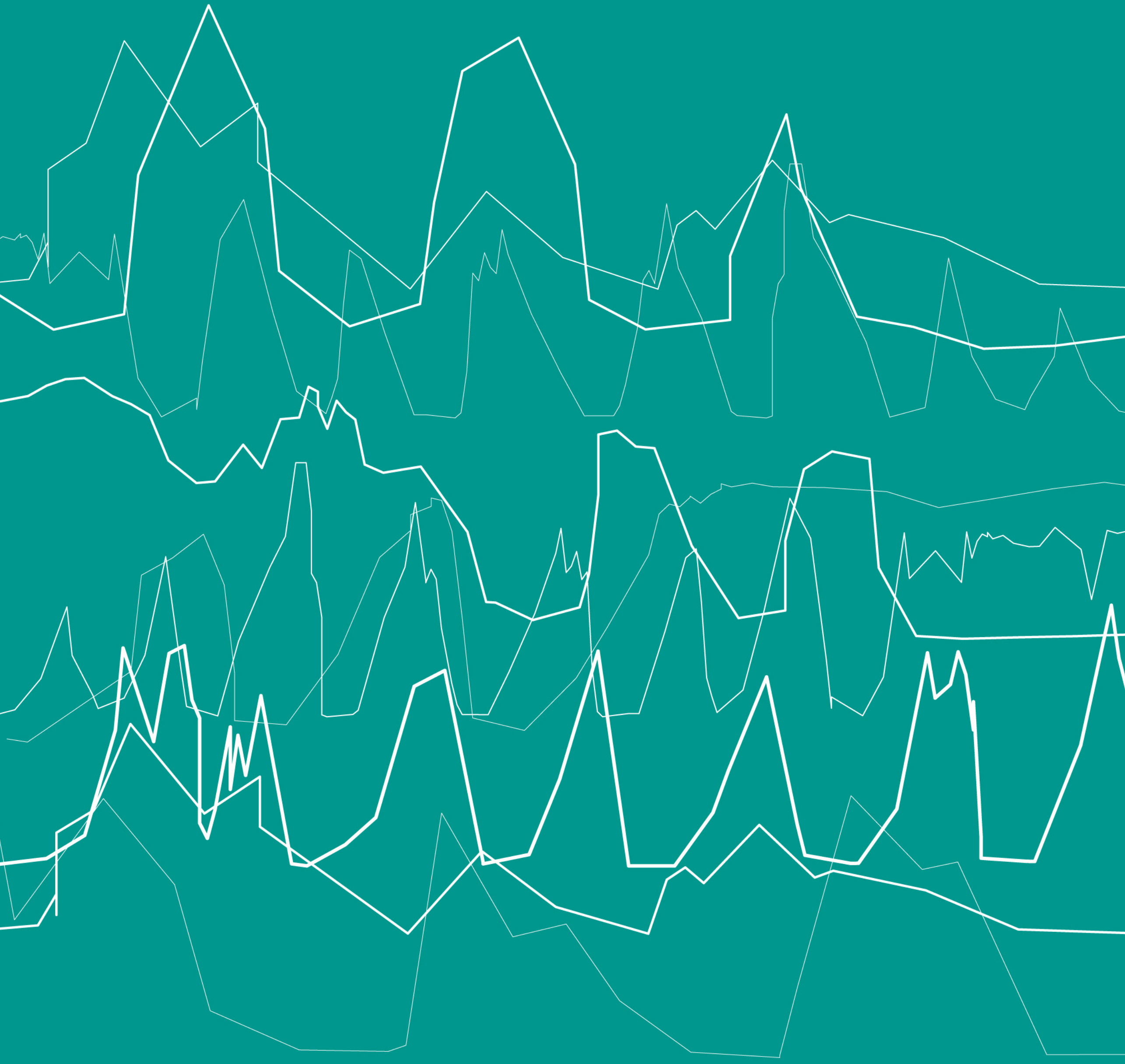
Milk progesterone measures to improve genomic selection for fertility in dairy cows

Amabel M. M. Tenghe

ACTA UNIVERSITATIS AGRICULTURAE SUECIAE
DOCTORAL THESIS No. 2017:37



ISSN 1652-6880
ISBN (print version): 978-91-576-8847-7
ISBN (electronic version): 978-91-576-8848-4

Milk progesterone measures to improve genomic selection for fertility in dairy cows

Amabel Manyu Mefru Tenghe

Thesis committee

Promotors
Prof. Dr R.F. Veerkamp
Professor of Numerical Genetics and Genomics
Wageningen University & Research

Prof. Dr B. Berglund
Professor in Animal Breeding and Genetics
Swedish University of Agricultural Sciences

Co-promotor
Prof. Dr D.J. de Koning
Professor in Animal Breeding and Genetics
Swedish University of Agricultural Sciences

Dr A.C. Bouwman
Researcher, Animal Breeding and Genomics Center
Wageningen University & Research

Other members
Prof. Dr B. Kemp, Wageningen University and Research
Dr S. van Beek, Cooperative Cattle Improvement Organization CRV BV
Dr B. Heringstad, Norwegian University of Life Sciences
Prof. U. Emanuelson, Swedish University of Agricultural Sciences

This research was conducted under the joint auspices of the Swedish University of Agricultural Sciences and the Graduate School Wageningen Institute of Animal Sciences of Wageningen University and is part of the Erasmus Mundus Joint Doctorate Program EGS-ABG.

Milk progesterone measures to improve genomic selection for fertility in dairy cows

Amabel Manyu Mefru Tenghe

ACTA UNIVERSITATIS AGRICULTURAE SUECIAE
DOCTORAL THESIS No. 2017:37

Thesis
submitted in fulfilment of the requirements for the joint degree of doctor between
Swedish University of Agricultural Sciences
by the authority of the board of the Faculty of Veterinary Medicine and Animal Sciences
and
Wageningen University
by the authority of the Rector Magnificus, Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board of Wageningen University and the Board of the Faculty of Veterinary Medicine and Animal Sciences at
the Swedish University of Agricultural Sciences
to be defended in public
on Friday 28 April 2017
at 4 p.m. in the Aula of Wageningen University.

Amabel M.M. Tenghe.
Milk progesterone measures to improve genomic selection for fertility in dairy cows

Joint PhD thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden and Wageningen University, the Netherlands (2017)
With references, with summary in English

ISSN: 1652-6880
ISBN: (print version): 978-91-576-8847-7
ISBN: (electronic version): 978-91-576-8848-4
ISBN: 978-94-6343-133-0
DOI: 10.18174/409680

Abstract

Tenghe, A.M.M. (2017). Milk progesterone measures to improve genomic selection for fertility in dairy cows. Joint PhD thesis, between Swedish University of Agricultural Sciences, Sweden and Wageningen University, the Netherlands

Improved reproductive performance has a substantial benefit for the overall profitability of dairy cattle farming by decreasing insemination and veterinary treatment costs, shortening calving intervals, and lowering the rate of involuntary culling. Unfortunately, the low heritability of classical fertility traits derived from calving and insemination data makes genetic improvement by traditional animal breeding slow. Therefore, there is an interest in finding novel measures of fertility that have a higher heritability or using genomic information to aid genetic selection for fertility. The overall objective of this thesis was to explore the use of milk progesterone (P4) records and genomic information to improve selection for fertility in dairy cows. In a first step, the use of in-line milk progesterone records to define endocrine fertility traits was investigated, and genetic parameters estimated. Several defined endocrine fertility traits were heritable, and showed a reasonable repeatability. Also, the genetic correlation of milk production traits with endocrine fertility traits were considerably lower than the correlations of milk production with classical fertility traits. In the next step 17 quantitative trait loci (QTL) associated with endocrine fertility traits, were identified on *Bos taurus* autosomes (BTA) 2, 3, 8, 12, 15, 17, 23, and 25 in a genome-wide association study with single nucleotide polymorphisms. Further, fine-mapping of target regions on BTA 2 and 3, identified several associated variants and potential candidate genes underlying endocrine fertility traits. Subsequently, the optimal use of endocrine fertility traits in genomic evaluations was investigated; using empirical and theoretical predictions for single-trait models, I showed that endocrine fertility traits have more predictive ability than classical fertility traits. The accuracy of genomic prediction was also substantially improved when endocrine and classical fertility traits were combined in multi-trait genomic prediction. Finally, using deterministic predictions, the potential accuracy of multi-trait genomic selection when combining a cow training population measured for the endocrine trait commencement of luteal activity (C-LA), with a training population of bulls with daughter observations for a classical fertility trait was investigated. Results showed that for prediction of fertility, there is no benefit of investing in a cow training population when the breeding goal is based on classical fertility traits. However, when considering a more biological breeding goal for fertility like C-LA, accuracy is substantially improved when endocrine traits are available from a limited number of farms.

List of publications

This thesis is based on the work contained in the following papers

I. **A.M.M. Tenghe,** B. Berglund, R.F. Veerkamp, and D.J. de Koning. 2016. Opportunities for genomic prediction for fertility using endocrine and classical fertility traits in dairy cattle. J. Anim. Sci.

II. **A.M.M. Tenghe,** A.C. Bouwman, B. Berglund, E. Strandberg, D.J. de Koning, and R.F. Veerkamp. 2016. Genome-wide association study for endocrine fertility traits using single nucleotide polymorphism arrays and sequence variants in dairy cattle. J. Dairy Sci. 99:5470–5485

III. **A.M.M. Tenghe,** A.C. Bouwman, B. Berglund, E. Strandberg, J.Y. Blom, and R.F. Veerkamp. 2015. Estimating genetic parameters for fertility in dairy cows from in-line milk progesterone profiles. J. Dairy Sci. 98:5763–73

IV. **A.M.M. Tenghe,** A.C. Bouwman, B. Berglund, D.J. de Koning and R.F. Veerkamp. Improving accuracy of bulls ‘predicted genomic breeding values for fertility using daughters’ milk progesterone profiles (Submitted to Journal of Dairy Science)

1

General introduction

1.1 Importance of fertility

Fertility is one of the non-yield traits which is of great economic importance in dairy herds. Poor fertility increases costs due to fertility treatments, multiple inseminations, prolongs calving interval, and leads to a high replacement rate due to involuntary culling (Boichard, 1990; Dekkers, 1991; González-Recio et al., 2004). The economic consequences of poor fertility have been widely studied. Inchaisri et al., (2010) studied the economic consequences of non-optimal fertility of dairy cows under Dutch conditions, where non-optimal fertility was defined as “average” or “poor” reproductive performance, using different fertility traits. They reported mean net economic losses of €34 and €231 per cow per year for “average” and “poor” reproductive performance respectively, compared to a “good” fertility. The losses were mainly caused by decreased milk production and increased number of non-pregnant cows, especially in the situation of poor fertility. Also, a net cost of a one day increase in calving interval was €0.57 per cow per day for “average” fertility, and €0.70 per cow per day for “poor” fertility (Inchaisri et al., 2010). Other studies have reported costs of one day extra for of €0.06 to €1.10 for days open (de Vries and Conlin, 2003; Groenendaal et al., 2004; Meadows et al., 2005) and €2.07 to €2.95 for calving interval (Plaizier et al., 1997). This economic importance is one of the reasons why there is an interest in genetic improvement of fertility.

