Slutrapport

Registreringar från automatiska mjölkningssystem som informationskälla i genomisk avelsvärdering

Projektnummer: V1330048
Projekttidsperiod: 1/10/2013-31/12/2018
Huvudsökande:
Erling Strandberg, institutionen för husdjursgenetik, SLU, Erling.Strandberg@slu.se
(tidigare: Freddy Fikse, institutionen för husdjursgenetik, SLU)

Medsökande:
Dirk-Jan de Koning, institutionen för husdjursgenetik, SLU
Forskare i projektet:Sandrine Duchemin och Sreten Andonov, institutionen för husdjursgenetik, SLU

Del 1: Utförlig sammanfattning

Syftet med studien var att härleda det effektivaste sättet att introducera genomisk selektion för nya egenskaper, som de egenskaper som kommer från mjölkrobotar (även kallade automatiska mjölkningssystem eller AMS).

Vi studerade två delar av denna utmaning: 1) Hur kan man förbättra imputeringen av genotyper i nordiska mjölkraser med hjälp av befintliga genotypdata? 2) Vilken säkerhet i den genomiska selektionen kan man uppnå baserat på tidigare resultat och nuvarande status för genotypning av besättningar i Sverige? Detta studerades med hjälp av omfattande simuleringsstudier för ett stort antal scenarier.

Resultaten visade det är möjligt att imputera genotyper med hög säkerhet med nordiska data, inklusive imputering av icke-genotypade djur. Med hjälp av en referenspopulation på 5000 kor med både genotyper och fenotyper kan genomiska avelsvärden skattas med en säkerhet > 0,4 för egenskaper med en måttlig arvbarhet (0,3), förutsatt att genotyperna är korrekt imputerade till medeltäthet (45000 SNP). När man använder ytterligare djur med bara fenotypinformation i en så kallades 'single step' modell (dvs utan ytterligare genotypning) blir noggrannheten för egenskaper med en arvbarhet omkring 0,15 också acceptabel (>0,4). För genomisk avelsvärdering av egenskaper med låg arvbarhet kommer ytterligare genotypning och fenotypning att krävas före genomförandet av genomiskt selektion. Denna studie har visat att genomiskt urval för egenskaper som mäts av mjölkrobotar är tekniskt möjligt i Sverige utan ytterligare investeringar. Lösningar för datautbyte mellan olika ekonomiska intressenter behöver emellertid vara enklare innan detta kan bli verklighet.

Projekt har fått finansiering genom:
Stiftelsen Lantbruksforskning

Del 2: Rapporten

Introduction

According to LRF Sweden and Swedish Agriculture board, there were 4048 dairy farms with a total of 337,000 cows in 2016. About 20% of the farms use a milking robot equating to 1/3 of all the cows. Around 220 farms are enrolled in the routine genotyping program of Viking Genetics and Växa Sverige, corresponding to about 10,000 genotyped cows. About half of these farms use a milking robot. These milking robots collect vast amounts of data that, at present, are mainly used for management advice to the farmers. These traits could also be valuable for selective breeding, in particular when it comes to traits that make animals better suited to robot milking. In an earlier project, we explored the heritabilities of a range of traits that can be derived from milking robot data (as presented in Carlström, 2014; Carlström et al., 2013, 2014, 2016a, 2016b). It was concluded that many of these traits have moderate heritiabilites and could be suitable for selective breeding.

Following the introduction of genomic selection by Meuwissen et al. (2001), the traditional pedigree-based breeding schemes in dairy cattle have shifted to genomic-based breeding schemes in Western societies. The benefits of genomic-based over traditional breeding schemes have been proposed by Schaeffer (2006), and are summarized as follows: a) genomic-based can double the genetic gain per year compared with traditional breeding schemes; b) genomic-based considerably reduce the average age of an animal when a replacement progeny is born (generation interval) of sires of bulls, sires of cows and dams of bulls, and c) genomic-based produce more than 90% savings in logistical costs compared with the traditional breeding schemes.

