Immunogenetics in dairy cattle:
somatic cell score and natural antibody levels

Immunogenetics in dairy cattle:

somatic cell count and natural antibody levels

Susan Wijga

Thesis

submitted in fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 18 October 2013
at 11 a.m. in the Aula.

Abstract

Wijga, S. (2013). Immunogenetics in dairy cattle: somatic cell count and natural antibody levels. PhD thesis, Wageningen University, the Netherlands

There remains a lot to be learned about the interpretation of genetic parameters and the biology of disease resistance and somatic cell count, also known as somatic cell score (SCS). This PhD thesis aimed to obtain additional insight in disease resistance and SCS by: 1) quantifying the impact of genetics on innate immunity, represented by natural antibodies (NAb), through estimation of heritabilities and genetic correlations; 2) identifying the genomic regions involved in SCS and NAb levels; 3) quantifying the impact of genetics on environmental sensitivity for SCS.

Natural antibody levels were found to be heritable with heritabilities ranging from 0.06 to 0.55 and in general, heritabilities for NAb isotypes were higher than heritabilities for total NAb levels, the latter making no distinction between isotypes. Genetic correlations suggest that isotypes IgA and IgM have a common genetic basis, but that the genetic basis for IgG1 differs from that for IgA or IgM. An additional genome-wide association study for NAb levels showed that information can be gained when total NAb levels are further subdivided into isotype levels. A region on chromosome 23 was significantly associated with genetic variation in isotype IgM levels. The bovine major histocompatibility complex (MHC) is located near this region, making this a region of candidate gene(s) involved in NAb expression in dairy cows both from a functional and positional perspective. Results from the study on genetic parameters and the genome-wide association study suggest that NAb isotypes may provide a better characterization of different elements of the immune response or immune competence and enable more effective decisions when breeding programs start to include innate immune parameters. A genome-wide association study was not only performed for NAb levels, but also for SCS. Relatively few associations, however, were found, which suggests that SCS is controlled by multiple loci, each with a relatively small effect, distributed across the genome.

Somatic cell score is partly under genetic control, but is also affected by the environment. Sensitivity to respond to environmental factors, however, can have a genetic origin. Environmental factors can be divided into known and unknown factors, referred to as macro- and micro environment, respectively. Macro-environmental sensitivity can be expressed as genetic variation in the slope of a reaction norm, whereas micro-environmental sensitivity can be expressed as differences in residual variance that have a genetic origin. Both macro- and micro-environmental sensitivity were found for SCS and these sensitivities were positively

correlated. Knowledge on both forms of sensitivity can aid in optimization of selection as correlations between the additive genetic variance in intercept, slope and environmental variance were all away from unity. Selection for reduced environmental sensitivity has the potential to reduce variability in animal performance due to environmental factors and herewith increase predictability of performance across and within environments.

Knowledge on disease biology is important to fully understand the processes involved when selecting for increased disease resistance, as a better understanding enables a better prediction of the consequences. In this context, the general discussion discussed the phenotype definition and statistical modeling, influence of maternal effects and genetic variation in the MHC region. The discussion contained three conclusions: 1) analyses of cell types (detailed phenotypes) rather than SCS can provide further insight in the genetic control of SCS and mastitis; 2) no evidence was found for maternal genetic effects on NAb levels in milk. Maternal environmental effects, however, could play a role in NAb levels; 3) genetic diversity in the MHC region is maintained by natural selection. Selective breeding and farm management practices may affect this genetic diversity, which could bring about negative effects on animal fitness, such as fertility problems. Selective breeding for specific MHC haplotypes may therefore impose a risk for negative effects on animal health.

Contents

1

General introduction

From early days and continued to this day, societies strongly depend on animal agriculture to not only provide food, fiber and other products, but also economic prosperity (Fisher and Mellor, 2008). Over time, a diversity of measures including nutrition, management and breeding have led to improved livestock productivity (Fisher and Mellor, 2008). Initially, breeding focused primarily on production traits such as milk yield, growth rate and number of eggs (Oltenacu and Broom, 2010). Today's breeding programs increasingly take into account functional traits such as fertility, survival and disease resistance to mitigate and improve the effects of breeding on animal behavior, physiology, health and welfare as some could become negatively affected when animal breeding primarily focuses on production traits (Fisher and Mellor, 2008), due to for instance negative correlations among traits.

This first chapter introduces the functional traits natural antibody (NAb) levels and somatic cell count (SCC). After an introduction of the innate and adaptive immune system, the structure and function of NAb are reviewed. Further the relation of SCC with mastitis is discussed and concepts of genome-wide association studies and environmental sensitivity are introduced.

1.1. Disease resistance

An important functional trait is disease resistance. Infectious diseases have a negative effect on livestock health and welfare and are an important source of economic loss (e.g. Glass, 2004). Further, disease causing pathogens in livestock and the substances used to eradicate these pathogens can provide a threat to biosecurity, public health, product quality and food safety (e.g. Glass et al., 2012). Potential means to increase disease resistance include vaccination and breeding, where the economically most effective means likely lies in the coupling of the development of vaccines with genetic selection for disease resistance. The optimal approach, however, depends on the host and the pathogen (Glass, 2004). Breeding for disease resistance has the potential to permanently improve resistance in a population. Disease resistance is, however, a complex phenomenon which not only involves various host-resistance mechanisms, but also depends strongly on environmental factors. Disease resistance remains a topic that warrants further study, as it may lead to new insights and approaches to reduce the impact of disease.
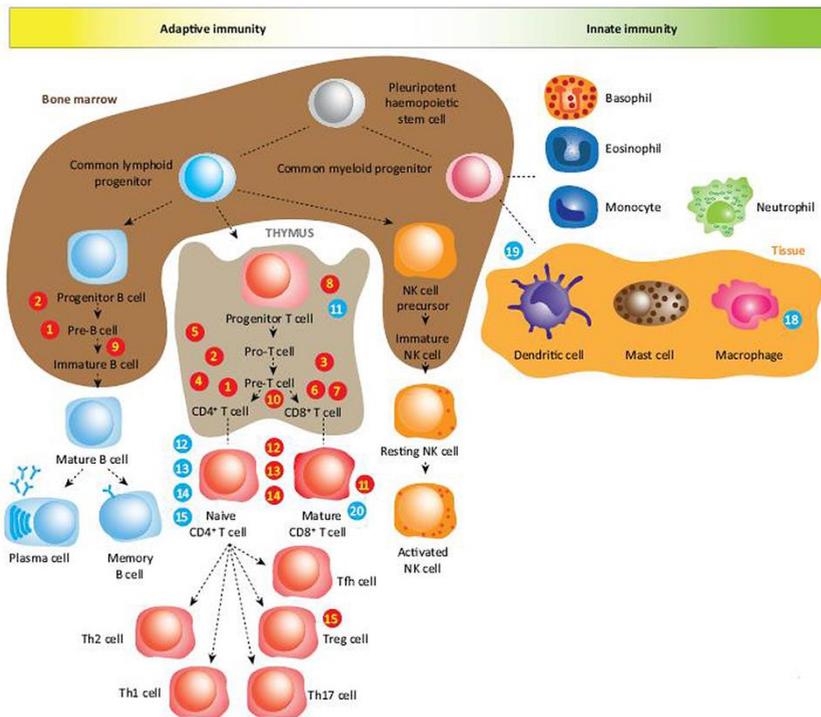
Figure 1.1 Overview of the (relation between) innate and adaptive immune system. Abbreviations: NK cell, natural killer cell; Tfh, follicular helper T cell; Th, T helper cell; Treg, regulatory T cell. (Adapted from Knight, J.C., 2013).

1.2. Immunity

The first physiological barrier that protects an individual from potentially pathogenic agents is an external barrier, such as skin, mucous membranes, body fluids or bacterial flora (Uthaisangsook et al., 2002). Once the potential pathogen has passed this first barrier, the second barrier is formed by the immune system.

Innate and adaptive immune system

A typical feature of the immune system is its ability to recognize and differentiate between the body's own (self) molecules, and non-self molecules (Sordillo and Streicher, 2002). In theory, the immune system can be divided into two components: the innate immune system and the adaptive immune system. In practice, however, the separation of innate and adaptive immunity is less obvious, because these systems are closely entangled and share many effector mechanisms (Rainard and Riollet, 2006). Innate immunity involves inborn resistance which is

expressed upon a very first pathogenic exposure, whereas adaptive immunity involves resistance acquired upon (repeated) pathogenic exposure. This adaptive immunity is initially weak, but strengthens with repeated encounters (Uthaisangsook et al., 2002). Figure 1 depicts an overview of the (relation between) innate and adaptive immune system. Note that this overview is for illustrating purposes and as such not exhaustive.

Innate immunity is a universal evolutionary conserved system, whereas adaptive immunity has evolved primarily in vertebrate species (Beutler, 2004, Medzhitov and Janeway, 1997). Recognition by the innate immune system occurs via conserved structures present on potential pathogens, called antigens, which most probably play a role in primary cell stability and cell preservation mechanisms. The benefit of recognizing conserved structures is that the immune system does not have to follow all mutational changes that occur in pathogens (Vollmers and Brändlein, 2009; Vollmers and Brändlein 2006). The innate immune response forms the foundation for activation of all adaptive immune responses by releasing effector cytokines, which is followed by an effector response by the adaptive immune system (Medzhitov and Janeway, 1997).

Natural antibodies

Natural antibodies are part of the innate immune system (Vollmers and Brändlein, 2009), and reside in serum, colostrum or milk, in the absence of apparent external antigenic stimulation. Natural antibodies are of importance in the activation of a primary immune response against (potential) pathogens. Upon invasion NAb trap the pathogens by binding to antigens present on the pathogens and transport them to lymphoid organs, herewith removing the pathogen from the circulation, and as such protecting vital organs from infection. The increased concentration of antigen in lymphoid organs triggers the T-cell and B-cell responses of the adaptive immune system (Kohler et al., 2003; Ochsenbein et al., 1999). Further roles of NAb may involve physiological 'housekeeping' tasks, and recognize and discard senescent cells, and other self-molecules (Vollmers and Brändlein, 2009).

Natural antibodies are measurable in bovine milk (Ploegaert et al., 2010; Van Knegsel et al., 2007). The bovine mammary gland has an active function in the regulation of the antibody concentration in colostrum and milk, although antibodies are not produced by the epithelium in the mammary gland itself (Stelwagen et al., 2009). Whereas some antibodies may "leak" from the blood serum into the colostrum or milk, most antibodies enter the udder via selective transport. These antibodies may originate from blood serum or are produced by intramammary plasma cells (Stelwagen et al., 2009). Little is yet known about the

exact role of NAb in bovine milk. In general, immune components in bovine milk are involved in conveying passive immunity to offspring and in providing host immunity to the mammary gland (Stelwagen et al., 2009). Most likely NAb play an eminent role in the innate immune defense of the mammary gland (Rainard and Riollet, 2006).

Antibodies, NAb but also antibodies that are part of the adaptive immune system, can be classified based on their binding specificity, but they can also be grouped according to the structure of their heavy chains. These groups, different in the structure of their heavy regions, are known as isotypes. Although NAb are mostly isotype immunoglobulin (Ig) M, IgG and IgA NAb were reported (Matter and Ochsenbein, 2008). An overview of antibody isotypes is provided in Figure 2.
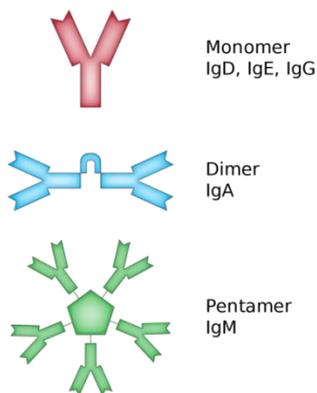
Figure 1.2 Antibody isotypes.

Which antigens an antibody is able to bind, i.e. the binding specificity of an antibody, is defined in the variable region of the antibody by the antigen binding site (Jerne, 1985). An overview of the antibody structure is provided in Figure 3. Innate immunity mostly functions through a transmitted germ-line coded pool of receptors that mainly bind conserved structures. A key characteristic of antibodies, however, is variability, which is acquired by combinatorial association (Vollmers and Brändlein, 2009), a process that generates antibody diversity by differential combination of antibody light chains and heavy chains. Additionally, recombination mistakes, deletions and additions in germ-line immunoglobulin genes provide a genetic and antigen binding site variability that is adequate to enable protection against a broad diversity of pathogenic agents (Vollmers and Brändlein, 2009).

Figure 1.3 Antibody structure.

1.3. Mastitis

Mastitis, an inflammation of the mammary gland, is regarded as the most frequent and costly infectious disease in dairy herds of developed countries, where costs can be attributed to revenues not earned due to for instance reduced milk yield and to real expenditures such as treatment costs and culling (Seegers et al., 2003). Many countries have included mastitis or the commo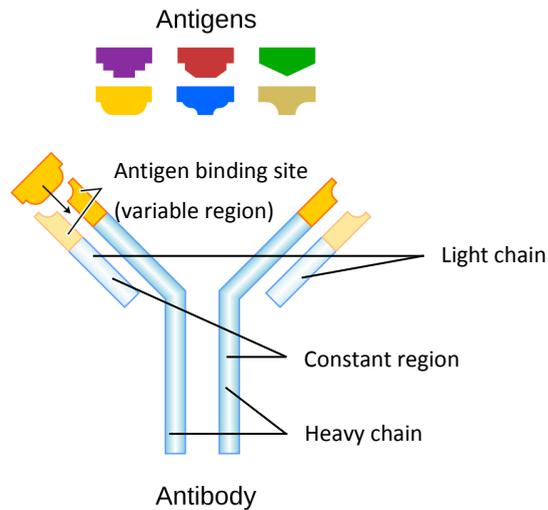nly used indicator trait SCC in their national genetic evaluations (Mark et al., 2002) in order to increase udder health.

Innate and adaptive immunity are required to be highly interactive and coordinated to enable optimal protection from mastitis (Sordillo and Streicher, 2002; Stelwagen et al., 2009).

Upon invasion of pathogens through the teat end into the mammary gland, resident cells already present in the mammary gland form the first line of defense. In most uninfected quarters SCC is below 10^5 cells/ mL (Rainard and Riollet, 2006). After entrance into the teat end, bacteria grow and divide rapidly, where bacterial numbers in milk from infected quarters often exceed 10^6/mL (Kehrli and Shuster, 1994). Quickly after pathogens are recognized by resident cells in the udder, additional cells are transported to the udder through the blood-milk barrier. Recently it was shown that the best immune protection against an intramammary infection is provided by a fast recruitment of novel immune cells that reside in the blood. This influx of novel immune cells is indicated by a rapid rise in SCC after pathogenic invasion (Baumert et al., 2009). Due to reduced mammary gland defense mechanisms, cows are particularly susceptible to mastitis during the

periparturient period (Sordillo and Streicher, 2002). Usually SCC exceed 1 million cells/mL of milk at parturition, and in the 7 to 10 days after calving SCC decrease to 10^5 cells/mL (Rainard and Riollet, 2006). The state of host resistance is affected not only by diverse physiological stressors, but also genetics (Kehrli and Shuster, 1994). Although heritability for mastitis or SCC is usually rather low, genetic variability is present. The heritability is low, however, because a very large environmental variance typically dilutes the genetic variance. Heritability estimates for clinical mastitis mostly range from 0.02 to 0.04, for SCC[1] these estimates range from 0.05 to 0.14 for single monthly test-days and from 0.10 to 0.18 for lactation-average SCC (Rupp and Boichard, 2003). Most estimates regarding the genetic correlation between mastitis and SCC originate from Scandinavian data as, unlike most countries, Scandinavian countries routinely record clinical mastitis incidence. These data show genetic correlations that range from 0.50 to 0.80, with an average 0.70, which suggests that SCC and mastitis partly have a common genetic basis (Rupp and Boichard, 2003).

Although the heritability and genetic correlation provide insight in the genetic control of traits, they do not reveal the disposition of the underlying genes, their number or their effect on the phenotype (Glass et al., 2012).

1.4. Genome-wide association studies

A genome-wide association study may aid in obtaining insight in the disposition of the underlying genes, their number or their effect on the phenotype. The objective of genome-wide association studies is to identify associations between one or multiple genetic markers and a trait (Cordell and Clayton, 2005) using a genome-wide dense marker map. Three marker types can be distinguished: 1) the causative mutation itself; 2) linkage disequilibrium markers, these markers are in population-wide linkage disequilibrium with the causative mutation; 3) linkage equilibrium markers, these markers are in population-wide linkage equilibrium with the causative mutation but within some families they are linked to the causative mutation (Dekkers, 2004). The primary difference between linkage disequilibrium and linkage equilibrium markers is that for linkage disequilibrium markers the same allele is associated with the trait in a comparative fashion across the entire population, whereas for linkage equilibrium markers the allele can be differently associated with the trait in different families (Cordell and Clayton, 2005).

---

[1] For genetic parameter estimation SCC are commonly log-transformed into somatic cell score because of the skewness of the SCC distribution.

Frequently used genetic markers are single nucleotide polymorphisms, which are based on linkage disequilibrium with the causative mutation. Discovery of thousands of single nucleotide polymorphisms markers in the genome coupled with a substantial reduction in genotyping costs has enabled genomic selection (Hayes et al., 2009) and identification of genetic control of traits.

Genetic control of quantitative traits, such as disease resistance, lies somewhere on a spectrum, with one the one end of this spectrum the polygenic hypothesis and on the other end the monogenic hypothesis. The polygenic hypothesis assumes genetic control through many genes with relatively small effects (Detilleux, 2009), whereas the monogenic hypothesis assumes genetic control through a single gene. Most quantitative traits are likely affected by a combination of genes with major and minor effects, coupled with environmental influences (Detilleux, 2009). Genome-wide association studies provide opportunities to not only increase insight in the location of a trait on the polygenic-monogenic spectrum, but also to identify candidate genes for the genetic control. Identification of regions on the genome or, preferably, candidate genes associated with a trait is important for an increased understanding of the trait biology.

1.5. Environmental influences

Phenotypes are not only influenced by genetics, but also by environmental factors. Different genotypes may, however, show differences in response to (alterations in) environmental factors. These differences in response by different genotypes are defined as genotype by environment (GxE) interaction, and indicate that there is genetic variation in environmental sensitivity (Falconer and Mackay, 1996). Environmental factors can be classified into two categories: macro- and micro-environmental factors. Macro-environmental refers to known environmental effects that can be categorized or quantified, such as temperature, diet etc. Macro-environmental sensitivity can be expressed either as the genetic variance in the slope of a reaction norm if the environmental parameter is a continuous trait, or as genetic covariances between environments if the environmental parameter is a categorical trait, where phenotypes in the separate environments are regarded as individual traits (Mulder et al., 2013). Micro-environmental refers to unknown environmental effects (Falconer and Mackay, 1996). In contrast to macro-environmental sensitivity, micro-environmental sensitivity cannot be extracted from the residual variance due to their unknown nature. Micro-environmental sensitivity can, however, be modeled by allowing for genetic effects in the residual

variance part of the model, such as in the double hierarchical generalized linear model (Rönnegård et al., 2010).

Recently, an extension of the double hierarchical generalized linear model with a reaction norm model was developed by Mulder et al. (2013) to analyze macro- and micro-environmental sensitivity simultaneously. Simultaneous analysis of macro- and micro-environmental sensitivity provides additional insight in the interactions of traits with environmental factors. Further, macro-micro analyses may aid to optimize estimation of variance-controlling breeding values. Selection on variance-controlling breeding values is relevant when trait uniformity is desired, when interest is in traits with an intermediate optimum value (Mulder et al., 2008) or when selecting for increased robustness.

1.6. Aim and outline of this thesis

Reduction of the impact of disease in livestock production becomes increasingly important, not only from an economical and environmental perspective, but also from a human and animal welfare viewpoint. Livestock breeding provides a means to permanently decrease the impact of disease. Mastitis is a major infectious disease in dairy cattle and SCC is an indirect measure of mastitis frequently used in dairy cattle breeding schemes. Still, there is much to be learned about the interpretation of genetic parameters and the biology of disease resistance and SCC. Insight in the genetics of disease resistance can be gained by detection of genomic regions involved in existing disease resistance traits, by exploration of potential novel indicator traits and by studying the interactions of disease traits with environmental factors.

This thesis has the objective to:
1)   quantify the impact of genetics on innate immunity, represented by NAb, through estimation of heritabilities and genetic correlations;
2)   identify the genomic regions involved in SCC and NAb levels;
3)   quantify the impact of genetics on environmental sensitivity for SCC.

The outline of this thesis is as follows:

Chapter 2 covers the estimation of heritabilities for natural antibody levels in milk. In addition, genetic correlations between different natural antibody isotype levels are estimated.

In Chapter 3 a genome-wide association study is performed on the natural antibody levels for which genetic parameters are estimated in chapter 2. Chapter 4 includes a genome-wide association study on SCC, where both lactation-average SCC and

the trait standard deviation in test-day SCC are analyzed. Chapter 5 studies the genetic variation in environmental sensitivity for SCC by means of a double hierarchical generalized linear model. Finally, chapter 6 will discuss some of the biology that is behind NAb and SCC and includes the conclusions of this PhD thesis.

2

Genetic parameters for natural antibody isotype titers in milk of Dutch Holstein-Friesians

S. Wijga[1], H. Bovenhuis[1], J.W.M. Bastiaansen[1], J.A.M. van Arendonk [1], T.C.W. Ploegaert[2], E. Tijhaar[2], and J.J. van der Poel[1]

[1] Animal Breeding and Genomics Centre, Wageningen University, P.O. Box 338, 6700 AH, Wageningen, The Netherlands; [2] Cell Biology and Immunology Group, Wageningen University, P.O. Box 338, 6700 AH, Wageningen, The Netherlands

Abstract

The objective of the present study was to estimate genetic parameters for natural antibody isotypes immunoglobulin (Ig)A, IgG1 and IgM titers[1] binding the bacterial antigens lipopolysaccharide, peptidoglycan and the model antigen keyhole limpet hemocyanin in Dutch Holstein-Friesian cows (n=1,695). Further, the present study included total natural antibody titers binding the antigens mentioned above, making no isotype distinction, as well as total natural antibody titers and natural antibody isotypes IgA, IgG1 and IgM titers binding lipoteichoic acid. The study showed that natural antibody isotype titers are heritable and that these heritabilities were generally higher than heritabilities for total natural antibody titers. Genetic correlations for nearly all possible combinations of total natural antibody titers and natural antibody isotype titers were positive. Strong genetic correlations were found between IgA and IgM. Genetic correlations were substantially weaker when they involved an IgG1 titer, indicating that IgA and IgM have a common genetic basis, but that the genetic basis for IgG1 differs from that for IgA or IgM. Results from this study indicate that natural antibody isotype titers show the potential for effective genetic selection. Further, natural antibody isotypes may provide a better characterization of different elements of the immune response or immune competence. As such, natural antibody isotypes may enable more effective decisions when breeding programs would start to include innate immune parameters.

Key words: immunoglobulin, dairy cattle, innate, immune, heritability, correlation, IgA, IgG1, IgM

---

[1] A measure of antibody level

2.1 Introduction

Natural antibodies (NAb) are part of the innate immune system (Matter and Ochsenbein, 2008) and exist without apparent antigenic stimulation (Ochsenbein and Zinkernagel, 2000). They are poly-reactive and show low affinity binding with various antigens (Casali and Notkins, 1989).

Antibodies can be grouped according to two mutually exclusive important structural resemblances. On the one hand, NAb can be classified according to the structure of their variable part, i.e. the 'binding specificity' (Jerne, 1985). Natural antibodies generally bind antigens shared by classes of pathogens (Medzhitov and Janeway, 1997). Major antigens include lipopolysaccharide (LPS), part of gram negative bacteria, lipoteichoic acid (LTA), part of gram-positive bacteria, peptidoglycan (PGN), part of both gram-positive and gram-negative bacteria and the model antigen keyhole limpet hemocyanin (KLH). On the other hand, NAb can be classified according to the molecular structure of their constant part, i.e. their isotype (Jerne, 1985). Each isotype has its own biochemical and biological properties (Schroeder and Cavacini, 2010). The availability of a broad diversity of binding specificities coupled with different isotypes allows for an effective immune response against a multitude of antigens.

Natural antibody titers can be measured in bovine blood plasma and milk (Van Knegsel et al., 2007; Ploegaert et al., 2011), where they are repeatable over time and show variation among cows (Ploegaert et al., 2011). Ploegaert et al. (2010) showed that NAb titers in bovine milk are heritable, with heritabilities ranging from 0.10 to 0.53, and that positive genetic correlations exist among them. These findings, together with the immediate and early protection and poly-reactivity of NAb (Baumgarth et al., 2005) make them potentially interesting indicators that can be useful in selection for disease resistance once genetic correlations with disease resistance have been found.

The study of Ploegaert et al. (2010) estimated genetic parameters for total NAb binding KLH, LPS, LTA, or PGN, making no isotype distinction. Further, the study estimated genetic parameters for NAb isotypes immunoglobulin (Ig)A, IgG1 and IgM binding LTA. More recently, IgA, IgG1 and IgM isotype titers binding KLH, LPS or PGN became available. Isotypes may be genetically different traits and may provide biologically relevant information when compared to total NAb titers. Knowing the genetic parameters of isotype specific titers binding an antigen, in addition to total NAb binding that antigen, will among other allow more effective decisions on inclusion of NAb titers in the breeding goals of cattle. Knowledge on genetic parameters can provide more insight in the genetic control of NAb isotypes,

and as such, provide information about the relevance of measuring individual immune components. The objective of the present study was to estimate genetic parameters for NAb isotypes IgA, IgG1 and IgM binding KLH, LPS or PGN.

2.2 Materials and methods

Cows and phenotypes

The present study was based on data collected as part of the Milk Genomics initiative. Natural antibody titers were available for 1,695 cows, which were previously included in the study by Ploegaert et al. (2010). Cows were located on 380 commercial dairy herds in the Netherlands and between 66 and 263 days in lactation on the day that milk samples were taken, with an average of 166 days. Cows were at least 87.5% Holstein-Friesian and were sired by 1 of 50 young bulls (n=761), 1 of 5 proven bulls (n=764) or 1 of 45 other proven bulls (n=170), where the latter were added to the dataset to obtain at least five cows per farm. In contrast to proven bulls, young bulls were still under evaluation regarding their genetic merit. The pedigree was made available by the Dutch cattle cooperative (CRV, Arnhem, the Netherlands), and contained 26,300 individuals.

Natural antibodies binding *Megathura crenulata*-derived KLH, *Escherichia coli*-derived LPS, *Staphylococcus aureus*-derived LTA, and *S. aureus*-derived PGN were measured in morning milk samples by means of indirect ELISA as outlined by Ploegaert et al. (2010) and Ploegaert (2010b). Milk antibody levels were expressed as titers. Titers were defined as log2 values of the dilutions that gave an extinction nearest to 50% of the highest average extinction of a standard positive milk sample (Ploegaert et al., 2010). This standard positive milk sample was present in duplicate on each microtiter plate and consisted of a milk sample known to contain detectable NAb levels.

This resulted in the following 16 traits: total NAb binding KLH (KLH-T), LPS (LPS-T), LTA (LTA-T) or PGN (PGN-T), NAb isotype IgA binding KLH (KLH-IgA), LPS (LPS-IgA), LTA (LTA-IgA) or PGN (PGN-IgA), NAb isotype IgG1 binding KLH (KLH-IgG1), LPS (LPS-IgG1), LTA (LTA-IgG1) or PGN (PGN-IgG1) and NAb isotype IgM binding KLH (KLH-IgM), LPS (LPS-IgM), LTA (LTA-IgM) or PGN (PGN-IgM). The sum of IgA, IgG1 and IgM isotype titers is not equivalent to total NAb, as, among others, titers are relative measures and as such not an absolute amount of binding antibody.

Statistical analyses

Variance components for the NAb titers were estimated using a linear animal model:

$$y_{ijklmn} = \mu + b_1 \times dim_{ijklmn} + b_2 \times e^{-0.05 \times DIM_{ijklmn}} + b_3 \times afc_{ijklmn} + b_4 \times afc_{ijklmn}^2 + season_k \\ + stype_l + herd_m + animal_n + e_{ijklmn}$$ (1)

where the response variable y represented the NAb titer; μ was the general mean; dim_{ijklmn} was the covariate describing the effect of days in milk, modeled with a Wilmink curve (Wilmink, 1987) with regression coefficients $b1$ and $b2$; afc_{ijklmn} was the covariate describing the effect of age at first calving with regression coefficients $b3$ and $b4$; $season_k$ represented the fixed effect of calving season in three classes: summer (June - August 2004), autumn (September - November 2004) and winter (December 2004 - February 2005); $stype_l$ was the fixed effect of sire type in three classes: proven bull, young bull or other proven bull, and accounted for the genetic level of the sire types; $herd_m$, distributed as ~$N(0, I\sigma^2_{herd})$, was the random effect accounting for farm effects; $animal_n$, distributed as ~$N(0, A\sigma^2_a)$, was the random additive genetic effect of the cow; and e_{ijklmn}, distributed as ~$N(0, I\sigma^2_e)$, was the random residual effect. Matrix $\mathbf{I}$ represented the identity matrix, and matrix $\mathbf{A}$ contained the additive genetic relationships between animals.

Variance components for estimating the heritability and the herd variance proportion were estimated by means of univariate analyses.

The intra-herd heritability (h^2_{in}) in the narrow sense represented the proportion of phenotypic variance attributable to genetic variation among cows within a herd and was calculated as

$$h^2_{in} = \frac{\sigma^2_a}{\sigma^2_a + \sigma^2_e} ,$$ (2)

where σ^2_a was the additive genetic variance, σ^2_e was the residual variance, and where the denumerator represented the phenotypic variance (σ^2_p).

The across-herd heritability (h^2_{ac}) in the narrow sense represented the proportion of phenotypic variance attributable to genetic variation among cows between herds and was calculated as

$$h^2{}_{ac} = \frac{\sigma_a^2}{\sigma_{herd}^2 + \sigma_a^2 + \sigma_e^2} \;,\qquad(3)$$

where σ_a^2 was the additive genetic variance, σ_{herd}^2 was the variance attributable to herd, and σ_e^2 was the residual variance.

The herd variance proportion (herd^{vp}) was calculated as

$$herd^{vp} = \frac{\sigma_{herd}^2}{\sigma_a^2 + \sigma_{herd}^2 + \sigma_e^2} \;,\qquad(4)$$

where σ_a^2 was the additive genetic variance, σ_{herd}^2 was the variance attributable to herd, and σ_e^2 was the residual variance.

(Co)variance components for estimating genetic correlations and correlations attributable to herd were estimated by means of bivariate analyses.

Genetic correlations were calculated as

$$r_g = \frac{\sigma_{a1,a2}}{\sqrt{\left(\sigma_{a1}^2 \times \sigma_{a2}^2\right)}} \;,\qquad(5)$$

where σ_{a1}^2 was the additive genetic variance for trait 1, σ_{a2}^2 was the additive genetic variance for trait 2 and $\sigma_{a1,a2}$ was the genetic covariance between trait 1 and trait 2.

Correlations attributable to herd were calculated as

$$r_{herd} = \frac{\sigma_{herd1,herd2}}{\sqrt{\left(\sigma_{herd1}^2 \times \sigma_{herd2}^2\right)}} \;,\qquad(6)$$

where σ_{herd1}^2 was the variance attributable to herd for trait 1, σ_{herd2}^2 was the variance attributable to herd for trait 2 and $\sigma_{herd1,herd2}$ was the covariance attributable to herd between trait 1 and trait 2.

All analyses were performed using the ASReml software package (Version 2; Gilmour et al., 2006).

2.3 Results

Descriptive Statistics

The present study included total NAb titers as well as NAb isotype IgA, IgG1 and IgM titers binding KLH, LPS, LTA or PGN. Means and corresponding standard deviations are shown in Table 1. On average, PGN-T had the highest mean titer (6.67), and PGN-IgA had the lowest mean (0.05). When compared with isotypes IgG1 and IgM within each binding specificity NAb isotype IgA consistently had the lowest mean titer. When NAb isotype titers were compared with total NAb titers within individual binding specificities, total NAb generally had higher means, with LTA-T as an exception. Standard deviations ranged from 0.76 for KLH-T to 1.49 for KLH-IgG1. Natural antibody isotype IgG1 consistently showed the highest standard deviation, regardless of binding specificity.