1.2 Fertility measures used in genetic evaluation

Overall fertility can be described and measured by several different traits. For example, it is useful to distinguish among traits which are affected by the cow (female fertility), and traits affected by the sire mated to the cow (male fertility), and traits affected by both. Traits like age at first puberty in heifers, estrous expression, and the time between calving and regular ovarian activity or first insemination measure female fertility, whereas traits like sperm count and semen quality measure male fertility. Traits that measure the results of insemination and conception rate are influenced by both male and female fertility, and a combination of them. They include non-return rate, number of inseminations per conception or service period, percentage of cows bred that are pregnant (conception rate), the period between calving and confirmed conception (days open), and pregnancy rate. This thesis focuses on female fertility. Female fertility is made up of different underlying traits (Figure 1.1). The traits used in genetic evaluation of female fertility can be categorized into two aspects of fertility. The first reflects the ability of the cow to return to cyclicity after calving, e.g., the interval from calving to first service. The second reflects the ability of the cow to conceive following insemination, become pregnant, and maintain pregnancy, e.g., the time between the first and last insemination, and non-return rate after service. Therefore, because female fertility can be measured by different indicators, it is important to identify the optimal measures to facilitate genetic improvement.

The classical fertility traits derived from insemination and calving data which are used in genetic evaluation of fertility have the disadvantage that they have a low heritability (Jansen, 1985; Berry et al., 2003), which makes genetic improvement difficult. Heritability estimates for classical fertility traits generally range from 0.01 to 0.10 (Hou et al., 2009; Sun et al., 2009; Berry et al., 2012). The low heritability may be explained by the fact that classical fertility traits are highly influenced by farm management decisions and poor recording practices (Hayes et al., 1992; Campos et al., 1994; Marti and Funk, 1994). For example, a planned extended CI will delay CFS, not because a cow has a late start of cyclic activity, but because of the farmer’s decision on when to inseminate. Such management practices are justified, but results in large residual variance and low heritability estimates. The low heritability sparks the interest for alternative approaches that might yield more accurate information for genetic evaluations of dairy cows’ fertility. For instance, fertility measures that more directly reflect the cows’ own reproductive physiology like endocrine fertility traits that are derived from progesterone concentration levels in milk, or the use of DNA information.

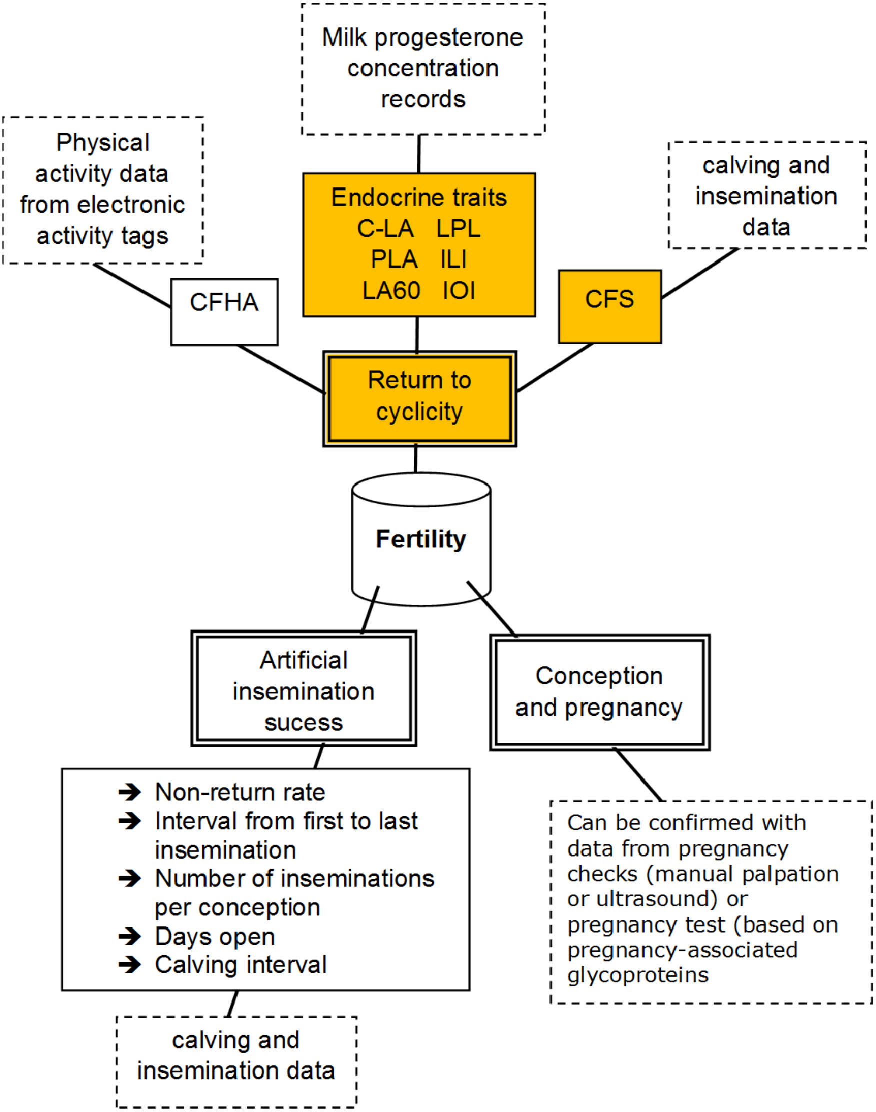


Figure 1.1 Schematic chart of different components of female fertility in dairy cattle (in double boxes), most widely used underlying phenotype measures associated with each component (in solid boxes), and source of data used to derive the phenotypes (in dashed boxes). This thesis investigated the traits in the orange boxes. C-LA = Interval from calving to commencement of luteal activity; PLA = Proportion of samples in luteal activity; LA60 = Occurrence of luteal activity during first 60 days in milk; LPL = Luteal phase length; ILI = Length of inter-luteal interval (ILI); IOI = Length of inter-ovulatory interval; CFHA = interval from calving to first high activity; CFS = interval from calving to first service

1.3 Progesterone defined fertility traits

Understanding the estrous cycle is important for managing and improving reproductive performance in dairy cattle. The estrous cycle is divided into two phases which are characterized by changes on the ovary (Figure 1.2). There is the follicular phase during which the pre-ovulatory follicle on the ovary, which contains the oocyte (or egg), produces estrogen. When estrogen concentrations are high enough, a surge of luteinizing hormone is released, initiating ovulation. The luteal phase begins after ovulation, when the follicle transforms into a corpus luteum which produces progesterone to maintain pregnancy (Senger, 2003). Early resumption of ovarian cyclicity postpartum facilitates a greater number of estrus cycles before insemination which, on average, increases the likelihood of subsequent conception (Darwash et al., 1997). Studies have shown that there is a high correlation between progesterone (P4) concentration in blood and milk (e.g., r = 0.88; Dobson and Fitzpatrick, 1975), hence, P4 analysis of milk samples can be used to study postpartum ovarian activity in dairy cows. A period of low P4 levels usually occurs after calving, when a cow exhibits anestrous (Lamming and Bulman, 1976). This period is followed by an increase in P4 levels, which is indicative of the first postpartum ovulation. The cavity of the ovulated follicle is gradually filled with progesterone-secreting luteal cells, which forms the corpus luteum. From about the fifth day after ovulation, the corpus luteum dominates the estrous cycle during the luteal phase with high P4 levels for about 14 days. After that, unless the cow becomes pregnant, the corpus luteum degenerates (luteolysis), and a new ovulation can occur. If the cow becomes pregnant, the corpus luteum is maintained during the pregnancy (Ball and Peters, 2004).

Progesterone and overall fertility of the dairy cow have been shown to be connected in several ways. Low probability of embryo survival was shown to be associated with both low and excessive P4 levels 5 to 7 days after insemination, which indicates that an optimum in P4 level is required for embryo survival after insemination (Stronge et al., 2005). In a study, repeat breeding heifers (i.e., heifers that fail to conceive from 3 or more regularly spaced services in the absence of detectable abnormalities) tended to have higher basal P4 concentrations at estrus, a late P4 rise in early luteal phase and low luteal concentrations of P4 compared to virgin heifers (Båge et al., 2002). In addition, increased basal P4 levels at insemination have been associated with increased probability of repeat-breeding for cows and higher return rate at insemination (Waldmann et al., 2001). The ovarian activity of a cow after calving also affects overall fertility. For example, early onset of estrous cyclicity after calving has been shown to increase probability of an early insemination after calving, shorten the interval from calving to conception, increase conception rate and reduce the number of services per conception (Darwash et al., 1997).

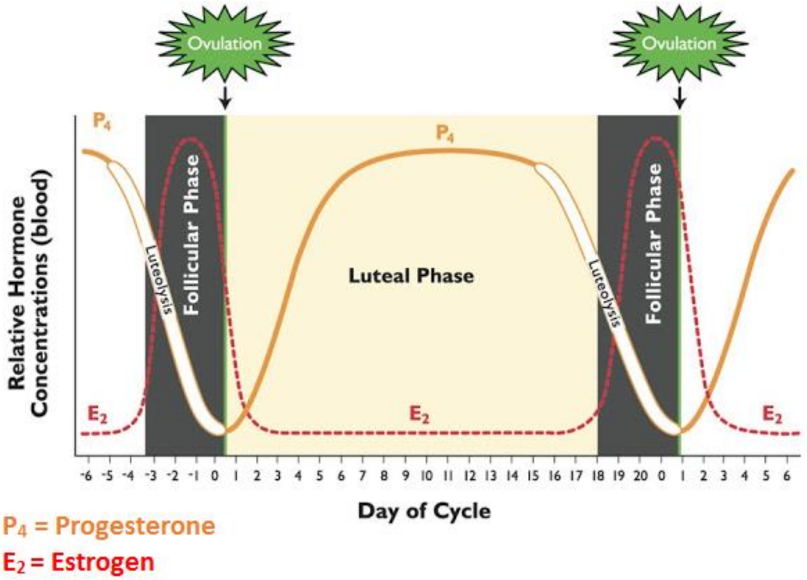


Figure 1.2 Stages within the follicular and luteal phases in association with respective changes in circulating concentrations of estrogen and progesterone during the bovine estrous cycle. Adapted from P.L. Senger Pathways to pregnancy and parturition Current Conceptions, Inc., Pullman, WA

1.4 Genomic regions associated to endocrine fertility traits

Several genome-wide association studies have attempted to locate genomic regions associated with reproductive performance in dairy cattle, see the cattle QTL (quantitative trait loci) database for an overview (Hu et al., 2016). To date, 81,653 QTL have been catalogued in cattle, for 521 traits of which 8,969 QTL are for fertility traits (http://www.animalgenome.org; 2016). For CFS, QTL have been reported on chromosomes 4, 7, 9, 13, 20, 23, and 25 (Druet et al., 2008; Sahana et al., 2010; Schulman et al., 2011; Höglund et al., 2014, 2015). Most QTL studies for fertility have been based on classical fertility traits, with few studies performed with endocrine fertility traits. In a genome-wide association study with 50,000 single nucleotide polymorphisms (SNP), Berry et al., (2012)reported QTL regions associated to C-LA on chromosomes 2 and 21 and these regions have been associated with reproductive performance in other studies (Huang et al., 2010; Sahana et al., 2010; Schulman et al., 2011). Also, most of the QTL studies used low density (e.g., 50,000) SNP panels which do not allow for precision mapping of quantitative trait loci. The use of genome sequence data for association studies has been enabled by advances in next generation sequencing techniques which have led to sequencing numerous animals in cattle, e.g., projects like the 1000 Bull Genome Consortium (Daetwyler et al., 2014). It is expected that associations with endocrine fertility traits can be targeted to smaller chromosomal regions with sequence compared to smaller genetic marker panels like 50,000 SNP markers.

