Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Faculty of Veterinary Medicine and Animal Science
Department of Animal Breeding and Genetics

Association of genomic breeding values and parental average breeding values with future phenotypic performance in Swedish Red cows

Christian Bengtsson

Examensarbete / Swedish University of Agricultural Sciences
Department of Animal Breeding and Genetics
489
Uppsala 2016

Master's Thesis, 30 hp

Agricultural Science programme
– Animal Science

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Faculty of Veterinary Medicine and Animal Science
Department of Animal Breeding and Genetics

Association of genomic breeding values and parental average breeding values with future phenotypic performance in Swedish Red cows

Hur väl förutsäger genomiska avelsvärden och härstamningsindex den framtida fenotypen hos SRB-kor?

Christian Bengtsson

Supervisor:
Freddy Fikse, SLU, Department of Animal Breeding and Genetics

Examiner:
Erling Strandberg, SLU, Department of Animal Breeding and Genetics

Credits: 30 hp
Course title: Degree project in Animal Science
Course code: EX0558
Programme: Agricultural Science programme - Animal Science
Level: Advanced, A2E

Place of publication: Uppsala
Year of publication: 2016
Name of series: Examensarbete / Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics, 489
On-line publicering: http://epsilon.slu.se

Key words: genomic selection, red dairy cattle, Sweden, genotyping

Abstract

The purpose of this study was to compare female genomically enhanced breeding values with parental average breeding values in how well those match the animal's future phenotypes. At the early years of genomic selection mainly bulls were genotyped, but when the costs decreases, genotyping of heifers becomes more and more interesting. Higher accuracy when selecting replacement animals and better mating plans are some of the main arguments for genotyping heifers.

Genotyping heifers in Sweden started in larger scale in 2012, with the start of VikingGenetics LD-project. The main reason was to get genotyped females into the reference population and thereby increase the accuracy of genomically enhanced breeding values. Since the start, over 10,000 females have been genotyped and production results from some of the animals' first lactations have been recorded. Production, fertility, conformation and functionality records were analyzed from 2637 genotyped females.

In general genomically enhanced breeding values and parental average breeding values worked best to predict future phenotypes for high heritability traits. Except for better genomic prediction for milkability there were no significant differences between indexes and their prediction of future phenotypes. There were tendencies of genomically enhanced breeding values functioned better then parental average breeding values for milk, fat and protein yield. Low accuracy of genomically enhanced breeding values and too few records for some traits could be some explanations of the results. Even though there were few significant differences between genomically enhanced breeding values and parental average breeding values the study indicated that also the conventional genetic evaluation, without genomic information works well for many of the studied traits. Furthermore, the study was made a bit early as some traits could not be analyzed fully because of few completed lactations. Future studies have to be made to confirm the results.

Sammanfattning

Syftet med denna studie var att jämföra genomiska avelsvärden med härstamningsindex i hur väl de förutser djurens framtida fenotyp. I början av den genomiska eran testades mestadels tjurar, men med lägre genotypningskostnader blir det allt mer intressant att testa hondjur. Högre säkerhet när rekryteringsdjur väljs ut och bättre parningsplaner är några av huvudargumenten för att testa sina hondjur.

Genotypning av hondjur startade i Sverige i samband med VikingGenetics LD-projekt. Huvudorsaken var att få in hondjur i referenspopulation och därmed öka säkerheten på de genomiska avelsvärderna. Sedan starten av projektet har nästan 10 000 hondjur blivit testade och produktionsresultat från några av deras första laktationer har dokumenterats. Produktionsresultat från 2637 genotypade hondjur fanns tillgängliga och egenskaperna som analyserades var avkastning, fertilitet, exteriör och funktionalitet.

Den generella trenden visade att genomiska avelsvärden och härstamningsindex var bättre på att förutse framtida fenotyper på egenskaper med höga arvbarheter. Genomiska avelsvärden fungerade bättre än härstamningsindex för mjölkbarhet men i övrigt fanns det inga signifikanta skillnader mellan de båda indexen. Det fanns dock tendenser att genomisk avelsvärden fungerade bättre än härstamningsindex för mjölk, fett och protein avkastning. Låg säkerhet på genomiska avelsvärden och få produktionsresultat från vissa egenskaper kan möjligen förklara resultaten. Även om det fanns få signifikanta skillnader mellan genomiska avelsvärden och härstamningsindex så indikerade studien att den traditionella avelsvärderingen, utan genomisk information, fungerar väl för många av de studerade egenskaperna. Vissa egenskaper kunde inte analyseras fullständigt på grund av för få fullständiga laktationer. Framtida studier måste göras för att bekräfta resultatet.

Table of content

Introduction

In dairy cattle production, genomic selection becomes more and more popular. The first step was the sequencing of the bovine genome, which further led to the detection of thousands of single-nucleotide polymorphism (SNP) markers (Meuwissen et al., 2001). Further, Meuwissen et al. (2001) showed that precise selection decisions could be made using genomically enhanced breeding values (GEBV) calculated from the marker effects. The genome is divided into small segments and the marker effects are estimated in a reference population where animals are both genotyped and phenotyped. Further, the next generation can be genotyped for the markers and thereafter the sum of the effects determine their predicted GEBV (Meuwissen et al., 2001). The main goal with genomic selection in dairy cattle is to increase the genetic gain (Schaeffer, 2006).

Genomic selection makes it possible to predict accurate breeding values for young animals (Hayes et al., 2009). Progeny testing has been used for decades to identify elite bulls. This method results in a long generation interval and with genomic selection elite bulls could be identified much earlier. Many of the important traits in dairy cattle breeding are shown late in life and can only be measured on females (de Roos et al, 2011). In simulation, Meuwissen et al. (2001) showed that accuracies for GEBV at birth for a bull calf can be as high as accuracies for estimated breeding values (EBV) after progeny testing. This has been suggested to potentially double the genetic gain and also save 92% of the costs for breeding companies if progeny testing was excluded (Schaeffer, 2006).

During the early years of genomic selection mainly bulls were genotyped, but when the costs decreases genotyping of heifers becomes more and more interesting. Higher accuracy when selecting replacement animals and better mating plans are some of the main arguments for genotyping heifers (Pryce & Hayes 2012).

Genotyping heifers in Sweden started in larger scale in 2012, with the start of VikingGenetics LD-project. The main reason was to get genotyped females into the reference population and thereby increase the accuracy of GEBV. Since the start over 10 thousands RDC females have been genotyped and milk production results from some of the animals first lactations have been recorded (VikingGenetics, 2015). Validation of GEBV and illustrating the relationship between genomic prediction and the future phenotype could be a key to increase confidence for the genomic technology.

The purpose of this study was to compare female genomically enhanced breeding values with parental average breeding values in how well those match the animal’s future phenotypes.

Literature Review

Genomic information in breeding programs

Studies from Canada & United States (VanRaden et al., 2009), New Zealand (Harris et al., 2008), and the Netherlands (Hayes et al., 2009) have reported greater reliabilities of GEBV than for parental average (PA) breeding values. Animals in all countries were genotyped with Illumina Bovine SNP50^{TM} chip. Reliabilities were greatest in United States and New Zealand, where there were more bulls in the reference population, 3,576 and 4,500 compared to 1,583 in the Netherlands. In New Zealand, the reliabilities for milk, protein, fat and protein yield, live body weight, SCC and fertility ranged from 50 to 67% for GEBV compared to an average of 34% for PA. In United States and Canada, the reliabilities for GEBV were 50% compared to an average of 27% for PA over all traits. The results from the Netherlands showed higher reliabilities for GEBV compared to an average of PA; 9% for fertility, 13% for udder depth and SCS, 15% for feet and legs, 19% for kilograms of protein and 33% over for fat percentage.

Yao et al. (2015) proposed that SNP genotypes and health data can be used to predict future phenotypes. Feed efficiency was studied through measurement of residual feed intake (RFI). The RFI was calculated as the difference between the actual intake and the expected feed intake. The study used SNP genotypes and health history for prediction of future dry matter intake (DMI), live body weight, RFI and milk yield. Accuracies were measured as correlations between predicted values and phenotypes. The accuracies without health history for RFI were 8.76% using random forests algorithm and 20.45% using support vector machine algorithm. In general adding health history improved accuracies slightly. There was no effect on adding health data for residual feed intake (Yao et al., 2015).

Pryce et al. (2014) validated two published studies of genomic prediction of RFI and DMI. The number of lactating cows used was 78 and an accuracy of 0.27 for RFI was achieved when the reference population consisted of 843 Australian and 939 New Zealand heifers. An average accuracy of 0.72 was achieved when a multicountry model was used, which included cows in lactation from two countries; 958 cows form the Netherlands and United Kingdom and also 843 growing Australia heifers (Pryce et al., 2014).

Female genomic information

During the early years of genomic selection mainly bulls were tested, but as the costs are reducing genotyping of heifers gets more and more interesting. One use of genotyping heifers is to find the best heifers for replacement (Pryce & Hayes 2012). The study assumed a herd of 100 cows where the heifers available for selection varied from 20 to 50. The replacement rate varied from 15% to 30%. Three different cost of genotyping was assumed; 5 Australian dollar (AU$5), AU$50 or AU$100. Comparison of genomic selection with PA information or no PA information was made. Genotyping heifers became profitable when the price of genotyping was AU$50 with no PA information and at AU$5 when PA information was included. The largest benefit was with a high number of candidates for a few replacement spots. However, their

comparison of costs and benefits of genotyping heifers did not take marketing into account. It might be profitable to market heifers or embryos from heifers with breeding values at birth with up to 60% reliability. Other advantages of genotyping heifers may more optimal mating plans and keeping recessive alleles under control (Pryce & Hayes 2012).