Furthermore, costs of genotyping cows have decreased considerably in the last years. The low density SNP chip routinely used in Sweden is about 30 € per animal. Although 30 € seems reasonable, it is still the farmer who has to pay. In Sweden, the average farms has 75 cows per herd. The current strategy consists in genotyping all cows in a herd, which can still be considered expensive in terms of time to collect blood samples, processing and genotyping all the cows. Nevertheless, genotyping all cows does not necessarily add relevant information in genomic-based breeding schemes. Only cows with relevant information will help improve genomic predictions and help optimize mating/selection strategies. Thus, genotyping fewer cows may reduce the costs paid by farmers, and this may be done without reducing the relevant information needed for the success of genomic-based breeding schemes.

In the last five years, around 10,000 cows have been genotyped in Sweden. These cows provide an important so-called reference population. A proposed cost-effective alternative to reduce the number of genotyped cows is imputation (Druet et al., 2014). Imputation statistically predicts the genotype of an animal based on a reference population. With imputation, a lower density of genetic markers (e.g., 7,000 SNP) can be imputed to the higher density of genetic markers (e.g., 50,000 SNP). Because imputation is a statistical prediction, uncertainty will be present. Nevertheless, at a certain

accuracy, imputation can be done without prejudice of the information needed for genomic-based breeding schemes. The guarantee that a non-genotyped cow will be accurately imputed will depend on some factors (Daetwyler et al., 2014), such as: a) how sampled and reference population are genetically related; b) that the reference population contains enough genotyped cows and bulls, and c) a not too low minor allele frequency of a genetic marker is needed. Therefore, imputation can increase the number of genetic markers without neither prejudice of information nor increasing costs for farmers.

The aim of the present study was to derive the most efficient way to introduce genomic selection for novel traits, such as those derived from milking robots (also called automatic or voluntary milking systems or AMS/VMS). We studied both the use of imputation and the implementation of genomic selection using a range of scenarios. This shows what accuracy of genomic prediction we can achieve now for novel traits, and how we can increase that level in the future.

Material and Methods

Imputation of genotypes in Nordic Red cattle. Originally, our reference population contained 400 genotyped Nordic Red cows for 777,000 SNP genotypes. To exemplify that imputation is a feasible approach, we performed a 5-fold cross-validation of an imputation. Our sampled population of cows was created by randomly selecting 20% of the reference population (i.e., 80 cows). This sampling was repeated 5 times until all the cows had been in the prediction set exactly once. To create a panel of genotypes with lower marker density, we masked the genotypes of our sampled population resulting in a 50,000 SNP genotypes panel in each prediction set. Subsequently, we imputed each set at a time (i.e., 80 cows with 50,000 SNP genotypes) based on a reference population of 320 cows with 777,000 SNP genotypes. This imputation was performed in Beagle software (Browning and Browning, 2009). Accuracy of imputation was estimated using the correlation of the imputed genotype with the true genotypes.

Imputation of genotypes in Nordic Holstein cattle. In a related study, we examined the efficiency of imputation in Nordic Holstein, in particular the imputation of the X chromosome and imputation of non-genotyped animals (Mao et al., 2016). There were 26,884 genotyped animals in our dataset, which spanned multiple countries. These animals were genotyped by Illumina BovineSNP50 BeadChip version 1 and 2 (50k). Imputation was carried out using FImpute V2.2 (Sargolzaei et al., 2014). Target groups were generated by randomly masking 26,884 50k animals to 2,000 non-genotyped (0k) animals or 2,000 Illumina BovineLD BeadChip typed (7k) animals, leaving 24,884 animals in the dataset as a reference set.