1.5 Genomic selection for fertility

Genomic selection is a breeding tool that uses SNP markers spread across the genome to predict genomic breeding values for individuals (Meuwissen et al., 2001; Hayes et al., 2009). Genomic selection has become the standard for dairy cattle breeding in most countries, because it increases the potential genetic gain by up to 80% due to reduced generation interval (Schaeffer, 2006). Two steps are required to perform genomic selection. The first step entails genotyping and phenotyping a set of animals to form the training population. The phenotypes and genotypes of the training population are then used to obtain prediction equations. In the second step, the prediction equations are used with genotypes of evaluation animals (e.g., selection candidates), to predict their genomic breeding values. The selection candidates usually are juveniles, and their genomic breeding values are more accurate than when estimated with traditional breeding tools that rely on parent average information (Meuwissen et al., 2001). The accuracy of genomic selection depends several factors including the number of animals in the training population, the heritability of the trait, and the genetic architecture of the trait, in particular the number of loci affecting the trait and distribution of their effects (Daetwyler et al., 2008; Meuwissen, 2009). An important feature of genomic selection which makes it interesting to apply to expensive or difficult to measure traits like endocrine fertility traits is that the traits do not have to be recorded on a routine scale and on a large number of daughters for each selection candidate. With a training population of limited size, genomic selection was shown as a promising tool for starting selection for scarcely recorded or difficult to measure traits (Calus et al., 2013).

1.6 This thesis

The main objective of this thesis was to explore the use of milk P4 records and genomic information to improve selection for fertility in dairy cows. This main objective was divided into two sub-objectives. The first sub-objective investigated the use of in-line milk P4 records to define endocrine fertility traits and identify genomic regions associated to the defined traits, using SNP markers and whole genome sequence variants. The second sub-objective aimed to determine the optimal use of endocrine fertility traits in genomic evaluations by investigating the impact of different phenotyping strategies on the accuracy of predicting genomic breeding values for fertility.

In **chapter 2** of this thesis, the use of in-line recorded milk progesterone levels to define endocrine fertility traits was investigated, genetic parameters were estimated and genetic correlations of the endocrine traits with classical fertility and milk production traits were estimated. In **chapter 3**, genomic regions associated to endocrine fertility traits were identified by performing a genome-wide association study with 85,485 SNP, and targeted QTL regions were fine-mapped using imputed sequence variants. **Chapter 4** investigated the added value in terms of accuracy, of using endocrine traits along with classical fertility traits in genomic prediction of fertility, by performing single-trait and multi-trait genomic predictions. The aims of **chapter 5** were: a) to investigate the potential accuracy of different scenarios when combining a cow training population measured for an endocrine trait with a training population of bulls with daughter observations for a classical trait for multi-trait genomic prediction of fertility, by using a deterministic prediction equation within and between populations, and b) to investigate recording strategies that optimally use the Herd Navigator for genomic prediction, in terms of, the number of farms, and recording period for endocrine fertility traits.

References

Båge, R., H. Gustafsson, B. Larsson, M. Forsberg, and H. Rodrıguez-Martınez. 2002. Repeat breeding in dairy heifers: follicular dynamics and estrous cycle characteristics in relation to sexual hormone patterns. *Theriogenology*. 57:2257–2269.

Ball, P.J.H., and A.R. Peters. 2004. Reproduction in cattle. Wiley Online Library.

Berry, D.P., J.W.M. Bastiaansen, R.F. Veerkamp, S. Wijga, E. Wall, B. Berglund, and

M.P.L. Calus. 2012. Genome-wide associations for fertility traits in Holstein-Friesian dairy cows using data from experimental research herds in four European countries. *Animal*. 6:1206–15.

Berry, D.P., F. Buckley, P. Dillon, R.D. Evans, M. Rath, and R.F. Veerkamp. 2003. Genetic relationships among body condition score, body weight, milk yield, and fertility in dairy cows. *J. Dairy Sci.* 86:2193–2204.

Boichard, D. 1990. Estimation of the economic value of conception rate in dairy cattle. *Livest. Prod. Sci.* 24:187–204.

Calus, M.P.L., Y. De Haas, M. Pszczola, and R.F. Veerkamp. 2013. Predicted accuracy of and response to genomic selection for new traits in dairy cattle. *Animal*. 7:183–191.

Campos, M.S., C.J. Wilcox, C.M. Becerril, and A. Diz. 1994. Genetic Parameters for Yield and Reproductive Traits of Holstein and Jersey Cattle in Florida. *J. Dairy Sci.* 77:867–873.

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

Daetwyler, H.D., B. Villanueva, and J.A. Woolliams. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS One*. 3:e3395.

Darwash, A.O., G.E. Lamming, and J.A. Woolliams. 1997. The phenotypic association between the interval to post-partum ovulation and traditional measures of fertility in dairy cattle. *Anim. Sci.* 65:9–16.

Dekkers, J.C.M. 1991. Estimation of economic values for dairy cattle breeding goals: bias due to sub-optimal management policies. *Livest. Prod. Sci.* 29:131–149.

Dobson, H., and R.J. Fitzpatrick. 1975. Clinical application of the progesterone-in-milk test. *Br. Vet. J.* 132:538–542.

Druet, T., S. Fritz, M. Boussaha, S. Ben-Jemaa, F. Guillaume, D. Derbala, D. Zelenika, D. Lechner, C. Charon, D. Boichard, I.G. Gut, A. Eggen, and M. Gautier. 2008. Fine Mapping of Quantitative Trait Loci Affecting Female Fertility in Dairy Cattle on BTA03 Using a Dense Single-Nucleotide Polymorphism Map. *Genetics*. 178:2227–2235.

González-Recio, O., M.A. Pérez-Cabal, and R. Alenda. 2004. Economic value of female fertility and its relationship with profit in Spanish dairy cattle. *J. Dairy Sci.* 87:3053–3061.

Groenendaal, H., D.T. Galligan, and H.A. Mulder. 2004. An Economic Spreadsheet Model to Determine Optimal Breeding and Replacement Decisions for Dairy Cattle. *J. Dairy Sci.* 87:2146–2157.

Hayes, B.J., H.D. Daetwyler, P.J. Bowman, G. Moser, B. Tier, R. Crump, M. Khatkar, H.W. Raadsma, and M.E. Goddard. 2009. Accuracy of genomic selection:

comparing theory and results. *In* Proc Assoc Advmt Anim Breed Genet. 34–37.

Hayes, J.F., R.I. Cue, and H.G. Monardes. 1992. Estimates of Repeatability of Reproductive Measures in Canadian Holsteins. *J. Dairy Sci.* 75:1701–1706.

Höglund, J.K., B. Guldbrandtsen, M.S. Lund, and G. Sahana. 2015. Identification of genomic regions associated with female fertility in Danish Jersey using whole genome sequence data. *BMC Genet.* 16:60.

Höglund, J.K., G. Sahana, R. Brøndum, B. Guldbrandtsen, B. Buitenhuis, and M.S. Lund. 2014. Fine mapping QTL for female fertility on BTA04 and BTA13 in dairy cattle using HD SNP and sequence data. *BMC Genomics*. 15:790.

Hou, Y., P. Madsen, R. Labouriau, Y. Zhang, M.S. Lund, and G. Su. 2009. Genetic analysis of days from calving to first insemination and days open in Danish Holsteins using different models and censoring scenarios. *J. Dairy Sci.* 92:1229–1239.

Hu, Z.-L., C.A. Park, and J.M. Reecy. 2016. Developmental progress and current status of the Animal QTLdb. *Nucleic Acids Res.* 44:D827–D833.

Huang, W., B.W. Kirkpatrick, G.J.M. Rosa, and H. Khatib. 2010. A genome-wide association study using selective DNA pooling identifies candidate markers for fertility in Holstein cattle. *Anim. Genet.* 41:570–578.

Inchaisri, C., R. Jorritsma, P. Vos, G.C. Van der Weijden, and H. Hogeveen. 2010. Economic consequences of reproductive performance in dairy cattle. *Theriogenology*. 74:835–846.

Jansen, J. 1985. Genetic aspects of fertility in dairy cattle based on analysis of A.I. data — A review with emphasis on areas for further research. *Livest. Prod. Sci.* 12:1–12.

Lamming, G.E., and D.C. Bulman. 1976. The use of milk progesterone radioimmunoassay in the diagnosis and treatment of subfertility in dairy cows. *Br. Vet. J.*

Marti, C.F., and D.A. Funk. 1994. Relationship Between Production and Days Open at Different Levels of Herd Production. *J. Dairy Sci.* 77:1682–1690.

Meadows, C., P.J. Rajala-Schultz, and G.S. Frazer. 2005. A Spreadsheet-Based Model Demonstrating the Nonuniform Economic Effects of Varying Reproductive Performance in Ohio Dairy Herds. *J. Dairy Sci.* 88:1244–1254.

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

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.

Plaizier, J.C.B., G.J. King, J.C.M. Dekkers, and K. Lissemore. 1997. Estimation of Economic Values of Indices for Reproductive Performance in Dairy Herds Using Computer Simulation. *J. Dairy Sci.* 80:2775–2783.

Sahana, G., B. Guldbrandtsen, C. Bendixen, and M.S. Lund. 2010. Genome-wide association mapping for female fertility traits in Danish and Swedish Holstein cattle. *Anim. Genet.* 41:579–588.

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

Schulman, N.F., G. Sahana, T. Iso-Touru, S.D. McKay, R.D. Schnabel, M.S. Lund, J.F. Taylor, J. Virta, and J.H. Vilkki. 2011. Mapping of fertility traits in Finnish Ayrshire by genome-wide association analysis. *Anim. Genet.* 42:263–269.

Senger, P.L. 2003. Placentation, the endocrinology of gestation and parturition. *Pathways to pregnancy parturition*. 2:304–325.

Stronge, A.J.H., J.M. Sreenan, M.G. Diskin, J.F. Mee, D.A. Kenny, and D.G. Morris. 2005. Post-insemination milk progesterone concentration and embryo survival in dairy cows. *Theriogenology*. 64:1212–1224.

Sun, C., P. Madsen, U.S. Nielsen, Y. Zhang, M.S. Lund, and G. Su. 2009. Comparison between a sire model and an animal model for genetic evaluation of fertility traits in Danish Holstein population. *J. Dairy Sci.* 92:4063–4071.

de Vries, A., and B.J. Conlin. 2003. Economic Value of Timely Determination of Unexpected Decreases in Detection of Estrus Using Control Charts1. *J. Dairy Sci.* 86:3516–3526.

Waldmann, A., O. Reksen, K. Landsverk, E. Kommisrud, E. Dahl, A.O. Refsdal, and E. Ropstad. 2001. Progesterone concentrations in milk fat at first insemination—effects on non-return and repeat-breeding. *Anim. Reprod. Sci.* 65:33–41.

6

General discussion

6.1 Introduction

The research in this thesis on exploring the use of milk progesterone (P4) measures to improve selection for fertility in dairy cows can be divided in to two main areas. The first focused on the genetic aspects of fertility as measured using P4, which have been dealt with in chapters 2 and 3. Chapter 2 investigated the use of in-line measured milk P4 concentrations to define endocrine fertility traits and estimate genetic parameters. In chapter 3, a study on identifying genomic regions associated to endocrine fertility traits, using single nucleotide polymorphisms (SNP) panels as well as and imputed whole genome sequence variants was performed. The second part focused on the use of endocrine fertility traits in genomic prediction, to improve genomic selection for fertility. On the one hand, classical fertility traits derived from calving and insemination data are available for numerous cows for national genetic evaluations. On the other hand, the availability of endocrine fertility traits presents the opportunity to improve accuracy of predicting genomic breeding values of cows' fertility by combining endocrine and classical fertility traits in multi-trait genomic predictions (chapter 4). Also, an important question is how many cows with endocrine fertility records in the training population will be required to improve the accuracy of bulls' genomic breeding values (chapter 5).