Calus et al. (2015) investigated the economic effects of prioritizing heifers depending of their GEBV. The number of available heifers for a herd of 100 cows varied from 15 to 45 and the replacement rate varied from 15 to 40%. Also use of sexed semen was considered which maximum resulted in twice as many available heifers. The used formula included number of lactating animals, difference in accuracy between PA breeding values and GEBV and the selection intensity as input. Genotyping heifers was profitable in most scenarios, when two or more extra candidates where available for selection. When sexed semen was used a preselection based on PA was done. The most beneficial proportion of preselection based on PA was 0.67 (Calus et al., 2015).

Koivula et al. (2014) studied the use of genotyped Nordic Red Dairy cows in the reference population. The study included in the evaluation 5,593 or 3,111 or 0 genotyped cows in the reference population. In all evaluations 4,188 genotyped bulls were used. The extra gain in accuracy from cows in the reference population varied from 0.8% to 2.6%-units (Koivula et al., 2014). Wiggans et al. (2011) pre-adjusted records from Jersey and Holstein genotyped females so they would be comparable with genotyped bulls. When females were included in the reference population an extra gains in reliabilities of 3.5%-units for Holstein and 0.9%-units for Jersey were achieved. Further, Pryce et al. 2012 demonstrated an increase of 8%-units reliabilities when 10,000 cows were added to an reference population of 3,000 bulls.

Hugh et al. (2011) investigated genomic breeding programs with female information. The study used a stimulation program and the population consisted of 100 males and 100 females and Fisher-Wright population model was used. The study showed that including females in genomic breeding programs could triple the genetic gain. The reason for the extra genetic gain was increased accuracy and also a shorter generation interval (Hugh et al., 2011).

At a study made at Allenstein Dairy herd at UW-Madison. Approximately 400 heifers were tested with Zoetis low-density chip (CLARIFIDE®). The study compared selections based on their own genomic results at 12 months of age or their sire’s current daughter performances. Their predicted performance was divided into quartiles and was compared with their actual results. The traits included were milk yield as 305-day mature equivalent (ME), days open (DO) in first lactation and somatic cell score (SCS) as actual average log SCC. The difference between bottom and top quartiles in milk was 2,366 pounds per lactation for sire sorted and 4,801 for own genomic results. For DO difference between bottom and top quartiles was 3.4 days for sire sorted and 21.0 days for own genomic results. For SCC the difference between bottom and top was 0.18 for sire sorted and 0.64 for own genomic results (Weigel et al., 2015).

Genotyping methods

The BovineSNP50 genotyping array with approximately 54 thousands SNP probes is widely used for dairy cattle genomic prediction around the world (Matukumalli et al., 2009). It first became available in 2007 and is used in cattle breeding to detect genomic regions contributing for variation in phenotype traits (Sherman et al., 2009). The density of SNP markers affects the accuracy (Meuwissen, 2009, Habier et al., 2009). In theory, a higher density should lead to a higher accuracy but it also leads to an increased cost for genotyping (Peipei, et al 2013). Some countries have genotyped bulls with a 777,000-markers high density chip (777K; high-density, HD), with the purpose of increasing the accuracy (Su et al., 2012). In addition, low density chips with 6,900-markers and 2,900-markers (BovineLD and Bovine3K) have been developed, those should be more suitable for a large scale and have a lower genotype cost (Boichard et al., 2012).

When several chips are used in genomic selection it is important to make use of all available marker data by imputation of missing genotypes. Imputation is also useful to increase the call rate of genotyped animals when the same chip is used (Peipei, et al 2013). Imputing from 3K to 54K gave lower imputation accuracies than imputing from 54K to 777K, 93.5 to 97.1% compared to 97.1 to 99.3% (Peipei, et al 2013).

Prediction with genomic data

There are several strategies to use genomic data for prediction (Koivula et al., 2012). The solving algorithm in the Nordic countries has just been changed from BLUP at individual level (G-BLUP) to single nucleotide polymorphism level (SNP-BLUP) (Nordic Cattle Genetic Evaluation, 2015). Koivula et al. (2012) compared genomic prediction methods in Nordic red bulls. Three different BLUP models were compared; SNP-BLUP and G-BLUP and the one-step approach (H-BLUP). The study showed that SNP-BLUP and G-BLUP resulted in the same direct genomic breeding values with correlation between SNP-BLUP and G-BLUP of 0.99. Correlation between H-BLUP and SNP-BLUP or G-BLUP was 0.96 (Koivula et al., (2012).

Breeding and genotyping in Sweden

In 2002 the Nordic Cattle Genetic Evaluation was established, which further led to a Swedish-Finnish-Danish AI cooperation. Kolmodin et al. (2003) found only small differences within and across the Nordic countries in the genotype-by-environment interaction **(**G×E**).** That means that most of the genes have the same effect in all the Nordic countries. This resulted in the joint breeding goal Nordic Total Merit (NTM). The Nordic cooperation enables a higher genetic gain as a result of a larger population and a higher selection intensity. It is also makes it easier for the farmers to compare bulls and cows from the different Nordic countries (Kargo et al., 2014). The Nordic Red Cattle (NRC) consist of the Swedish Red breeds, Danish Red and the Finnish Ayrshire. It is compared with the Holstein known as cow a with lower mastitis incidence, shorter calving interval, lower rate of stillbirths and lower production (Höglund et al., 2015).

In Sweden the genotyping of heifers started in larger scale in 2012, with the start of VikingGenetics LD-project. The main reason was to get genotyped females into the reference population and thereby increase the accuracy of GEBV. Possibilities to increase the reference population outside the Nordic countries is limited and to achieve moderately accurate GEBV females were introduced (VikingGenetics, 2015). Heritability for RDC traits used in the Nordic cattle genetic evaluation are presented in Table 1 and accuracies of GEBV of RDC bulls born in bulls are presented in Table 2 (Nordic Cattle Genetic Evaluation, 2015). Some of the indexes are combinations of several underlying component indexes. For example yield index describes genetic potential for milk, fat and protein production (Nordic Cattle Genetic Evaluation, 2015).

Table 1. *Heritability for RDC used in Nordic cattle genetic evaluation. Source: Nordic Cattle Genetic Evaluation (2015)*

Trait	Heritability
Milk	0.41
Fat	0.41
Protein	0.35
Cell count	0.12
FLS	0.02
CFS	0.04
NINS	0.025
Calvin ease	0.04
Milking speed	0.25
Temperament	0.15
Conformation	0.17-0.42
Clinical mastitis	0.04

Table 2 *Accuracies for GEBV of RDC bulls born in 2014. Source: Nordic Cattle Genetic Evaluation (2015)*

Index	Accuracy (%)
Yield	67
Fertility	47
Calving	47
Udder health	57
Survivial	38
Leg conformation	54
Udder conformation	55
Milkability	66
Temperament	53

Illustrating the relationship between genomic prediction and the future phenotype could be a key to increase confidence for the genomic technology (Pryce & Hayes 2012). In Sweden this has never been done before but could be an important step for the Swedish farmers to get confidence in the technology on home ground. To participate in the LD-project the farmers had to be a part of the Swedish milk recording scheme, register veterinary treatments, register claw health and be a part of "Individavel". "Individavel" includes conformation judging, documentation of functionally traits and mating plan guidance (VikingGenetics, 2015).

Materials and methods

Data

Data from herds with genotyped Swedish red cows were collected from the Swedish milk recording scheme. Animals were born between 2008 and 2013 and data from up to three parities were collected, with age at first calving between 18 to 38 months. In total there were 51,428 cows that had calved at least one time, of those 26,914 cows had calved twice and 10,090 cows had calved three times. The breed distribution was 30,170 RDC, 17,264 Holstein, 2,816 crossbred, 168 Jersey and 110 Swedish Polled Cattle SKB.

Cow GEBV were collected from the Nordic Cattle Genetic Evaluation. In total genomic breeding values from 15 breeding evaluations from September 2011 to August 2015 were used. In total there were 39,912 records with GEBV from 5,146 genotyped RDC that had calved at least one time.

PA breeding values were also collected from the Nordic Cattle Genic Evaluation. The PA breeding values used was the animal's breeding value published the year before their first calving.

Production traits

Standardized values for 305-d milk, fat and protein yield were used. Sampling at herd level is done up to twelve times per year in Sweden. The limit for calculation of 305-d yield is two test days per lactation. Values for milk, fat and protein yield more extreme than three standard deviation from the mean were set as missing.

Udder health

Mastitis treatments are reported by veterinarians. Two different traits for mastitis were used. The first trait showed if a mastitis was reported between -10 to 150 days or not, 1 respective 0. The second trait was number of mastitis cases up to 300 days of lactation.

Data for somatic cell scores (SCS) was collected in the same way as production traits.

Fertility

The fertility traits used was calving interval (CI), calving to first service (CFS), first to last service (FLS) and number of services (NINS). The different traits ranged from 280 to 700 days for CI, 21 to 290 days for CFS, 0 to 250 days for FLS, and 1 to 7 number of services for NINS.