Table 2.1 Mean and standard deviation (SD) for total natural antibody titers (T) and natural antibody isotype immunoglobulin (Ig)A, IgG1 and IgM titers binding keyhole limpet hemocyanin (KLH), lipopolysaccharide (LPS), lipoteichoic acid (LTA), or peptidoglycan (PGN).

Antibody[1]	Mean	SD
KLH-T	4.88	0.76
KLH-IgA	0.99	0.98
KLH-IgG1	2.96	1.49
KLH-IgM	2.31	0.99
LPS-T	4.20	0.97
LPS-IgA	1.95	1.13
LPS-IgG1	3.06	1.43
LPS-IgM	2.62	1.03
LTA-T	4.43	0.89
LTA-IgA	3.86	1.16
LTA-IgG1	4.39	1.38
LTA-IgM	5.26	0.89
PGN-T	6.67	1.04
PGN-IgA	0.05	0.96
PGN-IgG1	5.45	1.40
PGN-IgM	2.18	0.90

[1]Antibody names are a combination of antigen abbreviation and immunoglobulin isotype.

Heritabilities

Table 2 includes the estimates for the heritability and the phenotypic variance for all NAb titers studied. The following part of this study focused on the intra-herd heritability. Results for the across-herd heritability were similar to those for the intra-herd heritability, though across-herd heritabilities were somewhat lower than the intra-herd heritabilities due to the additional variance component in the denumerator. Heritabilities ranged from 0.08 for LTA-IgG1 to 0.55 for KLH-IgA. Of the different isotypes IgG1generally had high phenotypic variances and low heritabilities. In contrast to IgG1, isotype IgM had lower phenotypic variances and higher heritabilities. In general, isotype IgA had higher heritability estimates than IgG1 and IgM isotypes, except for PGN-IgM and additionally for LTA-IgM in case of the across-herd heritability. Heritabilities for LPS-T and PGN-T were low compared to heritabilities for the different isotypes and heritabilities for LTA-T and KLH-T were intermediate.

Genetic correlations

With eight exceptions, genetic correlations were positive and ranged from 0.01 to 0.99 (Table 3). A correlation was regarded strong when its absolute value was ≥ 0.70 and as weak when it was ≤ 0.35. Of the 120 correlations ~30% were weak, ~50% were strong and ~20% were intermediate between 0.35 and 0.70. All weak correlations, which included all negative correlations, involved an IgG1 titer as one of the traits. A heat map of the estimated genetic correlations between the 16 NAb traits shows that the IgG1 isotypes cluster in a separate branch, whereas the IgA and IgM isotypes cluster together in another branch (Figure 1). Within NAb binding specificity, i.e. binding KLH, LPS, LTA or PGN, correlations between NAb isotype IgA and IgM ranged from 0.87 to 0.95. The correlations between IgG1 and IgA or IgM were consistently weaker with correlations ranging from -0.23 to 0.75. Across binding specificity, isotype IgA antibody titers were strongly correlated, ranging from 0.89 to 0.98. The same was also found for isotype IgM antibody titers, with genetic correlations ranging from 0.92 to 0.96. The genetic correlations between isotype IgG1 antibodies for different specificities were substantially weaker, ranging from 0.05 to 0.92.
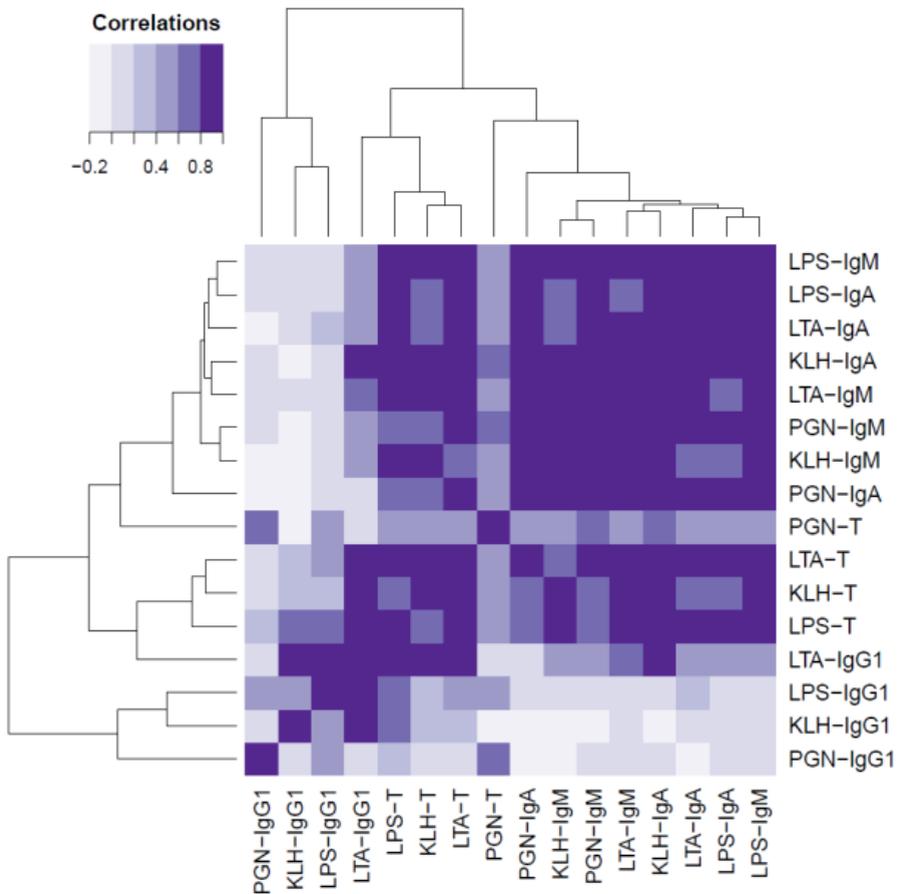
Figure 2.1 Heat map of genetic correlations between total natural antibody titers (T) and natural antibody isotype immunoglobulin (Ig)A, IgG1 and IgM titers binding keyhole limpet hemocyanin (KLH), lipopolysaccharide (LPS), lipoteichoic acid (LTA), or peptidoglycan (PGN).[1,2]

Herd variance proportions and correlations due to herd

Table 2 includes the herd variance proportion for all NAb titers studied. The herd variance proportions ranged from 0.08 for KLH-IgM to 0.37 for LTA-IgG1. Herd variance proportions were generally highest for IgG1 titers, ranging from 0.18 to 0.37, indicating that a relatively large proportion of the differences in IgG1 can be explained by differences between herds. Herd variance proportions for IgM were generally lowest ranging from 0.08 to 0.13. Herd variance proportions for isotype IgA were more variable with intermediate values for KLH-IgA and PGN-IgA, a low value for LPS-IgA and a high value for LTA-IgA. Correlations attributable to herd

were in general positive and ranged from -0.18 to 0.86 (results not shown). Of the 120 correlations attributable to herd ~70% were weak (≤ 0.35), ~5% were strong (≥ 0.70) and ~25% were intermediate between 0.35 and 0.70.

Table 2.2 Phenotypic variance $(\sigma^2_p)^1$, intra-herd $(h^2_{in})^1$ and across-herd heritability $(h^2_{ac})^1$, herd variance proportion $(herd^{vp})^1$ and corresponding standard errors in parentheses for total natural antibody titers (T) and natural antibody isotype immunoglobulin (Ig)A, IgG1 and IgM titers binding keyhole limpet hemocyanin (KLH), lipopolysaccharide (LPS), lipoteichoic acid (LTA), or peptidoglycan (PGN).

Antibody[2]	σ^2_p	h^2_{in}	h^2_{ac}	$herd^{vp}$
KLH-T	0.49 (0.02)	0.40 (0.09)	0.34 (0.08)	0.16 (0.03)
KLH-IgA	0.85 (0.05)	0.55 (0.12)	0.48 (0.10)	0.16 (0.03)
KLH-IgG1	1.72 (0.08)	0.40 (0.10)	0.32 (0.08)	0.25 (0.04)
KLH-IgM	0.92 (0.04)	0.45 (0.10)	0.41 (0.09)	0.08 (0.02)
LPS-T	0.82 (0.03)	0.15 (0.06)	0.13 (0.06)	0.13 (0.03)
LPS-IgA	1.23 (0.06)	0.52 (0.11)	0.48 (0.10)	0.08 (0.02)
LPS-IgG1	1.77 (0.08)	0.24 (0.08)	0.20 (0.07)	0.18 (0.03)
LPS-IgM	0.98 (0.05)	0.42 (0.10)	0.39 (0.09)	0.09 (0.02)
LTA-T	0.72 (0.03)	0.32 (0.09)	0.28 (0.08)	0.14 (0.03)
LTA-IgA	1.06 (0.05)	0.49 (0.11)	0.37 (0.09)	0.31 (0.04)
LTA-IgG1	1.39 (0.06)	0.08 (0.06)	0.06 (0.04)	0.37 (0.05)
LTA-IgM	0.70 (0.03)	0.44 (0.10)	0.39 (0.09)	0.13 (0.03)
PGN-T	0.78 (0.03)	0.13 (0.06)	0.10 (0.04)	0.36 (0.05)
PGN-IgA	0.81 (0.04)	0.30 (0.09)	0.26 (0.08)	0.16 (0.03)
PGN-IgG1	1.57 (0.07)	0.15 (0.07)	0.12 (0.05)	0.26 (0.04)
PGN-IgM	0.75 (0.04)	0.39 (0.10)	0.35 (0.09)	0.10 (0.03)

[1] $\sigma^2_p = \sigma^2_a + \sigma^2_e$; $h^2_{in} = \sigma^2_a/\sigma^2_p$; $h^2_{ac} = \sigma^2_a/(\sigma^2_p + \sigma^2_{herd})$; $herd^{vp} = \sigma^2_{herd}/(\sigma^2_p + \sigma^2_{herd})$, where σ^2_a was the additive genetic variance, σ^2_e was the environmental variance, and σ^2_{herd} was the variance attributable to herd.
[2] Antibody names are a combination of antigen abbreviation and immunoglobulin isotype.

Residual correlations

Residual correlations were all positive and ranged from 0.05 to 0.75 (results not shown). Of the 120 correlations attributable to residual aspects ~50% were weak, ~47% were intermediate and ~3% were strong.

Table 2.3 Genetic correlations between total natural antibody titers (T) and natural antibody isotype immunoglobulin (Ig)A, IgG1 and IgM titers binding keyhole limpet hemocyanin (KLH), lipopolysaccharide (LPS), lipoteichoic acid (LTA), or peptidoglycan (PGN).[1,2]

PGN-IgG1	KLH-IgG1	LPS-IgG1	LTA-IgG1	LPS-T	KLH-T	LTA-T	PGN-T	PGN-IgA	KLH-IgM	PGN-IgM	LTA-IgM	KLH-IgA	LTA-IgA	LPS-IgA	LPS-IgM	Trait
															1.00	LPS-IgM
														1.00	0.95	LPS-IgA
													1.00	0.89	0.82	LTA-IgA
												1.00	0.97	0.98	0.99	KLH-IgA
											1.00	0.92	0.87	0.80	0.94	LTA-IgM
										1.00	0.92	0.91	0.87	0.83	0.93	PGN-IgM
									1.00	0.96	0.95	0.93	0.80	0.80	0.96	KLH-IgM
								1.00	0.85	0.91	0.81	0.95	0.93	0.95	0.91	PGN-IgA
							1.00	0.43	0.56	0.66	0.54	0.63	0.57	0.60	0.55	PGN-T
						1.00	0.57	0.84	0.80	0.86	0.91	0.98	0.92[3]	0.85	0.81	LTA-T
					1.00	0.83	0.43	0.70	0.82	0.77	0.91	0.84	0.80	0.80	0.88	KLH-T
				1.00	0.77[3]	0.99	0.53	0.68	0.84	0.74	0.90	0.99	0.89	0.91	0.91	LPS-T
			1.00	0.91[3]	0.99[3]	0.92[3]	0.05	0.03	0.53	0.47	0.75	0.81	0.56	0.49	0.55	LTA-IgG1
		1.00	0.92[3]	0.65	0.33	0.42	0.44	0.10	0.02	0.08	0.19	0.16	0.31	0.16	0.13	LPS-IgG1
	1.00	0.56	0.91[3]	0.72	0.38	0.33	-0.11	-0.13	-0.10	-0.11	0.10	-0.06	0.12	0.10	0.07	KLH-IgG1
1.00	0.18	0.57	0.05	0.22	0.09	0.08	0.62	-0.23	-0.05	0.03	0.06	0.04	-0.08	0.01	0.04	PGN-IgG1

[1] Antibody names are a combination of antigen abbreviation and immunoglobulin isotype
[2] Standard errors ranged from 0.02 to 0.32, with an average of 0.15
[3] Genetic (co)variances liable to change from positive definite to fixed at boundary; log likelihood converged

1.4 Discussion

The present study aimed to generate insight in genetic control of NAb isotypes through estimation of heritabilities and genetic correlations. In general, heritabilities for NAb isotypes were high and higher than heritabilities for total NAb. Further, isotype IgG1 generally had the lowest heritability coupled with the highest herd variance proportions whereas isotype IgM generally had the higher heritabilities and the lowest herd variance proportions. Genetic correlations between studied NAb were nearly all positive; IgA and IgM were strongly correlated whereas the correlation of IgA or IgM with IgG1 was substantially weaker.

NAb and innate immunity

Natural antibody levels are expected to be maintained, throughout life, at relatively constant levels (Lutz and Miescher, 2008). Within and even between species, the NAb repertoire and reactivity is notably stable. This stability appears to originate from an evolutionary selection process which provides an innate legacy of specificities that enable protection against pathogens, malignant cells and other disadvantageous modifications (Vollmers and Brändlein, 2009). The present study included titers with binding specificities for LPS, LTA and PGN derived from *S. aureus* or *E. coli*, both of which are among the most prevalent causative bacteria for mastitis in dairy cattle (Barkema et al., 1998). As such, it is likely that cows have been previously exposed to LPS, LTA and PGN. The NAb titers binding these antigens may therefore, in part, reflect previous exposure and these NAb titers can consequently be regarded as a combination of authentic NAb and antibodies produced by the adaptive immune system. Although antibodies directed against LPS, LTA and PGN may, in part, have been produced under the influence of antigenic exposure, the present study, following Star et al. (2007), Van Knegsel et al. (2007) and Ploegaert et al. (2010), regarded LPS-, LTA- and PGN-binding antibodies as natural antibodies because cows were not intentionally nor controllably exposed to these antigens.

The model antigen KLH, however, is an antigen that cows do not encounter except for experimental conditions. Therefore antibodies binding KLH can be regarded as authentic NAb that reflect the ability to respond to novel antigens.

Ploegaert et al. (2011) showed that NAb in bovine milk generally give an accurate representation of serum NAb titers, herewith providing a less invasive and less laborious method of sampling to assess the NAb titers that are systemically available to these cows. Further, a high repeatability was reported for NAb titers, which suggests that a single milk sample per cow provides a representative

phenotype. The greater part of the NAb phenotypes was explained by effects of genetics and herd; effects of fixed model terms were small relative to genetics and herd.

Influence of genetics and management

Heritabilities for NAb ranged from 0.08 to 0.55 and herd variance proportion ranged from 0.08 to 0.37. Although all NAb titers are influenced by herd-environment, ~70% of the correlations attributable to herd were weak (≤ 0.35) suggesting that NAb titers are influenced by different aspects of the herd-environment. The same suggestion can be drawn from the residual correlations where ~50% of the correlations were weak.

Based on the relatively low heritability and the relatively high herd variance proportion, Ploegaert et al. (2010) suggested that LTA-IgG1 titers may depend more on environmental than genetic factors when compared to IgA and IgM titers. Results for KLH-IgG1, LPS-IgG1 and PGN-IgG1 were consistent with the findings for LTA-IgG1. Current findings are in line with Baumgarth et al. (2005), who indicated that isotype IgM was less dependent on external antigenic stimulation than IgA and especially IgG1. Thus, although the genetic composition of a cow is an important factor in antibody titers, there are also substantial herd effects, which are likely to include differences between herds in housing system, infection pressure, feed and management. This implies that even when genetic improvement programs would start to include parameters of the innate immune system, the health of dairy cattle will continue to be greatly influenced by environmental management, as was also pointed out by e.g. Burton et al. (1989).

Heritability

In addition to Ploegaert et al. (2010) the present study is, to our knowledge, the first to estimate genetic parameters for NAb titers in dairy cattle herewith making a distinction in NAb binding specificity as well as isotype. Kelm et al. (1997) and Mallard et al. (1983) studied IgM, IgG1 and IgG2 antibody isotype serum levels in non-immunized Holstein cows, making no distinction between binding specificities, and found heritabilities up to 0.74 (Kelm et al., 1997) and 0.85 (Mallard et al., 1983). In calves, the heritability of IgG1 and IgG2 serum levels upon natural exposure to nematodes ranged from 0 to 0.99 (Gasbarre et al., 1993). Genetic parameters for antibody isotypes have also shown to be heritable in other species. In Romney lambs, the heritability for total antibody levels and antibody isotype IgG1 upon natural exposure to nematodes ranged between 0.25 and 0.37 (Douch et al., 1995). Finally, Duah et al. (2009) studied antibody isotypes IgG1, IgG2, IgG3,

IgG4, IgM, IgA and IgE binding Plasmodium falciparum antigens in human serum and found heritabilities up to 0.98.

In the present study, heritabilities ranged from 0.08 to 0.55. Heritabilities of this magnitude show the potential for effective genetic selection. Heritabilities for NAb isotypes were generally equal to or higher than total NAb, although these differences were not always significant. Total NAb are a combination of different isotypes and, based on the genetic correlations found in the present study, a combination of genetically different traits. Therefore, NAb isotypes may be more specific measures that may allow for a better characterization of different aspects of the immune response or immune competence and as such allow for more effective decisions on inclusion of NAb titers in the breeding goals of dairy cattle. Combined with the higher heritabilities, NAb isotypes may be more useful in an animal breeding context than total NAb.

Genetic correlations

Natural antibody isotype titers IgA and IgM were strongly correlated, regardless of binding specificity. Correlations, however, were substantially weaker when they involved an IgG1 titer. These results suggest that IgA and IgM have a common genetic basis but that the genetic basis for NAb isotype IgG1 titers differs from that for IgA or IgM titers. The genetic distinction between IgG1 and IgA or IgM may originate from the functions these isotypes serve. Isotype IgM plays a role in the early, innate, immune response (e.g. Ehrenstein and Notley, 2010; Schroeder and Cavacini, 2010). IgA may play a critical role, mainly in mucosal tissue, in prevention of infections by bacteria or viruses, pathogenesis, or both (e.g. Woof and Kerr, 2006; Schroeder and Cavacini, 2010). Isotype IgG is mainly involved in the adaptive immune response and among other aids in neutralization of toxins and viruses (e.g. Schroeder and Cavacini, 2010). Thus, isotypes IgA and IgM play an important preventive role early in the immune response, whereas isotype IgG is of main importance in the adaptive, and frequently curative, immune response. A functional relationship between IgA and IgM was also suggested by findings of Harriman et al. (1996) who found altered immunoglobulin isotype expression in IgA deficient mice. In these mice secretion of IgM increased substantially, whereas secretion of IgG1 was also increased but to a lesser extent.

Further, the strong positive genetic correlations between IgA and IgM are supported by molecular similarities. For instance, the CH3 domains of both IgA and IgM have short tailpieces that bind with the J-chain by means of disulfide bonds (Schroeder and Cavacini, 2010).

Few studies on genetic correlations between natural antibodies have been published. A study in poultry by Siwek et al. (2006) included KLH-T, LPS-T and LTA-T. Genetic correlations were mostly positive and ranged from -0.03 to 0.92, depending on population and age. In general, genetic correlations were in line with the current findings.

The positive genetic correlations between nearly all NAb titers found in the present study indicate that genetic selection for an individual NAb titer is unlikely to be genetically detrimental to the other NAb titers.

NAb and health

Not much is known about NAb and their meaning for health. Somatic cell counts (SCC) are a commonly used measure for intramammary infections (e.g. Rupp and Boichard, 2003). Ploegaert et al. (2010) have studied the influence of elevated SCC (SCC > 80,000 cells/mL) on heritabilities for total NAb and isotypes IgA, IgG1 and IgM titers binding LTA by removing records with elevated SCC from the dataset. Similar to the study by Ploegaert et al. (2010), heritability estimates in the present study were generally not affected by the presence of elevated SCC.

Studies on the relations between NAb titers and disease status, prevalence or severity in livestock are limited, especially for NAb isotypes. Van Knegsel et al. (2007) reported a negative relation between energy balance and milk KLH-T and LPS-T titers. In poultry, Sun et al. (2011) have identified high serum KLH-IgM and KLH-IgG titers as a protective element enabling survival, where, depending on the genetic background, IgM seemed to be a stronger indicator for survival than IgG. Isotype IgA was not studied. Further, Star et al. (2007) have suggested that high KLH-T serum levels and low LPS-T serum levels are indicative of an increased probability to survive a laying period. Moreover, studies in mice have shown that NAb contribute to resistance against viral infections, bacterial infections and that they may also have a role in parasite clearance (Matter and Ochsenbein, 2008).

Although limited evidence is available on the relation between NAb titers and health, scientific results suggests that it is worthwhile to continue research in this direction.

1.5 Conclusion

This study showed that NAb isotypes IgA, IgG1 and IgM binding KLH, LPS and PGN are heritable traits and that NAb isotype titers generally are more heritable than total NAb titers. Further, strong positive genetic correlations were estimated between isotypes IgA and IgM, which indicates a common genetic basis for these isotypes. Correlations of IgA or IgM with IgG1 were considerably weaker, though

mostly positive. These findings indicate that, when compared to total NAb, isotypes may be more specific parameters that may allow for a better insight in different elements of the immune response or immune competence and, as such, allow for more effective decisions when genetic improvement programs would start to include parameters of the innate immune system. Further research is needed to assess the relations between NAb and disease resistance.

1.6 Acknowledgements

3

Genome-wide associations for natural antibodies in milk of Holstein-Friesians

S. Wijga[1], J.W.M. Bastiaansen[1], J.J. van der Poel[1], H.K. Parmentier[2], J.A.M. van Arendonk[1], and H. Bovenhuis[1]

[1] Animal Breeding and Genomics Centre, Wageningen University, P.O. Box 338, 6700 AH, Wageningen, The Netherlands; [2] Adaptation Physiology Group, Wageningen University, P.O. Box 338, 6700 AH, Wageningen, The Netherlands

Abstract

Natural antibodies (NAb), part of the innate immune system, are considered to provide a barrier at the onset of infection and to act as a stimulus for activation of the adaptive immune system. Variation in NAb levels can result from genetic or environmental factors. This study aimed to provide more insight in the largely unknown genomic basis of NAb levels in cow milk. Generally, NAb bind antigens shared by classes of pathogens. This study included NAb binding lipopolysaccharide (LPS), shared by gram-negative bacteria, lipoteichoic acid (LTA), shared by gram-positive bacteria, peptidoglycan (PGN), shared by gram-negative and gram-positive bacteria and the model antigen keyhole limpet hemocyanin (KLH). DNA and NAb levels from 1,939 cows were analyzed. Cows were genotyped for about 50,000 markers. Each individual marker was tested for detection of variation in NAb levels. Results show that information can be gained when total NAb levels are further subdivided into isotype levels, suggesting that analysis of detailed phenotypes can provide further insight in the genetic control of traits. Further, this study identified genomic regions associated with NAb levels on all bovine autosomes. A region on BTA23 was consistently found significantly associated with genetic variation in isotype IgM levels across the different binding specificities. The bovine major histocompatibility complex (MHC) is located near this region, making this a region of candidate gene(s) involved in NAb expression in dairy cows both from a functional and positional perspective.

Key words: genomic region, SNP, immunoglobulin, dairy cattle, innate, immune, IgA, IgG1, IgM

3.1 Introduction

Natural antibodies (NAb) reside in human and animal serum, colostrum or milk, in the absence of apparent external antigenic stimulation. Further, NAb are poly-reactive and show low affinity binding with various antigens (e.g. Casali and Notkins, 1989). Not only within but also between species the natural antibody repertoire and reactivity pattern is noteworthily stable. It appears that this stability evolves from an evolutionary selection process which offers an innate hereditament of binding specificities that provide protection against pathogenic agents, malignant cells and other detrimental alterations (Vollmers and Brändlein, 2009). Given this resemblance in repertoire and reactivity between species and the poly-reactivity of NAb, NAb levels are possible means to improve natural resistance in livestock which may reduce the need for antibiotics once effects on disease resistance are known.

Natural antibodies can be classified in two ways: 1) according to their binding specificity, which concerns the structure of the variable region of the antibody, and 2) according to their isotype, which concerns the constant region of the antibody (Jerne, 1985). The binding specificity refers to the antigens the antibodies bind. Generally they bind antigens shared by classes of pathogens (Medzhitov and Janeway , 1997); important antigens are lipopolysaccharide (LPS), lipoteichoic acid (LTA), peptidoglycan (PGN) and the model antigen keyhole limpet hemocyanin (KLH). The isotype of an antibody defines its biochemical and biological qualities (Schroeder and Cavacini, 2010).

It was shown that NAb of different binding specificities and isotypes are measurable in bovine milk (Ploegaert et al., 2010; Van Knegsel et al., 2007), that milk NAb levels vary between cows (Ploegaert et al., 2011) and that they are under substantial genetic control (Ploegaert et al., 2010, Wijga et al., 2013). Little, however, is known about genes underlying this genetic variation in NAb levels. Availability of dense genotypic data in the form of single nucleotide polymorphisms (SNP) makes it possible to perform a genome-wide association study (GWAS; e.g. Hirshorn and Daly, 2005). A GWAS provides opportunities to detect genomic regions involved in traits and aims at an increased understanding of the genetic control of traits, potentially leading to the identification of (candidate) genes (e.g. Pryce et al., 2010). Increased knowledge on genetic control of NAb not only gives insight in the immunological background of the trait, but such information can also be useful in selecting animals for breeding.

By means of a GWAS, we aimed to identify genomic regions associated with NAb levels measured in bovine milk. To obtain detailed insight in the genetic

background of NAb, the present study involved detailed phenotypes. Besides total NAb levels, making no isotype class distinction, also the distinct NAb isotypes immunoglobulin (Ig)A, IgG1 and IgM, each binding LPS, LTA, PGN or KLH were studied.

3.2 Materials and methods

A brief overview of the data is provided. Additional details are provided by Ploegaert et al. (2010) and Wijga et al. (2013), who used the same NAb phenotypes as used in the present study to estimate genetic parameters. Schopen et al. (2011) and Bouwman et al. (2011) used the same genotypes as used in the present study.

Cows and phenotypes

Milk samples and phenotypes for this study were collected as part of the Milk Genomics Initiative. The NAb levels were measured as titers. Phenotypes included in the present study were 1) total NAb titers, making no isotype distinction, binding the antigens KLH (KLH-T), LPS (LPS-T), LTA (LTA-T) or PGN (PGN-T) and 2) NAb isotype IgA, IgG1 and IgM titers binding KLH (KLH-IgA, KLH-IgG1 and KLH-IgM), LPS (LPS-IgA, LPS-IgG1 and LPS-IgM), LTA (LTA-IgA, LTA-IgG1 and LTA-IgM) and PGN (PGN-IgA, PGN-IgG1 and PGN-IgM). Rather than an absolute amount of binding antibody, titers are relative measures. As such, the sum of IgA, IgG1 and IgM is not equivalent to total NAb.

Titers were measured in morning milk samples using an indirect ELISA procedure (Ploegaert et al., 2010a; Ploegaert et al. 2010). Titers were available for 1,695 genotyped cows that were at least 87.5% Holstein-Friesian and that were housed in 380 commercial dairy herds across the Netherlands. Cows were on average 166 days in lactation when milk samples were taken, with a range from 66 to 263 days. The pedigree contained 26,300 animals and was made available by the Dutch cattle cooperative CRV (Arnhem, The Netherlands).

Genotypes

DNA was extracted from blood samples of cows. Blood was sampled in agreement with the directive for animal use and care as approved by the ethical committee on animal experiments of Wageningen University (protocol number 200523.b). Cows were genotyped in September 2008 for 50,905 SNP with a custom Infinium assay (Illumina, San Diego, CA, USA) designed by CRV. Quality control was applied to the SNP data from all 1,868 cows that were genotyped with a call rate ≥ 90%. A SNP was retained for analyses when the minor allele frequency (MAF) was > 1%, and when the SNP did not show a strong deviation from Hardy Weinberg equilibrium

($\chi 2$ values < 600). In total, 1,695 cows with both phenotypes and genotypes, consisting of 48,862 SNP, were retained and included in the analyses. The SNP positions were mapped according to the Baylor 4.0/bosTau4 assembly.

Statistical analyses

The association of each individual SNP with the individual NAb titers was estimated following a univariate linear animal model:

$$y_{ijklmn} = \mu + b_1 \times dim_{ijklmn} + b_2 \times e^{-0.05 \times DIM_{ijklmn}} + b_3 \times afc_{ijklmn} + b_4 \times afc^2_{ijklmn} + season_k$$
$$+ stype_l + herd_m + animal_n + e_{ijklmn}$$
, (1)

where the response variable y was the NAb titer; μ represented the general mean; dim_{ijklmn} was the covariate that defined the effect of lactation stage, modeled with a Wilmink curve (Wilmink, 1987) with regression coefficients $b1$ and $b2$; afc_{ijklmn} was the covariate that accounted for the effect of age at first calving with regression coefficients $b3$ and $b4$; $season_k$ accounted for the fixed effect of season of calving in three categories: summer (June - August 2004), autumn (September - November 2004) and winter (December 2004 - February 2005); $stype_l$ was the fixed effect of sire type in three classes: proven bull, young bull or other proven bull. The dataset consisted of five large paternal half sib families with about 200 daughters each, sired by proven bulls and 50 smaller families with about 20 daughters each, sired by young bulls. To obtain at least three cows per herd the dataset included additional cows sired by other proven bulls. Unlike proven bulls, young bulls were still under evaluation regarding their genetic merit; SNP_m referred to the fixed effect of SNP genotype; $herd_n$, distributed as $\sim N(0, I\sigma^2_{herd})$, was the random effect that accounted for farm effects where I was the identity matrix; $animal_o$, distributed as $\sim N(0, A\sigma^2_a)$, was the random additive genetic effect of the cow where A was the matrix containing the additive genetic relationships between animals based on the pedigree; and $e_{ijklmno}$, distributed as $\sim N(0, I\sigma^2_e)$, was the random residual effect.

To reduce computation time, the additive genetic variance and the herd variance were fixed at the estimate obtained from univariate analyses without the fixed effect of SNP. The additive genetic variance and herd variance were previously reported by Wijga et al. (2013). Analyses were performed using the ASReml software package (Version 2; Gilmour et al., 2006).

Multiple testing

Single SNP association studies result in multiple testing, which increases the chance for false positive results. To control the number of false positives a false discovery rate (FDR) threshold was applied. The FDR represents the expected proportion of SNP falsely denoted associated with a trait. The package "qvalue" (Story and Tibshirani, 2003) in R was used to calculate the FDR. In the present study SNP associations were considered suggestive if they surpassed an FDR of 0.20 and significant if they surpassed an FDR of 0.05. Suggestive associations were reported, significant associations were discussed in detail.

Regions

Multiple SNP in linkage disequilibrium with the causative mutation all can show an effect. Suggestive and significant associations were therefore not only reported individually, but they were also grouped into genomic regions; SNP were considered in the same genomic regions if they were less than 200kb apart. Each SNP that had a distance less than 200kb to the previous SNP was added to the same genomic region. This approach was applied to both suggestive and significant SNP. At 200kb the average linkage disequilibrium, as quantified by r2, has decreased to about 0.15 (Khatkar et al., 2008; de Roos et al., 2008).

Quality and sensitivity of association

Sensitivity of the association to a small number of observations per genotype class and to extreme phenotypic values was assessed for the significant (FDR 0.05) associations. This study applied a MAF threshold (MAF ≥ 1%), which may result in few cows in a particular genotype class. The minimal number of cows in each genotype class was set at 6. Genotype classes that consisted of 5 cows or less were removed from the data for that SNP. Subsequently, the SNP was retested to confirm the significant association. Further, for genotype classes containing relatively few animals, one or a few extreme phenotypes may have a large influence on the significance. In the present study, the sensitivity of the significant association was tested when a SNP genotype class consisted of 50 cows or less. Phenotypes were regarded as extreme when the sequential increase of the titers within a genotype class showed an interval of at least one titer and there were only one or two values above or below this interval. If extreme NAb titers were present, the SNP was retested without these extreme phenotypes to confirm the association.

Table 3.1 Chromosomal distribution of SNP that passed a false discovery rate threshold of 0.20 and 0.05 for total natural antibody titers (T) and natural antibody isotype immunoglobulin (Ig)A, IgG1 and IgM titers binding keyhole limpet hemocyanin (KLH), lipopolysaccharide (LPS), lipoteichoic acid (LTA), or peptidoglycan (PGN).[1,2,3]

BTA	KLH-T	KLH-IgG1	KLH-IgM	LPS-T	LPS-IgA	LPS-IgG1	LPS-IgM	LTA-T	LTA-IgA	LTA-IgG1	LTA-IgM	PGN-T	PGN-IgA	PGN-IgG1	PGN-IgM	All
1	8^5		2^2		2^2		2^2				3^3		2^1			19
2	30^{17}		25^{18}|**1**	1	8^2	3^3|**1**	8^7|**1**	2^1		1	14^{12}|**1**	1	1			94|**4**
3	39^{29}|**1**	4^2	16^{14}	1|1	12^{10}	3^1|**3^1**	10^7|**1**		1		2^2	1	3^2	6^5|**1**	2^2	100|**7**
4	4^3		10^8				4^4				7^6		1		1	27
5			10^5|**5^1**		2^1		5^1|**5^1**				6^2|**5^1**				6^2	29|**15**
6	3^3		3^3				3^3|**1**						4^3		1	14|**1**
7	3^3		20^{13}|**2^1**	2^2	1		11^5|**4^1**				4^3|**1**	3^2	1		2^2	44|**7**
8	12^6		14^{12}|**1**	2^1	2^2		4^3				8^7		3^2			45|**1**
9	1		15^{11}			1	4^3				4^3		3^2		6^5	34
10			3^2				2^2				4^2		1			10
11	15^8		6^3	1			3^2				3^3	1			4^4	33
12	4^3		3^3		3^2		2^2				7^5					19
13	3^2		7^7		2^1		3^2				6^4					21
14	26^{15}		17^{11}				19^{11}				22^{12}					84
15	6^6		7^5		3^1		17^9|**2^2**				1				1	35|**2**
16	3^3		8^4				1				4^3				2^1	18
17			4^4|**1**	1							6^5				2^2	13|**1**
18	2^2		10^7		2^2		5^3				10^{10}|**2^2**				7^5	36|**2**
19			4^2		2^2		2^2			3^3	7^5|**1**					18|**1**
20			5^5		1		5^5			3^2	10^5|**2^2**					24|**2**

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	Total
21	3^2	1	10^6			1	$2^2|1$				3^2			$18^6|\mathbf{14^6}$	1		$39|\mathbf{15}$
22	3^3						1		1	6^2							11
23	9^6		$86^{29}|\mathbf{17^6}$		10^5		$56^{25}|\mathbf{25^{11}}$		1		$29^{11}|\mathbf{5^2}$				$38^{19}|\mathbf{6^4}$		$229|\mathbf{53}$
24			1								3^2				1		5
25		2^2	3^3		1							2^2					8
26	45^{20}		9^3		1		$22^{12}|\mathbf{6^1}$		6^1		16^8				8^3		$107|\mathbf{6}$
27	7^5		7^3		4^2		4^3		2^1		$7^5|\mathbf{2^1}$				1		$32|\mathbf{2}$
28	4^4		8^6				4^3				1						17
29	11^8		12^7		$5^3|\mathbf{1}$		$7^7|\mathbf{1}$				$9^8|\mathbf{1}$			$1|\mathbf{1}$	2^2		$47|\mathbf{4}$
U^4	1		2				2				3				1		9
0.20^5	242^{154}	7^5	327^{197}	13^{10}	56^{36}	8^6	208^{127}	2^1	17^{10}	7^3	199^{130}	13^{11}	11^8	25^{12}	86^{53}	
0.05^6	**1**		$\mathbf{27^{11}}$	$\mathbf{2^2}$		$\mathbf{4^2}$	$\mathbf{47^{21}}$				$\mathbf{20^{12}}$			$\mathbf{16^8}$	$\mathbf{6^4}$	

[1]Antibody names are a combination of antigen abbreviation and immunoglobulin isotype.

[2]SNP that passed FDR 0.20, with regions in superscript, are reported before "|", SNP that passed FDR 0.05, with regions in superscript, are reported in bold after "|".

[3] KLH-IgA and the X chromosome were omitted from the table as no SNP passed FDR 0.20.

[4] Unassigned SNP

[5]Total number of SNP associations per trait that passed FDR 0.20, with regions in superscript. Regions do not include unassigned SNP.

[6]Total number of SNP associations per trait that passed FDR 0.05, with regions in superscript. Regions do not contain unassigned SNP. [Notes]

3.3 Results
Associations

Suggestive associations (FDR 0.20) were found for all traits, except for KLH-IgA. In total, 1,221 suggestive SNP associations were found on all bovine autosomes. Some of these associations overlapped between traits, after removal of duplicates 807 unique SNP remained. We found 123 significant associations (FDR 0.05) that involved 84 unique SNP distributed over 15 out of 29 bovine chromosomes. An overview of the suggestive and significant associations is provided in Table 1. Position, MAF, estimated effects of SNP genotype classes, and –log10 P-values of significant SNP are provided in Supplementary Table 1. Two significant SNP had a genotype class consisting of 5 animals or less and after excluding this genotype class these SNP associations were confirmed. In total, 15 significant associations involved a SNP genotype class that consisted of 50 cows or less. One of these associations, SNP ARS-BFGL-NGS-88208 with LPS-IgM, contained a genotype class consisting of 15 cows of which two had extreme phenotypes. Removal of the two extreme phenotypes caused a decrease of the –log10 P-value of association from 4.5 to 1.7, at which point this association fell below the 0.20 FDR threshold. Therefore this association was removed from the further discussion.

Total NAb titers

In total 154 genomic regions were found suggestively (FDR 0.20) associated with KLH-T, 10 with LPS-T, one with LTA-T and 11 for PGN-T. These regions contained few significant (FDR 0.05) SNP associations: for KLH-T one was found on chromosome (BTA) 3 and for LPS-T one was also found on BTA3 as well as one on BTA29 (Figure 1). No significant SNP associations were found for LTA-T and PGN-T (Figure 1).

NAb isotype IgA titers

For KLH-IgA no suggestive or significant regions were found. In total, 36 regions were found for LPS-IgA, 10 for LTA-IgA and eight for PGN-IgA. None of these regions contained significant associations (Figure 2). The smallest FDR estimates were 0.11 for LTA-IgA on BTA20, 0.13 for LPS-IgA on BTA27, 0.16 for PGN-IgA on BTA3 and 0.65 for KLH-IgA on BTA1.

Figure 3.1 –Log$_{10}$ P-values from single SNP analyses for total natural antibody binding keyhole limpet hemocyanin (KLH), lipopolysaccharide (LPS), lipoteichoic acid (LTA) and peptidoglycan (PGN). Chromosomes are arranged from left to right from chromosome 0 (unassigned SNP) to chromosome X. SNP above the black solid horizontal line passed the 0.05 false discovery threshold. For LTA-T and PGN-T no SNP surpassed the 0.05 false discovery threshold. SNP above the black dashed horizontal line surpassed the 0.20 false discovery threshold.

Figure 3.2 –Log$_{10}$ P-values from single SNP analyses for natural antibody isotype IgA binding keyhole limpet hemocyanin (KLH), lipopolysaccharide (LPS), lipoteichoic acid (LTA) and peptidoglycan (PGN). Chromosomes are arranged from left to right from chromosome 0 (unassigned SNP) to chromosome X. For all binding specificities no SNP surpassed the 0.05 false discovery threshold. SNP above the black dashed horizontal line surpassed the 0.20 false discovery threshold, for KLH-IgA no SNP surpassed the 0.05 false discovery threshold.
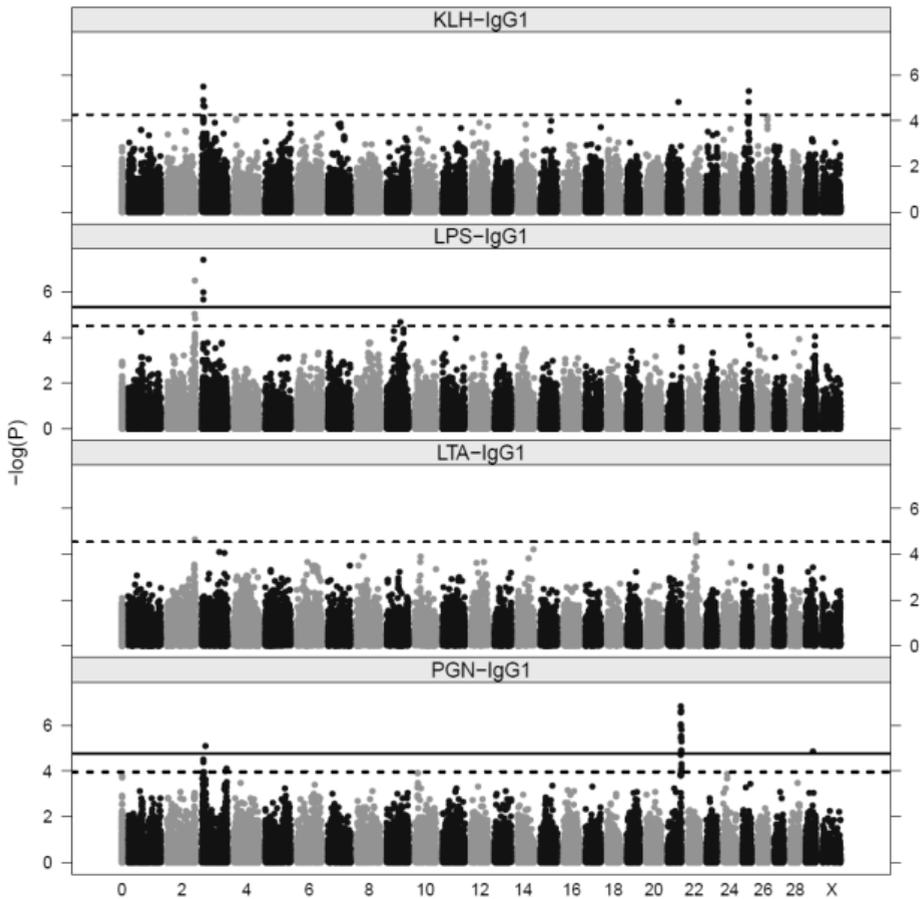
Figure 3.3 –Log$_{10}$ P-values from single SNP analyses for natural antibody isotype IgG1 binding keyhole limpet hemocyanin (KLH), lipopolysaccharide (LPS), lipoteichoic acid (LTA) and peptidoglycan (PGN). Chromosomes are arranged from left to right from chromosome 0 (unassigned SNP) to chromosome X. SNP above the black solid horizontal line surpassed the 0.05 false discovery threshold. For KLH-IgG1 and LTA-IgG1 no SNP surpassed the 0.05 false discovery threshold. SNP above the black dashed horizontal line surpassed the 0.20 false discovery threshold.
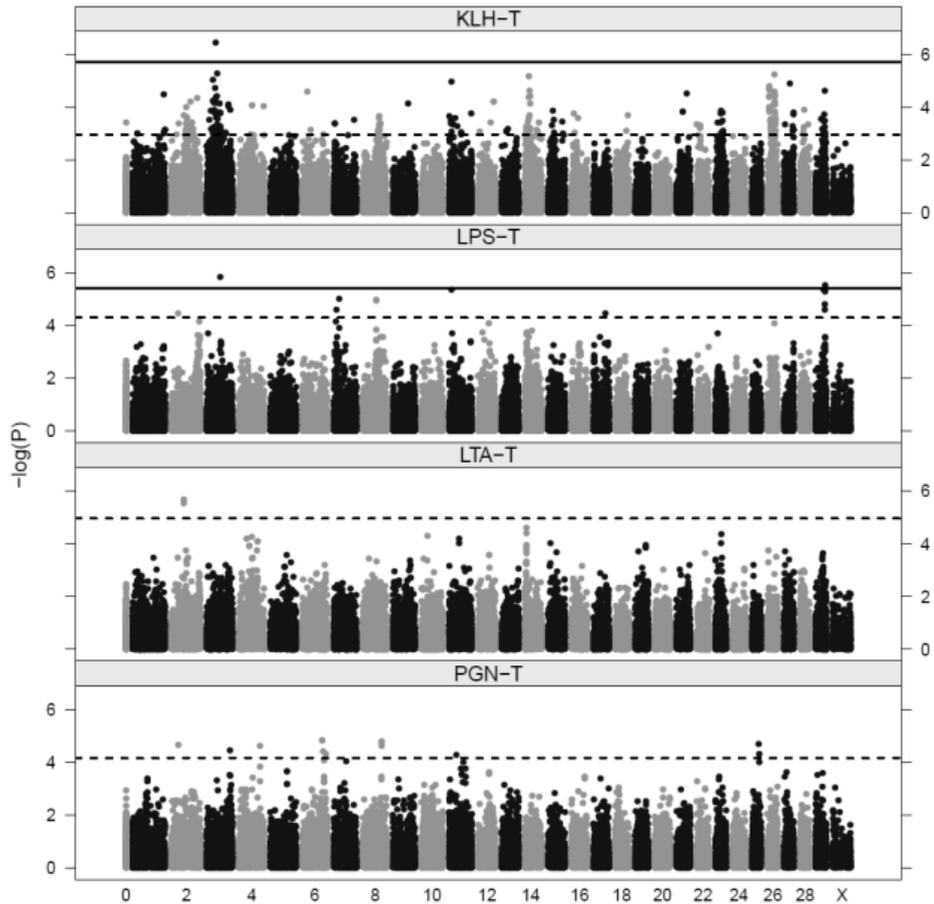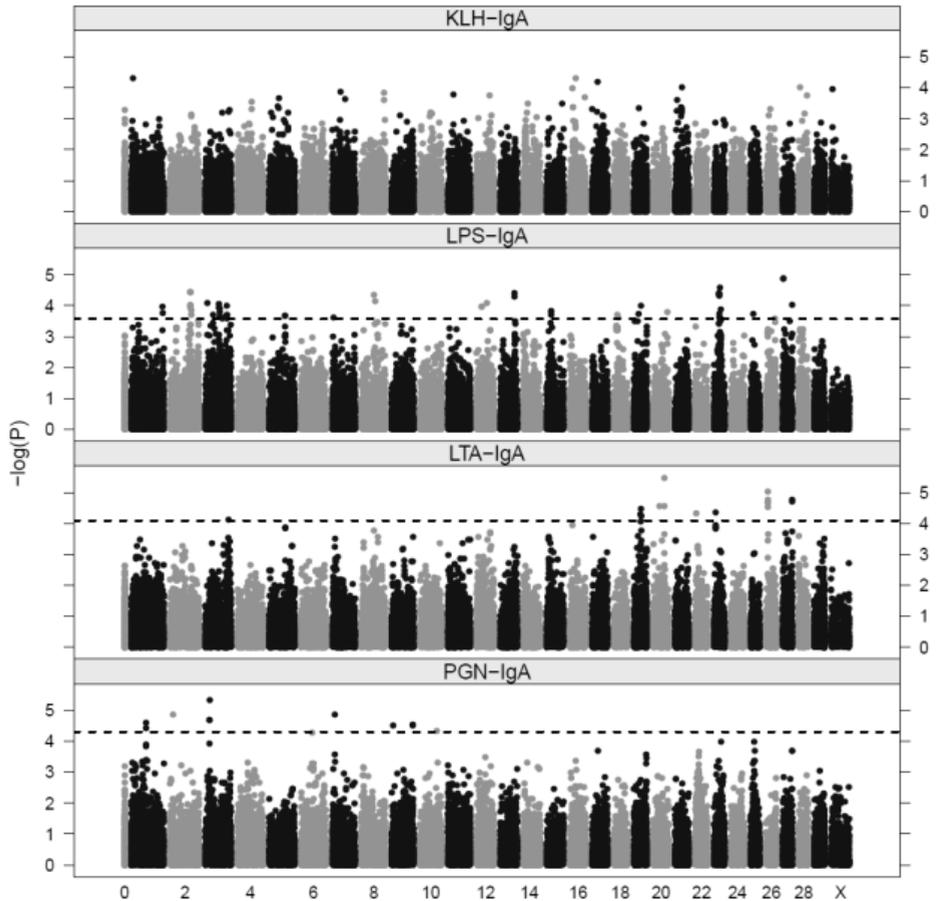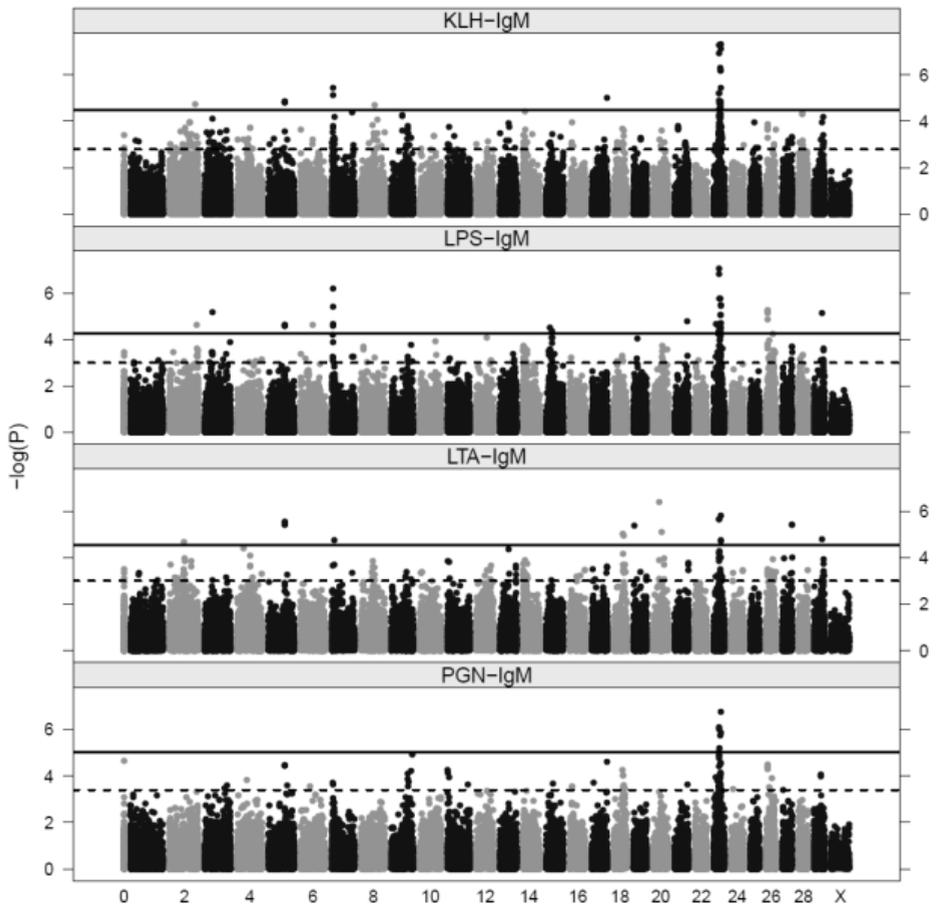
Figure 3.4 –Log$_{10}$ P-values from single SNP analyses for natural antibody isotype IgM binding keyhole limpet hemocyanin (KLH), lipopolysaccharide (LPS), lipoteichoic acid (LTA) and peptidoglycan (PGN). Chromosomes are arranged from left to right from chromosome 0 (unassigned SNP) to chromosome X. SNP above the black solid horizontal line surpassed the 0.05 false discovery threshold, SNP above the black dashed horizontal line surpassed the 0.20 false discovery threshold.

NAb isotype IgG1 titers

Five regions were detected for KLH-IgG1, six for LPS-IgG1, three for LTA-IgG1 and twelve for PGN-IgG1. Some of these regions contained significant associations. Four significant associations were detected for LPS-IgG1; one on BTA2 and three on BTA3. As the maximum distance between associations on BTA3 was about 63,000 base pairs, these three SNP were considered to be in one region.

Sixteen significant associations were detected for PGN-IgG1; one on BTA3, 14 on BTA21 and one on BTA29 (Figure 3). The 14 associations on BTA21 were located in six regions between about 65 and 69 million base pairs. No significant associations were detected for KLH-IgG1 and LTA-IgG1. For the most significant SNP an FDR of 0.11 was predicted for KLH-IgG1 and an FDR of 0.18 was predicted for LTA-IgG1.

NAb isotype IgM titers
We detected 197 regions for KLH-IgM, 127 for LPS-IgM, 130 for LTA-IgM and 53 for PGN-IgM, where each SNP that was > 200kb apart from the preceding SNP was assigned to a new genomic region. The 197 regions contained 100 significant SNP associations: 27 for KLH-IgM, 47 for LPS-IgM, 20 for LTA-IgM and six for PGN-IgM (Figure 4). The majority, 53, of these significant SNP associations were found on BTA23. For KLH-IgM additional significant associations were found on BTA2, 5, 7, 8 and 17, where the five associations found on BTA5 were located in one region. For LPS-IgM additional significant associations were found on BTA2, 3, 5, 6, 7, 15, 21, 26 and 29, where the five associations on BTA5, the four associations on BTA7 and the six associations on BTA26 were located in one region on the individual chromosomes. For LTA-IgM additional significant associations were found on BTA2, 5, 7, 18, 19, 20, 27 and 29, where the five associations on BTA5 were located in one region. For PGN-IgM no significant associations were found on chromosomes other than BTA23.

Pleiotropy
In multiple instances, a SNP significantly associated with one NAb trait was also found significantly associated with one or more additional NAb traits. Such overlapping associations were solely found for isotype IgM. In total, 100 significant SNP associations were detected for NAb isotype IgM, which, after removal of duplicates caused by pleiotropy, involved 60 unique SNP.
On BTA5 a region that contained five SNP was found associated with KLH-IgM, LPS-IgM and LTA-IgM. The maximum base pair distance between these SNP was 3,117. For PGN-IgM the −log10 P-values of this region just fell short of the 0.05 FDR threshold; for the most significant SNP an FDR of 0.07 was estimated. On BTA 7, a region that contained two SNP was found significantly associated with both KLH-IgM and LPS-IgM. Chromosome 23 contained the majority of significantly associated SNP. Four of these SNP were significantly associated with all four binding specificities, three with three binding specificities and eight with two binding specificities. Finally, one SNP on BTA29 was found associated with both LPS-IgM and LTA-IgM.

Although small differences existed, SNP effects for the overlapping SNP were in the same direction and in the same order of magnitude for the individual binding specificities (Supplementary Table 1).

3.4 Discussion

The present study aimed to identify genomic regions associated with genetic variance in NAb levels. The present study included not only total NAb levels binding KLH, LPS, LTA or PGN, making no isotype distinction, but also detailed phenotypes including NAb isotypes IgA, IgG1 and IgM binding KLH, LPS, LTA or PGN. Suggestive associations that surpassed a 0.20 FDR threshold were reported and significant SNP associations that surpassed a 0.05 FDR threshold were described in more detail. In total, the present study detected 1,221 suggestive associations of which 123, located in 61 regions, were significant. The majority of the associations was found for IgM, and most associations were located on BTA23.

Environmental influence on NAb

Antibodies are a product of innate ability and reactivity. The present study included NAb binding KLH, LPS, LTA and PGN. Except for KLH, these antigens are present on bacteria commonly present in the environment. Therefore, cows may have been exposed to these antigens prior to sampling. As such, the measured antibodies may partly find their origin in the adaptive immune response. The cows in this study, however, were not intentionally or controllably exposed and therefore the antibodies binding LPS LTA and PGN were regarded as natural antibodies which reflect the cow's natural ability to mount an immune response. Antibodies binding KLH were regarded as authentic NAb that reflect the capacity to respond to a novel antigen, as cows did not previously encounter KLH. This definition of NAb is in line with studies by Star et al. (2007) and Van Knegsel et al. (2007).

It was shown that a substantial proportion of the phenotypic variance found for NAb levels can be explained by herd effects, which are likely to reflect herd differences in among others housing system, infection pressure, feed and management (Wijga et al., 2013). These environmental effects were corrected for by including herd into the model.

Detailed phenotypes

The present study detected few significant associations for total NAb binding KLH, LPS, LTA or PGN while several were found for separate NAb isotypes, primarily for isotype IgM. The lack of associations with total NAb may be because the total is a mixture of several isotypes and these isotypes do not have a completely common

genetic basis, as indicated by non-unity genetic correlations (Wijga et al., 2013). It is expected that the differences in the genetic basis of isotypes increases the number of regions that affect total NAb compared to NAb isotypes. This increase would reduce the power to detect any of these regions, because the contribution of individual isotypes to total NAb dilutes the effect of genes that influence only one isotype.

Most associations were found for isotype IgM in each of the four binding specificities. Based on the high genetic correlation between IgM and IgA (Wijga et al., 2013) it was expected that similar associations would be found for IgA and for IgM. Although no significant associations were found for IgA, 29 suggestive associations found for IgA (out of 84) were located in regions with suggestive and significant associations for IgM. For all but one of these 29 associations the same SNP was found associated with both IgA and IgM. For IgG1 little overlap was found, not only between IgG1 and IgA or IgM, but also between IgG1 of different binding specificities. These results are in line with our earlier findings (Wijga et al., 2013) that suggested a similar genetic background for IgA and IgM, whereas the genetic background for IgG1 is different from the other isotypes.

Disease resistance
Few studies were conducted to assess the role of NAb in disease resistance. Most studies focus on intentionally induced adaptive immune responses (e.g. Thompson-Crispi et al., 2012, Rupp et al., 2007, Flori et al., 2011) or on adaptive immune products measured upon natural exposure (e.g. Gonda et al., 2006, Douch et al., 1995, White et al., 2012). There is a positive non-unity genetic correlation between antibodies produced by the adaptive immune system and NAb, which suggests differences in the genetic background of these types of antibodies (Siwek et al., 2006, Wijga et al., 2009, Wijga et al., 2013) which makes them genetically different traits.

Ploegaert et al. (2010b) studied the relation between NAb levels and the risk for mastitis and elevated somatic cell count, and found that NAb isotypes IgA and IgM may play a protective role against mastitis and elevated somatic cell count. For the levels of these NAb isotypes the current study identified a large number of associated SNP in the BoLA region which corresponds to other studies that identified the involvement of BoLA in resistance or susceptibility to disease, including mastitis (Rupp et al., 2007, Kelm et al.1997, Takeshima et al., 2008).

Further, Van Knegsel et al., (2007) showed a negative relation between energy balance and KLH-T and LPS-T levels in milk. A negative energy balance is associated with a suppressed immune function, which makes the cow more vulnerable to

metabolic and infectious diseases (Goff, 2006). In poultry, high serum KLH-IgM and KLH-IgG (Sun et al., 2011) and high KLH-T and low LPS-T (Star et al., 2007) levels seem to be associated with survival. Moreover, murine results have indicated that NAb are involved in parasite clearance and resistance to viruses and bacteria. The exact role of NAb in livestock disease resistance and survival, however, remains subject for further study.

Candidate genes

The present study used a SNP panel, rather than whole genome sequence information. As such, this SNP panel is unlikely to include causative mutations and any detected effects are assumed to be through LD. Given the bovine LD pattern (De Roos et al., 2008, Khatkar et al., 2008), a scan was applied to identify candidate genes within a 200kb window around significantly associated regions. Regions on BTA5, BTA7, BTA21, BTA23 and BTA26 are discussed because these regions contain significant associations that show pleiotropy or, in case of their involvement in a single NAb trait, multiple SNP show a significant association.

BTA5

A region with five SNP was significantly associated with KLH-IgM, LPS-IgM and LTA-IgM. The 200kb window around these SNP, which were located between 79,050,638 base pairs and 79,053,755 base pairs, included five genes: *HMGXB4*, *TOM1*, *HMOX1*, *MCM5* and *RASD2*. None of these genes have a known influence on NAb levels, but two serve functions relevant in immunity. The *TOM1* gene is of major importance in autophagy, functioning in lysosomal degradation of, among other, pathogens (Tumbarello et al., 2012). Gene products of *HMOX1* were found to be involved in inhibition of viral replication (Schmidt et al., 2012).

BTA7

A region with two SNP was significantly associated with KLH-IgM and LPS-IgM. The 200kb window around this region contained six genes: *F2RL3*, *NWD1*, *TMEM38a*, *MED26*, *CHERP*, and *CALR3*. The gene product of *F2RL3*, also known as *PAR4*, plays a role in platelet aggregation (Covic et al., 2000). Further, a GWAS for susceptibility to and control of ovine lentivirus by White et al., (2012) has previously identified *NWD1*, *TMEM38a*, *MED26* and *CHERP* as candidate genes for the provirus concentration in peripheral blood, but not for anti-ovine lentivirus antibodies. The exact function of these genes in hindering viral replication remains subject of further study (White et al., 2012). The antibodies in the study of White et al., although produced by natural exposure, are most likely part of the adaptive

immune system because proviral concentrations were present in peripheral blood in which case antibodies were probably mostly of the isotype IgG. The proviral concentrations in peripheral blood may be influenced by NAb levels, higher NAb levels may aid in lower proviral concentrations.

BTA21

On BTA21 14 SNP, located between about 66.6 and 67.6 million base pairs, were significantly associated with PGN-IgG1. The 200kb window around these 14 SNP contained eight genes: *DIO3*, *PPP2R5C*, *DYNC1H1*, *HSP90AA1*, *WDR20*, *CINP*, *TECPR2*, and *ANKRD9*. None of these genes have an obvious function in immunity. An indirect effect, however, may exist. For instance, *HSP90AA1* plays a role in signaling regulation of TGFβ, which acts in several developmental processes and a diversity of disease pathogeneses (Wrighton et al., 2008).

BTA23

Most associations were detected on BTA23 and the majority of them were for isotype IgM. An association with BTA23 was consistently found for all four binding specificities, which increases the confidence in this region. Combining the IgM with different binding specificities showed that the majority of the associations was located between 24 and 34.5 million base pairs. In total, 180 out of the 209 suggestive (FDR 0.20) SNP associations on BTA23 were located in this genomic region, including 52 out of the 53 significant (FDR 0.05) SNP associations. A major cluster of candidate genes for these associations in this region is of course the bovine leukocyte antigen (BoLA), the bovine variant of the major histocompatibility complex (MHC). Among mammalian species, the general organization of the MHC is a reasonably preserved structure of gene clusters where some loci are closely linked whereas others are fairly distant relatives (Takeshima and Aida, 2006). The 200kb window between 23.8 and 34.7 million base pairs contains the BoLA gene *BoLA-A*, a gene that belongs to the class I region (Amills et al., 1998), and *DQA5*, *DQB*, *DRA* and *DRB3*, genes of the class IIa region that belong to the DQ and DR clusters (Amills et al., 1998). Further, this window contains genes that belong to the class III region, such as *C4A* and *CYP21A2*. The MHC has previously been linked to antibody responsiveness to LPS in mice (Rodo et al., 2006), natural and adaptive serum antibody levels in chicken (Biscarini et al., 2010), serum NAb levels in cows (Dietz et al., 1997), and NAb levels in human serum (Pozsonyi et al., 2009). In humans, the major immunoregulatory MHC region, the -DR and -DQ clusters, were particularly thought to be causative for the association with NAb levels (Pozsonyi et

al., 2009). It may very well be, however, that the observed effects are caused by multiple genes or mutations rather than just one.

The MHC may influence antibody levels directly, but an indirect influence is more likely. A study by Fesel and Coutinho (1999) suggests that natural antibody repertoires, and more specifically IgM, are regulated by T lymphocytes that are restricted by genes from the MHC complex, which implies an indirect influence of the MHC. The exact mechanisms behind this regulation are not yet known (Zelenay et al., 2007). Further, *IL-17A* and *IL-17F* were located in the currently identified region. The *IL-17A* gene was previously identified as a candidate gene for NAb regulation in chicken by Biscarini et al. (2010). The role of *IL-17A* in NAb regulation is yet unknown, but the IL-17A cytokine is known to affect a variety of cells to coordinate tissue inflammation (Akdis et al., 2012). The IL-17 cytokine was, for instance, found to have a pro-inflammatory involvement in bovine mastitis (Tao and Mallard, 2007).

BTA26

On BTA26, a region consisting of six SNP was significantly associated with LPS-IgM. Further, these SNP were suggestively associated with KLH-IgM, LTA-IgA, LTA-IgM and PGN-IgM. The 200kb window, located between 11,381,155 and 11,391,593 base pairs, contained one gene, *LIPA*, which is involved in lipid metabolism (Zschenker et al., 2006). In previous studies by Bouwman et al., (2011, 2012) these SNP were found to have an effect on fatty acids 10:1, 14:1 and 16:1 in winter and summer milk samples. Previously, it was shown that lipid metabolism and immune response are strongly connected (Genini et al., 2011). During the early immune response against mastitis-causing pathogens there seems to be a deregulation of lipid metabolism; among other fatty acid metabolism was significantly affected (Genini et al., 2011).

The associations with fatty acids on BTA26 were largely attributed to SCD1 (Bouwman et al., 2011). On the one hand this makes SCD1 an additional candidate gene for the detected SNP effects, on the other hand it is possible that these SNP are in linkage disequilibrium with both LIPA and SCD1 or that the SNP effects are caused by neither LIPA or SCD1.

The current study performed a GWAS without making prior assumptions or prior selection of regions or SNP, except for the quality control prior to the GWAS. It was the first study to show by means of a GWAS that a region on BTA23 affects isotype IgM levels in bovine milk. The association of this region with isotype IgM was found for four different binding specificities, herewith increasing the confidence in this

region. The region on BTA23 contains among others genes belonging to the BoLA complex, which further strengthens findings of previous studies that identified BoLA as a major candidate gene involved in NAb levels. The present study provides a further step in unraveling the immunological background and genetic control of NAb, which is, given the resemblances in NAb between species, not only relevant for dairy cattle but also for many other species.

3.5 Conclusion

Increased resistance of pathogens against antibiotics, increased emphasis on animal welfare and the high costs associated with disease raise the importance of knowledge on the genetic basis of disease resistance. The current results show that information can be gained when total NAb levels are further subdivided into isotype levels, suggesting that analysis of detailed phenotypes can provide further insight in the genetic control of traits. Further, this study identified genomic regions associated with NAb levels on all bovine autosomes. A region on BTA23 was consistently found significantly associated with genetic variation in isotype IgM levels across the different binding specificities. Candidate genes in this region are BoLA and IL-17, from both a functional and positional perspective. These candidate genes were previously also identified for NAb levels in poultry, suggesting a possible evolutionary conserved role. Also regions on other bovine chromosomes were found to be associated with NAb levels, although their effects were not as pronounced as those found for IgM on BTA23. This suggests that NAb levels in cow milk are polygenic traits, where a region on BTA23 has an important effect on levels of isotype IgM.

3.6 Acknowledgements

Supplementary Table 3.1 Position (chromosome (BTA), base pair (BP)), minor allele frequency (MAF) with minor allele in parentheses, −log$_{10}$ P-values, and estimated effects of SNP genotype classes with standard error in parentheses of SNP that passed FDR 0.05 for total natural antibody titers (T) and natural antibody isotype immunoglobulin (Ig)A, IgG1 and IgM titers binding keyhole limpet hemocyanin (KLH), lipopolysaccharide (LPS), lipotechoic acid (LTA), or peptidoglycan (PGN)[1,2,3]. The SNP effects are shown in the order as presented in column 'Alleles'.

Trait	SNP	BTA	BP	MAF (%)	-log10 P	Alleles	Effect allele 1	Effect allele 2
KLH-IgM	ARS-BFGL-NGS-40444	2	125,843,459	31.5 (T)	4.7	CT	0.90 (1.18)	-4.61 (1.14)
KLH-IgM	ULGR_BTA-73956	5	79,050,638	43.3 (C)	4.8	CT	3.66 (0.90)	-1.29 (0.75)
KLH-IgM	ULGR_BTA-73959	5	79,051,169	43.3 (C)	4.8	CT	3.66 (0.90)	-1.29 (0.75)
KLH-IgM	ULGR_rs29017602	5	79,053,306	43.2 (T)	4.8	AT	-1.29 (0.75)	3.68 (0.90)
KLH-IgM	ULGR_rs29017601	5	79,053,570	43.2 (T)	4.8	CT	-1.29 (0.75)	3.68 (0.90)
KLH-IgM	ULGR_BTA-73955	5	79,053,755	43.3 (C)	4.8	CT	3.66 (0.90)	-1.29 (0.75)
KLH-IgM	ULGR_BTA-80302	7	6,696,843	26.8 (A)	5.4	AG	-6.68 (1.37)	-0.31 (0.67)
KLH-IgM	ULGR_BTA-94681	7	6,799,790	33.7 (G)	5.1	GT	-5.07 (1.09)	-0.64 (0.68)
KLH-IgM	ULGR_BTA-81715	8	66,637,734	6.4 (T)	4.9	CT	-4.67 (1.06)	-[3]
KLH-IgM	ULGR_rs41255651	17	74,610,587	16.4 (C)	5.0	CT	-1.31 (2.10)	3.39 (0.74)
KLH-IgM	ULGR_BTA-55976	23	25,400,136	19.4 (G)	5.2	AG	-2.43 (0.71)	4.61 (1.69)
KLH-IgM	ULGR_BTA-55977	23	25,400,154	20.9 (G)	4.9	AG	-2.07 (0.70)	4.94 (1.62)
KLH-IgM	ULGR_BTA-55986	23	25,404,687	19.4 (A)	5.2	AC	4.61 (1.69)	-2.43 (0.71)
KLH-IgM	ULGR_BTA-56015	23	25,735,114	23.1 (G)	7.3	AG	-2.62 (0.69)	5.30 (1.49)
KLH-IgM	ULGR_BTA-55999	23	25,821,757	23.2 (A)	6.9	AG	5.06 (1.50)	-2.61 (0.69)
KLH-IgM	ULGR_BTA-56062	23	27,541,568	49.4 (T)	4.8	CT	-0.75 (1.00)	3.38 (0.79)
KLH-IgM	ULGR_BTA-55829	23	28,077,464	25.0 (C)	4.5	CT	3.58 (1.41)	-2.21 (0.69)
KLH-IgM	ULGR_AAFC03027556_71020	23	32,502,273	21.4 (T)	6.3	CT	-1.23 (0.69)	7.51 (1.61)

KLH-IgM	ULGR_BTA-56309	23	33,355,369	27.0 (A)	4.7	AG	3.79 (1.26)	-1.87 (0.67)
KLH-IgM	ULGR_BTA-56314	23	33,355,913	26.0 (C)	4.9	CT	4.48 (1.327)	-1.69 (0.68)
KLH-IgM	ULGR_BTA-56316	23	33,355,951	26.0 (A)	4.8	AG	4.44 (1.33)	-1.69 (0.68)
KLH-IgM	ULGR_rs29014813	23	33,358,902	24.0 (C)	6.2	CT	5.57 (1.41)	-1.85 (0.69)
KLH-IgM	ULGR_BTA-56317	23	33,359,113	24.0 (G)	6.2	AG	-1.85 (0.69)	5.57 (1.41)
KLH-IgM	ULGR_BTA-56318	23	33,380,250	28.3 (G)	4.5	AG	-1.69 (0.67)	4.00 (1.27)
KLH-IgM	ULGR_AAFC03010340_60689	23	34,108,791	19.8 (G)	7.1	AG	-2.44 (0.72)	6.49 (1.68)
KLH-IgM	ULGR_AAFC03125015_28415	23	34,240,241	21.1 (T)	7.3	CT	-2.63 (0.72)	5.70 (1.56)
KLH-IgM	ULGR_BTA-56432	23	34,323,308	28.3 (G)	5.4	AG	-2.62 (0.70)	2.83 (1.19)
KLH-T	ULGR_BTA-91070	3	46,057,230	20.3% (T)	6.4	GT	1.87 (0.38)	-1.21 (0.90)
LPS-IgG1	ARS-BFGL-NGS-8557	2	137,307,256	32.8 (A)	6.5	AG	3.25 (1.36)	-3.47 (0.83)
LPS-IGg1	ULGR_BTA-66934	3	8,592,574	38.5 (T)	5.7	CT	-3.76 (0.86)	1.73 (1.24)
LPS-IgG1	ULGR_BTA-66937	3	8,594,930	32.2 (T)	6.0	CT	-3.91 (0.86)	1.68 (1.26)
LPS-IgG1	ULGR_BTA-66939	3	8,657,866	37.6 (A)	7.4	AG	1.28 (1.26)	-4.53 (0.86)
LPS-IgM	BFGL-NGS-112511	2	133,007,840	24.3 (A)	4.6	AG	1.70 (1.32)	-2.47 (0.61)
LPS-IgM	ULGR_BTA-67611	3	39,211,879	42.8 (T)	5.2	CT	1.33 (0.66)	-3.03 (0.79)
LPS-IgM	ULGR_BTA-73956	5	79,050,638	43.3 (C)	4.6	CT	3.57 (0.81)	-0.43 (0.68)
LPS-IgM	ULGR_BTA-73959	5	79,051,169	43.3 (C)	4.6	CT	3.57 (0.81)	-0.43 (0.68)
LPS-IgM	ULGR_rs29017602	5	79,053,306	43.2 (T)	4.6	AT	-0.43 (0.68)	3.59 (0.82)
LPS-IgM	ULGR_rs29017601	5	79,053,570	43.2 (T)	4.6	CT	-0.43 (0.68)	3.59 (0.82)
LPS-IgM	ULGR_BTA-73955	5	79,053,755	43.3 (C)	4.6	CT	3.57 (0.81)	-0.43 (0.68)
LPS-IgM	ULGR_BTA-76480	6	60,145,510	30.1 (A)	4.6	AC	0.84 (1.08)	2.86 (0.62)
LPS-IgM	ULGR_BTA-80302	7	6,696,843	26.8 (A)	6.2	AG	-6.38 (1.24)	-0.07 (0.61)

LPS-IgM	ULGR_BTA-80282	7	6,770,805	45.2 (A)	4.7	AG	-3.39 (0.88)	0.05 (0.68)
LPS-IgM	ULGR_BTA-94681	7	6,799,790	33.7 (G)	5.4	GT	-4.68 (0.99)	-0.33 (0.61)
LPS-IgM	ULGR_BTA-80228	7	6,844,949	35.6 (A)	4.6	AG	3.87 (0.96)	-0.10 (0.62)
LPS-IgM	ULGR_BTA-38017	15	19,401,333	36.0 (A)	4.5	AG	2.05 (0.95)	-1.99 (0.62)
LPS-IgM	ULGR_BTA-36271	15	28,497,636	9.9 (A)	4.4	AG	1.18 (2.83)	-3.36 (0.77)
LPS-IgM	ULGR_BTA-119906	21	61,592,669	31.8 (C)	4.8	AC	0.33 (0.62)	4.57 (1.03)
LPS-IgM	BFGL-NGS-117853	23	10,610,366	20.3 (T)	4.6	CT	2.00 (0.73)	-5.51 (1.70)
LPS-IgM	ULGR_BTA-81671	23	24,354,847	39.3 (T)	4.2	GT	-2.73 (0.92)	0.11 (0.85)
LPS-IgM	ULGR_BTA-81672	23	24,355,053	39.2 (A)	4.2	AG	0.12 (0.85)	-2.73 (0.93)
LPS-IgM	ULGR_BTA-81670	23	24,359,275	39.4 (G)	4.2	GT	0.09 (0.84)	-2.74 (0.92)
LPS-IgM	ULGR_BTA-81668	23	24,359,315	39.3 (A)	4.3	AC	0.17 (0.85)	-2.75 (0.92)
LPS-IgM	ULGR_BTA-81665	23	24,359,519	39.2 (T)	4.2	CT	-2.73 (0.93)	0.12 (0.85)
LPS-IgM	ULGR_BTA-55976	23	25,400,136	19.4 (G)	4.5	AG	-2.22 (0.64)	3.39 (1.53)
LPS-IgM	ULGR_BTA-55977	23	25,400,154	20.9 (G)	4.3	AG	-1.99 (0.64)	3.60 (1.47)
LPS-IgM	ULGR_BTA-55986	23	25,404,687	19.4 (A)	4.5	AC	3.39 (1.53)	-2.22 (0.64)
LPS-IgM	ULGR_BTA-56015	23	25,735,114	23.1 (G)	7.0	AG	-1.96 (0.63)	5.47 (1.35)
LPS-IgM	ULGR_BTA-55999	23	25,821,757	23.2 (A)	6.8	AG	5.31 (1.35)	-1.97 (0.63)
LPS-IgM	ULGR_AAFC03115293_6435	23	26,084,397	24.7 (T)	4.6	CT	-2.36 (0.63)	2.33 (1.25)
LPS-IgM	ULGR_BTA-56043	23	26,442,573	44.3 (A)	4.2	AC	3.25 (0.79)	-0.42 (0.91)
LPS-IgM	ULGR_BTA-56062	23	27,541,568	49.4 (T)	4.6	CT	-0.73 (0.91)	2.97 (0.71)
LPS-IgM	ULGR_BTA-55829	23	28,077,464	25.0 (C)	5.8	CT	3.84 (1.27)	-2.16 (0.62)
LPS-IgM	ULGR_AAFC03027556_71020	23	32,502,273	21.4 (T)	5.8	CT	-0.95 (0.63)	6.61 (1.46)
LPS-IgM	ULGR_BTA-56309	23	33,355,369	27.0 (A)	4.5	AG	3.12 (1.14)	-1.79 (0.61)

LPS-IgM	ULGR_BTA-56314	23	33,355,913	26.0 (C)	4.5	CT	3.19 (1.20)	-1.85 (0.62)
LPS-IgM	ULGR_BTA-56316	23	33,355,951	26.0 (A)	4.5	AG	3.22 (1.20)	-1.85 (0.62)
LPS-IgM	ULGR_rs29014813	23	33,358,902	24.0 (C)	5.0	CT	4.17 (1.28)	-1.73 (0.62)
LPS-IgM	ULGR_BTA-56317	23	33,359,113	24.0 (G)	5.0	AG	-1.73 (0.62)	4.17 (1.28)
LPS-IgM	ULGR_BTA-56321	23	33,379,805	31.9 (G)	4.4	AG	-1.45 (0.61)	3.19 (1.01)
LPS-IgM	ULGR_AAFC03010340_60689	23	34,108,791	19.8 (G)	5.5	AG	-1.95 (0.65)	5.08 (1.52)
LPS-IgM	ULGR_AAFC03125015_28415	23	34,240,241	21.1 (T)	5.5	CT	-2.03 (0.65)	4.51 (1.41)
LPS-IgM	ULGR_BTA-56432	23	34,323,308	28.3 (G)	4.7	AG	-1.61 (0.63)	3.46 (1.08)
LPS-IgM	ULGR_BTA-62031	26	11,181,155	27.4 (T)	5.2	CT	-2.69 (0.61)	1.60 (1.28)
LPS-IgM	ULGR_AAFC03087750_60000	26	11,182,685	30.4 (C)	4.7	CT	1.02 (1.14)	-2.64 (0.61)
LPS-IgM	ULGR_BTA-62035	26	11,184,582	30.1 (G)	5.2	AG	-2.77 (0.60)	0.99 (1.16)
LPS-IgM	ULGR_BTA-62034	26	11,184,612	30.1 (G)	5.2	AG	-2.77 (0.61)	0.93 (1.16)
LPS-IgM	ULGR_AAFC03087750_68833	26	11,191,518	30.1 (C)	5.2	CT	0.94 (1.15)	-2.77 (0.60)
LPS-IgM	ULGR_AAFC03087750_68908	26	11,191,593	30.0 (A)	5.1	AG	1.02 (1.16)	-2.73 (0.61)
LPS-IgM	ARS-USMARC-Parent-DQ404153-no-rs	29	39,386,276	40.0 (C)	5.1	CT	-1.22 (0.81)	2.72 (0.66)
LPS-T	BTB-00132025	3	68,187,867	39.2 (G)	5.8	GT	-3.48 (0.67)	-0.69 (0.50)
LPS-T	ULGR_BTA-65731	29	45,924,822	29.4 (A)	5.5	AG	-1.18 (0.91)	2.31 (0.51)
LTA-IgM	ULGR_BTA-113895	2	70,583,895	41.3 (A)	4.7	AG	-0.31 (0.63)	-1.96 (0.46)
LTA-IgM	ULGR_BTA-73956	5	79,050,638	43.3 (C)	5.5	CT	2.81 (0.58)	-0.34 (0.49)
LTA-IgM	ULGR_BTA-73959	5	79,051,169	43.3 (C)	5.5	CT	2.81 (0.58)	-0.34 (0.49)
LTA-IgM	ULGR_rs29017602	5	79,053,306	43.2 (T)	5.4	AT	-0.33 (0.49)	2.79 (0.58)
LTA-IgM	ULGR_rs29017601	5	79,053,570	43.2 (T)	5.4	CT	-0.33 (0.49)	2.79 (0.58)
LTA-IgM	ULGR_BTA-73955	5	79,053,755	43.3 (C)	5.5	CT	2.81 (0.58)	-0.34 (0.49)

LTA-IgM	ARS-BFGL-NGS-86836	7	12,133,276	15.6 (C)	4.8	CT	-0.32 (1.33)	-2.27 (0.49)
LTA-IgM	ULGR_AAFC03026780_34111	18	44,584,750	16.0 (T)	5.0	AT	-1.59 (0.47)	4.21 (1.45)
LTA-IgM	ARS-BFGL-NGS-34276	18	49,836,014	18.9 (A)	4.9	AG	0.73 (1.12)	-2.12 (0.47)
LTA-IgM	ARS-BFGL-NGS-88208	19	5,374,024	9.4 (C)	5.4	AC	-1.88 (0.58)	7.33 (2.20)
LTA-IgM	ULGR_BTA-25165	20	31,684,865	42.9 (T)	6.4	GT	-2.41 (0.47)	0.51 (0.67)
LTA-IgM	ARS-BFGL-NGS-56135	20	43,028,079	23.4 (G)	5.1	AG	-1.27 (0.44)	-4.50 (1.03)
LTA-IgM	ULGR_BTA-56015	23	25,735,114	23.1 (G)	5.6	AG	-1.55 (0.45)	2.91 (0.97)
LTA-IgM	ULGR_BTA-55999	23	25,821,757	23.2 (A)	5.6	AG	2.70 (0.97)	-1.64 (0.45)
LTA-IgM	ULGR_AAFC03010340_60689	23	34,108,791	19.8 (G)	4.7	AG	-1.56 (0.47)	2.76 (1.09)
LTA-IgM	ULGR_AAFC03125015_28415	23	34,240,241	21.1 (T)	5.8	CT	-1.51 (0.46)	3.30 (1.01)
LTA-IgM	ULGR_BTA-56432	23	34,323,308	28.3 (G)	4.8	AG	-1.50 (0.45)	1.91 (0.77)
LTA-IgM	ULGR_BTA-63143	27	46,630,117	17.4 (G)	5.4	GT	1.62 (1.28)	-2.18 (0.48)
LTA-IgM	ULGR_BTA-63141	27	46,630,287	17.4 (T)	5.4	CT	-2.18 (0.48)	1.62 (1.27)
LTA-IgM	ARS-USMARC-Parent-DQ404153-no-rs	29	39,386,276	40.0 (C)	4.8	CT	-0.59 (0.58)	1.99 (0.47)
PGN-IgG1	ULGR_BTA-67256	3	18,583,954	6.5 (A)	4.9	AG	_[3]	-4.86 (1.11)
PGN-IgG1	ULGR_BTA-53045	21	65,500,194	36.9 (A)	6.0	AT	3.65 (1.03)	-2.00 (0.74)
PGN-IgG1	ARS-BFGL-BAC-32514	21	65,896,414	47.8 (A)	5.9	AG	1.88 (0.85)	-3.15 (0.80)
PGN-IgG1	ULGR_BTA-53065	21	66,071,922	29.4 (A)	6.6	AG	2.17 (1.24)	-3.34 (0.73)
PGN-IgG1	ULGR_BTA-53067	21	66,077,073	29.4 (A)	6.6	AG	2.17 (1.24)	-3.36 (0.73)
PGN-IgG1	ULGR_BTA-53078	21	66,222,318	47.5 (T)	5.5	CT	-2.89 (0.81)	-2.06 (1.00)
PGN-IgG1	ULGR_BTA-53082	21	66,646,319	22.5 (T)	6.8	CT	3.67 (0.75)	2.14 (1.50)
PGN-IgG1	ULGR_BTA-53085	21	66,964,131	20.9 (A)	6.6	AG	4.16 (1.61)	-3.137 (0.76)
PGN-IgG1	ULGR_AAFC03034898_116623	21	66,998,226	25.6 (T)	5.5	CT	-2.624 (0.73)	3.65 (1.43)

PGN-IgG1	ULGR_AAFC03034898_116772	21	66,998,375	25.5 (A)	5.4	AG	3.64 (1.43)	-2.59 (0.73)
PGN-IgG1	ULGR_BTA-121321	21	67,644,994	22.9 (T)	4.9	GT	-2.59 (0.74)	3.50 (1.53)
PGN-IgG1	ULGR_BTA-121322	21	67,645,253	48.5 (G)	4.8	AG	-1.71 (0.83)	2.94 (1.00)
PGN-IgG1	ARS-BFGL-NGS-33371	21	68,636,089	30.5 (T)	5.3	CT	-2.62 (0.72)	2.91 (1.26)
PGN-IgG1	ULGR_rs29019410	21	68,645,328	33.1 (G)	4.9	CG	-2.57 (0.76)	2.99 (1.28)
PGN-IgG1	ULGR_rs29016374	21	68,809,305	41.4 (A)	5.8	AC	3.03 (0.94)	-2.48 (0.78)
PGN-IgG1	BTA-66489-no-rs	29	35,667,458	27.3 (C)	4.9	CT	-0.11 (1.30)	-3.35 (0.73)
PGN-IgM	ULGR_BTA-56015	23	25,735,114	23.1 (G)	6.1	AG	-2.04 (0.55)	3.58 (1.20)
PGN-IgM	ULGR_BTA-55999	23	25,821,757	23.2 (A)	6.0	AG	3.51 (1.20)	-2.05 (0.55)
PGN-IgM	ULGR_BTA-55829	23	28,077,464	25.0 (C)	5.2	CT	3.34 (1.12)	-1.76 (0.55)
PGN-IgM	ULGR_AAFC03027556_71020	23	32,502,273	21.4 (T)	5.8	CT	-1.30 (0.55)	5.19 (1.29)
PGN-IgM	ULGR_AAFC03010340_60689	23	34,108,791	19.8 (G)	5.9	AG	-1.90 (0.57)	4.39 (1.35)
PGN-IgM	ULGR_AAFC03125015_28415	23	34,240,241	21.1 (T)	6.8	CT	-1.89 (0.57)	4.64 (1.25)

[1]Antibody names are a combination of antigen abbreviation and immunoglobulin isotype.
[2]SNP effects presented correspond to the two homozygote genotype classes. The heterozygote genotype class was set to zero.
[3]Homozygote genotype class was not available for this SNP.

4

Genomic associations with somatic cell score in first lactation Holstein cows

S. Wijga[1], J. W. M. Bastiaansen[1], E. Wall[2], E. Strandberg[3], Y. de Haas[4], L. Giblin[5], and H. Bovenhuis[1]

[1] Animal Breeding and Genomics Centre, Wageningen University, P.O. Box 338, 6700 AH, Wageningen, The Netherlands; [2] Sustainable Livestock Systems Group, Scottish Agricultural College, Easter Bush, Midlothian, EH25, 9RG, United Kingdom; [3] Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, P.O. Box 7023, S-750 07 Uppsala, Sweden; [4] Animal Breeding and Genomic Centre, Wageningen UR Livestock Research, P.O. Box 65, 8200 AB, Lelystad, The Netherlands; [5] Teagasc, Moorepark Food Research Centre, Fermoy, Co. Cork, Ireland

Abstract

This genome-wide association study aimed to identify loci associated with lactation-average somatic cell score (LASCS) and the standard deviation of test-day somatic cell score (SCS-SD). It is one of the first studies to combine detailed phenotypic and genotypic cow data from research dairy herds located in different countries. The combined dataset contained up to 52 individual test-days per lactation and thereby aimed to capture temporary increases in somatic cell score associated with infection. Phenotypic data for analysis consisted of 46,882 test-day records on 1,484 cows and genotypic data consisted of 37,590 single nucleotide polymorphisms (SNP). Using an animal model, the association between each individual SNP and the phenotypic data was estimated. To account for the risk of false positives, a false discovery rate threshold of 0.20 was set. The analyses showed that LASCS significantly associated with a SNP on *Bos taurus* autosome (BTA) 4 and a SNP on BTA18. Likewise SCS-SD associated with this SNP on BTA18. In addition, SCS-SD significantly associated with a SNP on BTA6. Relatively few associations were found, suggesting that LASCS and SCS-SD are controlled by multiple loci, each with a relatively small effect, distributed across the genome. Increased knowledge on genetic regulation of LASCS and SCS-SD may aid in identification of genes that play a role in mastitis resistance. Such knowledge aids in understanding the genetic mechanisms leading to mastitis and in discovery of targets for mastitis therapeutics.

Key words: genome wide, dairy cow, single nucleotide polymorphism, mastitis

4.1 Introduction

In the last decade, mastitis resistance has become an important breeding objective in dairy cattle from an economical and animal welfare perspective (Miglior et al., 2005; Rupp et al., 2007). The SCC is the commonly used indirect measure of mastitis (e.g. Rupp and Boichard, 2003; Detilleux, 2009). Although there is relatively large genetic variation for SCC (Rupp and Boichard, 2003), heritabilities for SCC are low to moderate (Rupp and Boichard, 2003) and a genetic antagonism exists between SCC and production traits (e.g. Rupp and Boichard, 1999; Carlén et al., 2004). As such and combined with the fact that SCC can only be measured on cows and not on bulls, selection for mastitis resistance could benefit from genomic information. Genomic information may be particularly useful for increasing the accuracy of the estimated breeding values of bull dams, which is currently based on own performance.

The recent discovery of millions of SNP in livestock genomes (Matukumalli et al., 2009), forming dense marker maps, and a concurrent strong reduction in genotyping costs (Daetwyler, 2008; Hayes et al., 2009) have created opportunities for the use of genomic information, allowing for genome-wide association studies (GWAS) (Hirschorn and Daly, 2005). Results from GWAS may not only identify markers that enable more accurate breeding value estimation but may also aid in the understanding of genetic control of traits through identification of genes (e.g. Pryce et al., 2010). So far, GWAS reported for dairy cattle are limited and GWAS that are reported have been performed on traits routinely recorded in commercial dairy herds, using daughter yield deviations or estimated breeding values of progeny tested bulls. The present study used combined phenotypic and genotypic data of first lactation Holstein cows from four different European research herds. These herds provide more detailed SCC recordings, with up to 52 individual test-days per lactation, compared to commercial herds where SCC is usually recorded monthly. These detailed recordings provide a more comprehensive representation of the phenotype. In addition, the use of genotypic data on cows allows for estimation of dominance effects, which is not possible when using daughter yield deviations.

Generally, genetic selection is based on lactation-average SCC. Lactation-average SCC, however, does not capture variation in SCC levels during lactation (De Haas et al., 2003). As suggested by Urioste et al. (2010) and Green et al. (2004), the standard deviation of test-day SCC largely reflects this variation and aims to capture temporary increases in SCC associated with infection. The standard deviation of test-day SCC is genetically variable and strongly associated with clinical

mastitis (CM) (Urioste et al., 2010). Therefore it is an interesting candidate in selection for mastitis resistance (Urioste et al., 2010).

The present study aimed to identify loci associated with lactation-average SCS and the standard deviation of test-day SCS through a GWAS using combined phenotypic and genotypic data of first lactation Holstein cows from four different European research herds.

4.2 Materials and methods

The present study was part of the EU funded RobustMilk project, a collaboration between six European research institutes (www.robustmilk.eu). The RobustMilk project combined unique phenotypic and genotypic data from eight research dairy herds located in four different European countries: three herds in Ireland (McCarthy et al., 2007), two herds in the Netherlands (Veerkamp et al., 2000), two herds in Scotland (Bell et al., 2011) and one herd in Sweden (Petersson et al., 2006). The Scottish cows belonged to two genetic lines (Veerkamp et al., 1994) and were therefore treated as two separate herds. At the time of data recording, some cows included in the RobustMilk project were subjected to dietary treatments as part of other studies. Data were recorded between October 1989 and September 2009.

Genotypes

DNA was extracted from blood samples. Cows were subsequently genotyped for 54,001 SNP by a commercial genotyping company (ServiceXS, Leiden, the Netherlands) using the Illumina BovineSNP50 BeadChip (Illumina Inc., San Diego, CA). The SNP genotypes were scored using Illumina BeadStudio software (v3.3.4). Quality control was performed on the genotypic data of the separate countries, using criteria set by Hayes et al. (2009) with minor modifications. A SNP was included in the dataset when the following criteria were met: 1) the minor allele frequency (MAF) was > 1% in each country; 2) the percentage of missing genotypes for the SNP across all samples was < 5%; 3) the Gen Train score (statistical score for accuracy of clustering) was > 0.55 and the Gen Call score (statistical score for genotyping accuracy) was > 0.20; and 4) the SNP did not deviate strongly from Hardy Weinberg equilibrium (Hardy Weinberg $\chi2$ values < 600). A SNP that failed a criterion in at least one country was discarded from the complete dataset. Furthermore, animals with SNP call-rates < 95% were removed from the dataset (n = 70). A total of 37,590 SNP were retained and were thus available for analyses. These SNP were distributed over the *Bos taurus* genome as shown in Table 1.

All animals genotyped within the RobustMilk project were checked for pedigree inconsistencies using the methodology outlined by Calus et al. (2011).

Table 4.1 Distribution of SNP available for analyses over *Bos taurus* chromosomes (Chr) and the size of the individual chromosomes (Size) in million base pairs (Mbp)[1]**.**

Chr.	Size (Mbp)	No. SNP	Chr.	Size (Mbp)	No. SNP
NULL	Unassigned SNP	1,089	16	77.9	1,135
1	161.1	2,302	17	76.5	1,107
2	140.8	1,911	18	66.1	927
3	127.9	1,810	19	65.3	972
4	124.5	1,775	20	75.8	1,085
5	125.8	1,498	21	69.1	968
6	122.6	1,771	22	61.8	942
7	112.1	1,547	23	53.4	761
8	116.9	1,643	24	65.3	885
9	108.1	1,452	25	44.1	709
10	106.4	1,542	26	51.8	743
11	110.2	1,587	27	48.7	692
12	85.4	1,170	28	46.1	647
13	84.4	1,275	29	60.0	773
14	81.3	1,189	X	88.5	498
15	84.6	1,185		Total No. SNP	37,590

[1]Based on Baylor 4.0/bosTau4 assembly

Animals and phenotypes

Phenotypic data was compiled from cows that passed the genotype quality control and the pedigree check. This dataset consisted of 81,408 first lactation test-day records on 1,816 Holstein cows. Test-day records with SCC greater than zero recorded before 350 DIM were retained. Cows that had less than ten SCC test-day records were discarded in order to have sufficient records for estimating mean and standard deviations of SCC.

Therefore, the edited phenotypic data used for the association study consisted of 46,882 first lactation test-day records on 1,484 genotyped Holstein cows located in four countries: Ireland (n = 329 cows; 8,795 test-day records), the Netherlands (n = 574 cows; 17,024 test-day records), Scotland (n = 390 cows; 13,312 test-day records) and Sweden (n = 191 cows; 7,751 test-day records).

Lactation-average SCC and test-day SCC standard deviation were calculated for each cow based on her test-day records. The number of test-day records per cow ranged between ten and 52 with an average of 31 test-days. Lactation-average SCC was log-transformed (Ali and Shook, 1980) to lactation-average SCS (LASCS), where

$$LASCS = \log_2\left(\frac{\sum SCC}{n}\right),$$ (1)

Test-day SCC standard deviation was log-transformed in a similar manner into test-day SCS standard deviation (SCS-SD)

$$SCS - SD = \log_2\left(\sqrt{\frac{\sum(SCC - \mu)^2}{n-1}}\right),$$ (2)

where μ was the mean test-day SCC and n was the number of test-day records.

In summary, the data for analyses consisted of 46,882 SCC test-day records on 1,484 first lactation Holstein cows. For all cows genotypic information consisted of 37,590 SNP.

The approval of the Dutch Animal Care and Use committee was not needed, as data used in the present study were obtained from existing databases.

Statistical Analyses

Genetic and residual variances for LASCS and SCS-SD were estimated with a bivariate linear animal model using the ASReml software package (version 3; Gilmour et al., 2009). Data were analyzed with the following model

$$y_{ijk} = \mu + CHYST_i + Byear_j + Animal_k + e_{ijk},$$ (3)

where y_{ijk} is the response variable corresponding to LASCS and SCS-SD of cow k from CHYST-group i born in year j; μ is the overall mean; $CHYST_i$ was a fixed effect accounting for the combination i of country (C) and herd (H) in which the record was produced, year (Y) and season (S) of calving of the cow producing the record and dietary treatment (T) the cow received during lactation (i = 1 to 146). Seasons were defined as calendar quarters (January to March, April to June, July to September and October to December); CHYST groups containing <5 individuals were merged with adjacent CHYST groups; $Byear_j$ was a fixed effect accounting for the year of birth j of the cow (j = 1 to 23); $Animal_k$, was the random additive genetic effect of animal k distributed as ~N(0, $\mathbf{A}\sigma^2_a$) which accounted for (co)variances between animals due to genetic relationships by formation of an $\mathbf{A}$-matrix based on pedigree records; the pedigree consisted of 9,368 individuals over 19 generations;

e_{ijk}, distributed as ~N(0, $\mathbf{I}\sigma^2_e$) was the random residual effect, where $\mathbf{I}$ was the identity matrix.

Heritabilities were calculated as

$$h^2 = \frac{\sigma^2_a}{\sigma^2_a + \sigma^2_e},$$ (4)

where σ^2_a represented the additive genetic variance and σ^2_e the residual variance.

The genetic correlation was calculated as

$$r_g = \frac{\sigma_{aLASCS,aSCS-SD}}{\sqrt{\left(\sigma^2_{aLASCS} \times \sigma^2_{aSCS-SD}\right)}},$$

(5)

where σ^2_{aLASCS} was the additive genetic variance for LASCS, $\sigma^2_{aSCS-SD}$ was the additive genetic variance for SCS-SD and $\sigma_{aLASCS,aSCS-SD}$ was the additive genetic covariance between LASCS and SCS-SD.

The association of each individual SNP with the phenotype was estimated from a univariate linear animal model using ASReml (version 3; Gilmour et al., 2009). The data were analyzed following model 3, with the addition of a fixed single SNP effect to the model. The consequence of this analysis is that SNP in linkage disequilibrium with the causative mutation will all show an effect. The heritabilities for LASCS and SCS-SD were fixed at the estimate obtained from the preceding bivariate analyses, for ease of computation which is valid when SNP effects are relatively small.
To assess the effect of population substructure in the present dataset, the single SNP analysis was also performed without the random additive genetic effect of the animal included in the model.

SNP variances were calculated based on the genotype frequencies and the estimated genotype effects.

Multiple Testing

In the single SNP association analyses the significance of each individual SNP was tested, resulting in a multiple testing problem. Multiple testing increases the risk of false positives and to account for this, a false discovery rate (FDR) was used. The P-value threshold corresponding to an FDR of 0.20 was calculated, based on the P-value for each SNP obtained from the single SNP analysis using the package ´qvalue´ in the statistical environment R (Storey and Tibshirani, 2003). The SNP with P-values below or equal to the P-value corresponding to the FDR were considered significantly associated with the phenotype.

4.3 Results

Descriptive Statistics and Genetic Parameters

Mean LASCS, mean SCS-SD and the corresponding standard deviation for the combined dataset as well as the individual countries within the RobustMilk dataset are shown in Table 2. Standard deviations in the separate countries were within 30% of the mean LASCS and SCS-SD of the combined data.

The heritability estimated for LASCS was 0.17 (SE = 0.06) with a phenotypic variance of 1.62 and the heritability estimated for SCS-SD was 0.14 (SE = 0.06) with a phenotypic variance of 3.2. The phenotypic correlation between LASCS and SCS-SD was 0.89 (SE = 0.01) and the genetic correlation between LASCS and SCS-SD was estimated at 0.96 (SE = 0.04).

Table 4.2 Mean, standard deviation (SD) and number of animals (N) for lactation-average somatic cell score (LASCS) and the standard deviation for test-day somatic cell score (SCS-SD) for the combined data (RobustMilk) and for each separate country of origin of the cows included in the data.

Data	N	LASCS		SCS-SD	
		Mean	SD	Mean	SD
RobustMilk	1484	6.4	1.3	6.3	1.9
Ireland	329	6.7	1.3	6.7	1.7
The Netherlands	574	6.3	1.4	6.2	1.9
Scotland	390	6.0	1.2	5.9	1.9
Sweden	191	6.8	1.3	6.9	1.8

Association of SNP with LASCS

The GWAS showed a significant association with LASCS on BTA4 and on BTA18 (Figure 1). On BTA4, SNP BTB-01841922 (ss64858711) showed a significant effect with a –log10 P-value of 5.1. On BTA18, SNP ARS-BFGL-NGS-101491 (ss86330740) showed a significant effect with a –log10 P-value of 5.6.

For SNP BTB-01841922, the genotype class with the smallest number of observations (AA) consisted of 13 cows of which 12 were from Ireland. The MAF for this SNP was 0.11. As AA animals came primarily from one country, the sensitivity of the analysis was tested by omitting cows belonging to the AA genotype class. This retest resulted in an increase in –log10 P-value from 5.1 to 5.7. –Log10 P-values and the genetic variance explained by the two significantly associated SNP are given in Table 3. Estimated SNP effects for the LASCS associated SNP in the combined RobustMilk dataset as well as in the separate populations are shown in Table 4. In general, estimated SNP effects in individual countries were in the same direction for each genotype class as the combined RobustMilk data, although effect sizes differed. These differences were, however, not significant ($\alpha = 0.10$). For SNP BTB-01841922 on BTA4, the homozygote genotype class with the largest number of observations (BB) consistently resulted in a lower LASCS than the heterozygote genotype class. Analysis of the untransformed SCC data showed that the SNP effect of -0.44 for this homozygote genotype class corresponded to a lower LASCS of approximately 51,000 cells/mL compared to the heterozygote genotype class. For SNP ARS-BFGL-NGS-101491 on BTA18, the homozygous genotype class with the smallest number of observations (AA) resulted in the highest LASCS whereas the homozygous genotype class (BB) with the largest number of observations resulted in the lowest LASCS. Ireland and Scotland showed a slight deviation from this pattern (Table 4), but these deviations were not significant ($\alpha = 0.10$). Analyses of the untransformed combined RobustMilk data showed that the implication of the AA genotype for this SNP was an additional 87,000 cells/mL compared to the BB genotype.

The MAF for BTB-01841922 was 0.11 in the total dataset and remained relatively stable over birth years 1991 to 2007. The minor allele, allele A, was the unfavorable allele, resulting in an increased LASCS. The MAF for ARS-BFGL-NGS-101491, allele A, was 0.27 and fluctuated over birth years from a minimum of 0.05 to a maximum of 0.34. Also for this SNP the minor allele was the unfavorable allele.

Association of SNP with SCS-SD

The GWAS for SCS-SD showed two SNP (Figure 2) with an FDR ≤ 0.20. The SNP ARS-BFGL-NGS-101491 on BTA18, previously associated with LASCS, had a –log10 P-

value of 5.8 for SCS-SD. On BTA6, SNP BTB-02087354 (ss65101233) had a–log10 P-value of 5.3. As part of the genotype quality control, SNP with a MAF ≤ 1% were removed from the dataset. For SNP BTB-02087354, however, this could not prevent a genotype class (BB) with only one cow. For this reason, the analysis was repeated without this cow. As a result the –log10 P-value of association increased from 5.3 to 6.1. –Log10 P-values and the genetic variance explained by the two significantly associated SNP are given in Table 3. Estimated SNP effects for the SCS-SD associated SNP in the total RobustMilk dataset as well as in the separate populations are shown in Table 5. For SNP BTB-02087354 on BTA6 the heterozygous genotype class resulted in the highest SCS-SD compared to the homozygous genotype class, with the exception of Sweden. Analysis of the untransformed SCC data showed that SCS-SD was approximately 235,000 cells/mL higher in the heterozygous genotype class compared to the homozygous genotype class. For SNP ARS-BFGL-NGS-101491 on BTA18, the smallest homozygote genotype class (AA) resulted in the highest SCS-SD and the largest homozygote genotype class (BB) resulted in the lowest SCS-SD, with the exception of Scotland where the heterozygous genotype class resulted in the highest SCS-SD. Analyses of the untransformed SCC data, using the combined RobustMilk data, showed that the AA genotype resulted in a 197,000 cells/mL increase in SCS-SD compared to the BB genotype.

The MAF for SNP BTB-02087354 was 0.02 and remained stable over the birth years tested.

Table 4.3 SNP significantly associated (FDR < 0.20) with lactation-average somatic cell score (LASCS) and the standard deviation for test-day somatic cell score (SCS-SD), their location on the *Bos taurus* autosome (BTA), $-\log_{10}$ P-value ($-\log_{10}$ P) and the proportion of genetic variance explained by the SNP (Var).

BTA	SNP	bp	Trait	$-\log_{10}$ P	Var
4	BTB-01841922	11,369,399	LASCS	5.1	0.11
6[1]	BTB-02087354	8,658,351	SCS-SD	6.1	0.14
18	ARS-BFGL-NGS-101491	12,785,002	SCS-SD	5.8	0.16
18	ARS-BFGL-NGS-101491	12,785,002	LASCS	5.6	0.15

[1]SNP consisted of two genotype classes

Table 4.4 Effect sizes per genotype class for SNP BTB-01841922 and ARS-BFGL-NGS-101491 significantly associated with lactation-average somatic cell score with the corresponding standard error in parentheses for the combined RobustMilk data and for each separate country of origin**.**

Data	BTB-01841922			ARS-BFGL-NGS-101491		
	AA	AB	BB	AA	AB	BB
RobustMilk	-0.45 (0.37)	0	-0.44 (0.09)	0.29 (0.15)	0	-0.32 (0.08)
Ireland	0.19 (0.37)	0	-0.18 (0.15)	0.44 (0.26)	0	0.27 (0.16)
The Netherlands	-[1]	0	-0.64 (0.15)	0.41 (0.26)	0	-0.30 (0.26)
Scotland	-[2]	0	-0.34	-0.05 (0.30)	0	-0.19 (0.14)
Sweden	-[2]	0	-0.28	0.24 (0.49)	0	-0.47 (0.48)

[1] Genotype class consisted of one cow
[2] No cows were present in this genotype class

Table 4.5 Effect sizes per genotype class for SNP BTB-02087354 and ARS-BFGL-NGS-101491 significantly associated with the standard deviation for test-day somatic cell score with the corresponding standard errors in parentheses for the combined RobustMilk data and for each separate country of origin.

Data	ARS-BFGL-NGS-101491			BTB-02087354[1]	
	AA	AB	BB	AA	AB
RobustMilk	0.50 (0.21)	0	-0.42 (0.11)	0	1.20 (0.24)
Ireland	0.77 (0.34)	0	-0.30 (0.20)	0	1.65 (0.37)
The Netherlands	0.82 (0.36)	0	-0.35 (0.36)	0	1.45 (0.45)
Scotland	-0.19 (0.45)	0	-0.32 (0.21)	0	0.50 (0.57)
Sweden	0.10 (0.66)	0	-0.77 (0.65)	0	-0.14 (0.70)

[1] SNP consisted of two genotype classes

Polygenic Component

All analyses in the present study included a polygenic component to account for family relationships between cows and population substructure caused by these relationships. In contrast to the LASCS analysis including a polygenic component (Figure 3b), a quantile-quantile plot for the LASCS analysis without a polygenic component showed many spurious SNP effects as deviations from the expected null distribution of the P-values and did therefore show an effect of population substructure (Figure 3a). The same was true for SCS-SD (results not shown).
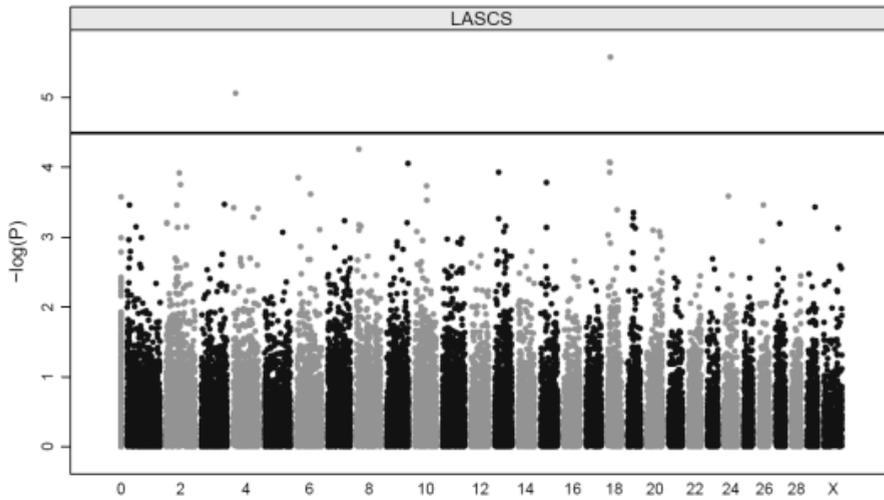
Figure 4.1 –$\log_{10}$ P-values from single SNP analysis for lactation-average somatic cell score (LASCS). Chromosomes are arranged from left to right from chromosome 0 (unassigned SNP) to chromosome X. SNP above the black horizontal line passed the 0.20 false discovery rate threshold.

Figure 4.2 –$\log_{10}$ P-values from single SNP analysis for test-day somatic cell score standard deviation (SCS-SD). Chromosomes are arranged from left to right from chromosome 0 (unassigned SNP) to chromosome X. SNP above the black horizontal line passed the 0.20 false discovery rate threshold.
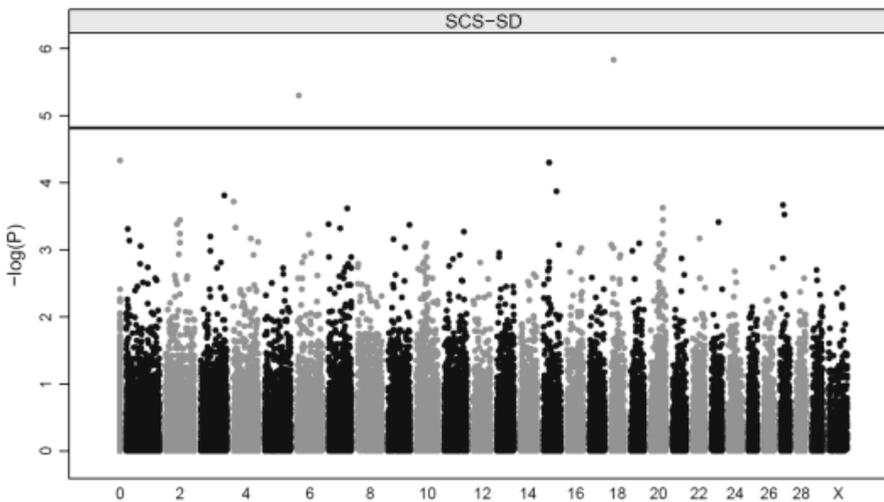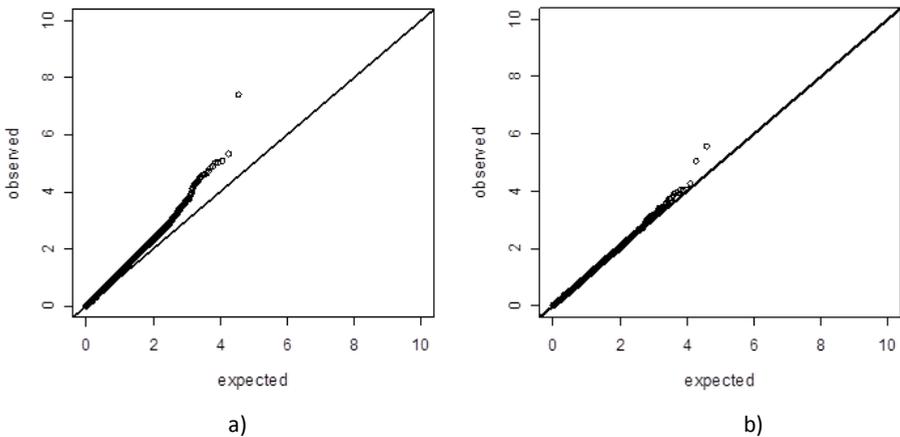
Figure 4.3 Quantile-quantile plot of the expected null distribution of the P-values versus the observed null distribution of the P-values for the lactation-average somatic cell score where a) relationships between animals were not accounted for in the model and b) relationships between animals were accounted for in the model.

4.4 Discussion

The objective of this GWAS was to identify loci associated with LASCS and SCS-SD using cow data. Few countries routinely record mastitis events, hindering direct selection for mastitis resistance (Rupp and Boichard, 1999). Instead the common practice is to use SCC or SCS as an indirect selection for mastitis (e.g. Rupp and Boichard, 2003). In commercial dairy herds, SCC is generally recorded monthly resulting in approximately 12 SCC test-day records per cow. Monthly SCC recordings may not always detect elevated somatic cells due to CM (Rupp and Boichard, 1999), because a mastitis infection can occur and the animal restored to health within a one month period (Vaarst and Enevoldsen, 1997). In the present study, the number of SCC test-day records per cow ranged from 10 to 52 with an average of 31 SCC records. This increases the likelihood that elevations in SCC caused by infections of short duration, such as CM caused by *Escherichia coli* (De Haas et al., 2002; Burvenich et al., 2003; Bannerman et al., 2004) were included.

LASCS and SCS-SD

The LASCS is often calculated as the average of the log-transformed test-day SCC (Mark et al., 2002). Our preference, however, was to average test-day SCC first and then log-transform the data, as the influence of elevated test-day SCC on lactation-average SCC and thus on LASCS is larger by this method (De Haas et al., 2008). Moreover, the genetic correlation between the log-transformed average of test-

day SCC and CM was found to be higher than the genetic correlation between the traditionally used average of the log-transformed test-day SCC and CM (De Haas et al., 2008). The SCS-SD was also chosen as a trait for analysis to capture variation in SCC levels of individual cows. For instance, cows can have the same lactation-average SCC but very different patterns of variation around the mean because of differences in infection status. Urioste et al. (2010) found that in general, cows without any CM episode tend to have a relatively low variation in SCC and that these animals show faster recovery when infected.

Genetic Parameters

The estimated heritabilities using the RobustMilk dataset were in good agreement with previously reported estimates. For example, the heritability of 0.17 for first lactation LASCS was also reported previously by Rupp and Boichard (1999) and Mrode and Swanson (2003). Heritabilities reported in other studies are of the same approximate magnitude (e.g. Carlén et al., 2004; Heringstad et al., 2008). The estimated heritability for SCS-SD in the present study was 0.14 and agreed well with the finding of Urioste et al. (2010), which was 0.10 for weekly recordings and 0.14 for monthly recordings. Although SCS-SD has previously been analyzed (Urioste et al., 2010), the present study is the first to report the genetic correlation coefficient between LASCS and SCS-SD which was 0.96 with an upper 95% limit of 1.0 and a lower 95% limit of 0.88. This indicates that the genetic component of these two traits was not significantly different.

GWAS

The present study detected few SNP associations with LASCS or SDS-SD which indicates that effects of most QTL involved in genetic control of LASCS and SCS-SD were not large enough to be detected with the traits defined in the present dataset. Firstly, this could suggest that the present dataset may not provide sufficient power to detect loci with relatively small effects for the currently defined traits. The detection power of GWAS, especially for low to moderately heritable polygenic traits, will likely be substantially increased by including a larger number of animals. Detection of few genomic regions involved in genetic regulation of SCS, or SCS related traits is supported by other studies (e.g. Heyen et al., 1999; Klungland et al., 2001; Meredith et al., 2010). This could also suggest that SCS may be influenced by multiple QTL dispersed throughout the genome, each with a relatively small effect which hinders detection of genomic regions.

In accordance with the genetic correlation between LASCS and SCS-SD, GWAS results identified a SNP significantly associating with both LASCS and SCS-SD. The

SNP ARS-BFGL-NGS-101491 passed the 0.20 FDR threshold for both LASCS and SCS-SD, with the same direction of the effect and a similar $-\log_{10}$ P-value for both traits. In addition, the SNP BTB-01841922 on BTA4 was found to be significantly associated with LASCS, with a $-\log_{10}$ P-value of 5.7. The direction of the effect of this SNP for SCS-SD was the same as for LASCS, but the $-\log_{10}$ P-value of 3.5 fell short of the 0.20 FDR threshold. Similarly, the SNP BTB-02087354 on BTA6 was found to be significantly associated with SCS-SD with a $-\log_{10}$ P-value of 6.1. The direction of the effect for this SNP for LASCS was the same as for SCS-SD but once again the $-\log_{10}$ P-value of 4.6 fell short of the 0.20 FDR threshold.

Candidate Genes

Genes within a window of 200,000 bp around the associated SNP were considered as candidates. At this distance linkage disequilibrium (r^2) decreases to a value of approximately 0.15 (Khatkar et al., 2008). The 200,000 bp window around SNP ARS-BFGL-NGS101491, which associated with both LASCS and SCS-SD, contained eight genes. Six of these genes, *SLC7A5*, *CA5A*, *RNF166*, *MVD*, *CTU2*, and *SNAI3*, have no apparent function in mastitis resistance. The bovine *SMAR1* gene, however, (Genebank BC119967.1) is located within this window. The gene product of *SMAR1* plays a central role in cell cycle, apoptosis and signaling pathways through its interaction with proteins such as NFκB, p53 and TGF-β (Malonia et al., 2011). Furthermore, the *p22-PHOX* gene (Genebank accession number AF036096.1) is also located in close proximity to SNP ARS-BFGL-NGS-101491. The protein product of this gene plays a role in phagocytosis as it is an essential component for an active phagocyte NADPH oxidase. This NADPH oxidase is required for the production of superoxide, a precursor of microbicidal oxidants (Sumimoto et al., 1996). A 200,000 bp window around SNP BTB-01841922 contained four genes, *TFPI2*, *GNGT1*, *GNG11* and *BET1*, of which three are involved in functions relating to mastitis resistance. The *GNGT1* and *GNG11* genes code for G proteins which function as signal transductors (Downes and Gautam, 1999) involved in regulation of cell migration and adhesion (Ahmed et al., 2010), key attributes of innate immune responses (Snyderman and Goetzl, 1981). The *TFPI2* gene (Genebank AY234861.1) codes for a protein that plays a role in proliferation and apoptosis of smooth muscle cells (Ekstand et al., 2010). The mammary gland contains smooth muscle-like myoepithelial cells whose protein expression resembles that of smooth muscle cells (Deugnier et al., 1995). Given this resemblance, the *TFPI2* gene product may also exert influence on mammary tissue. No genes were present within a 200,000 bp window around SNP BTB-02087354.

Comparison with Literature

Multiple studies have identified regions containing QTL underlying genetic variation for SCS on almost all bovine chromosomes (Rupp and Boichard, 2003; Khatkar et al., 2004; Hu and Reecy, 2007). Chromosomes one, 18, 21 and 23 are frequently reported to contain QTL for SCS (Rupp and Boichard, 2003; Khatkar et al, 2004). Previously, linkage studies have reported QTL for SCS on BTA4 (Zhang et al., 1998; Tal-Stein et al., 2010) and BTA 6 (Daetwyler et al., 2008; Lund et al., 2008). These QTL, however, do not appear to be in close proximity to SNP BTB-01841922 and SNP BTB-02087354. Further, linkage studies have reported a QTL for SCC and CM on BTA18 near the microsatellite marker TGLA227, located at approximately 65,000,000 bp (Bennewitz et al., 2003; Kühn et al., 2003; Schulman et al., 2004). This QTL could not be confirmed in the present GWAS.

From linkage studies, it seems that the magnitude of associations between phenotypes and genetic markers may be largely influenced by environmental factors (e.g. infection pressure) and the genetic background specific to studied populations and breeds (Rupp and Boichard, 2003).

GWAS results may be subjected to the same limitations. From GWAS with 1,341 SNP Kolbehdari et al. (2009) identified two SCS associated SNP on BTA18 located at 807,748 bp and 1,414,404 bp. The SNP ARS-BFGL-NGS101491 associating with LASCS and SCS-SD in the present study is not located in the same region given its position at 13,839,646 bp. A GWAS with 17,349 SNP performed by Sodeland et al. (2011) on Norwegian Red dairy cattle showed associations to daughter yield deviations for LASCS and CM on BTA12, BTA19 and BTA20. None of these associations were detected in the present study.

CM and Production Traits

The RobustMilk dataset contains 390 cows that were part of two genetic lines (Scottish data) (Veerkamp et al., 1994). One line was selected for kilograms of milk fat plus protein and the other line was selected to resemble the average genetic merit for milk fat plus protein for all UK evaluated cows. A genetic antagonism has been described between SCC and production traits (e.g. Rupp and Boichard, 1999; Carlén et al., 2004). Phenotypic differences in LASCS and SCS-SD were, however, small between the two lines and no differences in MAF could be detected for the significant SNP.

In contrast, SCC and the occurrence of mastitis show a strong positive genetic correlation (e.g. Rupp and Boichard, 1999; Carlén et al., 2004; Koivula et al., 2005). Clinical mastitis records were available for RobustMilk cows from Scotland and Sweden. In total 98 cows had a case of CM recorded during the first lactation

(mastitic cows); the remaining 483 cows from Sweden and Scotland without such a record were assumed free of CM during this period (non-mastitic cows). There were clear phenotypic differences in SCS between mastitic and non-mastitic cows; the average LASCS was 7.1 for mastitic and 6.1 for non-mastitic cows and the average SCS-SD was 7.8 for mastitic and 5.9 for non-mastitic cows. The MAF for the significant SNP, however, did not differ between mastitic and non-mastitic cows. For SNP ARS-BFGL-NGS-101491 on BTA18, the frequency of the unfavorable allele was 0.25 in mastitic cows and 0.24 in non-mastitic cows. For SNP BTB-02087354 on BTA4, the frequency of the unfavorable allele was 0.02 in both mastitic and non-mastitic cows. For SNP BTB-01841922 on BTA6, the favorable allele had a frequency of 0.09 in mastitic cows and 0.07 in non-mastitic cows. In agreement with other studies (Klungland et al., 2001; Sodeland et al., 2011), these results suggest that the SNP associated with LASCS or SCS-SD do not have major effects on CM. One explanation is that LASCS is an average value of multiple test-day records and therefore not directly comparable with CM records (Sodeland et al., 2011). In addition, the traits measure separate aspects of udder health (Lund et al., 2007), which results in a loss of statistical power (Lund et al., 2007; Sodeland et al., 2011). Furthermore, SNP effects on LASCS and SCS-SD in the present study were relatively small. So, if an association of these SNP exists with CM, the difference in allele frequency would have a minor impact on the phenotype.

4.5 Conclusion

The present study is one of the first studies to combine detailed phenotypic and genotypic cow data from research herds located in different countries. Relatively few associated SNP were found, which suggests that LASCS and SCS-SD are controlled by multiple loci, each with a relatively small effect, distributed across the genome, although the number of animals included in the study is relatively small. Findings from the present study need to be verified in subsequent independent studies. More knowledge on genetic control of LASCS and SCS-SD is needed to enable higher selection response for these traits and such knowledge could be particularly valuable for improvement of the accuracy of estimated breeding values for bull dams. Moreover, such knowledge contributes to the quest for genes for mastitis resistance and aids the understanding of the genetic mechanisms of mastitis and the discovery of targets for mastitis therapeutics.

4.6 Acknowledgements

5

Macro- and micro-environmental sensitivity for test-day somatic cell score in Swedish Holstein dairy cattle

S. Wijga[1], E. Strandberg[2], and H.A. Mulder[1]

[1] Animal Breeding and Genomics Centre, Wageningen University, P.O. Box 338, 6700 AH, Wageningen, The Netherlands; [2] Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, P.O. Box 7023, S-750 07 Uppsala, Sweden

Abstract

Sensitivity to respond to environmental factors can have a genetic origin. Environmental factors can be divided into known and unknown environmental factors, referred to as macro- and micro-environment, respectively. Macro-environmental sensitivity can be expressed as genetic variation in the slope of a reaction norm, whereas micro-environmental sensitivity can be expressed as differences in residual variance that have a genetic origin. The aim of this paper is to estimate macro- and micro-environmental sensitivity for somatic cell count in Holstein dairy cattle, where the macro-environment was defined as the herd-year average somatic cell count. A dataset containing over 1.6 million test-day records on about 177,000 cows was available. Variance components were estimated with a double hierarchical generalized linear model extended with a reaction norm using the ASReml software. Both macro- and micro-environmental sensitivity were found for somatic cell score and these two sensitivities were found to be positively correlated. Estimated variance components and genetic correlations obtained with the macro or micro model resembled those obtained with the macro-micro model, which suggests that use of macro or micro models does not lead to biased estimates resulting from micro models partly picking up macro-environmental sensitivity or vice versa. Knowledge on both forms of sensitivity, however, may aid in optimization of selection as correlations between the additive genetic variance in intercept, slope and environmental variance were all away from unity. We conclude that selection for reduced environmental sensitivity has the potential to reduce variability in animal performance due to known and unknown environmental factors and herewith increase predictability of performance across and within environments.

Key words: macro-environmental sensitivity, genotype environment interaction, environmental variance, double hierarchical generalized linear model, genetic parameters, heterogeneity, residual variance

5.1 Introduction

Individual animals may show genetically determined differences in their response to environmental influences, a phenomenon known as genotype by environment interaction (GxE) (Falconer and Mackay, 1996). The GxE can manifest itself in two forms: 1) reranking of animals across environments, 2) differential scaling of animal performance across environments (Falconer and Mackay, 1996).

A variety of factors contribute to the environment of an animal. The factors that are known and can be categorized (e.g. ad-libitum or restricted diet) or quantified (e.g. temperature) are referred to as macro-environmental factors (Falconer and Mackay, 1996). As such, genetic variation in macro-environmental sensitivity is genetic variation in response to macro-environmental factors. When macro-environmental factors can be quantified on a continuous scale, macro-environmental sensitivity can be expressed as genetic variation in the slope of a reaction norm. Genetic variances and covariances of a reaction norm can be translated into genetic correlations between pairs of environments, i.e. values of the environmental parameter (De Jong and Bijma, 2002), as measures of GxE interaction. When macro-environmental factors can be categorized, then the macro-environmental sensitivity can be expressed as the genetic covariance between environments. The combination of the unknown environmental factors constitutes the micro-environment (Falconer and Mackay, 1996). Genetic variation in micro-environmental sensitivity is genetic variation in response to micro-environmental factors. Micro-environmental sensitivity, also referred to as genetic heterogeneity of residual variance, can be expressed as differences in residual variance that have a genetic origin (Rönnegård et al., 2013; Mulder et al., 2013).

As livestock breeding is a global activity that serves a wide range of production systems, presence of macro- and micro-environmental sensitivity affects animal breeding. Increased knowledge on macro-and micro-environmental sensitivity may facilitate optimal performance of animals and selection for increased uniformity and robustness (Mulder, 2007).

Recently, Rönnegård et al. (2010) showed that application of a double hierarchical generalized linear model (DHGLM; Lee and Nelder, 2006) allows for estimation of micro-environmental sensitivity using standard variance component estimation programs such as ASReml. A subsequent study (Rönnegård et al., 2013) showed existence of moderate genetic variation for micro-environmental sensitivity for milk yield and somatic cell score in Swedish Holstein cattle. Mulder et al. (2013) have extended the DHGLM as proposed and used by Rönnegård et al. (2010; 2013) with a reaction norm model to study macro- and micro-environmental sensitivity

simultaneously in dairy cattle. By means of simulation it was shown that estimation of genetic parameters with sufficient precision required designs with at least 100 sires with each 100 half-sib offspring. Application of the model to lactation-average milk yield suggested presence of both macro- and micro-environmental sensitivity.

The present study builds on the study by Mulder et al. (2013) who applied the DHGLM extended with a reaction norm to simulated data and to lactation-average milk yield in field data. The model, however, has not been applied to test-day records, which are frequently used in dairy breeding. Test-day records provide a more comprehensive representation of the phenotype compared to lactation-averages; animals may have the same lactation-average SCS but patterns of variation around this mean may differ between these animals. Further, individual phenotypic measurements may manifest themselves under the influence of different environmental influences. These environmental influences may cancel each other out by using lactation averages. Therefore, the present study aimed to estimate genetic variation in macro- and micro-environmental sensitivity in Swedish Holstein cattle based on test-day records using a DHGLM extended with a reaction norm model. The model was applied to test-day somatic cell score (SCS). In addition, this study aimed to provide a biological and genetic interpretation of the results.

5.2 Materials and methods

Data

Data used in the current study were previously used in studies by Windig et al. (2013), Rönnegård et al. (2013) and Mulder et al. (2013). Data contained almost 1.7 million first lactation test-day records, obtained from 177,411 Swedish Holstein cows. On average 9.5 test-day records were available for each cow. Data were recorded between 2002 and 2009 and cows were located in 1,759 herds. Data included test-day records recorded between 7 and 366 days in milk for cows with age at calving between 19 and 38 months. Test-day somatic cell count (SCC) records were log-transformed (Ali and Shook, 1980) to SCS test-day phenotypes for analyses as SCS = $\log_2$ (SCC/100,000) + 3. Cows were sired by 762 sires, where the number of daughters per sire ranged from 46 to 7559. At least two generations of male ancestors were known for sires of all cows with records. In total, the pedigree contained 4,072 individuals.

Genetic model

The present study aimed to estimate genetic variation in macro- and micro-environmental sensitivity using a DHGLM (Rönnegård et al., 2010, 2013), which was

extended with a reaction norm model (Mulder et al., 2013). In brief, the genetic model that describes the genetic variance in macro- and micro-environmental sensitivity models the macro-environmental sensitivity as genetic variation in the slope of a linear reaction norm, whereas micro-environmental sensitivity is modeled as genetic variation in environmental variance according to an exponential model:

$$P = \mu + A_{int} + A_{sl}x + \exp(0.5\ln(\sigma_E^2) + 0.5A_v)\varepsilon ,\qquad(1)$$

where P represents the phenotype, μ is the population mean for the phenotype, A_{int} and A_{sl} are the additive genetic values for the intercept and for the slope of the reaction norm (= macro-environmental sensitivity), respectively, x is the environmental parameter that is responsible for the genotypes to respond differently, σ_E^2 is the environmental variance of the exponential model, A_v is the additive genetic value for the environmental variance (=micro-environmental sensitivity) and ε is a scaled environmental deviation with variance one. Variances and covariances for additive genetic values A_{int}, A_{sl}, and A_v are

$$\mathbf{G} = \begin{bmatrix} \sigma_{Aint}^2 & \sigma_{Aint,Asl} & \sigma_{Aint,Av} \\ & \sigma_{Asl}^2 & \sigma_{Asl,Av} \\ & & \sigma_{Av}^2 \end{bmatrix}$$

Statistical model
Data were analyzed with a sire model. The genetic model described above was translated in a statistical macro-micro model as follows

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{\psi}_s \end{bmatrix} = \begin{bmatrix} \mathbf{X} & 0 \\ 0 & \mathbf{X}_v \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \mathbf{b}_v \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_s & \mathbf{Z}_x & 0 \\ 0 & 0 & \mathbf{Z}_{sv} \end{bmatrix} \begin{bmatrix} \mathbf{s}_{int} \\ \mathbf{s}_{sl} \\ \mathbf{s}_v \end{bmatrix} + \begin{bmatrix} \mathbf{P} & 0 \\ 0 & \mathbf{P}_{sv} \end{bmatrix} \begin{bmatrix} \mathbf{pe} \\ \mathbf{pe}_{sv} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_s \\ \mathbf{e}_{sv} \end{bmatrix} ,\qquad(2)$$

where $\mathbf{y}$ was the vector with the phenotypic observations, $\mathbf{\psi}_s$ was the vector with the linearized values of transformed squared residuals calculated as in Rönnegård et al. (2010). $\mathbf{X}$ ($\mathbf{X}_v$) was the incidence matrix for fixed effects for $\mathbf{y}$ ($\mathbf{\psi}_s$), $\mathbf{b}$ ($\mathbf{b}_v$) was the vector with solutions for fixed effects for $\mathbf{y}$ ($\mathbf{\psi}_s$), $\mathbf{Z}_s$ and $\mathbf{Z}_{sv}$ were the incidence

matrices for the sire effects for the intercept of the reaction norm and the environmental variance, $\mathbf{Z_x}$ was the matrix with the environmental parameter x as a covariate for the sire effects for the slope of the reaction norm, $\mathbf{s_{int}}$, $\mathbf{s_{sl}}$ and $\mathbf{s_v}$ were the vectors with the estimated sire effects for intercept, slope and environmental variance, $\mathbf{P}$ ($\mathbf{P_{sv}}$) was the incidence matrix for the permanent environmental effects for $\mathbf{y}$ ($\mathbf{\psi_s}$), $\mathbf{pe}$ ($\mathbf{pe_{sv}}$) was the vector with solutions for permanent environmental effects for $\mathbf{y}$ ($\mathbf{\psi_s}$), and $\mathbf{e}$ and $\mathbf{e_{sv}}$ were the vectors with the residuals of y and ψ_s, respectively. Sire effects $\mathbf{s_{int}}$, $\mathbf{s_{sl}}$, and $\mathbf{s_v}$ were assumed distributed N~($\mathbf{0}$, $\frac{1}{4}\mathbf{G} \otimes \mathbf{A}$), where $\mathbf{A}$ represents the matrix with genetic relationships between animals based on pedigree records and where it was assumed that sire (co)variances represent a quarter of the additive genetic variance. Permanent environmental effects were assumed distributed as N~($\mathbf{0}$, $\mathbf{Q} \otimes \mathbf{I}$), where Q represents the matrix of permanent environmental (co)variances, $\mathbf{Q} = \begin{bmatrix} \sigma^2_{pe} & \sigma_{pe,pe_{sv}} \\ \sigma_{pe,pe_{sv}} & \sigma^2_{pe_{sv}} \end{bmatrix}$, and where $\mathbf{I}$ represents the identity matrix.

Residuals $\mathbf{e_s}$ and $\mathbf{e_{sv}}$ were assumed distributed N $\sim \left(\begin{matrix} \mathbf{0} \\ \mathbf{0} \end{matrix}, \begin{bmatrix} \mathbf{W_s^{-1}}\sigma^2_{\in} & 0 \\ 0 & \mathbf{W_{sv}^{-1}}\sigma^2_{\in_v} \end{bmatrix} \right)$, where $\sigma^2_{\in}$ and $\sigma^2_{\in_v}$ represent the scaling variances for the residual variances and where $\mathbf{W_s^{-1}}$ and $\mathbf{W_{sv}^{-1}}$ represent the weights corresponding to $\mathbf{y}$ and $\mathbf{\psi_s}$.

Weights $\mathbf{W_s^{-1}}$ and $\mathbf{W_{sv}^{-1}}$ were calculated as in Rönnegård et al. (2010), which was different from the calculation in Mulder et al. (2013), because in the current study, as in the study of Rönnegård et al. (2010), the residual variance does not include $\frac{3}{4}\sigma^2_A$. This $\frac{3}{4}\sigma^2_A$ is absorbed by the permanent environmental effect due to using repeated observations. The modification shown in Mulder et al. (2013) is, however, applicable for situations with single observations per animal.

The sire macro-micro model was applied by iterating this model for 50 runs, where every run was a converged ASReml (Version 3; Gilmour et al., 2009) analysis. Parameters were considered converged when the sum of the squared relative difference was below 10^{-4}. Each iteration generated REML-estimates of variance

components for the current values of $\boldsymbol{\psi_s}$, $\mathbf{W_s}$, and $\mathbf{W_{sv}}$. Subsequently, the vector $\boldsymbol{\psi_s}$, and the diagonals of $\mathbf{W_s}$ and $\mathbf{W_{sv}}$ were updated for the next iteration.

The fixed effects included in the mean model were the covariate lactation stage (DIM) modeled with a Wilmink curve (Wilmink, 1987), the covariate year-season of calving (ys), herd test-day effect (htd) consisting of 21,570 classes and the covariate age at calving (AgeC). Fixed effects included in the variance model were the covariate DIM modeled with a Wilmink curve (Wilmink, 1987), the covariate ys, and htd consisting of 21,570 classes.

Definition macro-environment

Matrix $\mathbf{Z_x}$ (see model 4) describes the environmental parameter x for the slope of the reaction norm. In the present study, two environmental parameters were estimated from the data and they were individually included in model 2. These parameters were herd-year-average (hy)SCC and herd-year-average milk yield (hyMY). Both were scaled to have mean zero and variance 1.0:

$$scaled\ hy_i = \frac{hy_i - \dfrac{\sum hy_i}{n}}{SD_{hy}}, \tag{3}$$

where hy_i was the hySCC or hyMY value belonging to record i, n was the number of records in the dataset, μ was the mean hySCC or mean hyMY over all records in the dataset, and SD_{hy} was the standard deviation over all herd-year classes which was calculated as $\sqrt{\dfrac{\sum (hy_i - \mu)^2}{n-1}}$. Two environmental parameters were used: the herd-year average SCC (hySCC) and the herd-year average milk yield (hyMY). The environmental parameter hySCC represents the disease incidence in the herd, where herds with high hySCC are expected to have a higher incidence of mastitis. The environmental parameter hyMY represents the management related to milk production, i.e. feeding regime or milking frequency. The environmental parameters hySCC and hyMY can be partly correlated as an environments with lower incidence of mastitis may have also higher milk yield.

Model selection

The goodness of fit of the macro-micro model was compared to three simpler models: 1) a macro model which only accounts for macro-environmental

sensitivity, 2) a micro model which only accounts for micro-environmental sensitivity and 3) a simple model which contained only one additive genetic effect for the phenotype. The goodness of fit of the four models was assessed to determine if simultaneously accounting for macro- and micro-environmental sensitivity substantially improved the fit of the model compared to either accounting for macro- or micro-environmental sensitivity or none of them. Further, the present study determined the fit of the macro-micro and the macro model when the used environmental parameter *x* was either hySCC or hyMY. The goodness of fit of each model was assessed using the Akaike's information criterion (AIC) using the h-likelihood framework (Mulder et al., 2013). The h-likelihood was approximated using the adjusted profile h-likelihood (APHL) (Mulder et al., 2013):

$$APHL = -2logL - \sum w_{sv_i}(\hat{\sigma}^2_{\in_{sv}})^{-1} - \sum_{\in_{sv}} \ln(\hat{\sigma}^2_{\in_{sv}})/w_{sv_i} ,$$ (4)

where w_{sv_i} was the weight for the variance model for observation *i*, which is the *i*[th] diagonal of $\mathbf{W}_{sv_i}$. AIC was calculated as (Mulder et al., 2013):

$$AIC = APHL + 2t ,$$ (5)

where *t* was the number of variance parameters estimated by the model. The model with the lowest AIC was considered as the model that had the best fit to the data.

Genetic parameters
Heritability
Heritabilities express the proportion of genetic variance relative to the phenotypic variance. Due to the use of a reaction norm model heritabilities for *y* were calculated as

$$h^2_{y,i} = \frac{\sigma^2_{Ay,i}}{\sigma^2_{Ay,i} + \sigma^2_{PEy} + \sigma^2_{Ey,i}} ,$$ (6)

where $\sigma^2_{Ay,i}$ was the additive genetic variance in *y* in environment *i*, σ^2_{PEy} was the variance of permanent environmental effects for *y*, and $\sigma^2_{Ey,i}$ was the residual

variance for *y* calculated as the average of the ratio of the residual variance (σ_{Ey}^2) and the weight (w_s) of the individual observations in environment *i*. The σ_{Ay}^2 in environment *i* was calculated as

$$\sigma_{Ay,i}^2 = \sigma_{Aint}^2 + 2x\sigma_{Aint,Asl} + x^2\sigma_{Asl}^2 ,$$

(7)

where σ_{Aint}^2 was the additive genetic variance of the breeding value for intercept, σ_{Asl}^2 was the additive genetic variance of the breeding value for slope variance, $\sigma_{Aint,Asl}$ was the covariance between the breeding value for intercept and the breeding value for slope variance and *x* was the value of the environmental parameter *x* in environment *i*.

For *ψ* the heritability was calculated as

$$h_\psi^2 = \frac{\sigma_{Av}^2}{\sigma_{Av}^2 + \sigma_{PE\psi}^2 + \sigma_{E\psi}^2} ,$$

(8)

where σ_{Av}^2 was the additive genetic variance in *ψ*, $\sigma_{PE\psi}^2$ was the permanent environmental effect of *ψ*, and $\sigma_{E\psi}^2$ was the residual variance expressed as the ratio of the residual variance of *ψ* and the average weight ($\overline{w_{sv}}$) of the individual observations ($\frac{\sigma_E^2}{\overline{w_{sv}}}$).

Genetic correlations
When reaction norms are fitted, the amount of reranking of genotypes between different macro-environments can be expressed by the genetic correlation between macro-environments 1 and 2:

$$r_g = \frac{\sigma_{Ay1,Ay2}}{\sqrt{(\sigma_{Ay1}^2 \, \sigma_{Ay2}^2)}} ,$$

(9)

where $\sigma_{Ay1,Ay2}$ was the genetic covariance between environment 1 and 2, and σ_{Ay1}^2 and σ_{Ay2}^2 were the additive genetic variances of the two environments (see equation 7 for calculation).

The covariance between the two macro-environments was calculated as

$$\sigma_{Ay1,Ay2} = \sigma_{Aint}^2 + (x_1 + x_2)\sigma_{Aint,Asl} + x_1 x_2 \sigma_{Asl}^2 , \tag{10}$$

where x_1 and x_2 correspond to the value of the environmental parameter x in environments 1 and 2.

Besides genetic correlations among macro-environments, genetic correlations between performance in macro-environment i and performance in micro-environment can exist. These correlations were calculated as

$$r_g = \frac{\sigma_{Ayi,Av}}{\sqrt{\sigma_{Ayi}^2 \, \sigma_{Av}^2}} , \tag{11}$$

where $\sigma_{Ayi,Av}$ was the covariance between the breeding value of y in environment i and the breeding value of ψ, σ_{Ayi}^2 was the additive genetic variance of the breeding value of y in environment i, and σ_{Av}^2 was the additive genetic variance of the breeding value for ψ.

The $\sigma_{Ayi,Av}$ was calculated as

$$\sigma_{Ayi,Av} = \sigma_{Aint,Av} + x\,\sigma_{Asl,Av} , \tag{12}$$

where $\sigma_{Aint,Av}$ was the covariance between the breeding value for intercept and ψ, $\sigma_{Asl,Av}$ was the covariance between the breeding value for slope and ψ, and x was the value of the environmental parameter in environment i.

5.3 Results

Model selection

Six models were run and the APHL and AIC of these models are shown in Table 1. The macro-micro model when using hySCC as *x* fitted the data best, followed by a macro model with the same *x*. The macro-micro model when using a different *x* was slightly better than using a micro model, which suggests that using a micro model would be sufficient. The simple model had the worst fit for both traits.

It can be concluded that the best fitting model was the macro-micro model using the environmental parameter corresponding to the response variable itself.

Table 5.1 Adjusted profile h-likelihood (APHL) and Akaike's information criterion (AIC) for the macro-micro[1], macro[1], micro[1] and the simple[1] model for response variable somatic cell score (SCS). The environmental parameter (*x*) (herd-year average somatic cell count (hySCC) or herd-year average milk yield (hyMY)) used for the reaction norm in the macro-micro and macro model is in parentheses. For readability, AIC of the models were scaled according to their deviation from the best performing model (AIC*)**.**

Trait (*x*)	Genetic model	APHL	AIC	AIC*
SCS (hySCC)	Macro-micro	2302930	2302948	0
SCS (hyMY)	Macro-micro	2304549	2304567	1619
SCS (hySCC)	Macro	2304239	2304251	1303
SCS (hyMY)	Macro	2305804	2305816	2868
SCS	Micro	2304556	2304568	1620
SCS	Simple	2305854	2305862	2914

[1]Macro-micro: model accounts for both macro- and micro-environmental sensitivity, macro: model accounts for macro-environmental sensitivity, micro: model accounts for micro-environmental sensitivity, simple: model accounts for a genetic effect for the phenotype, but not for macro- and micro-environmental sensitivity.

Estimated variance components

Variance components estimated for SCS are shown in Table 2. Only the macro-micro and macro model where *x* was the same trait as the response variable are shown in this table, because the fit of this model was better than that of the macro-micro and macro model where *x* was a different trait than the response variable.

Estimates for σ^2_{Aint} were between 0.21 and 0.25. The estimate for σ^2_{Av} was about 0.05 and the estimate for σ^2_{Asl} was about 0.03. Estimates for σ^2_{Aint} and σ^2_{Asl} obtained with the macro and micro model were similar to the estimates obtained with the macro-micro model. These similarities suggest that use of macro or micro models does not lead to biased estimates resulting from micro models partly

explaining macro-environmental sensitivity or vice versa. The average fitted residual variance was 0.60. The fitted permanent environmental effects were 1.20 for y and 0.61 for ψ and the correlation between the permanent environmental effects for y and ψ was 0.43.

Estimated genetic correlations between the variance components were positive and moderate to high, ranging from 0.46 to 0.78 (Table 2). Estimates obtained with the macro and micro model were of the same approximate magnitude as those obtained with the macro-micro model. The positive genetic correlations suggest that when selection decreases SCS (i.e. the trait level represented by the intercept) SCS is more stable across environments (represented by the slope) and that the variability within environments (represented by the environmental variance) will decrease as well. In short, animals with that are genetically prone to a high SCS are more sensitive to environmental changes.

It can be concluded that moderate to high positive genetic correlations exist between σ_{Aint}^2, σ_{Asl}^2 and σ_{Av}^2. Both variance component estimates and estimated genetic correlations obtained with the macro-micro model were similar to those obtained with the macro or micro model.

Genetic parameters

Heritability

When a reaction norm is applied, the different classes of x are regarded as individual traits. Therefore, a heritability can be calculated for each herd-year class. The macro-environment for SCS (hySCC) had 7,193 classes. For each class of the macro-environment the heritability was calculated according to equation 6. Here, $\sigma_{Ey,i}^2$ was defined as the average residual variance in each class. Results are shown in Figure 1. The average heritability was 0.10 and ranged from 0.02 to 0.68. The figure shows a quadratic trend which is the result of the quadratic component in σ_{Ay}^2 (see equation 7).

The micro-environment is defined as the genetic variation in environmental variance. Therefore, in contrast to macro-environmental sensitivity, one heritability can be calculated for micro-environmental sensitivity, which was 0.01.

We conclude that heritabilities of SCS differ greatly among herd-year classes. The heritability for micro-environmental sensitivity is relatively low.

Table 5.2 Estimated variance components[1] for response variable somatic cell score (SCS) estimated with different models[2]. Corresponding SE are in parentheses.

Trait	Model	σ^2_{Aint}	σ^2_{Asl}	σ^2_{Av}	$\rho_{Aint,Asl}$	$\rho_{Aint,Av}$	$\rho_{Asl,Av}$
SCS	Macro-micro	0.247 (0.004)	0.031 (0.001)	0.051 (0.001)	0.607 (0.054)	0.517 (0.044)	0.782 (0.049)
SCS	Macro	0.208 (0.004)	0.034 (0.008)	-	0.487 (0.061)	-	-
SCS	Micro	0.238 (0.004)	-	0.046 (0.012)	-	0.466 (0.048)	-
SCS	Simple	0.206 (0.004)	-	-	-	-	-

σ^2_{Aint} = additive genetic variance of the breeding value for intercept, σ^2_{Asl} = additive genetic variance of the breeding value for slope variance (macro-environmental sensitivity), σ^2_{Av} = additive genetic variance of the breeding value for environmental variance (micro-environmental sensitivity) $\rho_{Aint,Asl}$ = genetic correlation between intercept and slope, $\rho_{Aint,Av}$ = genetic correlation between intercept and environmental variance, and $\rho_{Asl,Av}$ = genetic correlation between slope and environmental variance.
[2]Macro-micro: model accounts for both macro- and micro-environmental sensitivity, macro: model accounts for macro-environmental sensitivity, micro: model accounts for micro-environmental sensitivity, simple: model accounts for a genetic effect for the phenotype, but not for macro- and micro-environmental sensitivity. The environmental parameter *x* used for the reaction norm in the macro-micro and macro model was the herd-year average somatic cell count.

Correlations macro-environmental sensitivity

The macro-environment is defined by environmental parameter x and genetic correlations with respect to animal performance can be estimated between pairs of environments, i.e. values of x. These correlations give an indication of reranking of animals across the different environments, and are as such a measure of GxE. Genetic correlations among different values of x (hySCC) were positive and generally high (Table 3). Correlations decreased as the difference between the values of x increased. Near unity correlations were found between adjacent values of x. The genetic correlation was lowest between x values -2 and 2, where the correlation had decreased to 0.40.

We conclude that genetic correlations among different values of x were generally positive and relatively high.

Correlations macro- and micro-environmental sensitivity

Genetic correlations can not only be estimated between different values of the environmental parameter x, but also between the different values of x and the micro-environment. Correlations between the micro-environment and different values of x (hySCC) ranged from -0.04 for $x = -2$ to 0.70 for $x = 2$.

Thus, genetic correlations between micro-environmental sensitivity and phenotypic level in different macro-environments were mostly positive, which suggests that an increase in phenotypic level often leads to an increase in micro-environmental sensitivity, i.e. an increased environmental variance.

Table 5.3 Genetic correlations between the different values of the macro-environment (x), and genetic correlations between the different values of the macro-environment (x) and the micro-environment for somatic cell count.

	Macro									Micro
x	-2.0	-1.5	-1.0	-0.5	0	0.5	1	1.5	2.0	
-2.0										-0.04
-1.5	0.98									0.13
-1	0.91	0.98								0.29
-0.5	0.81	0.92	0.98							0.42
0	0.71	0.85	0.94	0.99						0.52
0.5	0.61	0.77	0.89	0.96	0.99					0.59
1	0.53	0.71	0.84	0.93	0.97	0.99				0.64
1.5	0.46	0.65	0.79	0.89	0.95	0.98	1.00			0.67
2.0	0.40	0.59	0.75	0.86	0.93	0.97	0.99	1.00		0.70
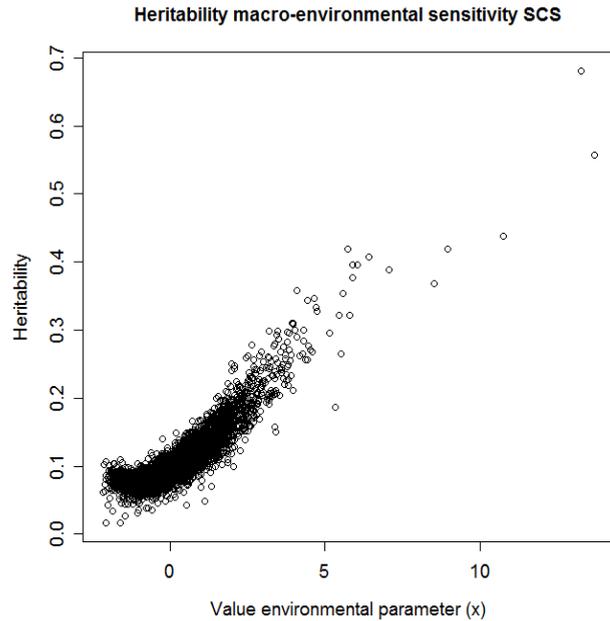
Figure 5.1 Heritability of the phenotype for each value of environmental parameter *x* (hySCC) for somatic cell score (SCS).

Figure 5.2 Estimated breeding values for intercept (EBV) plotted against estimated breeding values for micro-environmental sensitivity (vEBV) for somatic cell score (SCS).

Figure 5.3 Estimated breeding values for intercept (EBV) plotted against estimated breeding values for macro-environmental sensitivity (slEBV) for somatic cell score (SCS).

Figure 5.4 Estimated breeding values for micro-environmental sensitivity (vEBV) plotted against estimated breeding values for macro-environmental sensitivity (slEBV) for somatic cell score (SCS).

Figure 5.5 Illustration of the effect of the sire estimated breeding for slope (slEBV) on daughter phenotypes (SCS) along different values of the environmental parameter (hySCC). The figure shows observed daughter phenotypes for two example sires that were found to have a high and a low slEBV in the macro-micro model.

Breeding values

The macro-micro analyses performed in the present study not only provide knowledge on variance components for macro- and micro-environmental sensitivity, but also breeding values are obtained. In Figure 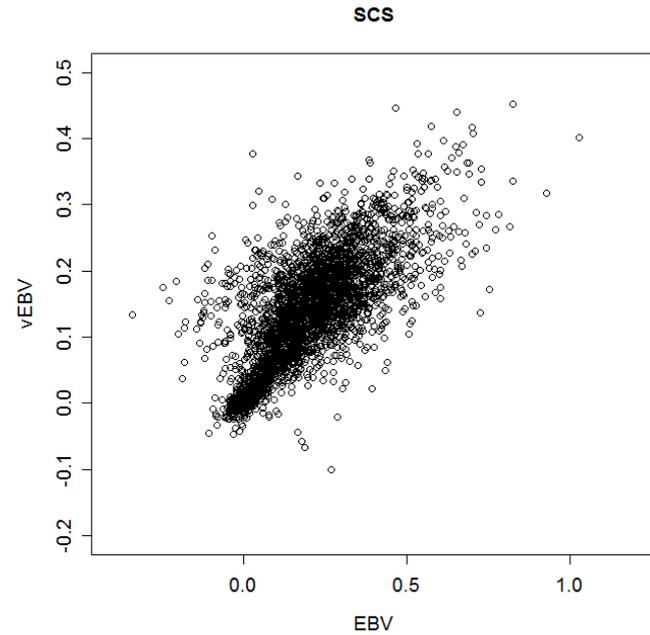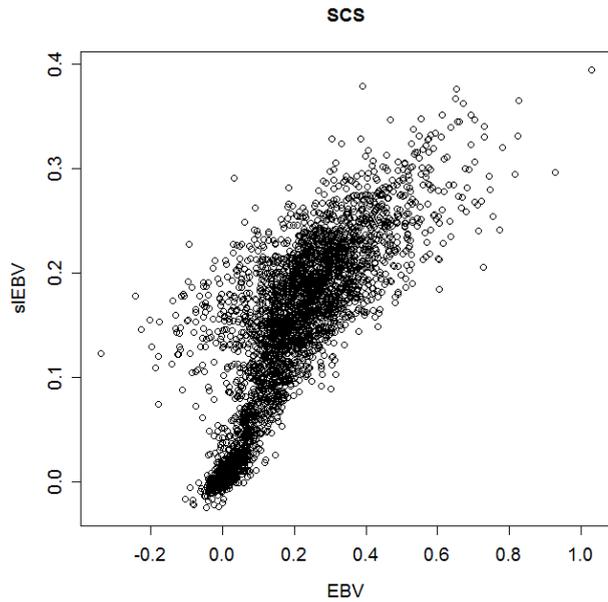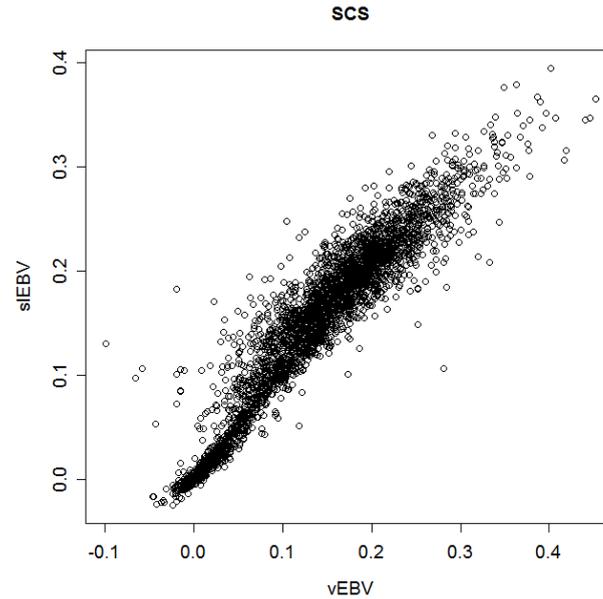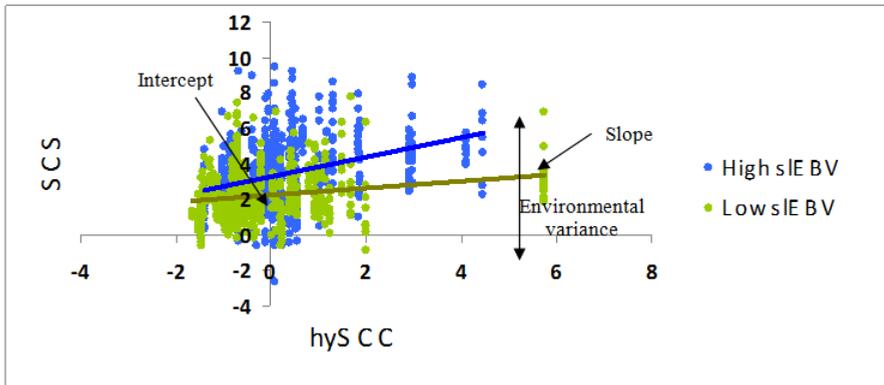2 estimated breeding values for intercept (A_{int}) are plotted against the estimated breeding values for micro-environmental sensitivity (A_v). In Figure 3 estimated breeding values for intercept (A_{int}) are plotted against the estimated breeding values for slope (A_{sl}) and in Figure 4 estimated breeding values for slope (A_{sl}) are plotted against the estimated breeding values for environmental variance (A_v). Due to the positive genetic correlations between A_{int}, A_{sl}, and A_v the studied population does not contain animals with a high estimated breeding value for intercept coupled with a low estimated breeding value for environmental variance. The population does contain multiple animals that have a low estimated breeding value for intercept coupled with a low estimated breeding value for environmental variance. Similar results were seen when estimated breeding values for intercept were plotted against estimated breeding values for slope and when estimated breeding values for environmental variance were plotted against estimated breeding values for slope, as was expected from the genetic correlations presented in Table 2. Figure 5 shows observed daughter phenotypes for two example sires that were found to either have a high or a low estimated breeding value for slope. As a result of the positive correlation between the intercept, the slope and the environmental

variance, the sire with a high slEBV also has a higher intercept and more variation around the slope compared to the sire with a low slEBV.

5.4 Discussion

The present study aimed to simultaneously estimate macro- and micro-environmental sensitivity for SCS based on test-day records using a DHGLM extended with a reaction norm model. Both macro- and micro-environmental sensitivity were detected. Moderate to high positive genetic correlations were found between σ^2_{Aint}, σ^2_{Asl} and σ^2_{Av}. Heritabilities (h^2_y, see equation 6) for SCS differed among herd-year classes. The heritability for micro-environmental sensitivity (h^2_ψ, see equation 8) was found to be relatively low. Genetic correlations for SCS between different values of x were generally positive and relatively high. Genetic correlations between micro-environmental sensitivity and phenotypic level in macro-environments were mostly positive and showed a larger range than genetic correlations for SCS between different macro-environments, i.e. different values of x.

Model selection

The present study tested six models and found that the fit of the macro-micro model that used x corresponding to the response variable itself was substantially better than the fit of the other five models. Mulder et al. (2013) made a similar model comparison for lactation-average milk yield and found the micro model to have the best fit, although the fit of the macro-micro model was almost as good. The substantially better fit of the macro-micro model in the present study, when compared to the simple model and the individual macro and micro model, may be due to the use of test-day records. Use of test-day records not only allows use of between-cow variation, but also within cow-variation in performance at different levels of environmental parameters (Hayes et al., 2003). The additional information obtained from use of test-day records may allow for a better estimation of environmental sensitivity. An additional advantage of use of test-day records is availability of more records per sire, which are potentially recorded across a wider range of production levels when compared to lactation-averages. As such, test-day information may yield more accurate predictions of GxE (Hayes et al., 2003).

The reaction norm model that used an environmental parameter (x) corresponding to the response variable was found to have a better fit than the model that used a different x. The difference in fit may be explained by the low correlation of -0.15

between hySCC and hyMY. This correlation suggests that the two environmental parameters explain different aspects of the management and environment, and it is therefore not surprising that the macro-micro and macro models that each apply a different environmental parameter show a different fit. The superior fit of hySCC is most likely due to the fact that hySCC directly originates from SCS and as such explains the response variable better than hyMY as the correlation of hySCC with SCS was 0.24, whereas the correlation of hyMY with SCS was -0.06. Benefits from using an environmental parameter x corresponding to the response variable itself was previously also reported by Calus et al. (2005), who found that environmental sensitivity or genotype by environment interaction was more often detected when x was the average of the analyzed trait or a characteristic closely related to the trait.

Environmental parameters

In the present study hySCC and hyMY were used as environmental parameters. These environmental parameters were scaled to allow for a similar interpretation of the parameters and, as such, to be able to compare the models. We preferred hySCC as an environmental parameter over hySCS because hySCC is the 'true' phenotype. When SCC is log-transformed into SCS extreme values are regressed toward the mean which decreases the influence of the extreme values relative to the effects of the intermediate values. As hySCC is an average value used as an environmental parameter and no regression was performed on this parameter log-transformation is less relevant than it would be if it were a (non-averaged) response variable. In addition, following the central limit theorem, the hySCC was approximately normally distributed with a small tail to the right.

Interpretation of results

Macro-environmental sensitivity

A reaction norm model analyses the phenotype as a function of the environmental parameter. As a result, for each value of the environmental parameter a heritability can be estimated and genetic correlations can be estimated between the different values of the environmental parameter. For SCS the average estimated heritability was 0.10 and ranged from 0.02 to 0.68. In general, the heritability for SCS increased as the hySCC increased. The average heritability of 0.10 is in agreement with findings from previous studies where heritabilities ranging from 0.05 to 0.14 for single monthly test-days were reported (e.g. Rupp and Boichard, 2003). The increase in heritability that is observed at increasing values of x agrees with the theory presented by Bishop and Wooliams (2010), that is, the heritability increases

at increasing prevalence (represented by a higher hySCC). When the prevalence in a herd increases the variation in disease exposure decreases which leads to higher heritability estimates. The high heritabilities for high hySCC classes should, however, be treated with caution because these heritabilities are probably biased upwards by the quadratic term in σ^2_{Ay}, although the extent of the bias is unknown.

In the present study genetic correlations for SCS between different values of the environmental parameter were found to range from unity to 0.40, which indicates presence of considerable GxE. Presence of GxE has consequences for optimal selection in breeding programs and on farm level presence of GxE implies that management and genetic selection should be coupled, rather than be regarded as separate components when reducing SCC (Calus et al., 2006). In a previous study by Calus et al. (2005), estimates for genetic correlations of SCS in different environments were near unity. In a later study, however, environmental sensitivity was assessed based on test-day records rather than lactation averages and greater effects of GxE were estimated (Calus et al., 2006). These results, coupled with the current findings compared to Mulder et al. (2013), suggest that analysis of test-day records would enable detection of more GxE than lactation records. A study by Windig et al. (2013), however, used the same dataset as was used for the present study to assess the prevalence of (sub)clinical mastitis as a measure of exposure; estimated genetic correlations across herd prevalences were above 0.92, much higher than in the present study. Most likely it is not only the size of the dataset but also the definition of the traits that affects the ability to estimate GxE. In the study of Windig et al. (2013) (sub)clinical mastitis was recorded as a binary trait. A 0/1 definition may limit the amount of detail in the phenotypic information, hereby hindering estimation of GxE.

Micro-environmental sensitivity

The present study builds on studies by Rönnegård et al. (2013) and Mulder et al. (2013). Rönnegård et al. (2013) focused solely on micro-environmental sensitivity, whereas Mulder et al. (2013) extended the model used by Rönnegård et al. (2013) with a reaction norm model. Micro-environmental sensitivity represents the sensitivity of an animal regarding an alteration in an unknown environmental factor and is quantified by additive genetic variance in the residual variance. Rönnegård et al. (2013) used the same dataset as was used in the present study to estimate micro-environmental sensitivity. As expected, the micro-model used in the present study gave results similar to those of Rönnegård et al. (2013).

Estimates for micro-environmental sensitivity obtained with the macro-micro model were similar to estimates obtained with the micro-model. According to the macro-micro model an increase of the average breeding value of ψ by $1\ \sigma^2_{Av}$ results in about 23% increase in residual variance. According to the micro model this increase in residual variance is about 22%. This similarity in estimates suggests that that the individual micro model does not lead to biased estimates when compared to the macro-micro model. The heritability found for micro-environmental sensitivity was relatively low with a value of 0.01 This estimate corresponds with heritabilities for micro-environmental sensitivity reviewed by Hill and Mulder (2010), although it should be noted that the definition of the heritability applied by Hill and Mulder (2010) is not identical to the definition applied in the present study.

Macro- and micro-environmental sensitivity
Extension of the micro model with a reaction norm model results in the macro-micro model. In addition to the additive genetic variance in intercept, which represents the trait mean, the additive genetic variance in the slope of the reaction norm and the environmental variance are estimated. The current study is the first to apply the macro-micro model to test-day records. Mulder et al. (2013) have previously applied the model to lactation-average MY. In the present study, as in the study by Mulder et al. (2013), both macro- and micro-environmental sensitivity were detected. This suggests that extension of the DHGLM with a reaction norm model provides information that would not be obtained when solely modeling micro-environmental sensitivity. In addition to a correlation between the intercept and environmental variance which can be estimated with the micro-model, the macro-micro model estimated genetic correlations between the intercept and the slope of the reaction norm and between the slope of the reaction norm and the environmental variance. Correlations were all away from unity, which suggests that accounting for both macro- and micro-environmental sensitivity may improve predictability of animal performance across environments.

Estimated breeding values macro- and micro-environmental sensitivity
The genetic parameters that were estimated in the current study as well as in previous studies (e.g. Mulder et al., 2013; Rönnegård et al., 2013) suggest that it is possible to breed for increased trait uniformity. Breeding for uniformity in SCS would require an animal to have a low estimated breeding value for micro- or macro-environmental sensitivity (Rönnegård et al., 2013), while the genetic level of SCS is maintained or reduced. Somatic cell score is a frequently used indicator trait

in selection against mastitis, as it is thought that selection for decreased SCS results in decreased susceptibility to mastitis (e.g. Rupp and Boichard, 2003). The estimated breeding values for the intercept were positively correlated with macro- and micro-environmental sensitivity, which suggests that selection for reduced SCS results in a correlated reduction in macro- and micro-environmental sensitivity. It makes sense that differences between sires with resistant and sires with less resistant daughters increase when the prevalence increases. Therefore, a macro-micro model may help to select more efficiently to reduce mastitis across environments.

5.5 Conclusion

It is important to obtain knowledge on livestock sensitivity to known (macro) and unknown (micro) environmental factors, among other because livestock breeding is a global activity that serves a wide range of production systems. We showed that both macro- and micro-environmental sensitivity exist for somatic cell score in dairy cattle and that these two forms are positively correlated. The fit of the macro-micro and macro model that used an environmental parameter corresponding to the response variable itself was better than the fit of the macro-micro and macro model that used an environmental parameter that did not correspond to the response variable. The use of macro or micro models does not lead to biased estimates resulting from micro models partly picking up macro-environmental sensitivity or vice versa. Knowledge on both forms of sensitivity can aid in optimization of selection because correlations between the additive genetic variance in intercept, slope and environmental variance were all away from unity. Selection for reduced environmental sensitivity provides opportunities to reduce variability in animal performance due to known and unknown environmental factors and herewith increase predictability of performance. In this way, selection for reduced environmental sensitivity may help to reduce mastitis incidence across environments.

5.6 Acknowledgements

6

General discussion

6.1 Introduction

Thesis objectives

Reduction of the impact of livestock diseases becomes increasingly important, not only for economic and environmental reasons, but also from a human and animal welfare perspective. For instance, improvement of disease resistance can contribute to a reduced use of antibiotics. One way to reduce the impact of disease is through incorporation of disease resistance into livestock breeding schemes. Livestock breeding has the potential to decrease the impact of disease on the long-term. Further, effects of breeding are cumulative and result in a permanent improvement of disease resistance, provided that the pathogen does not evolve in such a way that host resistance is bypassed. Besides production traits, breeding schemes increasingly include traits related to health and welfare, such as disease resistance (Mark and Sandøe, 2010, Fisher and Mellor, 2008, Miglior et al., 2005). Mastitis constitutes an important infectious disease in dairy cattle, where somatic cell score (SCS)[1] is frequently used in dairy cattle breeding as an indicator trait for mastitis. Despite the widespread use of SCS, still there is a lot to be learned about the interpretation of the trait and the corresponding genetic parameters and the biology of disease resistance and SCS. This PhD thesis aimed to obtain additional insight in disease resistance and SCS through the following objectives:

1) quantify the impact of genetics on innate immunity, represented by natural antibodies (NAb) measured in milk, through estimation of heritabilities and genetic correlations;
2) identify the genomic regions involved in SCS and NAb levels;
3) quantify the role of genetic factors on the environmental sensitivity of SCS.

Main results

In general, phenotypes are established under the influence of genetic and environmental factors, where the amount of genetic and environmental influences varies among traits. The extent of the genetic influence on the phenotype is represented by the heritability.

This thesis showed that NAb levels show heritable variation; heritabilities ranged from 0.08 to 0.55. Further, results suggest that different NAb isotypes are partly under different genetic control. Immunoglobulin (Ig)A and IgM levels were found to have a common genetic basis, but NAb isotype IgG1 levels were found to be

---

[1]For genetic parameter estimation SCC are commonly log-transformed into somatic cell score because of the skewness of the SCC distribution

genetically different from IgA or IgM levels (chapter 2). A genome-wide association study (GWAS) for NAb levels revealed that a region on chromosome 23 is strongly associated with genetic variation in NAb isotype IgM levels. Regions on other bovine chromosomes (14 out of the 29 remaining chromosomes) were also found to be associated with NAb levels, though their effects were not as pronounced as those found on BTA23 (chapter 3). A GWAS for SCS (chapter 4) detected two associations, one on chromosome 4 and one on chromosome 18. For the trait standard deviation in SCS, a measure for the variability of SCS, also two associations were detected, one on chromosome 6 and one on chromosome 18, which was the same region as found for SCS. Finally, in chapter 5 both macro- and micro-environmental sensitivity were found for somatic cell score in dairy cattle. In addition it was found that these two sensitivities are positively correlated.

Chapters 2 to 5 have focused on estimation of genetic parameters for NAb and SCS and on identification of genomic regions for these traits. The aim of this final chapter (chapter 6) is to discuss some of the biology behind NAb and SCS, as an understanding of trait biology may lead to alternative trait definition or alternative statistical analyses which may affect genetic parameter estimation, identification of candidate genes and breeding programs. Firstly, I will discuss the biology behind SCS with regard to the potential value of statistical models and detailed phenotyping in genetic analyses. Secondly, I will discuss the consequences an involvement of maternal effects on genetic parameters that are estimated for traits, with a focus on NAb. In addition, I will investigate if there are indications for maternal effects on NAb levels. Finally, as the MHC was identified as a major candidate gene involved in NAb levels, the final part of this general discussion discusses factors that affect MHC diversity. Further, possible implications of livestock breeding on genetic diversity of the MHC region are discussed.

6.2. SCS: definition and statistical analysis

Most countries routinely record SCS but not clinical mastitis (CM) (e.g. Rupp and Boichard, 2003). Given the relatively high genetic correlation between SCS and mastitis (Rupp and Boichard, 2003), SCS is frequently used as an indirect measure of mastitis in genetic evaluations (Mark et al., 2002). A potential difficulty in the use of SCS records in genetic analyses is that SCS records are a mixture of records on healthy and diseased animals. Statistical properties or the biological interpretation may differ between the healthy and diseased animals, possibly leading to suboptimal results of genetic analyses (Bishop and Woolliams, 2010; Riggio et al., 2010). Insufficiently accounting for the possible statistical and biological differences

between healthy and diseased animals may not only affect the estimated genetic parameters, but may also affect the ability to detect markers or genes associated with the trait. Possible approaches to deal with data potentially being a mixture of healthy and diseased animals include statistical modeling of SCS and detailed phenotyping. Statistical modeling of SCS involves models and SCS-derived traits that allow statistical properties to differ between healthy and diseased cows. Detailed phenotyping involves the possibility to further refine the SCS phenotype by differentiation of the cell types underlying SCS. Both approaches are discussed in the next two sections.

Modeling SCS

Lactation-average SCS and test-day SCS are frequently used traits in genetic selection for increased udder health, or, increased mastitis resistance. Lactation-average SCS and test-day SCS, however, may not optimally account for the dynamic nature of SCS caused by intra-mammary infections (Detilleux et al., 1997), as they either ignore variation or are not biologically related to mastitis infections (Windig et al., 2010). Therefore, alternative SCS traits have been suggested and studied by means of longitudinal SCS data. For instance, Franzén et al. (2012) proposed a method to account for changes in SCS during lactation by modeling transition probabilities between different states of mastitis (i.e. healthy or diseased), which considers both the mastitis liability and the recovery process. Including the entire disease course into the genetic evaluation offers possibilities for a substantial genetic gain (Franzén et al., 2012). A liability-normal mixture model presented by Madsen et al. (2008) allows individual animals to vary in the probability to develop mastitis. Results show differences in genetic parameters between SCS of healthy and diseased cows, and the model was suggested as a tool to select for liability of putative mastitis rather than SCS.

De Haas et al. (2003) found that pathogen-specific CM is genetically stronger correlated to patterns of peaks than to lactation-average SCS, suggesting an advantage of selection for patterns of peaks, rather than lactation-average SCS to reduce pathogen-specific CM. Heritabilities for the alternative SCS traits, however, were in the range of the heritability for lactation-average SCS (e.g. De Haas et al., 2008, De Haas et al., 2003) and also genetic correlations between alternative SCS traits and lactation-average SCS were high (De Haas et al., 2008). The high genetic correlations suggest that the alternative SCS traits and lactation-average SCS have a common genetic background. As such, when compared to lactation-average or test-day SCS, patterns of peaks, and maybe also other alternatively defined SCS traits, may provide limited additional information when the aim is to eventually

identify genes involved with udder disease to obtain a better understanding of the underlying biology.

Detailed phenotyping

Knowledge on the genetic control (e.g. heritability, correlated responses, and associated genes) of SCS may be obtained when phenotypes that describe the underlying physiological processes are measured. Somatic cell score is a combined trait, largely build up from four other traits, as somatic cells in milk are a blend of different cell types: polymorphonuclear neutrophils, macrophages, lymphocytes and, to a lesser extent, epithelial cells (Baumert et al., 2009). Although measurement of these four cell phenotypes requires additional effort compared to measurement of SCS, and therefore may not be feasible for a rather large number of animals, studying these four traits underlying SCS separately may provide insights that are not obtained when they are analyzed together as SCS because information on the individual traits may be diluted when they are combined.

In a healthy cow the majority of the somatic cells consist of macrophages and lymphocytes. Upon bacterial infiltration into the teat end an inflammatory response is elicited, which not only increases the number of cells, but also alters the predominant cell types. The inflammatory response generates a rapid influx of neutrophils from the blood stream into the udder. As a result, neutrophils become the predominant cell type (> 95%) in the infected udder (Kehrli and Shuster, 1994). As such, the somatic cell composition in healthy cows differs from the composition in diseased cows, which indicates that SCS in healthy and diseased cows requires a different biological interpretation.

Not only from a biological viewpoint, but also from a genetic perspective somatic cell score may be regarded as a different trait in healthy and diseased cows, as non-unity genetic correlations (0.78, 0.61) were found between SCS in healthy cows and cows with CM (Detilleux et al., 2009; Madsen et al., 2008). In sheep, the genetic correlation between SCS in healthy and diseased animals was 0.62 (Riggio et al., 2010). These findings suggest that SCS in healthy and diseased animals to some extent have a common genetic basis but that there are also genes involved in only healthy or diseased SCS (Detilleux et al., 2009). Genetic correlations between the different SCS cell types were, to my knowledge, not yet studied. The non-unity genetic correlation of SCS between healthy and diseased animals, however, suggests that these may differ from 1.

Heritability of disease outcome (healthy or diseased) or SCS are generally relatively low. Individual immune components, however, tend to show higher heritabilities. For instance, the heritability for the concentration of polymorphonuclear

neutrophils in bovine blood was estimated at 0.87 (Detilleux et al., 1994), heritabilities for adaptive immune components in porcine serum and plasma were above 0.40 for 15 out of 25 traits (Flori et al., 2011), and heritabilities for NAb levels ranged from 0.08 to 0.55 (chapter 2). Heritability estimates for NAb isotypes were generally equal to or higher than total NAb, the latter making no distinction between isotypes. These higher heritabilities for detailed phenotypes suggest a larger contribution of genetics to the phenotypic variance than is suggested by less detailed phenotypes. As such, less detailed phenotypes may provide an underestimated view on the role of genetics in disease resistance.

Chapter 3 studied the genetic control of NAb levels and not only included total NAb, but also NAb isotypes IgA, IgG1 and IgM. Each of these NAb levels was measured for four different binding specificities, resulting in 16 traits in total. For all binding specificities, few significant associations were identified for total NAb levels. Several, however, were detected for NAb isotypes, mainly for isotype IgM. The limited number of associations with total NAb could be caused by the trait definition. Total NAb are a blend of several isotypes. It was found that these isotypes do not all have a common genetic basis (chapter 2). As such, the number of genomic regions associated with total NAb is most likely larger than the number of genomic regions associated with individual NAb isotypes, whereas the power to detect them is most likely reduced. This reduction can be attributed to diluted effect of genes that affect a single isotype when not individual isotypes but total NAb are considered.

To summarize, statistical modeling and various SCS derived traits can provide useful information for selection. Currently, however, statistical models and SCS derived traits do not seem to be of additional value when compared to use of lactation-average or test-day SCS for detection of markers and genes associated with SCS when the aim is to obtain insight in the trait biology. Somatic cell score is composed of multiple cell types, which likely do not have the same genetic background. Studies (chapter 2 and 3) on NAb, which included not only the combined trait (total NAb) but also the individual component traits (NAb isotypes), suggest that analyses of detailed phenotypes rather than the combined trait can provide further insight in the genetic control of traits.

6.3. Maternal effects on antibodies

Maternal effects can be defined as maternal contributions to offspring phenotypes that go beyond the direct genetic contributions (Bernardo, 1996). Examples of

maternal contributions to offspring phenotypes include hormones, nutrients and antibodies (Hasselquist and Nilsson, 2009). Generally, maternal effects influence the offspring phenotype during early life, but in some instances their effects may be long-lasting. An enhanced immune response to an antigenic challenge in mice was found not only in the F1 generation but also the F2 generation, where only the grandmother received the challenge (Hasselquist and Nilsson, 2009). Further, maternally derived antibodies shape the antibody repertoire and B cell development in the offspring (Fink et al., 2008). Because maternal effects can exert a long-term effect on antibody phenotypes the relevance of accounting for maternal effects on NAb levels is assessed in this paragraph. Genetic parameters for NAb levels were estimated in chapter 2 using a model that did not take maternal effects into account. Presence of maternal effects, however, may affect results from genetic analyses by introducing a covariance between the maternally inherited genes of the offspring and the environment they experience. Therefore it is relevant to investigate whether or not this covariance exists. The first section of this paragraph focuses on the consequences of the presence of maternal effects on the estimation of genetic parameters, followed by a section on maternal antibodies. This second section first provides some physiological background on maternal antibodies, after which the possible involvement of maternal effects in NAb levels will be estimated and discussed.

Consequences of maternal effects for genetic parameters
Maternal effects can either be of maternal genetic or maternal environmental origin. In the case of a maternal genetic effect there is variation in the quality of the parental environment, where this variation has a genetic origin (Wolf et al., 1998). A dam can for instance have a genetic predisposition to be a good mother. Maternal effects can have relevant implications for breeding programs (Bernardo, 1996; Meyer, 1992). Models that solely account for direct genetic contributions to the phenotype, i.e. that ignore maternal effects, have the underlying assumption that the covariance between genotype and phenotype equals the additive genetic variance of the trait. When maternal effects affect the phenotype, offspring acquire not only the parental genes, but also (part of) the environment, which creates a covariance between the genes that they inherit and the environment that they experience (Wolf et al., 1998; Meyer, 1992). This covariance not only changes the relation between genotype and phenotype, but also alters the translation of genetic effects into the phenotype and hinders accurate estimation of heritabilities or additive genetic variances (Wolf et al., 1998). Thus, depending on the size and direction of the maternal effects, the response to selection in breeding programs

can differ from the expected response if maternal effects are present but unaccounted for in the statistical analysis (Wolf et al., 1998; Meyer, 1992).

Maternal antibodies

Background

Transfer of antibodies from mothers to offspring was already documented over a century ago (Ehrlich, 1892) and transfer of maternal antibodies can be defined as "the transfer of antibodies by an immunocompetent adult, typically a female, to an immunologically naïve neonate transplacentally or through colostrum, milk, yolk, etc." (Grindstaff et al., 2003). These maternal antibodies provide the neonate with passive immune protection which is fundamental in early life (Pastoret, 2007), as their immature adaptive immune system restricts them to an innate immune defense (Hasselquist and Nilsson, 2009). The maternal antibodies not only protect the neonate by coating pathogens which intervenes with pathogenicity, but also reduce the resource requirements on the neonate's immune system which secures availability of resources for growth or maintenance needs (Addison et al., 2009). The amount and repertoire of the maternal antibodies transmitted to offspring seem to reflect the maternal disease environment and to mirror the maternal systemic antibody population (Hasselquist and Nilsson, 2009).

In cattle, as in sheep, horses and pigs, antibodies are not transferred transplacentally (Pravieux et al., 2007). Rather, transmission of antibodies occurs via colostrum and milk, where the transmission depends on both the maternal secretion of antibodies into the colostrum or milk and the absorption of these secreted antibodies by the neonate (Baintner, 2007).

Antibodies in milk or colostrum may be blood-derived or they may be locally produced by plasma cells, the latter being pre-dominantly the case for isotype IgA (Stelwagen et al., 2009). Release of HCl and pepsinogen in the suckling stomach is limited, which allows passage of undegraded maternal antibodies into the small intestine. During 24 - 48 hours the bovine neonatal gut absorbs all kinds of proteins, including IgA, IgM and IgG, coupled with a negligible digestion of the antibodies. This short period that enables the neonate to absorb maternal antibodies is compensated for by an efficient absorption and a substantial gut length. The closure of the protein absorption is followed by a prolonged secretion of the absorbed antibodies in lymph and by a rapid digestion of newly acquired maternal proteins to cover the amino acid requirements of the neonate (Baintner, 2007). To optimize antibody absorption of neonates, calves should be fed one to one and a half liters colostrum directly (at the very most four to six hours (Heinrichs and Elizondo-Salazar, 2009)) after birth (Remmelink et al., 2012) and at least four

liters should be provided within the first eight hours after birth (Heinrichs and Elizondo-Salazar, 2009). Calves should be fed colostrum at least the first two days after birth. Preferably, calves are fed colostrum from cows that have been on the farm for some time and it is advised that farmers have some colostrum from older cows in storage, so that calves can always be provided with colostrum in case of insufficient production of the dam (Remmelink et al., 2012).

Maternal effects on Nab
It is yet unknown if NAb levels are affected by maternal effects, and, if so, to what extent levels are affected by maternal effects and whether levels are affected by maternal genetics or maternal environment. A limitation in analysis of maternal environmental effects is that unless embryo transfer or cross-fostering was applied, direct additive effects and maternal environmental effects are largely confounded (Meyer et al., 1992). Although the data is not optimal to estimate maternal effects, which were therefore not accounted for in the models of chapter 2 and 3, I tried to obtain some insight in whether maternal effects may be involved in antibody levels. To study whether maternal genetic effects affect NAb levels the model used to estimate genetic parameters for NAb levels, model 1 in chapter 2, was extended with a random dam effect which, in addition to the random animal effect, was linked to the pedigree:

$$
\begin{aligned}
y_{ijklmno} = \mu &+ b_1 \times \dim_{ijklmno} + b_2 \times e^{-0.05 \times \dim_{ijklmno}} + b_3 \times afc_{ijklmno} + b_4 \times \\
& afc_{ijklmno}^2 + season_k + stype_l + herd_m + animal_n + dam_o + e_{ijklmno}
\end{aligned}, \quad (1)
$$

where the response variable y was the NAb level; μ represented the population mean; $\dim_{ijklmno}$ was the covariate that defined the effect of lactation stage, modeled with a Wilmink curve (Wilmink, 1987) with regression coefficients b_1 and b_2; $afc_{ijklmno}$ was the covariate that accounted for the effect of age at first calving with regression coefficients b_3 and b_4; $season_k$ accounted for the fixed effect of season of calving in three categories: summer (June - August 2004), autumn (September – November 2004) and winter (December 2004 – February 2005); $stype_l$ was the fixed effect of sire type in three classes: proven bull, young bull or other proven bull. Unlike proven bulls, young bulls were still under evaluation regarding their genetic merit, to obtain at least three cows per herd the dataset included additional cows sired by other proven bulls; $herd_m$ was the random effect that accounted for farm effects; $animal_n$ was the random additive genetic effect of

the cow; dam_m was the random maternal genetic effect of the dam on the phenotype of the cow; and $e_{ijklmno}$ was the random residual effect.

The difference in log-likelihood between the extended model and the original model was tested for significance based on a χ_1^2 distribution. This was done for each individual NAb trait, as defined in chapter 2, as well as for fat, protein and milk yield. For none of the traits the extended model had a significantly better fit than the original model (results not shown). Therefore, I did not find evidence that NAb levels in milk are affected by maternal genetic effects.

An indication of maternal environmental effects on traits measured in cows may be obtained when the data are analyzed with a sire model and a dam model:

$$y_{ijklmn} = \mu + b_1 \times \dim_{ijklmn} + b_2 \times e^{-0.05 \times \dim_{ijklmn}} + b_3 \times afc_{ijklmn} + b_4 \times \\ afc_{ijklmn}^2 + season_k + stype_l + herd_m + sire_n + e_{ijklmn} \quad , \quad (2)$$

where model components were as formulated for model 1 and $sire_n$ was the random additive genetic effect of the sire, and

$$y_{ijklmn} = \mu + b_1 \times \dim_{ijklmn} + b_2 \times e^{-0.05 \times \dim_{ijklmn}} + b_3 \times afc_{ijklmn} + b_4 \times \\ afc_{ijklmn}^2 + season_k + stype_l + herd_m + dam_n + e_{ijklmn} \quad , \quad (3)$$

where model components were as formulated for model 1 and dam_n was the random additive genetic effect of the dam.

When maternal effects do not affect the concerning trait, the additive genetic variance component estimated with the sire model and with the dam model is expected to be of the same approximate magnitude:

$$\sigma_a^2 = \sigma_{sire}^2 + \sigma_{dam}^2 + MS \quad , \quad (4)$$

where σ_a^2 is the additive genetic variance, σ_{sire}^2 is the part of the additive genetic variance explained by the sire (i.e. $\frac{1}{4}\sigma_a^2$ in the absence of maternal effects), σ_{dam}^2 is the part of the additive genetic variance explained by the dam (i.e. $\frac{1}{4}\sigma_a^2$ in the absence of maternal effects), and MS is the Mendelian sampling term. When the

part of the additive genetic variance explained by the dam is substantially larger than that of the sire, this could suggest presence of maternal effects. Additive genetic variances for NAb levels obtained with an animal model, a sire model and a dam model are shown in Table 1.

Table 6.1 Additive genetic variances for NAb levels obtained with an animal model (σ_a^2), a sire model (σ_{sire}^2), and a dam model (σ_{dam}^2) for total natural antibody levels (T) and natural antibody isotype immunoglobulin (Ig)A, IgG1 and IgM levels binding keyhole limpet hemocyanin (KLH), lipopolysaccharide (LPS), lipoteichoic acid (LTA), or peptidoglycan (PGN).[1] Corresponding standard errors are in parentheses. The P-value reflects the significance of the difference in log-likelihood (DLog-L) between a model that estimated σ_{dam}^2 and a model that was fixed to the expected σ_{dam}^2 (= $\frac{1}{4}\sigma_a^2$).

Trait	σ_a^2	$4*\sigma_{sire}^2$ [2]	$4*\sigma_{dam}^2$ [2]	DLog-L	P-value
KLH-T	0.20 (0.05)	0.17 (0.01)	0.19 (0.03)	10.10	<0.01
KLH-IgA	0.47 (0.12)	0.35 (0.03)	0.70 (0.07)	17.18	<0.01
KLH-IgG1	0.68 (0.18)	0.70 (0.05)	0.24 (0.07)	0.44	0.51
KLH-IgM	0.41 (0.11)	0.28 (0.02)	0.79 (0.07)	28.58	<0.01
LPS-T	0.12 (0.05)	0.09 (0.01)	0.29 (0.04)	4.38	0.04
LPS-IgA	0.64 (0.16)	0.51 (0.04)	0.59 (0.07)	18.34	<0.01
LPS-IgG1	0.42 (0.15)	0.48 (0.05)	0.09 (0.06)	0.28	0.60
LPS-IgM	0.41 (0.11)	0.26 (0.02)	0.66 (0.07)	21.08	<0.01
LTA-T	0.23 (0.07)	0.17 (0.02)	0.42 (0.05)	13.52	<0.01
LTA-IgA	0.52 (0.13)	0.40 (0.03)	0.67 (0.07)	17.68	<0.01
LTA-IgG1	0.11 (0.08)	0.08 (0.02)	0.39 (0.07)	4.14	0.04
LTA-IgM	0.31 (0.08)	0.22 (0.02)	0.60 (0.05)	26.22	<0.01
PGN-T	0.10 (0.05)	0.11 (0.01)	0.13 (0.03)	0.90	0.34
PGN-IgA	0.25 (0.08)	0.10 (0.02)	0.52 (0.05)	23.56	<0.01
PGN-IgG1	0.24 (0.11)	0.32 (0.03)	0.02 (0.05)	0.00	1
PGN-IgM	0.29 (0.08)	0.15 (0.02)	0.71 (0.06)	28.90	<0.01
Fat yield	0.24 (0.06)	0.18 (0.01)	0.26 (0.14)	0.00	1
Milk yield	2.15 (0.62)	1.77 (0.15)	6.23E-07[3] (0.63E-08)	0.14	0.71
Protein yield	0.04 (0.01)	0.03 (0.002)	1.72E-09[3] (0.17^E-10)	0.10	0.75

[1] Antibody names are a combination of antigen abbreviation and immunoglobulin isotype

[2] Values were multiplied by 4 for ease of comparison to σ_a^2

[3] LogL not converged because the parameter value was close to boundary

For all but three of the NAb traits, it was found that the additive genetic variance explained by the dam was higher than the additive genetic variance explained by the sire. All three traits with higher sire variance were responses to IgG1. This was

not found for the production traits fat, milk and protein yield. For fat yield, the additive genetic variance explained by the dam was of the same approximate size as what was found for with the sire and animal model. For milk and protein yield the estimated genetic variance explained by the dam was close to zero, whereas the residual variance and herd variance for these two traits were higher when the dam model was applied than when the animal model was applied. This suggests that for milk and protein yield the dam model is not able to separate the individual variance components. To test whether the additive genetic variance explained by the dam differed significantly (according to χ^2_1 distribution) from her expected contribution (= $\frac{1}{4}\sigma^2_a$) the log-likelihood of the dam model was compared to the log-likelihood of an 'expected contributions' model. In this 'expected contributions' model the additive genetic variance explained by the dam was fixed to the ratio of the expected contribution, $\frac{1}{4}\sigma^2_a$, over σ^2_e where σ^2_a was obtained from the animal model and σ^2_e was obtained from the dam model. The P-values corresponding to the difference between the dam model and the 'expected contributions' model are provided in table 1. With five exceptions, additive genetic variances explained by the dam significantly differed from the expected contributions, which could indicate that the dam contributes more to the antibody phenotype than what is directly inherited by the offspring. Herd variances and phenotypic variances were of the same approximate magnitude for the dam model, the sire model and the animal model. Among the five non-significant exceptions were three out of four IgG1 levels. This may be due to the biology of IgG1. The antigens lipopolysaccharide, lipoteichoic acid and peptidoglycan are present on bacteria commonly present in the environment. Cows may have been previously exposed to these antigens. As such, antibodies measured against these antigens may partly find their origin in the adaptive immune system limiting the potential for maternal effects. Isotype IgG is of main importance in the adaptive immune response (Schroeder and Cavacini, 2010). Therefore, it is likely that the environment of the animal itself is much more important for the antibody levels than maternal effects. The importance of the animal's own environment for IgG levels is supported by results from chapter 2, which show that IgG levels are more dependent on environmental factors than IgA and IgM levels.

To summarize, maternal effects can have a maternal genetic and a maternal environmental origin. No evidence was found that NAb levels in milk are affected by maternal genetic effects. Maternal environmental effects, however, could play a

role in NAb levels as analyses suggested that in general dams contribute more to the offspring antibody phenotype than what was expected based on additive genetics alone. Heritabilities for most NAb levels (chapter 2) may therefore somewhat overestimated, but as the data structure is not optimal to estimate maternal effects it is not possible to determine the extent of the overestimation. These results warrant further study for the influence of maternal effects on NAb levels.

6.4. Genetic diversity and selection at the major histocompatibility complex

Chapter 3 identified the major histocompatibility complex (MHC), located on bovine autosome 23, as a candidate region with an important effect on the regulation of NAb isotype IgM levels. This information could be used in selective breeding by selecting for specific MHC haplotypes. The MHC plays a central role in the immune response and regulation, as suboptimal functioning of the MHC hinders an effective immune response and makes further infection likely (Knapp, 2005). The immune system is a highly interactive system, which functions through a cascade of responses that are directly or indirectly linked to each other, some more than others. It therefore seems almost inevitable that genetic selection for one immunological trait leads to correlated responses in other immunological traits, especially when selection acts on major genes of the immune system, such as the MHC. The MHC contains the most polymorphic genes in the vertebrate genome (e.g. Knapp, 2005; Bernatchez and Landry, 2003; Hughes, 2002), with co-dominant expression of alleles (Knapp, 2005). Several (evolutionary) processes may act to maintain this genetic diversity, including balancing selection (e.g. Bernatchez and Landry, 2003), recombination (e.g. Carrington, 1999) and sexual selection by mate choice (e.g. Knapp, 2005). Both the central role of the MHC in the immune system and the possible evolutionary aim to maintain genetic diversity at the MHC region may have consequences for breeding schemes. For instance, the central role of the MHC may lead to undesired correlated responses or selective breeding may be counteracted by natural selection. To prevent negative outcomes for the animal and the breeding scheme, it is important to obtain knowledge on the biology of the MHC region before pursuing genetic selection on this region. This paragraph about the MHC will first provide a short background in the possible forces acting on maintenance of genetic diversity in the MHC region, after which possible implications of selective breeding on MHC diversity are discussed.

Balancing selection

Maintenance of genetic diversity in the MHC may originate from host-pathogen co-evolution (e.g. Knapp, 2005), where two basic co-evolution types exist. The first type is the overdominance, or heterozygote advantage, where MHC heterozygote individuals have an advantage over homozygotes because they have the ability to resist a broader array of pathogens (Bernatchez and Landry, 2003; Hughes and Yeager, 1998). The second type is negative frequency dependent selection, where an inability of the most common host MHC genotypes to respond to one or more pathogens decreases the relative fitness of these genotypes. This provides a selective advantage to relatively rare MHC alleles. The time-lag character of such co-evolutionary responses could result in constant shifts in fitness values of different genotypes in both hosts and pathogens, and could thereby provide in the maintenance of a high genetic diversity (Bernatchez and Landry, 2003).

One of the two most frequently used methods to determine whether balancing selection is acting upon the MHC region is the *dn/ds* ratio test, which compares the number of non-synonymous (amino acid-altering) mutations to the number of synonymous mutations (e.g. Bernatchez and Landry, 2003; Hughes, 2002; Hughes and Yeager, 1998). In general, most non-synonymous mutations negatively affect protein structure and function and therefore natural selection rapidly eliminates such mutations. The *dn/ds* ratio, however, showed strong evidence of natural selection favoring non-synonymous mutations in the MHC region (Bernatchez and Landry, 2003; Hughes, 2002).

Recombination

Recombination events during meiosis allow for re-arrangements of genetic material which generates novel combinations of alleles in the offspring. In this way, recombination aids in maintaining genetic diversity and continuing genomic evolution (Carrington, 1999). Recombination, however, seems to occur in specific locations. In the murine MHC, four regions with high recombination rates were identified, encompassing several 'hotspots' (Carrington, 1999). Other regions show strong linkage disequilibrium (Carrington, 1999; White, 1989). This strong linkage disequilibrium suggests that in some instances insufficient time has elapsed for recombination to break up these regions (White, 1989), but it could also suggest that these regions have a selective advantage and recombination is suppressed by some mechanism (Carrington, 1999; White, 1989). Such suppression of recombination provides opportunities for polymorphisms to be retained for a substantial longer period of time than expected.

Sexual selection by mate choice

Another strategy to maintain genetic variation in the MHC is sexual selection by mate choice.

Sexual selection most likely has evolved for two reasons: inbreeding avoidance and increased disease resistance. The relative contribution of these two factors, however, is unknown (Potts et al., 1994). Although inbreeding avoidance driven as well as disease resistance driven sexual selection both aim to increase offspring fitness, the starting-point of these two forms of sexual selection differs. Inbreeding avoidance driven sexual selection tries to avoid mating of related individuals who are more likely to have certain (MHC) alleles in common (Bernatchez and Landry, 2003). This probably functions to increase overall fitness of individuals, rather than only disease resistance. Sexual selection driven by disease resistance, however, functions primarily to increase disease resistance in progeny. Choosy parents may have increased fitness for disease resistance because a disproportionate number of offspring have a heterozygote genotype (Bernatchez and Landry, 2003), which are likely able to bind a broader array of pathogens compared to offspring with a homozygote genotype (Bernatchez and Landry, 2003; Hughes and Yeager, 1998). Further, mating preferences may provide a means for a population to 'keep up' with the evolution of pathogens (Bernatchez and Landry, 2003), which evolve due to mutations.

It is hypothesized that sexual selection by mate choice functions through MHC-related olfactory signaling, which was found in a diversity of species, ranging from sticklebacks (Reusch et al., 2001) to humans (Chaix et al., 2008). More research, however, is needed to unravel the relationship between odor and MHC genes (Knapp, 2005).

Consequences of livestock production for MHC diversity

A key characteristic of the MHC is its high level of polymorphism (e.g. Bernatchez and Landry, 2003). Genetic diversity in the MHC region is maintained by natural selection through forces including balancing selection, recombination, and sexual selection by mate choice. Selective breeding may reduce genetic variation in a trait. This reduction in genetic variation may influence the ability of a population to cope with novel circumstances (Neeteson-van Nieuwenhoven et al., 2012) such as novel pathogens. Negative effects associated with reduced genetic variation in the MHC region have been reported in some species. For instance, it was shown recently that female Magellanic penguins that were homozygous for DRβ1 genes, in contrast to heterozygote females, did not fledge any chicks (Knafler et al., 2012). Negative effects of increased homozygosity in the MHC region were also suggested

by studies in humans which indicate that some homozygote offspring are selectively aborted. Further, in pigtailed macaques reproductively unsuccessful mates had a higher amount of MHC alleles in common than reproductively successful mates (Knapp, 2005).

The MHC diversity and dynamics may be affected by selective breeding on the MHC region, either via direct selection or indirect though correlated responses. It is unknown how or to what extent natural selection counteracts selective breeding or how and to what extent selective breeding affects the contribution of natural selection to genetic diversity in the MHC region. Further, it is unknown how important MHC diversity still is in current husbandry. In the remaining part of this paragraph I will briefly discuss how selective breeding may affect genetic diversity of the MHC and the role of natural selection in maintenance of diversity, followed by some suggestions to study MHC diversity in current livestock production.

The most obvious effect of selective breeding on the three described forces of natural selection is intervention in sexual selection by mate choice. One reason for the evolution of mate choice is inbreeding avoidance (Potts et al., 1994). In livestock production mate choice is a human decision, which may lead to a choice that differs from the choice the animals would have made themselves in the absence of selective breeding. In selective breeding it is common to pose restrictions on the relatedness between mates based on the average inbreeding coefficient. In this way we, unconsciously, allow for genetic diversity of the MHC as less related individuals are more likely to have fewer alleles in common. Under influence of natural selection, however, animals might specifically try to avoid inbreeding at the MHC locus rather than avoid average inbreeding.

Farm management practices may have reduced the importance of MHC diversity and the contribution of balancing selection. Hygienic measures can eliminate or diminish the risk of infection, whereas treatment and vaccination can prevent or diminish a reduction in fitness of infected animals that would have experienced a reduced fitness if these measures were not provided. As a result of these measures, our livestock populations may have a less dynamic MHC than they would have under the sole influence of natural selection.

The process of recombination is the least affected by current livestock production, as it is a process that occurs during meiosis (Carrington, 1999), which is the least likely to be under human control. The process of recombination will remain present every meiosis to generate new combinations of (MHC) alleles and as such contribute to genetic diversity. With more homozygosity in the MHC region of a population, however, fewer opportunities remain for meiosis to create novel combinations.

It is probable that MHC diversity has become less important for our livestock than it is for natural populations, as farm management practices may compensate for the consequences of reduced MHC diversity. Diversity could become, however, more important when the aim is to breed for more robust animals, i.e. animals that are more resistant to disease, require less treatment and that require less labour. A reduced importance of diversity does not mean that maintenance of diversity is not relevant in current livestock production. A large amount of allelic loss could result in negative effects, such as fertility problems (e.g. Knapp, 2005), that cannot be compensated for by treatments and vaccinations. In contrast, if there is indeed an evolutionary strive for maintenance of genetic diversity in the MHC region forces of natural selection may counteract selective breeding for this region.

Insight in the effects of human interference (e.g. selective breeding, husbandry) on MHC diversity can be obtained by comparing the diversity of the MHC region in bovine populations that differ in the extent of human interference. If MHC diversity is important for population fitness and survival, it can be expected that populations with and without human intervention are fairly similar in MHC diversity. Further, such comparative studies can provide insight in trans-species polymorphisms and recombination hotspots. Studies using MHC genotypes may provide insight in problems associated with a large amount of homozygosity the MHC region or with homozygosity in individual MHC genes. Such studies can aid in risk-assessment for loss of MHC diversity and provide insight in the importance of preserving MHC diversity in breeding schemes.

To summarize, in chapter 3 of this thesis I showed that the MHC is an important region affecting NAb levels in milk. Results from previous studies showed that natural selection strives for high genetic diversity of the MHC region by for instance balancing selection, recombination and mate choice. Negative effects of reduced MHC diversity have been reported for animal fitness, which means that direct or indirect selection on MHC can impose a risk. Reduced diversity, however, may be counteracted by natural selection. Although MHC diversity may have become less important in current animal husbandry due to farm management practices such as hygienic measures, treatment and vaccination, this does not mean that maintenance of MHC diversity is not relevant in current livestock production. Loss in MHC diversity may result in negative effects on animal health which cannot be compensated for by treatments and vaccinations. Further research, however, is needed to obtain more insight on MHC diversity and possible negative effects of loss of diversity.

6.5. Thesis conclusions

This thesis showed that genetic factors substantially contribute to innate immunity. Natural antibody levels are heritable traits and a common genetic basis was found for NAb isotypes IgA and IgM. Findings from chapters 2 and 3 indicate that, when compared to total NAb, isotypes are more detailed phenotypes that allow for a better insight in different elements of the immune response or immune competence. Genomic regions associated with NAb levels were identified on all bovine autosomes and a region on chromosome 23 was found to have an important effect on levels of isotype IgM. For SCS (chapter 4), however, few associated genomic regions were found. Chapter 5 detected macro- and micro-environmental sensitivity in test-day SCS, which indicates that a cow's sensitivity to changes in her environment is affected by genetics. Further, chapter 5 showed that macro- and micro-environmental sensitivity are positively correlated. The present chapter (chapter 6) discussed some of the biology behind NAb and SCS, as trait biology may affect genetic parameter estimation, identification of candidate genes and breeding programs. This chapter contained three conclusions: firstly, analyses of cell types (detailed phenotypes) rather than SCS can provide further insight in the genetic control of SCS and mastitis. Secondly, no evidence was found for maternal genetic effects on NAb levels in milk. Maternal environmental effects, however, could play a role in NAb levels. Finally, genetic diversity in the MHC region is maintained by natural selection. Selective breeding and farm management practices may affect this genetic diversity, which could bring about negative effects on animal fitness, such as fertility problems. Selective breeding for specific MHC haplotypes may therefore impose a risk for negative effects on animal health.

References

A

Addison, B., Klasing, K.C., Douglas Robinson, W., Austin, S.H., and R.E. Ricklefs. 2009. Ecological and life-history factors influencing the evolution of maternal antibody allocation: a phylogenetic comparison. Proceedings of the Royal Society B: Biological Sciences. 276: 3979-3987.

Ahmed, S.M., Daulat, A.M., Meunier, A., and S. Angers. 2010. G protein βγ subunits regulate cell adhesion through Rap1a and its effector Radil. Journal of Biological Chemistry. 285: 6538-6551.

Akdis M., Palomares, O., Van de Veen, W., Van Splunter, M., and C.A. Akdis. 2012. TH17 and TH22 cells: A confusion of antimicrobial response with tissue inflammation versus protection. Journal of Allergy and Clinical Immunology. 129: 1438-1449.

Ali, A.K.A., and G.E. Shook. 1980. An optimum transformation for somatic cell concentration in milk. Journal of Dairy Science. 63: 487-490.

Amills M., Ramiya, V., Norimine, J., and H.A. Lewin. 1998. The major histocompatibility complex of ruminants. OIE Revue Scientifique et Technique. 17: 108-120.

B

Baintner, K. 2007. Transmission of antibodies from mother to young: evolutionary strategies in a proteolytic environment. Veterinary Immunology and Immunopathology. 117: 153-161.

Bannerman, D.D., Paape, M.J., Lee, J., Zhao, X., Hope J.C., and P. Rainard. 2004. Escherichia coli and Staphylococcus aureus elicit differential innate immune response following intramammary infection. Clinical and Diagnostic Laboratory Immunology. 11: 463-472.

Barkema H.W., Schukken, Y.H., Lam, T.J.G.M., Beiboer, M.L., Wilmink, H., Benedictus, G., and A. Brand. 1998. Incidence of clinical mastitis in dairy herds grouped in three categories by bulk milk somatic cell counts. Journal of Dairy Science. 81: 411-419.

Baumert, A., Bruckmayer, R.M., and O. Wellnitz. 2009. Cell population, viability, and some key immunomodulatory molecules in different milk somatic cell samples in dairy cows. Journal of Dairy Research. 76: 356-364.

Baumgarth, N., Tung, J.W., and L.A. Herzenberg. 2005. Inherent specificities in natural antibodies: a key to immune defense against pathogen invasion. Springer Seminars in Immunopathology. 26: 347-362.

Beutler, B. 2004. Innate immunity: an overview. Molecular Immunology. 40: 845-859.

Bell, M.J., Wall, E., Russell, G., Simm, G., and A.W. Stott. 2011. The effect of improving cow productivity, fertility and involuntary culling on the global warming potential of dairy systems. Journal of Dairy Science. 94: 3662-3678.

Bennewitz, J., Reinsch, N., Grohs, C., Levéziel, H., Malafosse, A., Thomsen, H., Xu, N., Looft, C., Kühn, C., Brockmann, G.A., Schwerin, M., Weinmann, C., Hiendleder, S., Erhardt, G., Medjugorac, I., Russ, I., Förster, M., Brenig, B., Reinhardt, F., Reents, R., Averdunk, G., Blümel, J., Boichard, D., and E. Kalm. 2003. Combined analysis of data from two granddaughter designs: a simple strategy for QTL confirmation and increasing experimental power in dairy cattle. Genetics Selection Evolution. 35: 319-338.

Bernardo, J. 1996. Maternal effects in animal ecology. American Zoologist. 36: 83-105.

Bernatchez, L., and C. Landry. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? Journal of Evolutionary Biology. 16: 363-277.

Biscarini, F., Bovenhuis, H., Van Arendonk, J.A.M., Parmentier, H.K., Jungerius, A.P., and J.J. van der Poel. 2010. Across-line SNP association study of innate and adaptive immune response in laying hens. Animal Genetics. 41: 26-38.

Bishop, S.C., and J.A. Woolliams. 2010. On the genetic interpretation of disease data. PLoS ONE. 5: e8940. doi:10.1371/journal.pone.0008940

Bouwman, A.C., Bovenhuis, H., Visker, M.H.P.W., and J.A.M. van Arendonk. 2011. Genome-wide association of milk fatty acids in Dutch dairy cattle. BMC Genetics. 12: 43. doi: 10.1186/1471-2156-12-43

Bouwman, A.C., Visker, M.H.P.W., Van Arendonk, J.A.M., and H. Bovenhuis. 2012. Genomic regions associated with bovine milk fatty acids in both summer and winter milk samples. BMC Genetics. 13: 93. doi:10.1186/1471-2156-13-93

Burton, J.L., Kennedy, B.W., Burnside, E.B., Wilkie, B.N., and J.H. Burton. 1989. Variation in serum concentrations of immunoglobulins G, A, and M in Canadian Holstein-Friesian calves. Journal of Dairy Science. 72: 135-149.

Burvenich, C., Van Merris, V., Mehrzad, J., Diez-Fraile, A., and L. Duchateau. 2003. Severity of E. coli mastitis is mainly determined by cow factors. Veterinary Research. 34: 521-564.

C

Calus, M.P.L., Windig, J.J., and R.F. Veerkamp. 2005. Associations among descriptors of herd management and phenotypic and genetic levels of health and fertility. Journal of Dairy Science. 88: 2178-2189.

Calus, M.P.L., Janss, L.L.G., and R.F. Veerkamp. 2006. Genotype by environment interaction of somatic cell score across bulk milk somatic cell count and days in milk. Journal of Dairy Science. 89: 4846-4857.

Calus, M.P.L., Mulder, H.A., and J.W.M. Bastiaansen. 2011. Identification of Mendelian inconsistencies between SNP and pedigree information of sibs. Genetics Selection Evolution. 43: 34

Carlén, E., Strandberg, E., and A. Roth. 2004. Genetic parameters for clinical mastitis, somatic cell score, and production in the first three lactations of Swedish Holstein cows. Journal of Dairy Science. 87: 3062-3070.

Carrington., M. 1999. Recombination within the human MHC. Immunological Reviews. 167: 245-256.

Casali P., and A.L. Notkins. 1989. CD5+ B lymphocytes, polyreactive antibodies and the human B-cell repertoire. Immunology Today. 10: 364-368.

Chaix, R., Cao, C., and P. Donnelly. 2008. Is mate choice in humans MHC-dependent? PLoS Genetics. 4: e1000184.

Cordell, H.J., and D.G. Clayton. 2005. Genetic association studies. Lancet. 366: 1121-1131.

Covic, L., Gresser, A.L., and A. Kuliopulos. 2000. Biphasic Kinetics of Activation and Signaling for PAR1 and PAR4 Thrombin Receptors in Platelets. Biochemistry. 39: 5458-5467.

D

Daetwyler, H.D., Schenkel, F.S., Sargolzaei, M., and J.A.B. Robinson. 2008. A genome scan to detect quantitative trait loci for economically important traits in Holstein cattle using two methods and a dense single nucleotide polymorphism map. Journal of Dairy Science. 91: 3225-3236.

References

De Haas, Y., Barkema, H.W., and R. F. Veerkamp. 2002. Effect of pathogen-specific clinical mastitis on lactation curves for somatic cell count. Journal of Dairy Science. 85: 1314-1323.

De Haas, Y., Barkema, H.W., Schukken, Y.H., and R.F. Veerkamp. 2003. Genetic associations for pathogen-specific clinical mastitis and patterns of peaks in somatic cell count. Animal Science. 77: 187-195.

De Haas, Y., Ouweltjes, W., Ten Napel, J., Windig, J.J., and G. de Jong. 2008. Alternative somatic cell count traits as mastitis indicators for genetic selection. Journal of Dairy Science. 91: 2501-2511.

De Jong, G., and P. Bijma. 2002. Selection and phenotypic plasticity in evolutionary biology and animal breeding. Livestock Production Science. 78: 195-214.

Dekkers, J.C.M. 2004. Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. Journal of Animal Science. 82(E. Suppl.): E313–E328.

De Roos, A.P.W., Hayes, B.J., Spelman, R.J., and M.E. Goddard. 2008. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. Genetics. 179: 1503-1512.

Detilleux, J.C., Koehler, K.J., Freeman, A.E., Kehrli Jr., M.E., and D.H. Kelley. 1994. Immunological parameters of periparturient Holstein cattle: genetic variation. Journal of Dairy Science. 77: 2640–2650.

Detilleux, J., Leroy, P., and D. Volckaert. 1997. Alternative use of somatic cell counts in genetic selection for mastitis resistance. Proceedings of the international workshop on genetic improvement of functional traits in cattle; health, Uppsala, Sweden, pages 34-44.

Detilleux, J.C. 2009. Genetic factors affecting susceptibility to udder pathogens. Veterinary Microbiology 134: 157-164.

Deugnier, M., Moiseyeva, E.P., Thiery, J.P., and M. Glukhova. 1995. Myoepithelial cell differentiation in the developing mammary gland: progressive acquisition of smooth muscle phenotype. Developmental Dynamics. 204: 107-117.

Dietz, A.B., Detilleux, J.C., Freeman, A.E., Kelley, D.H., Stabel, J.R., and M.E. Kehrli Jr. 1997. Genetic association of bovine lymphocyte antigen DRB3 alleles with immunological traits of Holstein cattle. Journal of Dairy Science. 80: 400-405.

Douch, P.G.C., Green, R.S., Morris, C.A., and S.M. Hickey. 1995. Genetic factors affecting antibody responses to four species of nematode parasite in Romney ewe lambs. International Journal for Parasitology. 25: 823-828.

Downes, G.B., and N. Gautam. 1999. The G protein subunit gene families. Genomics. 62: 544-552.

Duah N.O., Weiss H.A., Jepson A., Tetteh K.K.A., Whittle H.C., and D.J. Conway. 2009. Heritability of antibody isotype and subclass responses to Plasmodium falciparum antigens. PLoS ONE 4, e7381. doi:10.1371/journal.pone.0007381

E

Ehrenstein M.R., and C.A. Notley. 2010. The importance of natural IgM: scavenger, protector and regulator. Nature Reviews Immunology. 10: 778-786.

Ehrlich, P. 1892. Ueber Immunität durch Vererbung und Säugung. Zeitschrift für Hygiene und Infectionskrankheiten 12: 183-203.

Ekstrand, J., Razuvaev, A., Folkersen, L., Roy, J., and U. Hedin. 2010. Tissue factor pathway inhibitor-2 is induced by fluid shear stress in vascular smooth muscle cells and affects cell proliferation and survival. Journal of Vascular Surgery. 52: 167-175.

F

Falconer, D.S., and T.F.C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Pearson Education Limited, Essex, United Kingdom.

Fesel, C., and A. Coutinho. 1999. Structured reactions of serum IgM repertoires to immunization are dependent on major histocompatibility complex genes. Scandinavian Journal of Immunology. 49: 251-257.

Fink, K., Zellweger, R., Weber, J., Manjarrez-Orduno, N., Holdener, M., Senn, B.M., Hengartner, H., Zinkernagel, R.M., and A.J. Macpherson. 2008. Long-term maternal imprinting of the specific B cell repertoire by maternal antibodies. European Journal of Immunology. 38: 90-101.

Fisher, M.W., and D.J. Mellor. 2008. Developing a systemic strategy incorporating ethical, animal welfare and practical principles to guide the genetic improvement of dairy cattle. New Zealand Veterinary Journal 56: 100-106.

Flori, L., Gao, Y., Laloë, D., Lemonnier, G., Leplat, J., Teillaud, A., Cossalter, A., Laffitte, J., Pinton, P., De Vaureix, C., Bouffaud, M., Mercat, M., Lefèvre, F., Oswald, I.P., Bidanel, J., and C. Rogel-Gaillard. 2011. Immunity traits in

pigs: substantial genetic variation and limited covariation. PLoS ONE. 6: e22717.

Franzén, J., Thorburn, D., Urioste, J.I., and E. Strandberg. 2012. Genetic evaluation of mastitis liability and recovery through longitudinal analysis of transition probabilities. Genetics Selection Evolution. 44: 10.

G

Gasbarre L.C., Leighton E.A., and C.J. Davies. 1993. Influence of host genetics upon antibody responses against gastrointestinal nematode infections in cattle. Veterinary Parasitology. 46: 81-91.

Genini, S., Badaoui, B., Sclep, G., Bishop, S.C., Waddington, D., Pinard van der Laan, M., Klopp, C., Cabau, C., Seyfert, H., Petzl, W., Jensen, K., Glass, E.J., De Greeff, A., Smith, H.E., Smits, M.A., Olsaker, I., Boman, G.M., Pisoni, G., Moroni, P., Castiglioni, B., Cremonesi, P., Del Corvo, M., Foulon, E., Foucras, G., Rupp, R., and E. Giuffra. 2011. Strengthening insights into host responses to mastitis infection in ruminants by combining heterogeneous microarray data sources. BMC Genomics. 12: 225. doi: 10.1186/1471-2164-12-225

Gilmour A.R., Gogel, B.J., Culis, B.R., and R. Thompson. 2006. ASReml User Guide. Release 2.0. VSN International Ltd., HP1 1ES, Hemel Hempstead, UK.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., and R. Thompson. 2009. ASReml User Guide Release 3.0. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Glass, E.J. 2004. Genetic variation and responses to vaccines. Animal Health Research Reviews 5: 197-208.

Glass, E.J., Baxter, R., Leach, R.J., and O.C. Jann. 2012. Genes controlling vaccine responses and disease resistance to respiratory viral pathogens in cattle. Veterinary Immunology and Immunopathology. 148: 90-99.

Goff, J.P. 2006. Major advances in our understanding of nutritional influences on bovine health. Journal of Dairy Science. 89: 1292-1301.

Gonda, M.G., Chang, Y.M., Shook, G.E., Collins, M.T., and B.W. Kirkpatrick. 2006. Genetic variation of Mycobacterium avium ssp. Paratuberculosis in US Holsteins. Journal of Dairy Science. 89: 1804-1812.

Green, M.J., Green, L.E., Schukken, Y H., Bradley, A.J., Peeler, E.J., Barkema, H.W., De Haas, Y., Collis, V.J., and G.F. Medley. 2004. Somatic cell count distributions during lactation predict clinical mastitis. Journal of Dairy Science. 87: 1256-1264.

Grindstaff, J.L., Brodie III, E.D., and E.D. Ketterson. 2003. Immune function across generations: integrating mechanism and evolutionary process in maternal antibody transmission. Proceedings of the Royal Society B: Biological Sciences. 270: 2309-2319.

H

Hasselquist, D., and J. Nilsson. 2009. Maternal transfer of antibodies in vertebrates: trans-generational effects on offspring immunity. Philosophical Transactions of the Royal Society B: Biological Sciences. 364: 51-60.

Hayes, B.J., Carrick, M., Bowman, P., and M.E. Goddard. 2003. Genotype x environment interaction for milk production of daughters of Australian dairy sires from test-day records. Journal of Dairy Science. 86: 3736-3744.

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

Harriman G.R., Bogue, M., Rogers-Fani, P., Finegold, M., Bradley, A., and S. Pacheco. 1996. IgA deficient mice show altered immunoglobulin isotype expression and mucosal antibody responses. FASEB Journal. 10: A1418.

Heinrichs, A.J., and J.A. Elizondo-Salazar. 2009. Reducing failure of passive immunoglobulin transfer in dairy calves. Revue de Medecine Veterinaire. 160: 436-440.

Heringstad, B., Sehested, E., and T. Steine. 2008. Short communication: Correlated selection responses in somatic cell count from selection against clinical mastitis. Journal of Dairy Science. 91: 4437-4439.

Heyen, D.W., Weller, J.I., Ron, M., Band, M., Beever, J E., Feldmesser, E., Da, Y., Wiggans, G.R., VanRaden, P.M., and H.A Lewin. 1999. A genome scan for QTL influencing milk production and health traits in dairy cattle. Physiological Genomics. 1: 165-175.

Hill, W.G., and H.A. Mulder. 2010. Genetic analysis of environmental variation. Genet. Res. Camb. 92: 381-395.

Hirshorn, J.N., and M.J. Daly. 2005. Genome-wide association studies for common diseases and complex traits. Nature Reviews Genetics. 6: 95-108.

Hu, Z., and J.M. Reecy. 2007. Animal QTLdb: beyond a repository. A public platform for QTL comparisons and integration with diverse types of structural genomic information. Mammalian Genome. 18: 1-4.

Hughes, A.L., and M. Yeager. 1998. Natural selection at major histocompatibility complex loci of vertebrates. Annual Review of Genetics. 32: 415-435.

Hughes, A.L. 2002. Natural selection and the diversification of vertebrate immune effectors. Immunological reviews. 190: 161-168.

J

Jerne, N.K. 1985. The generative grammar of the immune system. Science. 13: 1057-1059.

K

Kehrli, M.E. Jr, and D.E. Shuster. 1994. Factors affecting milk somatic cells and their role in health of the bovine mammary gland. Journal of Dairy Science. 77: 619-627.

Kelm S.C., Detilleux J.C., Freeman A.E., Kehrli Jr. M.E., Dietz A.B., Fox L.K., Butler J.E., Kasckovics I., and D.H. Kelley. 1997. Genetic association between parameters of innate immunity and measures of mastitis in periparturient Holstein cattle. Journal of Dairy Science. 80: 1767-1775.

Khatkar, M.S., Thomson, P.C., Tammen, I., and H.W. Raadsma. 2004. Quantitative trait loci mapping in dairy cattle: review and meta-analysis. Genetics Selection Evolution. 36: 163-190.

Khatkar, M.S., Nicholas, F.W., Collins, A.R., Zenger, K.R., Cavanagh, J.A.L., Barris, W., Schnabel, R.D., Taylor, J.F., 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. doi: 10.1186/1471-2164-9-187

Klungland, H., Sabry, A., Heringstad, B., Olsen, H.G., Gomez-Raya, L., Våge, D.I., Olsaker, I., Ødegård, J., Klemetsdal, G., Schulman, N., