on genetic relationships, NTM, and monogenic traits were available, linear programming using a regular laptop maximized the economic score within seconds for the herds studied. Hence, the method is suitable for implementation in mating software that farmers or advisors can use. Genotype phasing and extracting the genomic segments was the most time-consuming calculation, and required a more powerful computer. Further, estimating allele frequencies and genotype phasing require information from more than one farm. Therefore, we suggest that this be done at a central level, like today's breeding value estimation, and that genetic relationships could then be made available for downloading to the mating program. In this study, g_{SNP} was the fastest genomic relationship to calculate and it was relatively good at keeping the segment-based relationships low, making it an efficient implementation alternative. However, computation time aside, a segment-based relationship should be considered, because it is most likely better in prediction of inbreeding depression.

In this study, we optimized matings with a within-herd focus and only looked one generation ahead. Future studies should address how this type of mating allocation would affect a population over several generations. Matings optimal at the herd level are not necessarily optimal for the population. Hence, the mating allocation suggested in this study should not be seen as a replacement for optimum contribution selection for breeding organizations.

Breeders of other livestock species could also adopt the mating scenarios presented here, but they would need to be adopted to each specific situation. Further, including genomic relationships and information about genetic defects, as in this study, requires genotypes from both females and males. An economic score could also be developed for crossbred animals where the focus is to maximize heterosis instead of minimizing parent relationships. In this study, we did not consider ungenotyped animals. An option for ungenotyped animals could be to impute their genotype, as done by Carthy et al. (2019) using the method described by Gengler et al. (2007). Sun et al. (2013) suggested use of the **H** matrix in single-step genomic evaluation. However, farmers who do not genotype their females might have to avoid using carrier bulls to completely avoid at-risk mating for known genetic defects.

CONCLUSIONS

We explored mating allocations at the herd level with real data and found that it was possible to reduce genetic relationships and eliminate expression of genetic defects with minimal effect on the genetic level for NTM. It was also possible to increase the percentage of polled and β-CN homozygous (A2A2) offspring substantially in one generation when competitive bulls were available, without any significant negative effect on other mating criteria. Compared with maximizing only NTM index, the frequency of polled offspring increased from 13.5 to 22.5%, and that of offspring homozygous for β-CN (A2A2) from 66.7 to 75.0%, in one generation, without any substantial negative effect on other comparison criteria. Using only semen from polled bulls, which might be necessary if dehorning is banned, considerably affected the genetic level. We also found that animals in the data set carrying the polled allele were less likely to be homozygous for β-CN (A2A2) and more likely to be carriers of the genetic defect HH1. Based on this, we recommend monitoring of the comparison criteria used in this study in a modern genomic mating program.

ACKNOWLEDGMENTS

This study was funded by the Swedish project LivsID, with externally employed doctoral students within food-related research. We offer our special thanks to FABA (Hollola, Finland), Nordic Cattle Genetic Evalu-

ation (Aarhus, Denmark), SEGES (Aarhus, Denmark), and Växa Sverige (Uppsala, Sweden) for providing us with data. We also thank Bérodier et al. (2021) for providing their mating R script. The authors have not stated any conflicts of interest.

REFERENCES

Adams, H. A., T. S. Sonstegard, P. M. VanRaden, D. J. Null, C. P. Van Tassell, D. M. Larkin, and H. A. Lewin. 2016. Identification of a nonsense mutation in APAF1 that is likely causal for a decrease in reproductive efficiency in Holstein dairy cattle. J. Dairy Sci. 99:6693–6701. https://doi.org/10.3168/jds.2015-10517.

Bengtsson, C., H. Stålhammar, J. R. Thomasen, S. Eriksson, W. F. Fikse, and E. Strandberg. 2022. Mating allocations in Nordic Red Dairy Cattle using genomic information. J. Dairy Sci. 105:1281–1297. https://doi.org/10.3168/jds.2021-20849.

Berg, P., J. Nielsen, and M. K. Sørensen. 2006. EVA: Realized and predicted optimal genetic contributions. Page 27-09 in Proc. 8th World Congr. Genet. Appl. Livest. Prod. Belo Horizonte, Brazil. Instituto Prociência, Minas Gerais, Brazil.

Berkelaar, M. 2020. lpSolve: interface to "Lp_solve" v. 5.5 to solve linear/integer programs. https://cran.r-project.org/web/packages/lpSolve/index.html.

Bérodier, M., P. Berg, T. Meuwissen, D. Boichard, M. Brochard, and V. Ducrocq. 2021. Improved dairy cattle mating plans at herd level using genomic information. Animal 15:100016. https://doi.org/10.1016/j.animal.2020.100016.

Bisutti, V., S. Pegolo, D. Giannuzzi, L. F. M. Mota, A. Vanzin, A. Toscano, E. Trevisi, P. A. Marsan, M. Brasca, and A. Cecchinato. 2022. The β-casein (CSN2) A2 allelic variant alters milk protein profile and slightly worsens coagulation properties in Holstein cows. J. Dairy Sci. 105:3794–3809. https://doi.org/10.3168/jds.2021-21537.

Borchersen, S. 2019. Launch of Improved EuroGenomics Genotyping Beadchip: EuroG MD. Accessed Feb. 5, 2022. https://www.eurogenomics.com/actualites/ eurogenomics-new-eurog-md-beadchip.html.

Bradley, R., S. Terlecki, and F. G. Clegg. 1982. The pathology of a retinal degeneration in Friesian cows. J. Comp. Pathol. 92:69–83. https://doi.org/10.1016/0021-9975(82)90043-3.

Browning, B. L., and S. R. Browning. 2013. Improving the accuracy and efficiency of identity-by-descent detection in population data. Genetics 194:459–471. https://doi.org/10.1534/genetics.113.150029.

Browning, S. R., and B. L. Browning. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am. J. Hum. Genet. 81:1084–1097. https://doi.org/10.1086/521987.

Burnell, M. 2019. The use of sexed semen in dairy herds. Livestock (Lond.) 24:282–286. https://doi.org/10.12968/live.2019.24.6.282.

Carthy, T. R., J. McCarthy, and D. P. Berry. 2019. A mating advice system in dairy cattle incorporating genomic information. J. Dairy Sci. 102:8210–8220. https://doi.org/10.3168/jds.2019-16283.

Charlesworth, D., and J. H. Willis. 2009. The genetics of inbreeding depression. Nat. Rev. Genet. 10:783–796. https://doi.org/10.1038/nrg2664.

Chessa, S., S. Gattolin, P. Cremonesi, D. Soglia, R. Finocchiaro, J.-T. Van Kaam, M. Marusi, and G. Civati. 2020. The effect of selection on casein genetic polymorphisms and haplotypes in Italian Holstein cattle. Ital. J. Anim. Sci. 19:833–839. https://doi.org/10.1080/1828051X.2020.1802356.

Cole, J. B. 2015. A simple strategy for managing many recessive disorders in a dairy cattle breeding program. Genet. Sel. Evol. 47:94. https://doi.org/10.1186/s12711-015-0174-9.

Daetwyler, H. D., A. Capitan, H. Pausch, P. Stothard, R. van Binsbergen, R. F. Brøndum, X. Liao, A. Djari, S. C. Rodriguez, C. Grohs, D. Esquerré, O. Bouchez, M.-N. Rossignol, C. Klopp, D. Rocha, S. Fritz, A. Eggen, P. J. Bowman, D. Coote, A. J. Chamberlain, C. Anderson, C. P. VanTassell, I. Hulsegge, M. E. Goddard, B. Guldbrandtsen, M. S. Lund, R. F. Veerkamp, D. A. Boichard, R. Fries, and B. J. Hayes. 2014. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nat. Genet. 46:858–865. https://doi.org/10.1038/ng.3034.

de Cara, M. Á. R., B. Villanueva, M. A. Toro, and J. Fernández. 2013. Using genomic tools to maintain diversity and fitness in conservation programmes. Mol. Ecol. 22:6091–6099. https://doi.org/10.1111/mec.12560.

Doekes, H. P., P. Bijma, and J. J. Windig. 2021. How depressing is inbreeding? A meta-analysis of 30 years of research on the effects of inbreeding in livestock. Genes (Basel) 12:926. https://doi.org/10.3390/genes12060926.

EFFAB (European Forum of Farm Animal Breeders). 2020. Code of Good Practice. Accessed Feb. 2, 2022. https://www.responsiblebreeding.eu/uploads/2/3/1/3/23133976/01_general_document_2020_final-code_efabar.pdf.

EU (European Union). 2008. European Commission Regulation No. 889/2008. Products with Regard to Organic Production, Labelling and Control.

Fikse, F., and M. Kargo. 2020. The Economic Value of NTM. Accessed Feb. 2, 2022. https://www.nordicebv.info/wp-content/uploads/2020/11/The-economic-value-of-NTM_final.pdf.

Forutan, M., S. Ansari Mahyari, C. Baes, N. Melzer, F. S. Schenkel, and M. Sargolzaei. 2018. Inbreeding and runs of homozygosity before and after genomic selection in North American Holstein cattle. BMC Genomics 19:98. https://doi.org/10.1186/s12864-018-4453-z.

Freyer, B. G., Z. Liu, G. Erhardt, and L. Panicke. 1999. Casein polymorphism and relation between milk production traits. J. Anim. Breed. Genet. 116:87–97. https://doi.org/10.1046/j.1439-0388.1999.00181.x.

Fritz, S., C. Hoze, E. Rebours, A. Barbat, M. Bizard, A. Chamberlain, C. Escouflaire, C. Vander Jagt, M. Boussaha, C. Grohs, A. Allais-Bonnet, M. Philippe, A. Vallée, Y. Amigues, B. J. Hayes, D. Boichard, and A. Capitan. 2018. An initiator codon mutation in SDE2 causes recessive embryonic lethality in Holstein cattle. J. Dairy Sci. 101:6220–6231. https://doi.org/10.3168/jds.2017-14119.

Gallinat, J. L., S. Qanbari, C. Drögemüller, E. C. G. Pimentel, G. Thaller, and J. Tetens. 2013. DNA-based identification of novel bovine casein gene variants. J. Dairy Sci. 96:699–709. https://doi.org/10.3168/jds.2012-5908.

Gautier, M., T. Faraut, K. Moazami-Goudarzi, V. Navratil, M. Foglio, C. Grohs, A. Boland, J.-G. Garnier, D. Boichard, G. M. Lathrop, I. G. Gut, and A. Eggen. 2007. Genetic and haplotypic structure in 14 European and African cattle breeds. Genetics 177:1059–1070. https://doi.org/10.1534/genetics.107.075804.

Gengler, N., P. Mayeres, and M. Szydlowski. 2007. A simple method to approximate gene content in large pedigree populations: Application to the myostatin gene in dual-purpose Belgian Blue cattle. Animal 1:21–28. https://doi.org/10.1017/S1751731107392628.

Hill, W. G., and B. S. Weir. 2011. Variation in actual relationship as a consequence of Mendelian sampling and linkage. Genet. Res. (Camb.) 93:47–64. https://doi.org/10.1017/S0016672310000480.

Jansen, G. B., and J. W. Wilton. 1985. Selecting mating pairs with linear programming techniques. J. Dairy Sci. 68:1302–1305. https://doi.org/10.3168/jds.S0022-0302(85)80961-9.

Maignel, L., D. Boichard, and E. Verrier. 1996. Genetic variability of French dairy breeds estimated from pedigree information. No. 14 in Proc. Open Sess. Interbull Annu. Mtg. Interbull.

Makanjuola, B. O., F. Miglior, E. A. Abdalla, C. Maltecca, F. S. Schenkel, and C. F. Baes. 2020. Effect of genomic selection on rate of inbreeding and coancestry and effective population size of Holstein and Jersey cattle populations. J. Dairy Sci. 103:5183–5199. https://doi.org/10.3168/jds.2019-18013.

Martikainen, K., A. M. Tyrisevä, K. Matilainen, J. Pösö, and P. Uimari. 2017. Estimation of inbreeding depression on female fertility in the Finnish Ayrshire population. J. Anim. Breed. Genet. 134:383–392. https://doi.org/10.1111/jbg.12285.

Medugorac, I., D. Seichter, A. Graf, I. Russ, H. Blum, K. H. Göpel, S. Rothammer, M. Förster, and S. Krebs. 2012. Bovine polledness –

An autosomal dominant trait with allelic heterogeneity. PLoS One 7:e39477. https://doi.org/10.1371/journal.pone.0039477.
Mueller, M. L., J. B. Cole, T. S. Sonstegard, and A. L. Van Eenennaam. 2019. Comparison of gene editing versus conventional breeding to introgress the POLLED allele into the US dairy cattle population. J. Dairy Sci. 102:4215–4226. https://doi.org/10.3168/jds.2018-15892.
NAV. 2019. NAV Routine Genetic Evaluation of Dairy Cattle – Data and Genetic Models. Accessed Aug. 22, 2019. https://www.nordicebv.info/wp-content/uploads/2019/06/NAV-routine-genetic-evaluation-04062019.pdf.
Olenski, K., S. Kamiński, J. Szyda, and A. Cieslinska. 2010. Polymorphism of the beta-casein gene and its associations with breeding value for production traits of Holstein–Friesian bulls. Livest. Sci. 131:137–140. https://doi.org/10.1016/j.livsci.2010.02.023.
Olsen, H. F., S. Tenhunen, N. I. Dolvik, D. I. Våge, and G. Klemetsdal. 2020. Segment-based coancestry, additive relationship and genetic variance within and between the Norwegian and the Swedish Fjord horse populations. Acta Agric. Scand. A Anim. Sci. 69:118–126. https://doi.org/10.1080/09064702.2019.1711155.
Oskarsson, M., and E. Engelbrekts. 2015. Kostnader för hälsostörningar hos mjölkkor. Accessed Feb. 2, 2022. https://docplayer.se/28330924-Kostnader-for-halsostorningar-hos-mjolkkor.html.
Pryce, J. E., M. Haile-Mariam, M. E. Goddard, and B. J. Hayes. 2014. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. Genet. Sel. Evol. 46:71. https://doi.org/10.1186/s12711-014-0071-7.
Pryce, J. E., B. J. Hayes, and M. E. Goddard. 2012. Novel strategies to minimize progeny inbreeding while maximizing genetic gain using genomic information. J. Dairy Sci. 95:377–388. https://doi.org/10.3168/jds.2011-4254.
Sargolzaei, M. 2014. SNP1101 user's guide. Version 1.0. HiggsGene Solutions Inc. https://ovc.uoguelph.ca/pathobiology/people/faculty/Mehdi-Sargolzaei.
Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15:478. https://doi.org/10.1186/1471-2164-15-478.
Scheper, C., M. Wensch-Dorendorf, T. Yin, H. Dressel, H. Swalve, and S. König. 2016. Evaluation of breeding strategies for polledness in dairy cattle using a newly developed simulation framework for quantitative and Mendelian traits. Genet. Sel. Evol. 48:50. https://doi.org/10.1186/s12711-016-0228-7.
Schuster, D. E., M. E. Kehrli, M. R. Ackermann, and R. O. Gilbert. 1992. Identification and prevalence of a genetic defect that causes leukocyte adhesion deficiency in Holstein cattle. Proc. Natl. Acad. Sci. USA 89:9225–9229.
Segelke, D., H. Täubert, F. Reinhardt, and G. Thaller. 2016. Considering genetic characteristics in German Holstein breeding programs. J. Dairy Sci. 99:458–467. https://doi.org/10.3168/jds.2015-9764.
Smith, L. A., B. G. Cassell, and R. E. Pearson. 1998. The effects of inbreeding on the lifetime performance of dairy cattle. J. Dairy Sci. 81:2729–2737. https://doi.org/10.3168/jds.S0022-0302(98)75830-8.
Sørensen, L. P., J. Pedersen, M. Kargo, U. S. Nielsen, F. Fikse, J.-Å. Eriksson, J. Pösö, R. S. Stephansen, and P. Aamand. 2018. Review of Nordic Total Merit Index: Full Report November 2018. https://www.nordicebv.info/wp-content/uploads/2018/11/2018.11.06-NTM-2018-report-Full.pdf.
Spurlock, D. M., M. L. Stock, and J. F. Coetzee. 2014. The impact of 3 strategies for incorporating polled genetics into a dairy cattle breeding program on the overall herd genetic merit. J. Dairy Sci. 97:5265–5274. https://doi.org/10.3168/jds.2013-7746.
Stock, M. L., S. L. Baldridge, D. Griffin, and J. F. Coetzee. 2013. Bovine dehorning. Vet. Clin. North Am. Food Anim. Pract. 29:103–133. https://doi.org/10.1016/j.cvfa.2012.11.001.
Stoffel, M. A., S. E. Johnston, J. G. Pilkington, and J. M. Pemberton. 2021. Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. Nat. Commun. 12:2972. https://doi.org/10.1038/s41467-021-23222-9.
Summer, A., F. Di Frangia, P. Ajmone Marsan, I. De Noni, and M. Malacarne. 2020. Occurrence, biological properties and potential effects on human health of β-casomorphin 7: Current knowledge and concerns. Crit. Rev. Food Sci. Nutr. 60:3705–3723. https://doi.org/10.1080/10408398.2019.1707157.
Sun, C., P. M. VanRaden, J. R. O'Connell, K. A. Weigel, and D. Gianola. 2013. Mating programs including genomic relationships and dominance effects. J. Dairy Sci. 96:8014–8023. https://doi.org/10.3168/jds.2013-6969.
Thompson, N. M., N. O. Widmar, M. M. Schutz, J. B. Cole, and C. A. Wolf. 2017. Economic considerations of breeding for polled dairy cows versus dehorning in the United States. J. Dairy Sci. 100:4941–4952. https://doi.org/10.3168/jds.2016-12099.
VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423. https://doi.org/10.3168/jds.2007-0980.
VanRaden, P. M., K. M. Olson, G. R. Wiggans, J. B. Cole, and M. E. Tooker. 2011. Genomic inbreeding and relationships among Holsteins, Jerseys, and Brown Swiss. J. Dairy Sci. 94:5673–5682. https://doi.org/10.3168/jds.2011-4500.
Wang, H., I. Misztal, and A. Legarra. 2014. Differences between genomic-based and pedigree-based relationships in a chicken population, as a function of quality control and pedigree links among individuals. J. Anim. Breed. Genet. 131:445–451. https://doi.org/10.1111/jbg.12109.
Weigel, K. A., and S. W. Lin. 2000. Use of computerized mate selection programs to control inbreeding of Holstein and Jersey cattle in the next generation. J. Dairy Sci. 83:822–828. https://doi.org/10.3168/jds.S0022-0302(00)74945-9.
Woolliams, J. A., and E. A. Mäntysaari. 1995. Genetic contributions of Finnish Ayrshire bulls over four generations. Anim. Sci. 61:177–187. https://doi.org/10.1017/S1357729800013709.
Zhang, Q., B. Guldbrandtsen, M. Bosse, M. S. Lund, and G. Sahana. 2015. Runs of homozygosity and distribution of functional variants in the cattle genome. BMC Genomics 16:542. https://doi.org/10.1186/s12864-015-1715-x.

ORCIDS

C. Bengtsson https://orcid.org/0000-0001-5784-2557
H. Stålhammar https://orcid.org/0000-0002-7354-651X
J. R. Thomasen https://orcid.org/0000-0001-7068-0572
W. F. Fikse https://orcid.org/0000-0003-0578-0184
E. Strandberg https://orcid.org/0000-0001-5154-8146
S. Eriksson https://orcid.org/0000-0003-3357-5065

Acta Universitatis Agriculturae Sueciae

Doctoral Thesis No. 2024:6

Dairy farmers benefit from genotyping dairy cattle by increased accuracy of breeding values and improved mating plans at herd level. Genomically enhanced breeding values for virgin heifers predicted cow performance significantly more accurately than parent average breeding values. Genomic mating plans were investigated, both in a short-term and long-term perspective. Optimising economic score with linear programming is instrumental to avoid expression of genetic defects, increase favourable traits like polledness, and reduce genetic relatedness.

Christian Bengtsson received his postgraduate education at the Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences (SLU), Uppsala. He received his undergraduate degree from SLU in 2016.

Acta Universitatis Agriculturae Sueciae presents doctoral theses from the Swedish University of Agricultural Sciences (SLU).

SLU generates knowledge for the sustainable use of biological natural resources. Research, education, extension, as well as environmental monitoring and assessment are used to achieve this goal.

ISSN 1652-6880
ISBN (print version) 978-91-8046-276-1
ISBN (electronic version) 978-91-8046-277-8