Survival and stillbirths

Survival from first to second lactation was analyzed. The value was set equal to 1 if the cow survived and 0 if not. Two classes for stillbirths were used, early calving and difficult calving were set as 1, otherwise 0.

Conformation, milkability and temperament

Conformation traits used were udder, legs and body. Conformation data came from official classifiers. The overall conformation traits were studied and the observations varied from 60 to 93. Milkability and temperament records were based on owner's assessment. The scale for both milkability and temperament is 1 to 9 where a higher value is better.

Statistical analysis

Cows with GEBV were divided into two groups. The first group contained cows that had been genotyped before their first calving; for this group the PA breeding values and GEBV closest before calving was used. The result section focuses on this group because it was the only group were a cow's own performance did not affect her GEBV. The second group contained cows that did not have a GEBV before their first calving. The first genomic breeding values available for the second group was used. The first group consisted of 2,637 cows and the second group consisted of 2,515 cows. There was also a group with all available Swedish Red (SR) animals (genotyped and not) with PA breeding values consisting of 26,601 animals.

The program used for analysis was Statistical Analysis Software (SAS) version 9.4. PROC MEANS and PROC FREQ procedures were used for descriptive statistics. PROC HP MIXED was used to adjust cow phenotypes for systematic environmental factors using model [1] below. The adjusted phenotypes (cow and residual effect in model [1]) were named according to respective trait, for example: milk yield was named MilkE. PROC RANKS procedure was used to rank cows into four quartiles for GEBV or PA breeding value. Because GEBV were from several runs separated in time, they were not directly comparable and were corrected for genetic trends by running PROC REG in SAS. PROC CORR was used to calculate the correlation between breeding values and adjusted phenotypes. The correlations between PA breeding values or GEBV on one hand and adjusted phenotypes on the other for each of the groups, and a 95% confidence interval was used to assess significance of the difference correlations. PROC SQPLOT was used to plot the average adjusted phenotypes for each quartile.

Model

$Y_{ijklmno} = \mu + HY_{ij} + YM_{jk} + B_l + P_{m+} b_1*CA + b_2*CA^2 + C_n + e_{ijklmno}$

$Y_{ijklnmo}$ = the observed value

μ = mean of the population

HY_{ij} = Fixed class effect of herd *i* and calving year *j*: 2008, …., 2013.

YM_{jk} = Fixed class effect of calving year *j* and month *k*: 1, …., 12.

B_l = Fixed class effect of breed *l*: RDC, Holstein, Jersey, SKB, Crossbred.

P_m = Fixed class effect of parity *m:* 1, 2, 3.

CA_m = Calving age

b_1, b_2 = regression coefficients for CA and CA^2

C_n = Random effect of cow *n*, ~ND(0, σ_C^2)

$e_{ijklmnop}$ = Random residual, ~ND(0, σ_e^2).

Results

All presented results were from first lactation. In the text below, correlations are followed by their 95% confidence intervals in brackets. A summary of the main correlations are given in Table 3.

Production traits

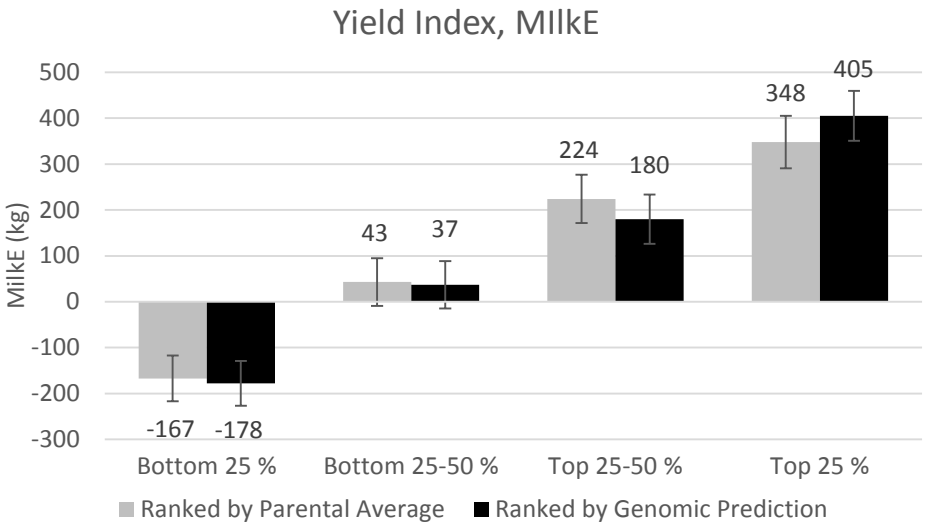

Figure 1. MilkE, 305-d adjusted milk yield for 2,489 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

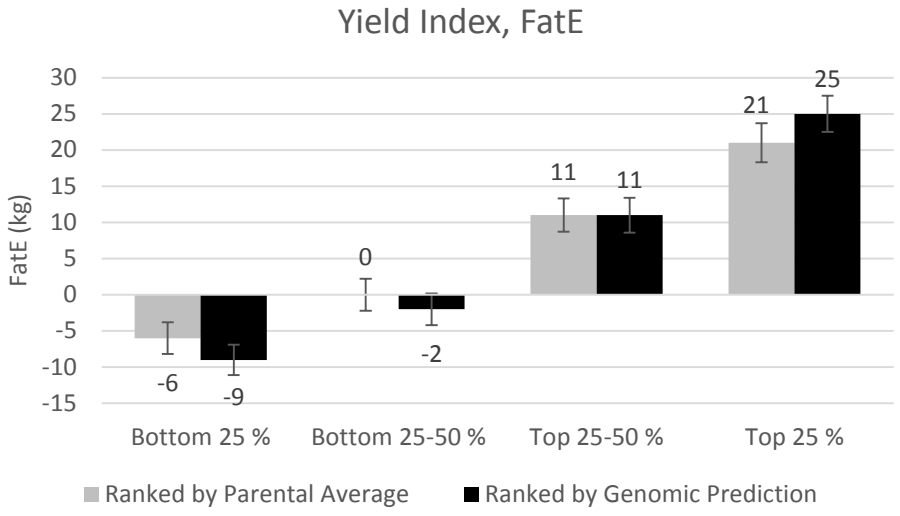

Figure 2. FatE, 305-d adjusted fat for 2,481 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

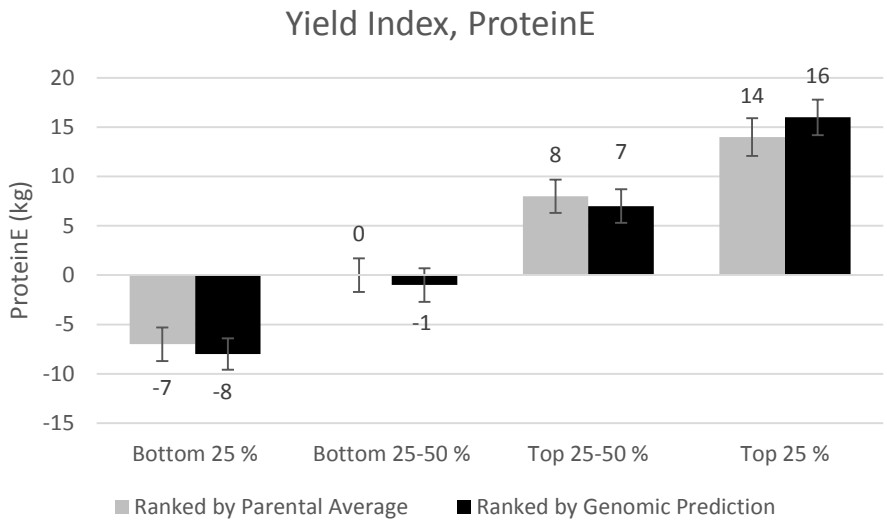

Figure 3. ProteinE, 305-d adjusted protein yield for 2,488 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

For milk yield the difference between the bottom and top quartile was 515 kg milk for PA prediction and 583 kg milk for genomic prediction (*Figure 1*). In the genotyped group the correlation between PA yield index and MilkE was 0.154 (0.116 – 0.193) and the correlation between genomic yield index and MilkE was 0.169 (0.130 – 0.207). The correlation was 0.152 (0.140 – 0.164) between PA yield index and MilkE for all available SR animals. The yield index is a combination of yields of milk, fat and protein, and the pure milk index was only available for genomic prediction; the correlation between pure milk index and MilkE was 0.290 (0.253 – 0.326).

The difference between the bottom and the top quartile for fat yield was 27 kg fat for PA prediction and 34 kg fat for genomic prediction (*Figure 2*). In the genotyped group the correlation between PA yield index and FatE was 0.191 (0.152 – 0.229) and the correlation between genomic yield index and FatE was 0.229 (0.191 – 0.266). The correlation was 0.173 (0.162 – 0.185) between PA yield index and FatE for all available SR animals. Pure fat index was only available for genomic prediction and the correlation between fat index and FatE was 0.290 (0.254 – 0.326).

For protein yield the difference between the bottom and the top quartile was 21 kg fat for PA prediction and 24 kg fat for genomic prediction (*Figure 3*). In the genotyped group the correlation between PA yield index and ProteinE was 0.194 (0.155 – 0.232) and the correlation between genomic yield index and ProteinE was 0.214 (0.177 – 0.252). The correlation was 0.192 (0.181 – 0.204) between PA yield index and ProteinE for all available SR animals. Pure protein index was only available for genomic prediction and the correlation between protein index and ProteinE was 0.223 (0.185 – 0.260).