Genomic prediction for a novel trait. At present around 220 farms comprising around 10,000 cows are enrolled in the genotyping program of VikingGenetics in Sweden. About half of these have a milking robot, so we can assume that about 5000 cows have both genotypes and potential access to AMS phenotypes (Växa Sverige, personal communication). Here we explored via simulation what accuracy for genomic selection could be expected if we started genomic selection for new traits, based on the AMS data

using different scenarios. Traits were divided into three groups based on their heritability (h^2) (as presented in Carlström, 2014; Carlström et al., 2013, 2014, 2016a, 2016b). Group 1 (with a simulated $h^2 = 0.30$) is representative for traits with moderate heritability like average flow rate (0.27-0.54); milking time (0.33-0.44); box time (0.21-0.44); proportion attachment (0.21-0.31); and udder conformation (0.18-0.45). Group 2 (simulated $h^2 = 0.15$) represents traits like handling time (0.05-0.15); milking interval (0.09-0.26); general temperament (0.08-0.15), and lactation average somatic cell score (SCS, 0.11-0.17). The third group are traits with a low heritability (h^2=0.05) such as number of milkings (0.02-0.07) and proportion incomplete milkings (0.02-0.06).

To mimic a Nordic cattle population, data were generated with QMSim simulation software (Sargolzaei and Schenkel, 2009) with 6 replicates. At the beginning, in order to produce a realistic level of linkage disequilibrium (LD close to 0.5), 800 generations of a historical population were generated with initially 3,000 animals, increasing to 40,000 at generation 500, and then decreasing to 30,000 by the last generation; one-tenth of animals were sires. Then complete data was generated for 20,000 cows and 2000 bulls over 20 generations. Single records were generated for dams, and heritabilities of 0.3, 0.15 and 0.05. The replacement rate for sires and dams was 90% and 30%, respectively. Due to intention to imitate the beginning of the breeding program where those traits will be introduced, in the simulation process the mating was defined random and the selection criterion was based on high values for phenotypes.

Genotypes were simulated for 45,000 biallelic SNP markers distributed along 29 chromosomes with a total length of 2,319 cM, which mimicked the bovine genome without sex chromosomes. A total of 450 biallelic and randomly distributed QTL affected the trait, with effects sampled from a normal distribution. Genomic information was generated for generations 16 through 20. Also, data for low density SNP chip was created by taking every 7th SNP, resulting in 5,959 markers. In the simulations, the reference population was always 5,000 cows with genotypes and phenotypes.

We first wanted to test how well this reference population would predict the breeding values of other animals using a range of scenarios: 1) for within-herd selection we tested how well the 5,000 cows in the reference population predicted the estimated breeding value (EBV) of their own daughters that only have genotype information (scenario DD_G). 2) In order to mimic across herd selection, we estimated how well the reference cows predicted for their half-sibs in other herd (scenario HS_G). 3) We also tested how well the current reference population would predict 20,000 random cows from the population (RAND_G).

Another approach to reduce genotyping costs, beside imputation, is to combine animals that have genotyping data and phenotyping data with those animals that have only phenotype data. With the development of so-called ‘Single Step’ genomic evaluation we can use all the animals that have phenotypic data for genetic evaluation, regardless of their genotyping status (Legarra et al., 2014). We evaluated several single-step scenarios, where we combine the current reference population of 5000 cows with the phenotypic data of 1) 5,000 daughters (DD_P), 2) 5,000 half-sibs (HS_P) or 3) 20,000 random cows

that are currently not genotyped (RAND_P). Finally we looked at a scenario where the 20,000 random animals had both phenotypes and genotypes (RAND_GP). The scenario would use existing phenotype data but require and additional genotyping effort.
For each scenario, we tested the accuracy of the predicted breeding values when using a) pedigree information only, b) a low density (LD) SNP chip with 5959 SNP and c) a medium density (MD) SNP chip with 45K SNP.