In each of the chapters, the main results have already been discussed. The general discussion will focus on aspects related to breeding for fertility in dairy cows, optimizing accuracy of genomic prediction for endocrine fertility traits, practical integration of endocrine fertility traits (and novel traits in general) in routine genetic evaluations, and prospects for future research will be given.

6.2 Breeding for fertility

This thesis specifically looked at estrous related traits that reflect the ability of a cow to return to cyclicity after calving (see Figure 1, chapter 1), and how they can be used in genetic evaluations to improve selection for fertility. Different traits derived from different sources of information have been proposed as measures of the ability of a cow to resume cyclicity postpartum. In this thesis, we investigated endocrine fertility traits derived from milk P4 concentrations and classical fertility traits derived from insemination and calving data. Endocrine fertility traits are expected to be more objective measures that are closer to the physiology underlying fertility than the classical traits. For genetic evaluations, the most commonly used traits are classical fertility traits like calving to first service (CFS) and calving interval (CI). Although the use of classical fertility traits for breeding has been facilitated by their ease of recording on a large scale and at feasible cost, the drawbacks are that classical traits are highly influenced by farm management decisions, are available late in life (e.g., CI), and have low heritability. Reported heritability estimates for CFS range from 0.05 to 0.10 (Roxström et al., 2001; Andersen-Ranberg et al., 2005; Sun et al., 2009; Berry et al., 2012), though estimates for CFS in this thesis were in the upper bound. The estimated heritability of CI in this thesis was low (0.03 to 0.05), and in agreement with previous studies (Wall et al., 2003; Berry et al., 2012). The low heritability of classical fertility traits makes selection for fertility less effective. Studies of physical activity derived traits e.g., the interval from calving to the first sign of high activity (CFHA) have reported higher heritability, ranging from 0.12 to 0.18 (Løvendahl and Chagunda, 2009; Ismael et al., 2015). Though CFHA has been suggested as a measure for resumption of ovarian activity, Ismael et al., (2016) showed that a drawback is that activity monitor derived fertility traits are seasonally sensitive.

The advantage of endocrine fertility traits is that they are less biased by farm management decisions and have higher heritability. In this thesis, several endocrine fertility traits were derived, but not all the traits were heritable. The traits with highest heritability were commencement of luteal activity (C-LA), and proportion of samples in luteal activity (PLA). The heritability of C-LA ranged from 0.12 to 0.14, whereas that of PLA ranged from 0.12 to 0.15. Reported heritability estimates for C-LA or transformed C-LA range from 0.16 to 0.30 (Veerkamp et al., 2000; Royal et al., 2002; Petersson et al., 2007), whereas Petersson et al., (2007) reported a heritability estimate of 0.30 for PLA. Thus, heritability estimates of C-LA and PLA in this thesis were in the lower bound of those reported in literature. Another advantage of endocrine fertility traits is that they more directly reflect a cow's reproductive physiology than classical traits. For example, in some breeding programs, CFS is used in a fertility index as an indirect indicator of calving to first ovulation, but C-LA is a more objective indicator of first ovulation because it more directly reflects a cows physiology (Bulman and Lamming, 1978; Lamming and Darwash, 1998; Darwash et al., 1999). In addition, for animals with delayed insemination, CFS records will be available much later, whereas C-LA is independent of farm management decisions. Thus, C-LA is a better trait to use in genetic evaluations.

Whether a trait is suitable for use in breeding depends not only on that it has substantial genetic variation, but also its genetic correlation with other traits. As discussed in 6.2.2, results in chapter 2 demonstrated that the genetic correlation of endocrine fertility traits with milk production traits were considerably lower than those observed between milk yield and classical fertility traits. The genetic correlation of C-LA was 0.18 with milk yield, and there was a favorable correlation of C-LA with fat yield (- 0.12) and with protein yield (- 0.04). This is another advantage of using endocrine traits compared with the classical traits. In this thesis, the genetic correlation between C-LA and CFS was positive (0.37 to 0.58), which indicates that selection for shorter C-LA will also reduce interval from CFS. Royal et al. (2003) reported a genetic correlation of 0.53 between CFS and C-LA. Furthermore, they found genetic correlations between a long C-LA and a long CI (0.39). In this thesis, the genetic correlation between C-LA and CI varied from 0.26 to 0.31. Ismael et al., (2015) also found a positive correlation between CFHA and CFS (0.96). This indicates that selection for a shorter interval from calving to resumption of ovarian activity would benefit overall fertility.

6.2.1 New tools and new breeding goals to improve selection for fertility in dairy cows

New breeding tools like genomic selection and in-line recording technologies like the Herd Navigator provide an opportunity to implement endocrine fertility traits in genetic evaluations to improve fertility in dairy cows. However, to benefit from the added value of endocrine traits in genetic evaluations, breeding goals for fertility will need to be redefined. For a long time, the use of endocrine fertility traits in genetic evaluations has been hindered by the fact that it is laborious and costly to measure sufficient milk P4 records per cow on a large scale. Also, all research on the use of endocrine fertility traits has focused on application in breeding schemes that generally target a classical breeding goal (e.g., cows with shorter calving intervals and shorter intervals from calving to first service). However, now that measuring sufficient P4 records on a large scale has been facilitated by in-line technologies, we now realize that endocrine traits will be useful for breeding only if they are important for the breeding goal. As shown in chapter 5, if the question is “can we improve selection for the biological breeding goal for fertility?”, that is, where the breeding goal trait is an endocrine trait like C-LA, then it is valuable to include endocrine fertility traits in genetic evaluations. However, if the question remains to improve the classical breeding goal for fertility (e.g., CI or CFS), there is little to no benefit of including endocrine traits in addition to the cheap classical traits in genetic evaluations. This means that if breeding programs want to make valuable use of endocrine fertility traits to improve selection for fertility, they will need to defining new breeding goals for fertility that include the endocrine traits.

A disadvantage of using the endocrine traits in progeny testing schemes is that it is unrealistic, as progeny testing schemes usually require a large number (100 - 150) of daughters with phenotypes of each selection candidate bull, and these daughters are often milked in many different herds. Another drawback of progeny testing is that the long generation interval between phenotyping and proving bulls hampers genetic gain. Genomic selection has become the standard in cattle breeding because it increases the potential genetic gain by up to 80% due to reduced generation interval (Schaeffer, 2006). Genomic selection has the added benefit that it also disconnects the phenotype recording in a training population from the selection of the candidates. This new feature of genomic selection provides a great opportunity for novel traits such as endocrine traits, because a trait can be selected as soon as several thousand of animals are recorded for the trait (Misztal, 2011). For endocrine fertility traits, deterministic predictions with the equation in chapter 5 show that for within population genomic prediction of C-LA with heritability of 0.13, and considering 1,566 for the effective number of chromosomes (*Me*), 15,000 animals in a cow training population with own C-LA records are needed to achieve 0.74 accuracy; to achieve 0.84 accuracy of prediction, 30,000 animals are required. Although these numbers are still high, it is easier to achieve than collecting records on progeny of every test bull. Also, across country collaboration can help. This is especially beneficial when one country has very limited number of animals in the training population. However, as shown in chapter 5, the accuracy of prediction when combining training populations across countries is largely influenced by the genetic correlation between the countries for the target trait. For example, in chapter 5, when the genetic correlation between Sweden and the Netherlands was high, fewer animals from the Netherlands were needed to obtain the same accuracy of genomic breeding values in Sweden than with lower genetic correlations. That is, with 200 animals from Sweden and a genetic correlation of 0.5 between countries, 15,000 animals from the Netherlands were needed to achieve 0.31 accuracy of predicting C-LA, but when the correlation was 0.9, only 3,000 animals from the Netherlands were needed to achieve the same accuracy.

Furthermore, recording strategies for endocrine fertility traits can be optimized to improve accuracy of genomic prediction. For example, in chapter 5, it was shown that it is more important to have more animals with phenotypes than more lactations per animal in the training population. Also, to maximize accuracy of genomic prediction, it is important to optimize the design of the training population. An optimal design of the training population should maximize the relationships between evaluated animals and animals in the training population, and minimize relationships between training animals (Habier et al., 2007; Meuwissen, 2009; Pszczola et al., 2012). Because animals in the training population do not have to be progeny or relatives of evaluated bulls, this conveniently allows for optimizing recording strategies for endocrine fertility traits as shown in chapter 5, while at the same time maintaining an optimal design of the cow training population. Optimizing the training population can be done by choosing animals for phenotyping based on their genomic relationship with evaluated animals.

Therefore, genomic selection and in-line recording technologies provide a great opportunity to reconsider the use of endocrine fertility traits to further improve selection for fertility. However, investing in novel traits for a cow training population should only be done when the trait is in the breeding goal.

6.2.2 Association of fertility traits with milk yield and genetic progress

There is accumulating evidence of the antagonistic effect of increased milk yield on fertility in dairy cattle, for a review see Lucy, (2001), Veerkamp and Beerda, (2007), and Walsh et al., (2011). For several decades, breeding objectives in dairy cattle focused mainly on increasing milk yield and its components, with the exception of the Nordic countries that included fertility since the 1970s. The unfavorable genetic correlation between milk yield and fertility traits meant that fertility decreased with increasing milk yield, which led to a long-term decline in fertility. In recent years, selection indices worldwide have changed to a more balanced breeding approach that includes longevity, udder health, and fertility (Miglior et al., 2005, 2012). This has in turn reversed the undesirable genetic trend for fertility, while genetic progress for production has continued at about the same rate as before. For example, Figure 6.1 shows the genetic trend in Holstein Friesian cows from the Netherlands, from 1990 to 2014. In the Netherlands, fertility traits were included in the total merit index from 2006. As can be seen in Figure 1, the genetic trend has reversed after the inclusion of fertility in the breeding goal. This indicates that it is possible to improve fertility genetically, without ceasing selection for milk yield, as milk production traits still have a major impact on the revenue of each farm, and are still worth improving.

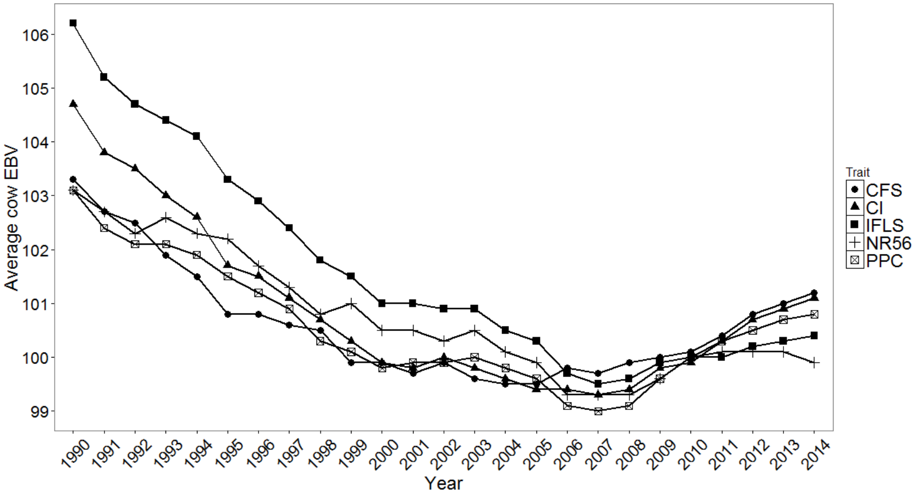


Figure 6.1 Genetic trend of fertility for Holstein Friesian cows from commercial farms in the Netherlands. CFS = calving to first service, CI = calving interval, IFLS = interval from first to last service, NR56 = non-return rate within 56 days, PPC = percentage of pregnant cows, EBV = estimated breeding value. Data from the Dutch herd book (https://crvnl-be6.kxcdn.com/wp-content/uploads/2016/09/gen_trend_koe_nl_20160816.pdf).

The traits currently used for fertility in the total merit index of the Netherlands are classical fertility traits derived from insemination and calving data. The current rate of genetic progress for fertility, as seen in Figure 6.1, is influenced by the unfavorable genetic correlation between milk production and classical fertility traits. In this thesis, the genetic correlation varied from 0.07 to 0.24 between milk production traits and CFS, and from 0.56 to 0.69 between milk production traits and CI, while phenotypic correlations of milk production traits with CFS and CI varied from 0.13 to 0.21. However, in chapter 2, the genetic correlations of endocrine fertility traits with milk production traits were considerably lower than those observed between milk yield and classical fertility traits. The genetic correlation of C-LA was 0.18 with milk yield, and there was a favorable correlation of C-LA with fat yield (- 0.12) and with protein yield (- 0.04). In addition, the phenotypic correlations of the endocrine fertility traits with milk production traits were close to zero (0.01 to 0.07). This indicates that novel fertility traits that are not (or less) unfavorably correlated with milk yield can improve genetic progress for fertility further. Higher genetic correlation estimates of C-LA with milk yield have been reported, but all unfavorable. Veerkamp et al., (2000) reported an unfavorable genetic correlation of 0.51 between C-LA and milk yield, and Nyman et al., 2014 found a correlation of 0.45. Although further research is needed to confirm the low genetic correlation between in-line endocrine fertility traits and milk production traits, this thesis shows that the genetic progress of fertility can be improved by considering C-LA as a breeding goal trait in genetic evaluations, or index trait in a fertility index.

Another approach of improving the genetic progress for fertility is to put a higher economic weight (breeding goal weight) on fertility to get faster genetic improvement, compared with the other traits in the breeding. Miglior et al 2005, analyzed national selection indices for 15 major dairy countries and reported that average relative emphasis was 59.5% for production, 28% for durability, and 12.5% for health and reproduction across all countries. In another study, Miglior et al., (2012) reported interesting changes in relative emphasis on production traits in selection indices of four countries taken as examples. This suggest that perhaps we could put more emphasis on fertility traits to improve genetic progress. However, such a scenario would reduce genetic progress in milk yield due to its unfavorable correlation with fertility, but again this higher emphasis on fertility could be put on endocrine fertility traits like C-LA, as it is less unfavorably correlated with milk production traits.

6.3 Optimal use of genomic information in selection for fertility

The main interest in genome-wide association studies (GWAS) in dairy cattle breeding is to find markers to improve the accuracy of predicting breeding values and to increase the understanding of the genetic control of economically important traits. Identifying the genes that affect traits such as fertility will also provide insight into the likely effect of selection on these mutations on other traits in the breeding goal. With the development of methods that allow to perform genomic prediction based on a large number of markers (Meuwissen et al., 2001), and the availability of commercial SNP chips, genomic selection has become the standard tool for animal and plant breeders. The accuracy of genomic prediction is influenced by the number of phenotypes in the training population used to derive the prediction equation, the heritability of the trait, the effective population size, the size of the genome, the density of markers, and the genetic architecture of the trait, in particular the number of loci affecting the trait and distribution of their effects (Daetwyler et al., 2008; Meuwissen, 2009).

In this thesis, I investigated the heritability of fertility by defining different endocrine fertility traits in chapter 2. In chapter 3, the genetic architecture of fertility was investigated by identifying quantitative trait loci (QTL) or genomic regions genes associated to fertility traits using SNP data. Using sequence variants, target QTL regions were also fine-mapped in chapter 3, and several candidate genes were identified. Chapter 4 evaluated empirical accuracy of predicting fertility using endocrine and classical fertility traits in multi-trait genomic prediction models, and in chapter 5 I evaluated the impact of the size of a cow training population with endocrine traits on accuracy of genomic prediction. In this section, I discuss the impact of the investigated factors on the accuracy of genomic prediction for fertility, and how to optimally use genomic information to improve accuracy of genomic prediction for endocrine traits.

The properties of QTL that control a trait like allele frequency spectra of QTL, and distribution of QTL effects are key factors that determine the accuracy of genomic predictions (Wientjes et al., 2015a). When the minor allele frequency (MAF) of QTL is on average lower than that of SNP marker, the accuracy of genomic prediction is reduced (Daetwyler et al., 2013; Wientjes et al., 2015a). This is because, a lower MAF of QTL than for SNPs, results in decreased strength of linkage disequilibrium (LD) between QTL and SNP markers (Khatkar et al., 2008; Yan et al., 2009; Wientjes et al., 2015b), hence reducing the proportion of the genetic variance captured by the SNP markers. The MAF of QTL underlying complex traits is expected to be lower than the MAF of SNP markers (Goddard, 2009; Yang et al., 2010; Kemper and Goddard, 2012), therefore, it is highly likely that not all the genetic variance can be captured by SNP marker panels in real data. This also indicates that, there is a probability of underestimating the heritability of complex traits using SNP markers.

In chapter 3, a genome-wide association study (GWAS) for endocrine and classical fertility traits was performed with 85,485 SNPs. The GWAS identified 17 QTL regions for 6 endocrine fertility traits and 3 regions for one classical trait. The average proportion of variation explained by SNPs in the QTL regions identified for each trait are in Figure 6.2. The proportion of variation was calculated as in Pryce et al., (2010). That is, the proportion of variation for any SNP can be calculated simply as, F/a, where F is the F statistic for that SNP based on number of animals (a) used in analysis. In the absence of any real effect, F statistics have an expected value of 1, so an unbiased estimate of the proportion of variation can be made by $(\mathrm{F}\text{-}1)/\mathrm{a}$. This estimate of the proportion of variation (R^2) was averaged over all N SNP that were statistically significant in each identified QTL region for each trait by calculating $\mathrm{R}^2 = \left(\frac{\sum_{i=1}^{N}\mathrm{F}}{\mathrm{N}}\text{-}1\right) \times 100\% \times \frac{1}{\mathrm{a}}$. The proportion of genetic variance explained by the QTL was 0.70% for CLA, 0.72% for PLA, 0.72% for LA60, 0.71% for LPL, 0.60% for IOI, 0.77% for CLAFS, and 0.74% for CFS. In general, the variance explained by the QTL was very minimal, and there were no QTL with large effects. Also, the average MAF of the QTL detected was 0.27. This suggest that the remaining genetic variance of endocrine fertility traits is influenced by QTL with low MAF that may not have been detected in the GWAS due to incomplete LD between the SNP markers and underlying QTL. Consequently, genomic prediction for endocrine fertility traits will be more beneficial with genomic best linear unbiased (GBLUP) models that assume that all SNP effects are drawn from the same distribution and explain equal amount of genetic variance, than Bayesian variable selection models that accommodate for SNPs explaining a larger part of the genetic variance compared to other SNPs. For instance, in chapter 4, empirical accuracies of genomic prediction for endocrine fertility traits were estimated with GBLUP, and theoretical accuracies were estimated with the prediction equation of (Daetwyler et al., 2008, 2010) which was derived with the assumptions of a GBLUP model. The theoretical accuracies were in line with the empirical estimates. Therefore, based on the genetic architecture of endocrine fertility traits, I do not expect an improvement in accuracy of genomic prediction from approaches such as Bayesian variable selection.

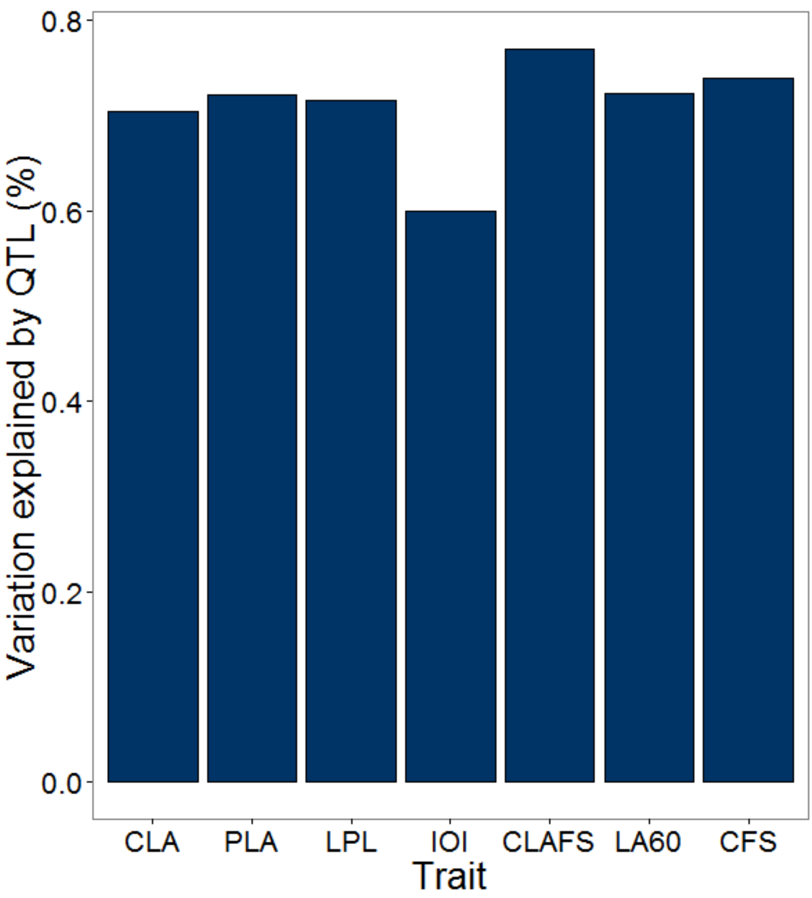


Figure 6.2 Average genetic variation explained by SNP markers in QTL regions identified in a genome-wide association study using 85,485 SNPs and 2,447 cows for endocrine and classical fertility traits in Holstein. CLA = interval from calving to commencement of luteal activity; PLA = proportion of samples in luteal activity during the first 60 days in milk; LPL = length of first luteal phase; IOI = length of first inter-ovulatory interval; CLAFS = interval from commencement of luteal activity to first service; LA60 = occurrence of luteal activity during the first 60 days in milk; CFS = interval from calving to first service

One option to increase the likelihood of identifying causal mutations is to increase marker density and (or) number of genotyped and phenotyped animals. Developments in genotyping technology have resulted in a reduction of costs, enabling the production of commercial high-density (HD) SNP chips (e.g., Illumina Bovine HD 770k SNP chip). With more animals genotyped, which increases the sample size, and with the genome more densely covered with markers, which leads to a smaller distance between the SNP and the causative mutation, a more precise detection of QTL is expected. The ultimate level of genotypic information though is sequence data. Sequencing tries to determine the order of all nucleotides of the DNA of a given organism. Therefore, sequence data should contain the causative mutations of a trait. Hence a GWAS using sequence data is expected to find the causative mutation (Meuwissen and Goddard, 2010). There have been efforts to increase the number of sequenced animals (Daetwyler et al., 2014), but sequencing a large number of animals for GWAS is still expensive. Therefore, imputation from lower density genotypes to whole genome sequence using a sequenced training population offers a good alternative. The approach that has been taken in chapter 3 was to perform a GWAS with a low density (e.g., here 85k) SNP chip panel, and then focus on the identified peaks, performing a region-wise association study (RWAS) using imputed sequence data (e.g., Höglund et al., 2014; Sahana et al., 2014; Wu et al., 2015). In chapter 3, significant QTL regions from the GWAS with 85k SNP were fine-mapped for endocrine fertility traits using imputed sequence variants. The RWAS was able to refine the QTL regions from the GWAS, but it was not possible to identify the causative mutation, mainly because of long-range LD that exist in cattle due to low effective population size and strong selection. Similar observations have been reported in previous studies (Höglund et al., 2014; Wu et al., 2015). Another factor that might be hampering identification of the causative mutation is that imputation is not 100% accurate, especially for rare variants and small training populations. The RWAS was able to identify several candidate genes associated to endocrine fertility traits that can help to learn more about the genetic architecture and underlying biology of fertility. In addition, variants significantly associated to endocrine fertility traits detected in the RWAS can be integrated in genomic selection to improve accuracy of prediction. As these significant regions on the genome are continually being found and described, it is of interest to integrate the significant markers in genomic evaluations. This integration is relevant because while the causative mutations are not detected, these significant markers provide knowledge regarding the genetic architecture of the trait. Even though the effects found are not large, they might add to the prediction accuracy, and thus should be exploited. Although there is ongoing research to develop methods that efficiently make use of sequence variants in genomic predictions, at the moment, there is no consensus on how to benefit from the use of sequence information in genomic predictions. Wientjes et al., (2015a) demonstrated in a simulation study that adding causal QTL to SNP panels increased accuracy of genomic prediction, with a much larger increase achieved when the initial number of SNPs is lower. Because the significant variants from the RWAS are expected to be in LD with the underlying causal mutation, an option would be to add the identified variants to a low-density SNP chip to improve accuracy of genomic prediction. Brondum et al., (2015) investigated this option in cattle wherein they performed a GWAS in Nordic cattle for three separate breeds using sequence variants for different traits. They then selected QTL and three to five variants to tag each QTL and combined 1,623 variants with a 54k SNP panel. Their results showed that the reliability of genomic prediction was improved when using the combined SNP panel compared to only the 54k panel, with the largest gains of up to 5 percentage points for production traits and lowest of 0.5 percentage point increase for classical fertility traits. Their results also showed that when using a Bayesian model accuracies were generally higher with only 54k data compared with the genomic BLUP. In chapter 4 of this thesis, I showed that there was more predictive ability for endocrine fertility traits than classical fertility traits. This suggests that combining the significant markers from the RWAS for endocrine fertility traits to a low SNP panel like 50k SNP might improve the accuracy of prediction further, compared to the 0.5 percentage point increase for classical fertility traits.

In general, the accuracy of genomic selection can be improved by spending time on trait definition to improve heritability, improving statistical modelling of the trait, or by increasing the size of the training population. Although all these factors are important to improve accuracy of prediction, for endocrine fertility traits and novel traits in general, the critical factor that needs to be improved is the size of the training population.

6.4 Practical integration of novel traits in routine genetic evaluations

Chapter 5 showed that for genomic prediction, it is more important to have phenotypes from more cows than more lactations with phenotypes per cow in the training population. That is, optimizing recording strategies for endocrine fertility traits will maximize the accuracy of genomic prediction for fertility. This is true for endocrine fertility traits, but will generally apply to novel traits that are expensive or difficult to measure, e.g., feed efficiency and methane emission. For endocrine fertility traits, some important questions that arise are “can we use the Herd Navigator for the national breeding goal?”, “who should pay for the Herd Navigator?”, and “should farmers be subsidized for collecting extra records?” There is no single answer to these questions, as this depends on several factors. In this section, I will discuss the pros and cons of using the Herd Navigator in national breeding evaluations, phenotyping strategies for novel traits, and in general, challenges for integrating novel traits in genetic evaluation schemes.

6.4.1 Use of the Herd Navigator in genetic evaluations

In this thesis, we specifically looked at estrous related fertility traits derived from in-line milk P4 levels measured with the Herd Navigator, and how to include these traits in genetic evaluations. One downside of using the Herd Navigator for the national breeding goal, is that it is not realistic for progeny testing schemes. Also, chapter 5 showed that it is only worthwhile to invest in endocrine traits if the endocrine traits are themselves breeding goal traits. A key requirement for the recording of data is the motivation of the stakeholders involved. That is, recording needs to have benefits beyond genetic improvement, and the additional effort required for recording must result in added value, also for the farmers. The added advantage of the Herd Navigator is that famers can also use the Herd Navigator for other management purposes like estrus detection, and monitoring of mastitis and ketosis. However, there is a lot of competition with other fertility systems e.g., fertility systems based on heat detection and movement like physical activity monitors; these other systems are generally cheaper than the Herd Navigator. Furthermore, fertility traits based on physical activity monitors like interval from calving to first high activity have been shown to have substantial genetic variation, and have the potential to improve genetic selection for fertility (Løvendahl and Chagunda, 2009; Ismael et al., 2015). However, comparisons of pedometer measures with progesterone-determined estrus has shown that activity monitors cannot detect silent behavioral estrus (Løvendahl and Chagunda, 2016). Studies show that body condition score has a favorable relationship with fertility (e.g., Pryce et al., 2001; Berry et al., 2003) and is growing in popularity as a novel predictor for fertility (Fogh et al., 2013), but body condition score cannot be used for estrus detection. Research on the use of mid-infrared predicted fatty acids as indicator traits for fertility (Bastin et al., 2012) and of mid-infrared for pregnancy status testing is also underway (Gengler, 2014). None the less, such novel traits are not designed to detect estrus, except for the activity meters, and P4 remains the ‘gold standard’ for estrus detection.

For the farms that have the Herd Navigator, P4 levels are usually recorded until pregnancy is established, and these P4 measure were available in the data used in this thesis. Therefore, these profiles could also be used to derive endocrine fertility traits that reflect the ability of a cow to conceive following insemination, and become pregnant. This will be useful, as these endocrine traits might be more informative and objective than some of the classical fertility traits currently used to reflect conception and pregnancy in dairy cows e.g., non-return rate (see Figure 1, chapter 1). For example, perhaps describing the length of the last luteal phase before pregnancy might be a more informative trait that reflects the ability of a cow to get pregnant after insemination. The last luteal phase before pregnancy might be more informative because in general, after ovulation, the luteal phase is the period during which the endometrium (uterine lining) is prepared for implantation of an eventual fertilized egg (or embryo). That is, estrogen and progesterone together promote thickening of the endometrium. The length of the last luteal phase before pregnancy might be informative on whether a pregnancy is successful, as shorter luteal phase lengths might mean that there is not enough time for the endometrium to develop, which might contribute to embryo death and pregnancy failure. The last luteal phase length before pregnancy might be a more objective measure for the ability of a cow to get pregnant after insemination. Yet another option would be to describe lengths of all luteal phases before pregnancy, as animals with regular luteal phase lengths might have a higher probability of getting pregnant. Also, the interval from commencement of luteal activity to first service was investigated in chapter 2, but perhaps it will be more useful to look at the interval from C-LA to the last service before pregnancy, as this might be a more objective measure of the interval from calving to conception. For future studies, it would be interesting to consider defining and investigating this group of endocrine fertility traits that reflect the ability of a cow to conceive and get pregnant. Therefore, the Herd Navigator also makes it possible to exploit P4 measures further to improve fertility, because ideas that were considered unrealistic in the past may become feasible in the near future.

6.4.2 Phenotyping strategies for novel traits

A major benefit of genomic selection in dairy cattle breeding is that it is no longer required to measure a large number of phenotypes from progeny groups for each male selection candidate. This significantly decreases generation intervals and increases genetic gain per year for all breeding goal traits (Pryce et al., 2016; Lund et al., 2011). For traits that have been part of the breeding goal, current training populations consist of bulls with highly accurate phenotypes based on the average phenotype of hundreds of daughters (for reviews see: Hayes et al., 2009; Calus, 2010) and may contain up to as much as 16, 000 bulls (Lund et al., 2011). For novel traits, depending on the cost of a single phenotypic measurement, composing training populations may only be feasible for a few thousand cows. Therefore, there is the need to optimize phenotyping strategies for novel phenotypes.

Genomic predictions with novel traits can be trained within a set of animals, representative for the whole population and then applied in the general population. In sheep or beef cattle breeding, selection for carcass and meat quality traits is implemented in a centralized approach, in which test animals are housed in information nucleus herds, and phenotypes are recorded on the breeding animals themselves (i.e., own phenotypes). One advantage of an information nucleus is that animals are well identified and similarly managed, and fixed effects are fully recorded. In its current form, dairy cattle breeding companies have contracts with farmers, where intensive recording of phenotypes is performed. However, this system results in lots of records but with lots of lactations per animal. In this thesis, I showed that for genomic prediction of novel traits, it is more important to have more cows with phenotypes than more lactations per cow in the training population (chapter 5). Perhaps the dairy industry could learn from phenotyping structures established for breeding in other species. That is, the dairy industry should think of optimizing phenotyping structures by establishing nucleus farms for phenotype recording, especially for novel phenotypes that need new technologies for optimal recording.

6.4.3 Phenotyping farms and labs

Advances in technology will have an impact on the future definition and availability of phenotypes. The challenge for dairy breeding companies is to start thinking of ways to optimize phenotyping strategies, which is something that has been lacking behind in the dairy industry compared to the meat industry. Some dairy companies (http://www.eaap.org/Annual_Meeting/2015_warsaw/S22_12.pdf), have currently set up partner herds called “DataPlus” farms for recording specific phenotypes. A next step for dairy breeding companies would be to optimize the phenotyping structure by establishing phenotyping labs where animals are phenotyped in batches for one lactation and dispatched. This will make space for the next batch of cows and leads to phenotypes recorded on as many animals as possible. Based on the results of chapter 5, I argue that this is an optimal way of collecting as many phenotypes on novel traits as possible to set up a training population for genomic prediction. In the case of endocrine traits, the Herd Navigator would be a great investment for phenotyping labs. Another advantage of having phenotyping labs is that animals to be measured can be selected in terms of their relatedness to important bulls (or offspring of the sires of the young bulls) so that there is always a close relationship between the cow training population and the candidate bulls, as this will increase the accuracy of genomic prediction. Studies have shown that a close relationship between evaluated animals and animals in the training population is expected to give more reliable prediction (Habier et al., 2007; Meuwissen, 2009; Pszczola et al., 2012). These relationships are especially important for small training populations (Wientjes et al., 2013). That is, optimally, all evaluated animals should have at least some closely related animals in the training population. Therefore, it will be more interesting to phenotype heifers (i.e., first parity cows) for novel traits than later parity cows. It is also important to have an idea of genetic correlation estimates between endocrine fertility traits across parities, but that requires large datasets, which were not available in this thesis. However, genetic correlation estimates of classical fertility traits across lactations can give an idea of what to expect from endocrine traits. Reported genetic correlation of classical fertility traits between first and second lactation vary between 0.7 and 0.9 (Roxström et al., 2001; Haile-Mariam et al., 2003), and are 0.9 or higher between the second and third lactation. Based on these, I do not expect a large difference for the correlation of endocrine fertility traits across lactations. To enable the immediate use of endocrine fertility traits in genetic evaluations, these correlation estimates for classical traits could be used as pointers rather than waiting to first obtain correlation estimates of endocrine traits. Therefore, rather than the partner farms, maybe the dairy breeding companies should be thinking about phenotyping labs.

On the one hand, investing in phenotyping labs with the Herd Navigator will optimize the recording strategy for endocrine fertility traits in dairy cattle, but this investment will not be worthwhile if the only traits recorded are fertility traits. On the other hand, if investments for recording other novel traits e.g., feed intake and methane emission are included, this might become a realistic option for the future, where breeding companies establish high-tech farms with equipment for measuring different types of data, and where animals are kept for one lactation only. As an example, to optimize phenotyping strategies, a breeding company could invest in 10 phenotyping labs where all novel phenotypes of interest are collected. Heifers are bought, genotyped, phenotypes are recorded for one lactation, and then animals are sold. Another advantage of phenotyping labs is that all traits can be recorded on the same animals. The number of phenotyping labs needed to optimize phenotyping can be decided based on predicted accuracy and expected size of the training population, as shown in chapter 5. A critical question that arises concerns genotype-by-environment (G X E) interactions. With limited training populations in a given environment, selection will be more sensitive to G X E interactions than with traditional progeny tests where daughters of candidate bulls are distributed over a large number of herds. Therefore, it is recommended to establish the phenotyping labs over a range of environments.

6.5 Concluding remarks

The research in this thesis investigated the use of in-line milk progesterone concentrations to define novel fertility traits that can be used in genetic evaluations to improve genomic selection for fertility in dairy cows. This was inspired by the availability of new in-line recording technologies like the Herd Navigator, and the opportunity due to genomic selection, that makes it feasible to implement novel traits in breeding programs for dairy cattle. In the first part of this thesis, I show how in-line milk progesterone records can be used to define several heritable endocrine fertility traits, and describe research that detected genetic markers significantly associated with endocrine fertility traits, fine-mapped target QTL regions and identified potential candidate genes. The less unfavorable genetic correlation of endocrine fertility traits like and their heritability makes them better alternatives for use in genetic improvement of fertility than classical fertility traits like CFS. I expect that GWAS will continue to be performed because they provide scientifically relevant results, especially with greater statistical power when more animals will be sequenced or genotyped using high density SNP chips. With more markers, the physical distance between markers and the causative mutation will be shortened, allowing QTL regions to be fine-mapped. However, finding the causal mutation will require more than just a GWAS using denser genotyping or sequence data. Linkage disequilibrium plays a major role in GWAS, and one may require additional evidence to distinguish associated variants. The results of GWAS with sequence variants can be used to augment low-density SNP chip panels like 50k, to increase the accuracy of genomic prediction.

In the second part of this thesis, I describe genomic prediction using endocrine fertility traits in dairy cattle. This is the first study on the use of endocrine traits in genomic prediction. I have shown that endocrine fertility traits have more predictive ability than classical fertility traits, and that combining endocrine fertility traits in multi-trait genomic prediction can substantially increase the accuracy of genomic prediction. Because endocrine traits are novel fertility traits, there is bound to be a limited training population size for these traits. I showed that for prediction, it is more important to phenotype more animals than more lactations per animal in the training population. Therefore, breeding companies can optimize phenotyping strategies for endocrine fertility traits (and novel traits in general) by establishing phenotyping labs where animals are phenotyped in batches for one lactation and dispatched, as leads to phenotypes recorded on as many animals as possible. However, as shown in this thesis, investing in a cow training population for endocrine fertility traits (and novel traits in general), is relevant only when the novel trait is in the breeding goal.

In the final discussion, I show that new tools like, genomic selection and in-line recording technologies provide a great opportunity to reconsider the use of endocrine fertility traits to further improve selection for fertility, and should be explored further. These tools make it possible to optimize phenotype recording and increase the size of the training population, which is the main limiting factor of improving accuracy of prediction with novel traits. Also, the availability of more and more new electronic technologies for farm management and phenotype recording means that in the future, dairy farming might shift from the traditional system to a more electronic technology oriented farming system. Therefore, the dairy industry might be faced with challenges on how to handle/integrate new phenotypes from these new technologies in to genetic evaluations. The results from this thesis can server as pointers.

References

Andersen-Ranberg, I.M., G. Klemetsdal, B. Heringstad, and T. Steine. 2005. Heritabilities, genetic correlations, and genetic change for female fertility and protein yield in Norwegian dairy cattle. *J. Dairy Sci.* 88:348–355.

Bastin, C., D.P. Berry, H. Soyeurt, and N. Gengler. 2012. Genetic correlations of days open with production traits and contents in milk of major fatty acids predicted by mid-infrared spectrometry. *J. Dairy Sci.* 95:6113–6121.

van den Berg, I., D. Boichard, and M.S. Lund. 2016. Sequence variants selected from a multi-breed GWAS can improve the reliability of genomic predictions in dairy cattle. *Genet. Sel. Evol.* 48:83.

Berry, D.P., J.W.M. Bastiaansen, R.F. Veerkamp, S. Wijga, E. Wall, B. Berglund, and M.P.L. Calus. 2012. Genome-wide associations for fertility traits in Holstein-Friesian dairy cows using data from experimental research herds in four European countries. *Animal*. 6:1206–15.

Berry, D.P., F. Buckley, P. Dillon, R.D. Evans, M. Rath, and R.F. Veerkamp. 2003. Genetic relationships among body condition score, body weight, milk yield, and fertility in dairy cows. *J. Dairy Sci.* 86:2193–2204.

Brøndum, R.F., G. Su, L. Janss, G. Sahana, B. Guldbrandtsen, and D. Boichard. 2015. Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. *J Dairy Sci*. 98. doi:10.3168/jds.2014-9005.

Bulman, D.C., and G.E. Lamming. 1978. Milk progesterone levels in relation to conception, repeat breeding and factors influencing acyclicity in dairy cows. *J. Reprod. Fertil.* 54:447–458.

Calus, M.P.L. 2010. Genomic breeding value prediction: methods and procedures. *Animal*. 4:157–164.

Daetwyler, H.D., M.P.L. Calus, R. Pong-Wong, G. los Campos, and J.M. Hickey. 2013. Genomic prediction in animals and plants: Simulation of data, validation, reporting, and benchmarking. *Genetics*. 193.

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

Daetwyler, H.D., R. Pong-Wong, B. Villanueva, and J.A. Woolliams. 2010. The Impact of Genetic Architecture on Genome-Wide Evaluation Methods. *Genetics*. 185:1021 LP-1031.

Daetwyler, H.D., B. Villanueva, and J.A. Woolliams. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS One*. 3:e3395.

Darwash, A.O., G.L. Ward, G.E. Lamming, and J.A. Woolliams. 1999. The effects of raising post-oestrus progesterone concentrations on luteal activity in post-partum dairy cows. *Anim. Sci.* 68:527–532.

Duchemin, S.I., M. Glantz, D.-J. de Koning, M. Paulsson, and W.F. Fikse. 2016. Identification of QTL on Chromosome 18 Associated with Non-Coagulating Milk in Swedish Red Cows. *Front. Genet.* 7:57.

Duchemin, S.I., M.H.P.W. Visker, J.A.M. Van Arendonk, and H. Bovenhuis. 2014. A quantitative trait locus on Bos taurus autosome 17 explains a large proportion of the genetic variation in de novo synthesized milk fatty acids. *J. Dairy Sci.* 97:7276–7285.

Fogh, A., G.P. Aamand, L. Hjorto, and U. Lauritsen. 2013. Recording of data and identification issues. *In* ICAR-Technical Workshop 2013.

Gengler, N. 2014. Mitigation strategies versus adaptation strategies. Invited presentation at the ASGGN Workshop.

Goddard, M. 2009. Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica*. 136:245–257.

Habier, D., R.L. Fernando, and J.C.M. Dekkers. 2007. The impact of genetic relationship information on genome-assisted breeding values. *Genetics*. 177:2389–2397.

Haile-Mariam, M., P.J. Bowman, and M.E. Goddard. 2003. Genetic and environmental relationship among calving interval, survival, persistency of milk yield and somatic cell count in dairy cattle. *Livest. Prod. Sci.* 80:189–200.

Hayes, B.J., P.J. Bowman, A.C. Chamberlain, K. Verbyla, and M.E. Goddard. 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol.* 41:1.

Höglund, J.K., G. Sahana, R. Brøndum, B. Guldbrandtsen, B. Buitenhuis, and M.S. Lund. 2014. Fine mapping QTL for female fertility on BTA04 and BTA13 in dairy cattle using HD SNP and sequence data. *BMC Genomics*. 15:790.

Ismael, A., E. Strandberg, B. Berglund, A. Fogh, and P. Løvendahl. 2016. Seasonality of fertility measured by physical activity traits in Holstein cows. *J. Dairy Sci.* 99:2837–2848.

Ismael, A., E. Strandberg, M. Kargo, A. Fogh, and P. Løvendahl. 2015. Estrus traits derived from activity measurements are heritable and closely related to the time from calving to first insemination. *J. Dairy Sci.* 98:3470–3477.

Kemper, K.E., and M.E. Goddard. 2012. Understanding and predicting complex traits: Knowledge from cattle. *Hum Mol Genet*. 21.

Khatkar, M.S., F.W. Nicholas, A.R. Collins, K.R. Zenger, J.A.L. Cavanagh, W. Barris, R.D. Schnabel, J.F. Taylor, and H.W. Raadsma. 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. *BMC Genomics*. 9:1.

Lamming, G.E., and a O. Darwash. 1998. The use of milk progesterone profiles to characterise components of subfertility in milked dairy cows. *Anim. Reprod. Sci.* 52:175–90.

Løvendahl, P., and M.G.G. Chagunda. 2009. Short communication: Genetic variation in estrus activity traits. *J. Dairy Sci.* 92:4683–4688.

Løvendahl, P., and M.G.G. Chagunda. 2016. On the use of physical activity monitoring for estrus detection in dairy cows. *J. Dairy Sci.* 93:249–259.

Lucy, M.C. 2001. Reproductive loss in high-producing dairy cattle: where will it end? *J. Dairy Sci.* 84:1277–1293.

Lund, M.S., A.P.W. De Roos, A.G. De Vries, T. Druet, V. Ducrocq, S. Fritz, F. Guillaume, B. Guldbrandtsen, Z. Liu, and R. Reents. 2011. A common reference population from four European Holstein populations increases reliability of genomic predictions. *Genet Sel Evol*. 43:43.

Meuwissen, T., and M. Goddard. 2010. Accurate prediction of genetic values for complex traits by whole-genome resequencing. *Genetics*. 185:623–631.

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

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.

Miglior, F., J. Chesnais, and B.J. Van Doormaal. 2012. Genetic improvement: a major component of increased dairy farm profitability. *38th Ann Meet ICAR*. 28.

Miglior, F., B.L. Muir, and B.J. Van Doormaal. 2005. Selection indices in Holstein cattle of various countries. *J. Dairy Sci.* 88:1255–1263.

Misztal, I. 2011. FAQ for genomic selection. *J. Anim. Breed. Genet.* 128:245–246.

Petersson, K.-J., B. Berglund, E. Strandberg, H. Gustafsson, A.P.F. Flint, J.A. Woolliams, and M.D. Royal. 2007. Genetic analysis of postpartum measures of luteal activity in dairy cows. *J. Dairy Sci.* 90:427–434.

Pryce, J.E., S. Bolormaa, A.J. Chamberlain, P.J. Bowman, K. Savin, M.E. Goddard, and B.J. Hayes. 2010. A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using variable length haplotypes. *J. Dairy Sci.* 93:3331–3345.

Pryce, J.E., M.P. Coffey, and G. Simm. 2001. The relationship between body condition score and reproductive performance. *J. Dairy Sci.* 84:1508–1515.

Pryce, J.E., M.E. Goddard, H.W. Raadsma, and B.J. Hayes. 2016. Deterministic models of breeding scheme designs that incorporate genomic selection. *J. Dairy Sci.* 93:5455–5466.

Pszczola, M., T. Strabel, H.A. Mulder, and M.P.L. Calus. 2012. Reliability of direct genomic values for animals with different relationships within and to the reference population. *J. Dairy Sci.* 95:389–400.

Roxström, A., E. Strandberg, B. Berglund, U. Emanuelson, and J. Philipsson. 2001. Genetic and Environmental Correlations Among Female Fertility Traits and Milk Production in Different Parities of Swedish Red and White Dairy Cattle. *Acta Agric. Scand. Sect. A — Anim. Sci.* 51:7–14.

Royal, M.D., A.P.F. Flint, and J.A. Woolliams. 2002. Genetic and Phenotypic Relationships Among Endocrine and Traditional Fertility Traits and Production Traits in Holstein-Friesian Dairy Cows. *J. Dairy Sci.* 85:958–967.

Royal, M.D., E. Wall, and A.P.F. Flint. 2003. The genetic relationship between interval to commencement of luteal activity postpartum and UK national fertility proofs for dairy cattle. *Proc. Br. Soc. Anim. Sci.(BSAS), York, UK. Br. Soc. Anim. Sci., Edinburgh, UK*. 54.

Sahana, G., B. Guldbrandtsen, B. Thomsen, L.E. Holm, F. Panitz, R.F. Brøndum, C. Bendixen, and M.S. Lund. 2014. Genome-wide association study using high-density single nucleotide polymorphism arrays and whole-genome sequences for clinical mastitis traits in dairy cattle1. *J. Dairy Sci.* 97:7258–7275.

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

Sun, C., P. Madsen, U.S. Nielsen, Y. Zhang, M.S. Lund, and G. Su. 2009. Comparison between a sire model and an animal model for genetic evaluation of fertility traits in Danish Holstein population. *J. Dairy Sci.* 92:4063–4071.

Veerkamp, R.F., and B. Beerda. 2007. Genetics and genomics to improve fertility in high producing dairy cows TT - Proceedings of the International Conference on Farm Animal Reproduction “From Egg to Embryo” International Conference on Farm Animal Reproduction. *Theriogenology*. 68, Supple: S266–S273.

Veerkamp, R.F., J.K. Oldenbroek, H.J. Van Der Gaast, and J.H.J. Van Der Werf. 2000. Genetic Correlation Between Days Until Start of Luteal Activity and Milk Yield, Energy Balance, and Live Weights. *J. Dairy Sci.* 83:577–583.

Wall, E., S. Brotherstone, J.A. Woolliams, G. Banos, and M.P. Coffey. 2003. Genetic evaluation of fertility using direct and correlated traits. *J. Dairy Sci.* 86:4093–4102.

Walsh, S.W., E.J. Williams, and A.C.O. Evans. 2011. A review of the causes of poor fertility in high milk producing dairy cows. *Anim. Reprod. Sci.* 123:127–138.

Wientjes, Y.C., M.P. Calus, M.E. Goddard, and B.J. Hayes. 2015a. Impact of QTL properties on the accuracy of multi-breed genomic prediction. *Genet Sel Evol*. 47.

Wientjes, Y.C.J., R.F. Veerkamp, and M.P.L. Calus. 2015b. Using selection index theory to estimate consistency of multi-locus linkage disequilibrium across populations. *BMC Genet.* 16:1.

Wu, X., M.S. Lund, G. Sahana, B. Guldbrandtsen, D. Sun, Q. Zhang, and G. Su. 2015. Association analysis for udder health based on SNP-panel and sequence data in Danish Holsteins. *Genet. Sel. Evol.* 47:1.

Yan, J., T. Shah, M.L. Warburton, E.S. Buckler, M.D. McMullen, and J. Crouch. 2009. Genetic characterization and linkage disequilibrium estimation of a global maize collection using SNP markers. *PLoS One*. 4: e8451.

Yang, J., B. Benyamin, B.P. McEvoy, S. Gordon, A.K. Henders, D.R. Nyholt, P.A. Madden, A.C. Heath, N.G. Martin, and G.W. Montgomery. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42:565–569.

Summary

Summary

Improved reproductive performance has a substantial benefit for the overall profitability of dairy cattle farming by decreasing insemination and veterinary treatment costs, shortening calving intervals, and lowering the rate of involuntary culling. Unfortunately, the low heritability of classical fertility traits derived from calving and insemination data makes genetic improvement by traditional animal breeding slow. Therefore, there is an interest in finding novel measures of fertility that have a higher heritability or using genomic information to aid genetic selection for fertility. Endocrine fertility traits based on progesterone concentration in milk have been widely accepted as valid indicators for fertility because they are not biased by farm management decisions and more directly reflect a cow's reproductive physiology than classical fertility traits. However, the use of endocrine traits in routine genetic evaluations has been constrained by the high cost associated with collecting progesterone records on sufficient number of animals. Nonetheless, in-line technology like the Herd Navigator now exists to automatically measure progesterone concentration during milking, enabling sampling of more animals at reasonable cost.

In this thesis, I explored the use of milk progesterone concentrations and genomic information to improve genetic selection for fertility in dairy cattle. In chapter 2, I investigated the use of in-line milk progesterone records to define endocrine fertility traits, and estimated genetic parameters. Several defined endocrine fertility traits were heritable, and showed a reasonable repeatability. The genetic correlation of milk production traits with endocrine fertility traits were considerably lower than the correlations of milk production with classical fertility traits. In chapter 3, genomic regions associated to endocrine fertility traits were identified on the cattle genome by genome-wide association and fine-mapping.

The genome-wide association study identified 17 quantitative trait loci (QTL) associated with endocrine fertility traits, on *Bos taurus* autosomes (BTA) 2, 3, 8, 12, 15, 17, 23, and 25. Overlapping QTL regions were found between endocrine traits on BTA 2, 3, and 17. For the classical trait calving to first service, three QTL regions were identified on BTA 3, 15, and 23, and an overlapping region on BTA23 with endocrine traits. Fine-mapping target regions for the endocrine traits on BTA 2 and 3 confirmed the QTL from the genome-wide association study, and identified several associated variants that can contribute to an index of markers for genetic improvement of

fertility. Several potential candidate genes underlying endocrine fertility traits were also identified in the target regions.

In the next two chapters, I investigated the optimal use of endocrine fertility traits in genomic evaluations. In chapter 4, using empirical and theoretical predictions for single-trait models, I showed that endocrine fertility traits have more predictive ability than classical fertility traits. The accuracy of genomic prediction was also substantially improved when endocrine and classical fertility traits were combined in multi-trait genomic prediction. Across country predictions were also evaluated in univariate predictions, and some predictive ability was observed. Because we were limited by sample size in chapter 4, in chapter 5, using a deterministic equation, I investigated the potential accuracy of multi-trait genomic selection when combining a cow training population measured for the endocrine fertility trait C-LA, with a training population of bulls with daughter observations for classical fertility traits. The results showed that for prediction of fertility, there is no benefit of investing in a cow training population when the breeding goal is based on classical fertility traits. However, when considering a more biological breeding goal for fertility like C-LA, accuracy is substantially improved when endocrine traits are available from a limited number of farms. Evaluation of the potential accuracy of across country genomic prediction showed that when the training population from one country (e.g., Sweden) is small, substantial increase in accuracy can be achieved by adding animals from another country (e.g., The Netherlands), however, the accuracy was highly dependent on the genetic correlation between countries. In chapter 5, recording strategies that optimally use the Herd Navigator for genomic prediction, in terms of, the number of farms, and recording period for endocrine traits were investigated. The results showed that for prediction of fertility using C-LA, it is more important to have more animals with C-LA records than more C-LA records per animal. That is, for prediction accuracy with novel traits in general, it is more beneficial to phenotype more animals than more lactations per animal in the training population.

Finally, in chapter 6, the relevance of the findings was discussed, i.e. how breeders can benefit from combining endocrine fertility traits with genomic information to further improve selection for fertility. Suggestions for future studies and how breeders can make use of the results generated from this study were given. I finalized by suggesting practical phenotyping strategies like phenotyping labs, that can by established by breeding companies to optimize phenotype recording for endocrine fertility traits (and novel traits in general).

Acknowledgements

Acknowledgements

The work in this thesis truly has resulted from the contribution of many. The list of individuals to whom I owe the utmost gratitude is lengthy. Without their support, this work would simply not have been possible. I first give honor to God, who has provided this vision and purpose in my life, and endowed me with the resources and facilities to fulfill it. These pass 4 years have been a worthy journey and adventure.

A special thank you to my supervisor team for their guidance and support throughout this journey. Roel, thank you for your critical and dynamic attitude, helping to push my thoughts in the right direction. I admire your dynamic skill, which I aspire to acquire. Aniek, thank you for the very detailed and useful remarks to my papers, even the minor details, they were very much appreciated. Britt, you always found a way to encourage me, and I learnt a lot about fertility from you, thank you. Erling, I am very grateful to have had you as a supervisor for the first half of my PhD, thank you for your always very insightful comments. DJ, although you joined the team half way, you immediately took interest in my work, and I very much enjoyed working with you on my last papers. Thank you for all the discussions, especially during the time spent in Uppsala.

Thank you, Sander de Roos, Donagh Berry, and Mike Coffey, for being in my thesis committee. Thank you for interesting and valuable discussions. Thanks to my opponents for reading the thesis, I look forward to the discussion during the public defense.

Mathijs, thank you for sharing the office with me at the start of my thesis, and for all the nice discussions we had. Your super-duper organized skills inspired me to want to improve on mine. Ernest, thank you for your friendship, and all the laughs :). Both of you, thanks for taking the responsibility of being my paranymphs. Marvis, thank for your hospitality, and for offering me a place to stay whenever I needed one. Thank you, Johan and Helena Valema, for opening your home to me, and always encouraging me. Special thanks to Coralia, Yvonne, Rianne, Sonia, Mathijs, Yvette, Mario, Ina, Eef, Marcin, and all at the Triton building for helping me fit in at the start of my PhD.

To my ICF family in Wageningen, thank you for receiving me always with warmth, for the great worships and fellowships. Special thanks to the worship team members; Yuni, Peter, Gerson, Koenraad, Adnan, Sebastiaan, Ebenezer for all the worship preparation time shared. Special thanks also go to the “Connect group Ede”, thank you for your companionship and elaborate discussion of the word.

This work was carried out at Wageningen University (WUR) and the Swedish University of Agricultural Sciences (SLU). To researchers and my colleagues at both Universities, thank you for your helpfulness, and for creating an atmosphere that I enjoyed a lot while working on my thesis. A special thank you to Ina, Zih-Hua, Mandy, Nancy, Henk, Jack, Jovanna, Juan, John, Lucia, Mahlet, Marcos, Marinus, Sandrine, Sandrine, Tessa, Hamed, Pascal, Hadi, Andre, Gabriel, Jeremie, Floor, Kasper, Sanne, Piter, Hans, Han, Bart Maria, Esther, Gareth, Biaty, Anouk, Anoop, Tom, Sabine, Shuwen, Robert, Asaah, Rita, from WUR. Thanks to Agnese, Merina, Chrissy, Berihu, Sofia, Emilie, Josh, Shizi, Thu, Cano, Suvi, Kim, Freddy, lotta, Lisa, Valentina, from SLU, for all the Firka and nice dinners together. I would also like to acknowledge all administrative staff members for all the help I received from them. Thank you to Ada, Lisette, and Maya at WUR, Monika, and Helena at SLU, and Helena Fuchs and Isabelle Laissy at the EGS-ABG secretary.

I would like to acknowledge members of the EU Prolific project, of which this work was a subset of the larger project, thank you for lots of valuable interactions during the annual project meetings.

To Coralia, Claudia, and Merina, thank you for all the awesome trips we had together. These diversions made the journey more fun and exciting, and I am honored to have shared them with you girls, and to get to know you more. Thank you, Sonia, Zih-Hua, Maria, and Coralia for the ladies’ cocktail nights. I admit, you girls got the best cocktail skills ;).

Special thanks to my family for their patience and support given to me over the last years. Mom and Dad, thank you for your love, and for being the role models in my life. My dearest sisters; Makah, Aza, Bella, Glory, and brothers; Akum, big & small Muluh, Nda, thank you for cracking me up with jokes whenever I called or was home. You guys are the best, you’re too “nyanga”

Dear reader, thank you for taking the effort to read this thesis. I hope you enjoyed it and consider the time spent valuable.

Amabel