Udder health

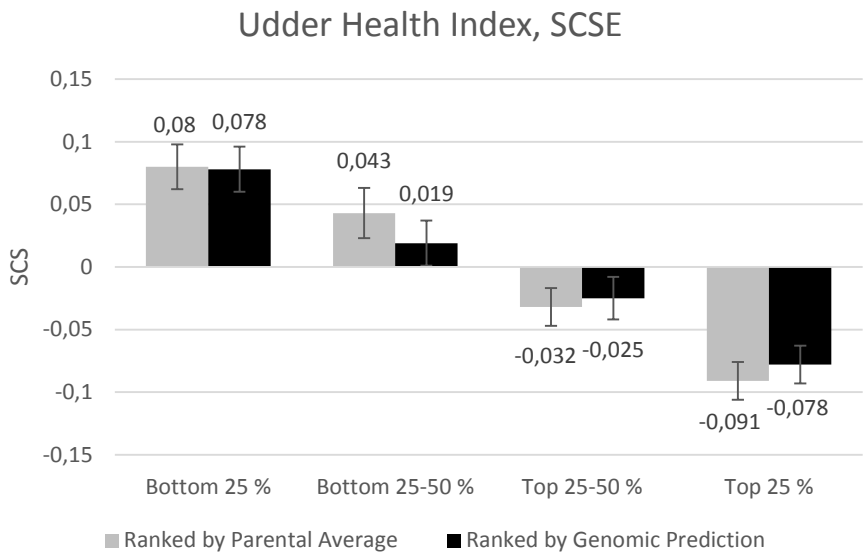

Figure 4. SCS, 305-d adjusted Somatic Cell Score for first lactation for 1,863 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

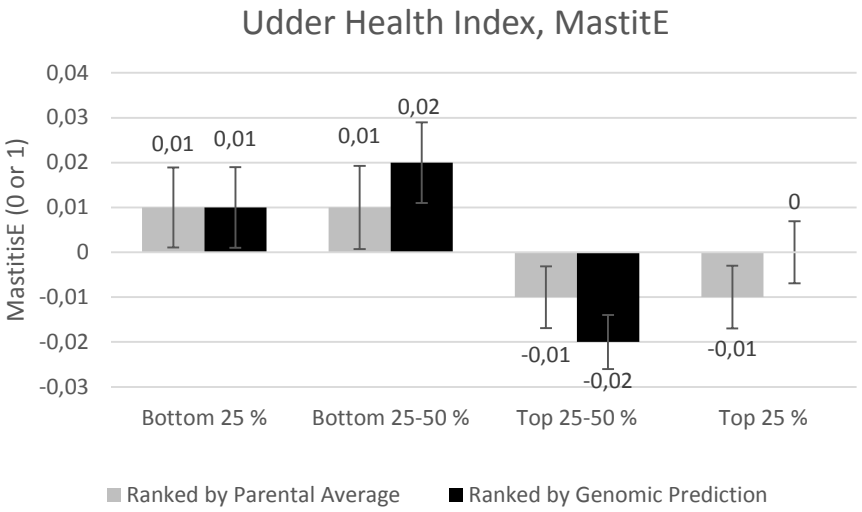

Figure 5. MasitisE, adjusted mastitis or not between -10 to 150 days of first lactation for 2,637 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

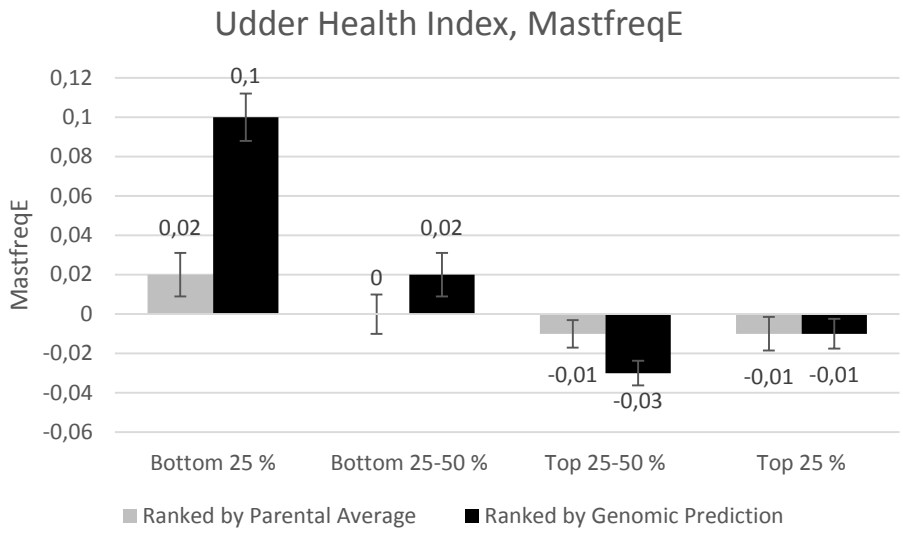

Figure 6. MasitisfreqE, adjusted number of mastitis up to 300 days for 2,637 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

The difference between the bottom and top quartile for SCS was 0.171 for PA prediction and 0.156 for genomic prediction (*Figure 4*). In the genotyped group the correlation between PA udder health index and SCSE was -0.173 (-0.217 – -0.128) and the correlation between genomic udder health index and SCSE was -0.164 (-0.201 – -0.120). The correlation was -0.124 (-0.137 – -0.111) between PA udder health index and SCSE for all available SR animals.

In the case of a mastitis in -10 to 150 days of first lactation there was not a significant difference between the quartiles (*Figure 5*). In the genotyped group the correlation between PA udder health index and MastitisE was -0.033 (-0.07 – 0.001) and the correlation between genomic udder health index and MastitisE was -0.049 (-0.097 – -0.011). The correlation was -0.04 (-0.048 – -0.024) between PA udder health index and MastitisE for all available SR animals.

For mastitis frequency up to 300 days in first lactation there was a significant difference between the bottom 25% and the rest of the quartiles (*Figure 6*). In the genotyped group the correlation between PA udder health index and MasitisfreqE was -0.038 (-0.076 – -0.000) and the correlation between genomic udder health index and MasitisfreqE was -0.059 (-0.097 – -0.021). The correlation was -0.03 (-0.046 – -0.023) between PA udder health index and MasitisfreqE for all available SR animals.

Fertility traits

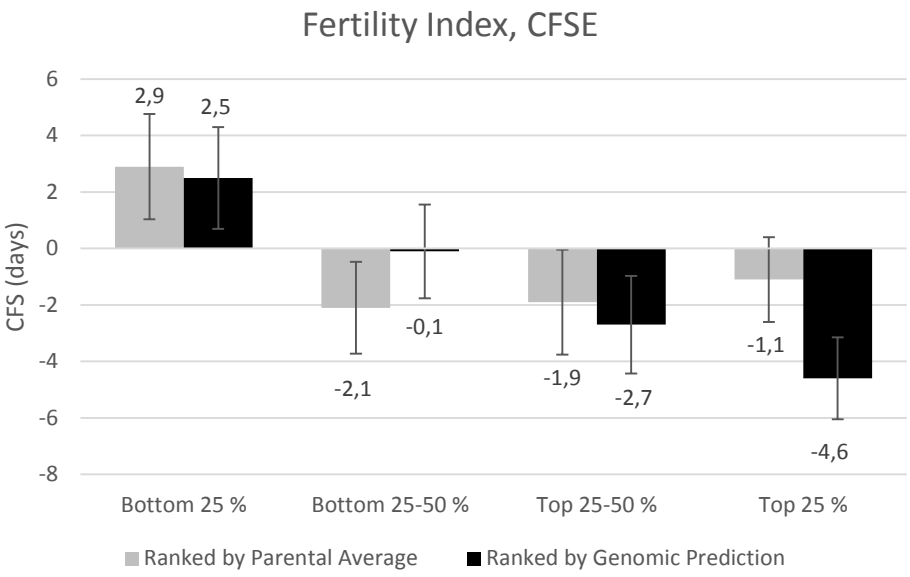

Figure 7. CFSE, adjusted days from Calving to First Service in first lactation for 663 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

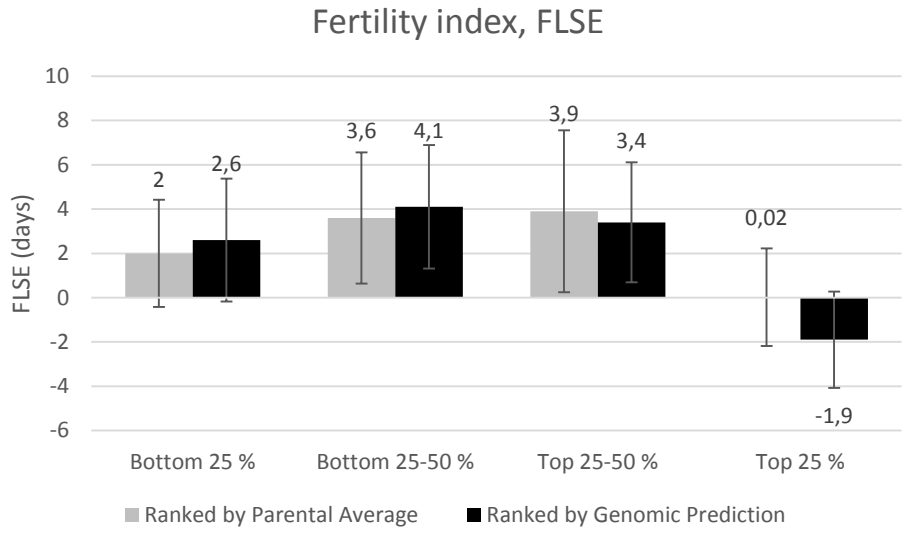

Figure 8. FLSE, adjusted days from First to Last Service in first lactation for 663 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

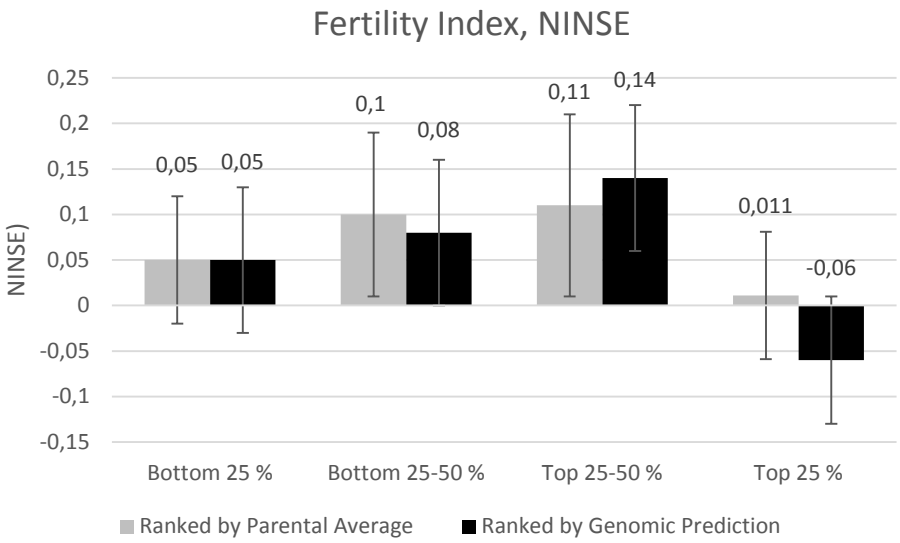

Figure 9. NINS, adjusted Number of Inseminations for 663 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

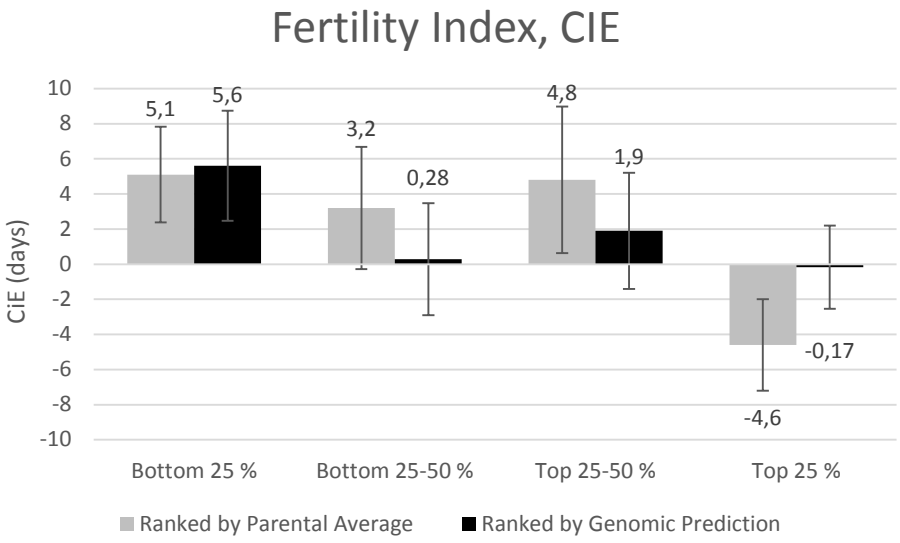

Figure 10. CIE, adjusted Calving Interval from first to second lactation for 664 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

There were no significant differences between the single quartiles for any of the fertility traits (*Figure 7, 8, 9, 10*). In the genotyped group the correlation between PA fertility index and CFSE was -0.064 (-0.149 – 0.062) and the correlation between genomic fertility index and CFSE was -0,069 (-0.144 – 0.007). The correlation was -0.032 (-0.049 – -0.015) between PA fertility index and CFSE for all available SR animals.

In the genotyped group the correlation between PA fertility index and FLSE was -0.0641 (-0.142 – 0.015) and the correlation between genomic fertility index and FLSE was -0.034 (-0.110 – 0.043). The correlation was -0.023 (-0.043 – -0.001) between PA fertility index and FLSE for all available SR animals.

For the genotyped group the correlation between PA fertility index and NINSE was -0.018 (-0.096 – 0.061) and the correlation between genomic fertility index and NINSE was -0.028 (-0.104 – 0.485). The correlation was -0.020 (-0.037 – -0.003) between PA fertility index and NINSE for all available SR animals.

In the genotyped group the correlation between PA fertility index and CIE was -0.046 (-0.124 – 0.032) and the correlation between genomic fertility index and CIE was -0.064 (-0.139 – 0.012). The correlation was -0.0447 (-0.061 – -0.028) between PA fertility index and CIE for all available SR animals.

Functional traits: Milkability, Temperament and Stillbirths

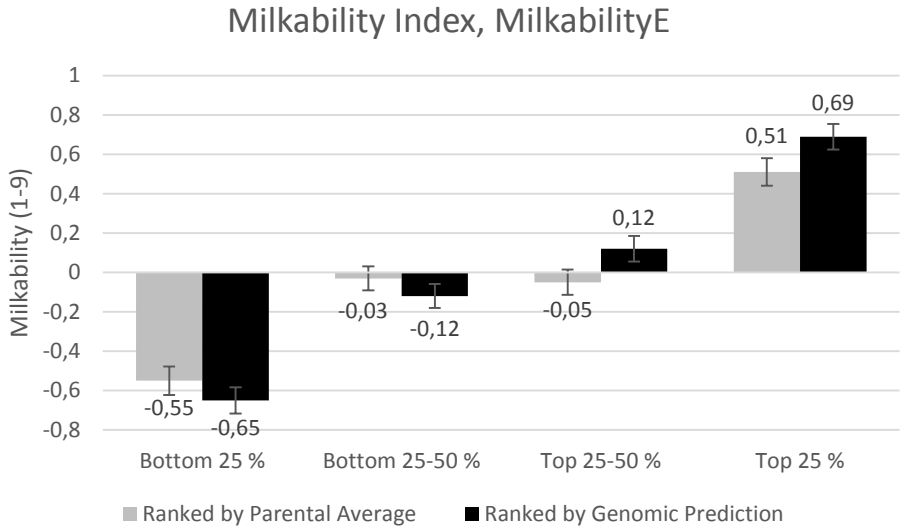

Figure 10. MilkabilityE, adjusted milkability for 1,669 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

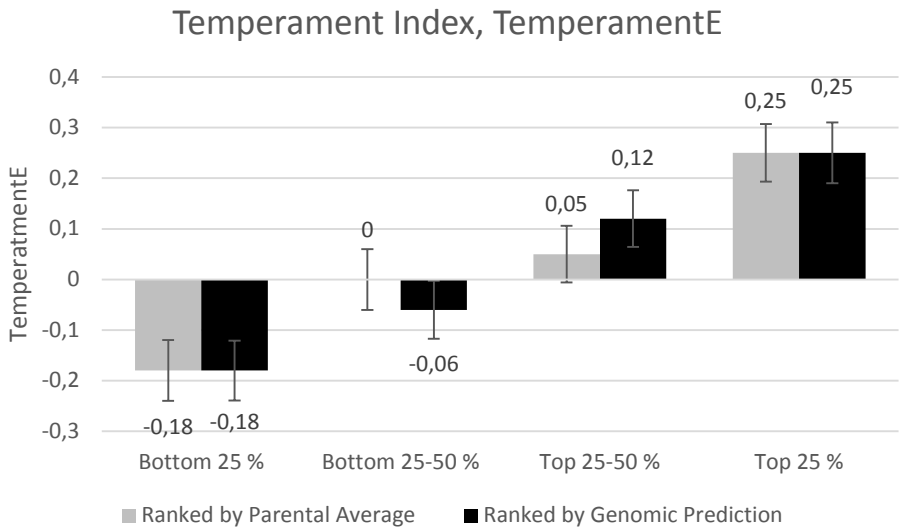

Figure 11. TemperamentE, adjusted temperament for 1,863 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

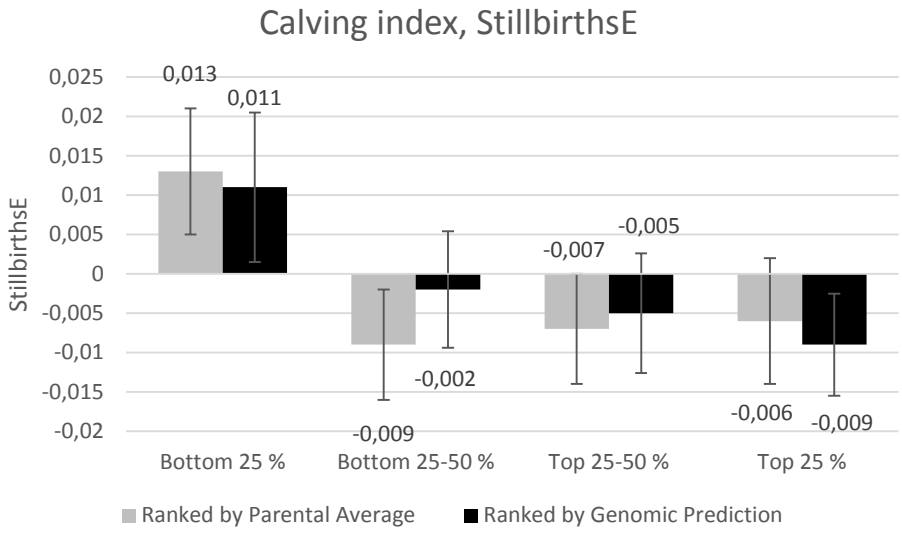

Figure 12. StillbirthsE, adjusted stillbirths of first calving for 2,557 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

For milkability the difference between the bottom and top quartile was 1.06 units and for PA prediction and 1.36 units for genomic prediction (*Figure 10*). In the genotyped group the

correlation between PA milkability index and MilkabilityE was 0.281 (0.235 – 0.325 and the correlation between genomic milkability index and MilkabilityE was 0.33 (0.293 – 0.377). The confidence intervals showed that there was a significant difference between the two correlations PA milkability index and MilkabilityE and genomic milkability index and MilkabilityE. The correlation between PA milkability index and MilkabilityE for all available SR animals was 0.245 (0.225 – 0.256).

The differences between the bottom and top quartile for temperament was equal to 0.43 units for both genomic and PA prediction (*Figure 11*). In the genotyped group the correlation between PA temperament index and TemperamentE was 0.132 (0.087 – 0.177) and the correlation between genomic temperament index and TemperamentE was 0.135 (0.091 – 0.180). The correlation between PA temperament index and TemperamentE for all available SR animals was 0.14 (0.126 -0.157).

For stillbirths at first calving there was no significant difference between PA and genomic prediction concerning the differences between bottom and top quartiles (*Figure 12*). In the genotyped group the correlation between PA calving index and StillbirthsE was -0.042 (-0.081 – -0.003) and the correlation between genomic calving index and StillbirthsE was -0.020 (-0.059 – 0.018). The correlation between PA calving index and StillbirthsE for all SR available animals was -0.033 (-0.045 – 0.020).

Conformation traits

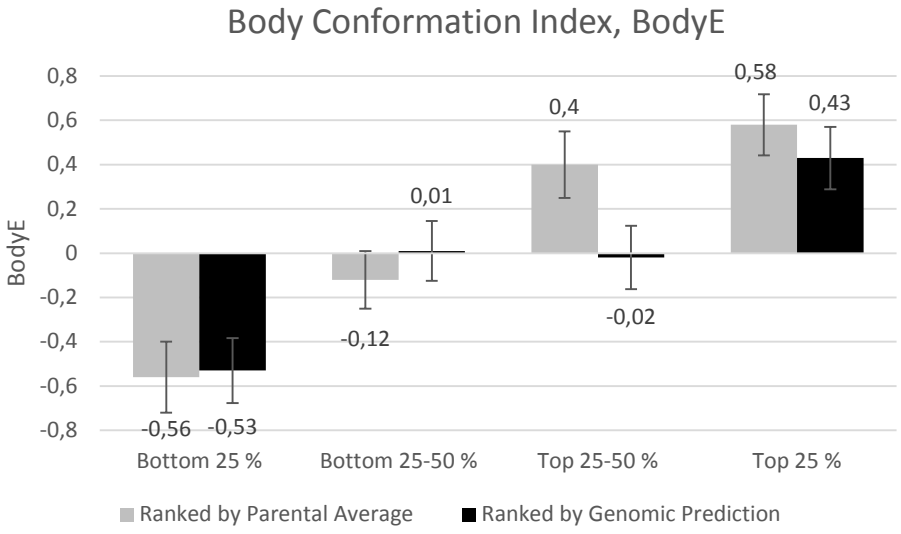

Figure 13. BodyE, adjusted body conformation for 1,897 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

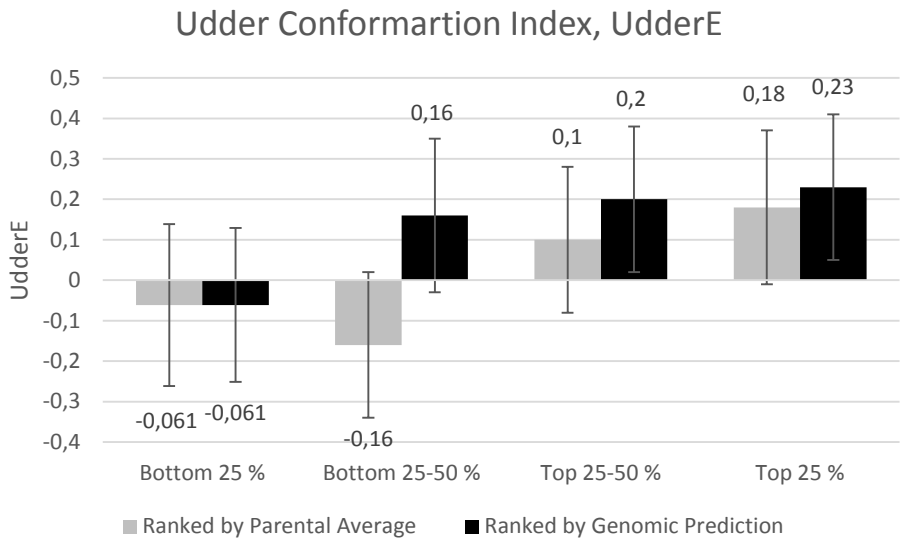

Figure 14. UdderE, adjusted udder conformation for 1,897 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

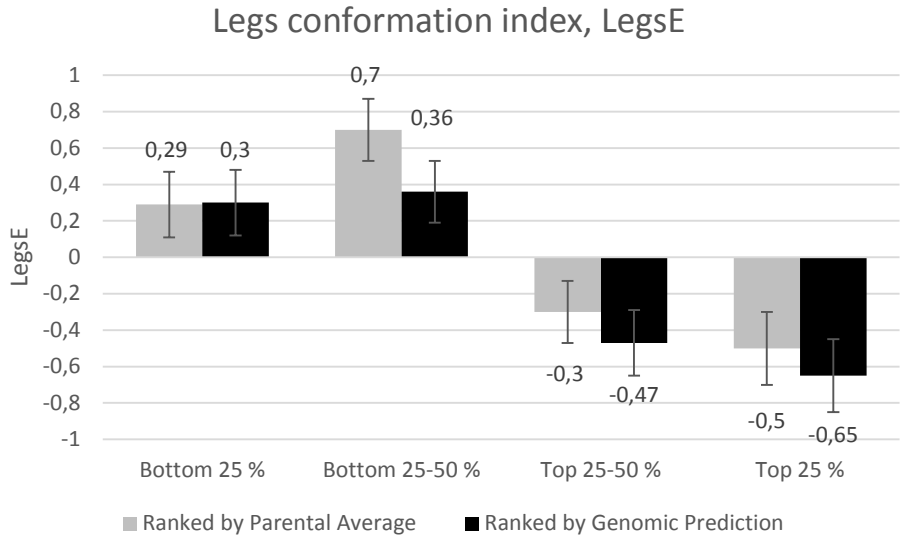

Figure 15. LegsE, adjusted leg conformations for 1,897 Swedish Red Dairy Cattle, according to quartile for parental average prediction the year before calving and quartile for genomic prediction closest before calving.

For body conformation the difference between the bottom and top quartile was 1.14 units for PA prediction and 0.96 units for genomic prediction (*Figure 13*). For the genotyped group the correlation between PA body conformation index and BodyE was 0.133 (0.088 – 0.177) and the correlation between genomic body conformation index and BodyE was 0.112 (0.067 – 0.156).

The correlation was 0.16 between PA body conformation index and BodyE for all available SR animals.

For udder conformation there were no significant differences (*Figure 14*). In the genotyped group the correlation between PA udder conformation index and UdderE was 0.017 (-0.028 – 0.063) and the correlation between genomic udder conformation index and UdderE was 0.024 (-0.021 – 0.069). The correlation between PA udder conformation index and UdderE was 0.005 (0.033 – 0.064) for all available SR animals.

Figure 15 looks the opposite way as expected although there were no significant differences. In the genotyped group the correlation between PA leg conformation index and LegsE was -0.102 (-0.147 – -0.057) and the correlation between genomic leg conformation index and LegsE was -0.077 (-0.122 – -0.032). The correlation between PA leg conformation index and LegsE was 0.001 (-0.015 – 0.016) for all available SR animals.

Survival

Survival was difficult to analyze due to the low number of animals that had had the chance to survive to the second lactation. Most of the animals that had had the chance to survive were genotyped after calving. However, the correlation between PA longevity index and SurvivalE (survived first lactation) was 0.091 (0.068 – 0.114) for SR animals born 2011 and earlier.

Summarizing table

Table 3. *Summarizing table with correlations between adjusted phenotypes and respective GEBV and parental average indexes*

Trait	Correlation with GEBV	Correlation with PA	Difference GEBV-PA
MilkE	0.169	0.154	0.015
FatE	0.229	0.191	0.038
ProteinE	0.214	0.194	0.02
SCSE	-0.164	-0.173	0.009
MastitisE	-0.049	-0.033	-0.016
MastitisfreqE	-0.059	-0.038	-0.021
CFSE	-0.069	-0.064	-0.005
FLSE	-0.034	-0.064	0.030
NINSE	-0.020	-0.018	-0.002
CIE	-0.064	-0.046	-0.018
MilkabilityE	0.330	0.281	0.049
TemperamentE	0.135	0.132	0.003
StillbirthsE	-0.020	-0.042	0.022
BodyE	0.112	0.133	-0.021
UdderE	0.024	0.017	0.007
LegsE	-0.077	-0.102	0.025

Discussion

Data

Almost half of the animals who had started production and had GEBV were removed. This was due to their breeding values were affected by own performance. For those animals the correlations between GEBV and phenotypes were much stronger and did not answer the problem statements of this thesis. The phenotype data set also lacked records, mainly for fertility traits but also for conformation traits, which might have affected the results of those traits (Appendix 3). It also made it harder to achieve significant differences for those traits.

The herds that participate in the LD-project might not be representative for all Swedish herds which could affect the results. For example the average production in this study was 8,640 kg in first lactation for the RDC with GEBV not affected by own performance in (Appendix. 3) and in the Swedish milk recording scheme the production was 8,682 kg for RDC over all lactations (Växa, 2014). The differences between first and second lactation was for all animals 1,548 kg milk (Appendix. 1) and even more between first and third lactation. . There were also lower mastitis incidences in LD herds (Appendix. 1) compared to Swedish milk recording (Växa, 2014). For stillbirths the definition of the analyzed trait differed from what official was published. The national average conformation records were not official published so it was hard to assess whether or not LD herds were representative regarding conformations traits. Together, this meant that it could not be excluded that LD-herds were not representative for all Swedish herds.

Trait analysis

For the three production traits there were stronger correlations between genomic yield index and the respective adjusted phenotypes than between PA yield index and respective adjusted phenotypes. However, the difference was not statistically significant. The differences between quartiles were slightly lower than what was published in Weigel et al. (2015), but their study was made on Holstein and their accuracies of GEBV were higher. Their accuracy for bull GEBV yield were 0.87 compared to 0.64 for yield GEBV in the Nordic cattle genetic evaluation for RDC bulls (Nordic Cattle Genetic Evaluation, 2015). Figures 1 to 3 indicates that selection based on both PA indexes and GEBV would have led to more production. Even though the differences were not statistically significant in this study a selection based on genomic ranking would most likely had led to more production for the farmer or a more accurate choice of replacement and also bull mothers for breeding companies for those traits. The correlations between genomic pure milk, fat and protein indexes and respective traits were stronger which could indicate that the SNP effects for respective traits were captured. It would have been interesting to compare the pure genomic trait indexes with pure PA traits indexes which were not available for this study. The correlation between PA indexes and respective trait for the genotyped group were also in line with the same correlation for all available animals.

The strongest correlation between udder heath index and respective trait was for SCS. This was also the udder health trait with the highest heritability (Nordic Cattle Genetic Evaluation, 2015). The correlation between PA udder index and SCSE was slightly stronger than the correlation between genomic udder index and SCSE. However, the correlation between PA udder index and SCSE for all animals were lower than for the genotyped group, -0.123 compared to -0.171. This indicates that PA udder health, as compared to GEBV was a better predicator of SCS for the genotyped group. In Weigel et al. (2015) the SCS showed almost no difference between quartiles for PA, but their study had bigger differences between quartiles for genomic values. The accuracy for udder heath GEBV was 0.57 in the Nordic cattle genetic evaluation for RDC bull calves born in 2014 (Växa, 2014). The accuracy for bull GEBV udder health was not published in Weigel et al. (2015).

Regarding the two mastitis traits there were no significant differences between correlations (PA vs GEBV and MastitisE/MastitisfreqE) in the genotyped group. It should again be noticed that the mastitis incidences for the three fist lactation in the phenotype data set varied from 6% to 11%. The average mastitis incidences per lactation from the Swedish milk recording scheme was 14% (Växa, 2014). This probably contributed to some of the small differences seen in Figure 5, where mastitis or not between -10 to 150 days were shown. In Figure 6 where mastitis frequency were shown the differences are slightly bigger, for this trait more than one mastitis per cow and lactation could be reported which led to a larger variation. In Figure 6 it seemed like the worst 25% who got the highest risk for mastitis could be identified with GEBV. In practice the worst ones would be of high interest to find, so extra preventive work could be deployed.

Four fertility traits were analyzed. In the genotyped group there were no significant differences between correlations between PA fertility index or genomic fertility index and the four analyzed fertility traits. The correlations were in general weak. Fertility traits have low heritability (Växa, 2014) which means that environment influences the trait to a large extent. Treating animals different depending on their expected result could for example have influenced the results. For example a cow expected to be better than average might have gotten more inseminations before culling decision. CFS was the trait with strongest correlation in the genotyped group and it might be a fertility trait with relative little influenced by environment. However, the correlation between CFS index and fertility index was the lowest 0.60 compared to highest correlation which was 0.97 between FLS index and fertility index (Nordic Cattle Genetic Evaluation, 2015) There are also studies and recommendations on prolonged calving interval which can impact the results (Österman & Bertilsson, 2003). Österman & Bertilsson (2003) suggested to wait with first insemination until after peak lactation and this would then affect many of the fertility traits. The distribution of the trait number of inseminations could also have had an impact on the results. Most of the animals have 1 to 3 inseminations but some have up to 7 which means that the data were not normally distributed. It also should be noticed that a quite big part of the animals lacked fertility records which furthermore could have impacted the results (Appendix 3).

The correlation between genomic milkability index and MilkabilityE was the strongest correlation achieved in this study. There was also a significant stronger correlation between genomic milkability index and MilkabilityE than for PA milkability index and MilkabilityE for the genotyped group. The correlation was also strong between PA milkability index and MilkabilityE for both the genotyped group and for all available animals. The heritability for milkability is relatively high and the accuracy of milkability GEBV was one of the highest for bull calves born in 2014 (Växa, 2014). Milkability is also a relatively clean trait: many of the other traits are combinations of several underlying component traits. For example, the fertility index is a combination of FLS, CI, FLS and NINS indexes. This could be the reason why it was harder to find clear pattern for example fertility. This was nicely illustrated for the yield traits for which there was stronger correlations between pure genomic milk index and MilkE compared to the correlation between yield index and MilkE in the genotyped group.

For temperament the figures and correlation were almost the same for both genomic and PA indexes. Even though there was no significant differences between GEBV and parental average breeding value for many of the analyzed trait, temperament was a good example that the conventional breeding evaluation, without genomic information works. The results for stillbirths showed for the genotyped group a slightly weaker correlation between calving PA index and StillbirthsE than for correlation between genomic index and StillbirthsE. The correlation between PA calving index and StillbirthsE for all available SR animals was also slightly lower than for correlation between PA calving index and StillbirthsE in the genotyped group.

For the three conformation traits in the genotyped group there were results that were difficult to interpret for udder and leg conformation. There were very weak or almost zero correlations for those two traits for any of the groups: this was also shown by the large standard errors bars in Figures 14 and 15. For body conformation there was a slightly stronger correlation for all groups compared to the other conformation traits. Some animals lacked conformation records which could have affected the figures and correlations (Appendix 3). Inconsistent judgement from the official classifiers could possibly be another explanation.

Practical usage

The general trend in this study was that GEBV worked slightly better than parental average for high heritability traits like production traits and milking speed. Overall there were small differences especially for the low heritability traits between and within genomic and parental prediction. When the environments affects the phenotype in such considerable way the remaining proportion of additive genetic effects is very low and a high number of animals is needed to see significant differences. This study was most likely also a bit early as almost half of all genotyped females in Sweden have not calved yet. Furthermore very few cows had records from second and third lactation and thereby also longevity traits were hard to analyze.

The small differences between the predicted best and the predicted worst animals in low heritability traits highlights the importance of good management. The RDC also lack accuracy of

GEBV due to the small reference population. This together leads to a lower genetic progress which was seen in the small difference between bottom and top quartiles in those traits. In the Nordic countries females were included in the reference population in July 2014. The goal was to increase the accuracy and as most of the animals used in this study had GEBV before 2014 and therefor the accuracies were most likely lower then animals born after the introduction. According to the literature females in the reference population increased the reliabilities somewhere from 0.9% to 8%-units (Koivula et al., 2014; Wiggans et al., 2011; Pryce et al., 2012). However, there were still advantageous correlations between indexes and adjusted phenotypes captured for low heritability traits which means that breeding of those traits can help to improve them.

On herd level finding the right replacement is one of the main usage of GEBV (Pryce & Hayes 2012; Calus et al., 2015). The selection intensity can be increased by usage of sexed semen and thereby increase the amount of available heifers for replacement (Calus et al., 2015). The results from this study showed only significantly better prediction for GEBV compared to PA breeding value for milkability, but there were strong tendencies for production traits as well. For example there would be a higher genetic progress for those traits if the bottom 25-50% also were not used as replacement compared to if only the bottom 25% were not used. In Calus et al. (2015) a herd with 100 cows required two more available candidates for replacement than needed per year to make genotyping profitable. Although, with lower replacement rate of cows some losses of yearly genetic trend will occur (Calus et al., 2015).

Another suggested advantage with GEBV was better mating plans (Pryce & Hayes 2012). There could for the traits with the highest differences be better mating plans, although there were also many traits for which no differences or slightly less differences were achieved. Another advantage is keeping control of genetic defects. This was not studied in the present study but could be one important factor when calculating the profitability of genotyping. Still the defect has to be captured on the chip (Pryce & Hayes 2012).

Conclusions

In general GEBV and PA breeding values worked best to predict future phenotypes for high heritability traits. Except for better genomic prediction for milkability there were no significant differences between indexes and their prediction of future phenotypes. There were tendencies that GEBV functioned better than PA breeding values for milk, fat and protein yield. Low accuracy of GEBV and too few records for some traits could be some explanations of the results. Even though there were few significant differences between GEBV and PA breeding values the study indicated that also the conventional genetic evaluation, without genomic information works well for many of the studied traits. Furthermore, the study was made a bit early as some traits could not be analyzed fully because of few completed lactations. Future studies have to be made to confirm the results.

References

Boichard, D., Chung, H., Dassonneville, R., David, X., Eggen, A., Fritz, S., ... & Wiggans, G. R. (2012). Design of a bovine low-density SNP array optimized for imputation. *PloS one*, *7*(3), e34130.

De Roos, A.P.W., Schrooten, C., Veerkamp, R.F., van Arendonk, J.A.M. 2011. Effects of genomic selection on genetic improvement, inbreeding, and merit of young versus proven bulls. Journal of Dairy Science 94, 1559-1567.

Habier, D., Fernando, R. L., & Dekkers, J. C. (2009). Genomic selection using low-density marker panels. *Genetics*, *182*(1), 343-353.

Calus, M. P. L., Bijma, P., & Veerkamp, R. F. (2015). Evaluation of genomic selection for replacement strategies using selection index theory. *Journal of dairy science*, *98*(9), 6499-6509.

Harris, B. L., D. L. Johnson, and R. J. Spelman. (2008). Genomic selection in New Zealand and the implications for national genetic evaluation. Proc. Interbull Meeting, Niagara Falls, Canada.

Hayes, B. J., and M. E. Goddard. (2001). The distribution of the effects of genes affecting quantitative traits in livestock. Genet. Sel. Evol.33:209–229.

Hayes, B.J., Bowman, P.J., Chamberlain, A.C, Goddard, M.E. (2009). Invited review: Genomic selection in dairy cattle: Progress and challenges. Journal of Diary Science 92, 433-443.

Höglund, J. K., Buitenhuis, B., Guldbrandtsen, B., Lund, M. S., & Sahana, G. (2015). Genome-wide association study for female fertility in Nordic Red cattle. *BMC genetics*, *16*(1), 110.

Koivula, M., Strandén, I., Su, G., & Mäntysaari, E. A. (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). *Journal of dairy science*, *95*(7), 4065-4073.

Koivula, M., Strandén, I., Aamand, G. P., & Mäntysaari, E. A. (2014, August). Effect of cow reference group on validation accuracy of genomic evaluation. In *Proceedings, 10th world congress of genetics applied to livestock production*. [Vancouver].

Kargo, M., Hjortø, L., Toivonen, M., Eriksson, J. A., Aamand, G. P., & Pedersen, J. (2014). Economic basis for the Nordic Total Merit Index. *Journal of dairy science*, *97*(12), 7879-7888.

MA, Peipei, et al. Comparison of different methods for imputing genome-wide marker genotypes in Swedish and Finnish Red Cattle. *Journal of dairy science*, 2013, 96.7: 4666-4677.

Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. (2001). Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.

Meuwissen, T. H. (2009). Accuracy of breeding values of 'unrelated'individuals predicted by dense SNP genotyping. *Genet Sel Evol*, *41*(1), 35.

Nordic Cattle Genetic Evaluation. 2015. NAV routine genetic evaluation of Dairy Cattle – data and genetic models. http://www.nordicebv.info / Routine evaluation / NAV routine evaluation – general description / NAV routine evaluation – general description [2015-12-02]

Pryce, J. E., Gonzalez-Recio, O., Thornhill, J. B., Marett, L. C., Wales, W. J., Coffey, M. P., ... & Hayes, B. J. (2014). Short communication: Validation of genomic breeding value predictions for feed intake and feed efficiency traits. *Journal of dairy science*, *97*(1), 537-542.

Pryce, J. E., B. J. Hayes, and M. E. Goddard. "Genotyping dairy females can improve the reliability of genomic selection for young bulls and heifers and provide farmers with new management tools." *Proceedings of the 38th ICAR Session* 28 (2012).

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

Su, G., Brøndum, R. F., Ma, P., Guldbrandtsen, B., Aamand, G. P., & Lund, M. S. (2012). Comparison of genomic predictions using medium-density (~ 54,000) and high-density (~ 777,000) single nucleotide polymorphism marker panels in Nordic Holstein and Red Dairy Cattle populations. *Journal of dairy science*, *95*(8), 4657-4665.

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. *Journal of dairy science*, *98*(3), 2027-2032

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

VikingGenetics. 2015. http://www.vikinggenetics.se/ om-oss/ nyhetsarkiv / [2015-11-12]

Växa. 2014. Husdjursstatistik 2014. http://www.vxa.se / Om oss / Forskning och statistik / Husdjursstatistik 2014 [2015-11-12]

Weigel K. A., Ashley A., Mikshowsky A., & V. E., (2015) Cabrera Effective Use of Genomics in Sire Selection and Replacement Heifer Management. Department of Dairy Science University of Wisconsin, Madison.

Wiggans, G. R., Cooper, T. A., VanRaden, P. M., & Cole, J. B. (2011). Technical note: Adjustment of traditional cow evaluations to improve accuracy of genomic predictions. *Journal of dairy science*, *94*(12), 6188-6193.

Österman, S., & Bertilsson, J. (2003). Extended calving interval in combination with milking two or three times per day: effects on milk production and milk composition. *Livestock production science*, *82*(2), 139-149.

Appendices

Appendix 1.

Descriptive statistics (number of observations (N) and average) of all available animals. Calving interval = CI, calving to first service = CFS, first to last service = FLS, number of services = NINS and somatic cell score = SCS

	Lactation 1		Lactation 2		Lactation 3	
Trait	**N**	**Average**	**N**	**Average**	**N**	**Average**
Milk (kg)	47839	8696.73	24324	10244.63	7129	10716.68
Protein (kg)	47829	304.75	24359	359.26	7145	373.86
Fat (kg)	47796	367.08	24341	429.14	7126	447.69
Stillbirths	49753	0.04	25589	0.02	9518	0.02
CI	26405	384.00	9945	378.46	1157	364.57
CFS	26320	77.92	9905	74.65	1151	70.01
NINS	26320	1.83	9905	1.79	1151	1.57
FLS	26320	28.18	9905	25.65	1151	16.23
Milkability	24308	5.12	597	5.68	982	5.57
Temperament	25555	5.58	605	6.14	1160	5.94
Body Conformation	27277	80.82	741	82.51	1516	82.40
Udder Conformation	27277	80.59	741	82.00	1516	82.33
Leg conformation	27277	81.19	741	81.83	1516	81.77
Mastitis (0/1)	51428	0.05	26914	0.07	10090	0.09
Mastitis frequency	51428	0.06	26914	0.08	10090	0.11
SCS	41724	0.74	20923	0.86	6423	0.98

Appendix 2.

Descriptive statistics (number of observations (N) and average) of all genotyped animals. Calving interval = CI, calving to first service = CFS, first to last service = FLS, number of services = NINS and somatic cell score = SCS.

Trait	N	Average
Milk (kg)	4983	8482.42
Protein (kg)	4981	303.21
Fat (kg)	4976	368.68
Stillbirths	5007	0.0469
CI	2727	380.83
CFS	2725	75.11
NINS	2725	1.87
FLS	2725	28.04
Milkability	3986	5.06
Temperament	4131	5.74
Body Conformation	4239	81.23
Udder Conformation	4239	80.86
Leg conformation	4239	81.24
Mastitis (0/1)	5152	0.045
Mastitis frequency	5152	0.051
SCS	4302	0.69

Appendix 3.

Descriptive statistics (number of observations (N) and average) of all genotyped animals with genomic breeding values without including no own performance. Calving interval = CI, calving to first service = CFS, first to last service = FLS, number of services = NINS and somatic cell score = SCS.

	N	Average
Milk (kg)	2489	8640.33
Protein (kg)	2488	308.64
Fat (kg)	2481	375.45
Stillbirths	2557	0.039
CI	664	373.00
CFI	663	72.59
NINS	663	1.73
FLS	663	22.47
Milkability	1699	5.03
Temperament	1863	5.72
Body Conformation	1897	81.18
Udder Conformation	1897	81.05
Leg conformation	1897	81.13
Mastitis (0/1)	2637	0.047
Mastitis frequency	2637	0.052
SCS	1863	0.69