Accuracy of EBVs was calculated as the correlation between true and estimated breeding values. All analyses were done using software from the BLUPF90 family (Misztal et al., 2015). Genomic EBVs were estimated using ssGBLUP (Aguilar et al., 2010). For simplicity and because all data were simulated, the model used for all scenarios accounted for 1 fixed effect (overall mean), a random animal genetic effect, and a random residual:

$$\mathbf{y}=\mathbf{1}\boldsymbol{\mu}+\mathbf{Z}\mathbf{a}+\mathbf{e}$$

where y is the vector of phenotypes, $\boldsymbol{\mu}$ is a vector of fixed effect (overall mean), **a** is the vector of additive animal effect, **e** is the vector of random residual effect and **Z** is the incidence matrix for the random effect in **a**. It was assumed that $\mathbf{a} \sim N(\mathbf{0}, \mathbf{H}\sigma^2_a)$; in which σ^2_a is the additive genetic variance and **H** is the matrix that combines pedigree and genomic relationships. In all analyses we included five generations of pedigree information.

Results and Discussion

Imputation in Nordic Red cattle. The accuracy of imputation for this 5-fold cross-validation is equal to 0.805 (Table 1). In practice, this means that >80% of the imputed data is accurately imputed at a higher density of genotypes.

Table 1. Accuracy of a 5-fold cross-validation of an imputation* from lower density (50,000 SNP) to higher density of genetic markers (777,000 SNP)

FOLD	N**	AVERAGE	MINIMUM	MAXIMUM
1	80	0.807	0.765	0.861
2	80	0.804	0.774	0.828
3	80	0.803	0.763	0.862
4	80	0.804	0.763	0.861
5	80	0.806	0.765	0.862
TOTAL	80	0.805		

*The reference population of each fold is composed of 320 genotyped cows for 777,000 SNP genotypes.
**N= sampled population

Imputation in Nordic Holstein cattle. In Mao et al (2016), we showed that imputation for non-genotyped animals was possible when close family relationships were present. The effects of different male/female ratio in the reference group on the imputation accuracy of the X chromosome, BTAs, and BTA2 are shown in Table 2. The results showed that in general, more females in the reference group increased the imputation accuracy, especially for the X chromosome. For SNP-wise accuracy, a reference group

of 8,000 males gave an accuracy of 0.74 for the X chromosome in the scenario of imputing from 0k to 50k, while 8,000 females resulted in an accuracy of 0.97. In general, imputation from 0k to 50k was less accurate than from 7k to 50k. For example, SNP-wise accuracy of 0k to 50k for BTAs was on average 10% lower than 7k to 50k (Table 2).

Table 2. SNP-wise imputation accuracy in Nordic Holstein for different scenarios. (adapted from Mao et al., 2016)

		0k -> 50k[1]				7k -> 50k[2]			
	No. of animals	BTA[3]	Chr2[4]	X[5]	PAR[6]	BTA[3]	Chr2[4]	X[5]	PAR[6]
Size of reference	24884	0.89	0.89	0.95	0.81	0.98	0.97	0.95	0.85
	20000	0.88	0.88	0.95	0.79	0.97	0.97	0.94	0.80
	15000	0.87	0.87	0.94	0.78	0.96	0.96	0.94	0.77
Male/Female	8000/0	0.85	0.84	0.74	0.68	0.95	0.95	0.92	0.67
	6000/2000	0.86	0.85	0.95	0.75	0.96	0.96	0.93	0.72
	4000/4000	0.85	0.84	0.96	0.77	0.95	0.95	0.94	0.69
	2000/6000	0.86	0.86	0.97	0.78	0.95	0.96	0.94	0.68
	0/8000	0.88	0.87	0.97	0.78	0.96	0.96	0.96	0.63

[1]0k ->50k: scenario of imputing non-genotyped animals to 50k
[2]7k -> 50k: scenario of imputing 7k animals to 50k
[3]BTA: average of 29 *Bos taurus* autosomes
[4]Chr2: Chromosome 2
[5]X: X chromosome
[6]PAR: pseudo-autosome region

Furthermore, the accuracy of imputation increases when more cows are present in the reference population. According to Pausch et al. (2013), an imputation accuracy of about 99% can be obtained when a reference population of >400 animals is used. This is clearly illustrated by the differences between the Nordic Red (Table1) and the Nordic Holstein (Table 2). The latter shows imputation accuracy > 0.85 for non-genotyped individuals using a large reference population. The combination of imputation with a reduced number of genotyped cows at farm level can be an efficient way of reducing the farmers' costs. By using this cost-effective alternative, we help optimize the decision-making process of the farmer in deciding to genotype less cows.

Genomic prediction for a novel trait. The estimated accuracies for the simulated data are summarized in Table 3 and Table 4. The first striking observation is that with the low density genotyping, the genomic prediction of the breeding values is hardly better than using only pedigree data. The importance of imputation is well illustrated by the compelling increase to a higher density as this shows clear improvement of accuracy across all scenarios. Those findings are in agreement with studies by Solberg et al. (2008), Erbe et al. (2012), and He et al. (2017). When predicting young animals that have only genotypes but no phenotypes, the accuracy for predicting daughters' breeding value is the

same as predicting half-sibs (Table 3). In all cases slightly better prediction of performance of young animals were obtained when within and across herd selection was performed compared to random selection of animals (Table 3). The reason for this is due to closely related reference population to young animals (dam –daughter or half sibs), compared to randomly chosen animals. For the traits with a moderate heritability of 0.30, the accuracy using the medium density SNP information was around 0.4 across scenarios. This level has been suggested as a decent starting point for genomic selection. One example is the recent introduction of a new breeding value for 'feed saved' in Australia, which was introduced with an initial accuracy of 0.4 (Dr. Jennie Pryce, LaTrobe University, personal communication). For traits with lower heritabilities it is clearly desirable to expand the reference population (Table 3).

When using single-step genomic prediction and including phenotypes of daughters (DD_P), half sibs (HS_P), or random animals (RANDOM_P) to the reference population, the accuracies improve for all scenarios (Table 4). As expected, estimates of traits with higher heritability were more accurately predicated than those with lower heritability. The phenotypic information from daughters contributes almost as much as taking phenotypic information from half-sibs. (Table 4). The latter has the additional benefit that this data is readily available at an earlier time. The highest accuracy comes from including phenotypes from 20,000 random animals (Table 4).

When we have genotyping and phenotyping on an additional 20,000 cows, randomly selected from the population, the accuracies are > 0,40 for all heritabilities, provided the data are accurately imputed to medium density (Figure 1).

Table 3. Estimated accuracy (± empirical SE) as a correlation between true breeding values and estimated breeding values for different scenarios using a reference population of 5,000 cows.

	DD_G		
	MD	LD	PED
h^2=0.30	0.41±0.02	0.33±0.01	0.34±0.01
h^2=0.15	0.25±0.01	0.22±0.01	0.22±0.01
h^2=0.05	0.18±0.02	0.16±0.03	0.15±0.02
	HS_G		
	MD	LD	PED
h^2=0.30	0.41±0.02	0.33±0.01	0.34±0.01
h^2=0.15	0.25±0.01	0.22±0.01	0.22±0.01
h^2=0.05	0.18±0.02	0.16±0.03	0.15±0.02
	RAND_G		
	MD	LD	PED
h^2=0.30	0.38±0.01	0.30±0.01	0.30±0.01
h^2=0.15	0.27±0.01	0.21±0.01	0.20±0.01
h^2=0.05	0.15±0.00	0.14±0.01	0.14±0.01

DD_G is a scenario where 5,000 reference cows with genotypes and phenotypes are used to predict breeding values of their 5,000 daughters with genotypes only. HS_G is the scenario where the 5,000 reference cows are used to predict 5,000 of their half sibs with genotypes only (across herd design). RAND_G is a scenario where the 5,000 reference cows are used to predict 20,000 randomly selected cows with genotypes only (across population). PED shows the accuracy of the EBV that can be obtained using only the pedigree information. LD refers to genomic selection using 5,959 SNP and MD refers to genomic selection with 45K SNP.

Table 4. Estimated accuracy (± empirical SE) as a correlation between true breeding values and estimated breeding values using single step genomic prediction for different scenarios.

	DD_P		
	MD	LD	PED
h^2=0.30	0.55±0.00	0.51±0.01	0.51±0.01
h^2=0.15	0.40±0.01	0.37±0.01	0.37±0.01
h^2=0.05	0.26±0.01	0.24±0.01	0.24±0.01
	HS_P		
	MD	LD	PED
h^2=0.30	0.56±0.00	0.52±0.01	0.52±0.01
h^2=0.15	0.42±0.00	0.38±0.01	0.39±0.01
h^2=0.05	0.27±0.01	0.25±0.01	0.25±0.01
	RAND_P		
	MD	LD	PED
h^2=0.30	0.60±0.00	0.56±0.00	0.56±0.00
h^2=0.15	0.46±0.00	0.43±0.01	0.43±0.00
h^2=0.05	0.30±0.01	0.28±0.01	0.28±0.01

DD_P is a scenario where the reference population with genotypes and phenotypes of the 5,000 reference cows is combined with phenotypes of 5,000 daughters. HS_P is the scenario where the 5,000 reference cows are combined with phenotypes of 5,000 of their half sibs (across herd design). RAND_P is a scenario where the 5000 reference cows are combined with phenotypes from 20,000 randomly selected cows (across population). PED shows the accuracy of the EBV that can be obtained using only the pedigree information. LD refers to genomic selection using 5,959 SNP and MD refers to genomic selection with 45K SNP.

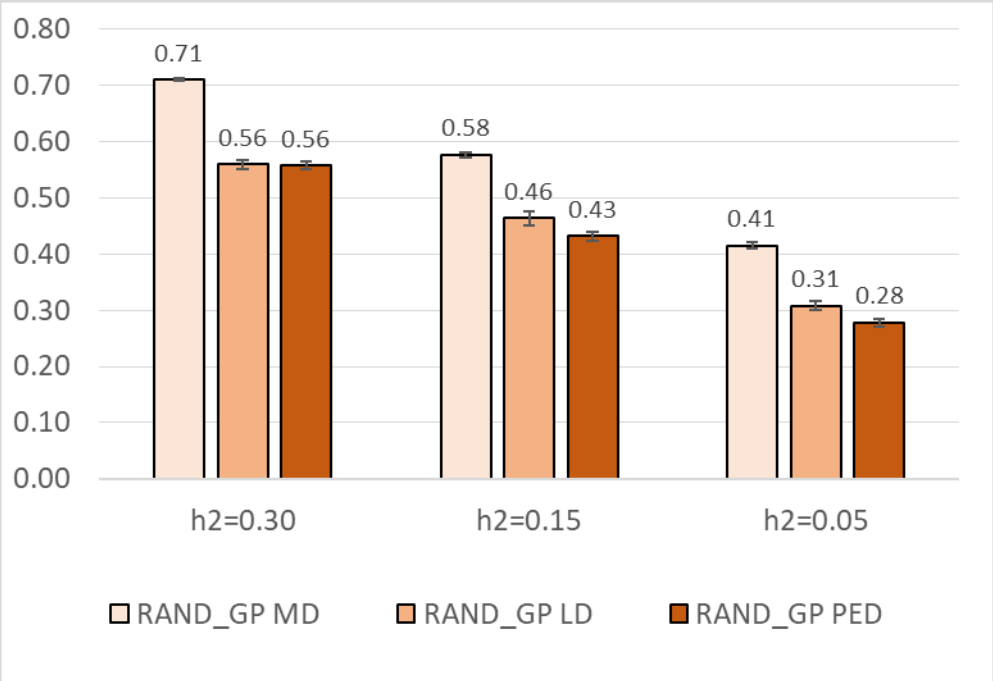

Figure 1. The expected accuracies when adding 20,000 random cows with phenotypes and genotypes to the current reference population of 5000 cows (scenario RAND_GP). PED shows the accuracy of the EBV that can be obtained using only the pedigree information. LD refers to genomic selection using 5,959 SNP and MD refers to genomic selection with 45K SNP.

Conclusions

This study shows that with the current level of genotyping and use of AMS data, genomic selection with acceptable accuracies can be realized for novel traits with low to moderate heritabilities. Accuracies around 0.4 can be achieved using the current data from farms that have both milking robots and are enrolled in the genotyping program by Viking Genetics. Genotyping data from these farms are routinely imputed to 50K SNP density using the large reference population (several tens of thousands across the Nordic countries) at Viking Genetics. Higher accuracies of genomic prediction can be achieved by using single step genomic selection using, potentially, all cows with milking robot data. It is important to note that is can be achieved **now**, without further investments by either the breeding companies, the robot manufacturers, or the farmers themselves.

The main challenge to start using this kind of data in genomic evaluation is that the data from milking robots reside with the companies that sell the milking robots and the associated management software. In order to share this kind of data for genomic evaluations, there needs to be a clear incentive for these companies.

In conclusion we can state that the implementation of genomic selection for data derived from milking robots is technically possible, but requires 'political' solutions to become a reality.

Recommendations for industry

In order to move genomic selection using AMS data from a technical possibility to a reality, all which is required is to develop protocols for data exchange between the robot providers and the breeding companies. We have started the required dialogue within this project but could not bring it to a stage where we could use real data to pilot this approach. The dialogue should be continued by the main stakeholders: Växa Sverige, Viking Genetics, De Laval and possibly also Lely. LRF could have a crucial role.

References

Aguilar I., Misztal I., Johnson D. L., Legarra A., Tsuruta S., Lawlor T.J.. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy Sci. 93:743–752. http://dx.doi.org/10.3168/jds.2009-2730

Browning, B. L., and S. R. Browning. 2009. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am. J. Hum. Genet. 84:210–223

Carlström, 2014. Genetic Variation of In-line Recorded Milkability Traits and Associations with Udder Conformation and Health in Swedish Dairy Cattle. Doctoral Thesis. Swedish University of Agricultural Sciences. Uppsala.

Carlström C., Pettersson G., Johansson K., Strandberg E., Stålhammar H., Philipsson J. 2013. Feasibility of using automatic milking system data from commercial herds for genetic analysis of milkability. J. Dairy Sci. 96 :5324–5332 http://dx.doi.org/ 10.3168/jds.2012-6221.

Carlström C., Strandberg E., Johansson K., Pettersson G., Stålhammar H., Philipsson J. 2014. Genetic evaluation of in-line recorded milkability from milking parlors and automatic milking systems. J. Dairy Sci. 97 :497–506. http://dx.doi.org/ 10.3168/jds.2013-6948

Carlström C., Strandberg E., Johansson K., Pettersson G., Stålhammar H., Philipsson J. 2016a. Genetic associations of in-line recorded milkability traits and udder conformation with udder health. Acta Agriculturae Scandinavica, Section A — Animal Science, 66:2, 84-91, DOI: 10.1080/09064702.2016.1260154.

Carlström C., Strandberg E., Pettersson G., Johansson K., Stålhammar H., Philipsson J. 2016b. Genetic associations of teat cup attachment failures, incomplete milkings, and handling time in automatic milking systems with milkability, temperament, and udder conformation. Acta Agriculturae Scandinavica, Section A — Animal Science, 66:2, 75-83, DOI: 10.1080/09064702.2016.1260153

Daetwyler, H. D., Capitan, A., Pausch, H., Stothard, P., van Binsbergen, R., Brøndum, R. F., et al. (2014). Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nat. Genet. 46, 858–865. doi: 10.1038/ng.3034

Druet, T., Macleod, I. M., and Hayes, B. J. (2014). Toward genomic prediction from whole-genome sequence data: impact of sequencing design on genotype imputation and accuracy of predictions. Heredity (Edinb). 112, 39–47. doi: 10.1038/hdy.2013.13

Erbe M., Hayes B.J., Matukumalli L.K., Goswami S., Bowman P.J., Reich C.M., Mason B.A., Goddard M.E. 2012. Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. J Dairy Sci 95:4114–4129. http://dx.doi.org/ 10.3168/jds.2011-5019.

He J., Xu J., Wu X.L., Bauck S., Lee J., Morota G., Kachman S.D., Spangler M.L. 2018. Comparing strategies for selection of low-density SNPs for imputation-mediated genomic prediction in U.S. Holsteins. Genetica. 146(2):137-149. doi: 10.1007/s10709-017-0004-9.

Legarra, A., Christensen, O.F., Aguilar, I., and Misztal, I. 2014. Single Step, a general approach for genomic selection. Livestock Science 166: 54-65

Mao, X., Johansson, A.M., Sahana, G., Guldbrandtsen, B. and De Koning, D.J. 2016. Imputation of markers on the bovine X chromosome. J. Dairy Sci. 99 (9), 7313-7318

Meuwissen, T. H., Hayes, B. J., and M. E. Goddard. 2001.Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, (4): 1819-1829.

Misztal, I., S. Tsuruta, D. Lourenco, I. Aguilar, A. Legarra, and Z. Vitezica. 2015. Manual for BLUPF90 family of programs. http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf.

Pausch, H., B. Aigner, R. Emmerling, C. Edel, K.-U. Götz, and R. Fries. 2013. Imputation of high-density genotypes in the Fleckvieh cattle population. Genet. Sel. Evol. 45:3.

Sargolzaei, M., and F. S. Schenkel. 2009. QMSIM: A large-scale genome simulator for livestock. Bioinformatics 25:680–681. http://dx.doi.org/10.1093/bioinformatics/btp045.

Sargolzaei, M, Chesnais, JP, Schenkel, FS. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15: 478.

Schaeffer LR. Strategy for applying genome-wide selection in dairy cattle. Journal of animal Breeding and genetics. 2006 Aug;123(4):218-23.

Solberg T.R., Sonesson A.K., Woolliams J.A., Meuwissen T.H.E. 2008. Genomic selection using different marker types and densities. J Anim Sci 2008.86:2447-2454. doi:10.2527/jas.2007-0010.

Del 3: Resultatförmedling

Vetenskapliga publiceringar	Carlström C., Strandberg E., Johansson K., Pettersson G., Stålhammar H., Philipsson J. 2016a. Genetic associations of in-line recorded milkability traits and udder conformation with udder health. Acta Agriculturae Scandinavica, Section A — Animal Science, 66:2, 84-91, DOI: 10.1080/09064702.2016.1260154.
	Andonov, S. and De Koning, D.J. Bias and accuracy in genomic prediction for new breeding goal traits., abstract for Gordon Conference in Quantitative Genetics and Genomics, February 10 - 15, 2019, Barga Italy
	Andonov, S. and De Koning, D.J. 2019. Expected accuracy in genomic prediction for new breeding goal traits in dairy cattle, a simulation approach. To be submitted to Acta Agriculturae Scandinavica
Studentarbete	
	Milk Symposium at SLU, 28 January 2019, with invited presentations from DeLaval, Växa Sverige and Viking Genetics. Chair of organizing committee: Prof. DJ de Koning.
Övrigt	**GIGACOW**, a new research infrastructure by SLU to enable precision dairy farming and genetics research. Led by Prof. DJ de Koning. This infrastructure will collect detailed phenotypes and genotypes on around 3000 dairy cows using milkrobot data and sensor technology.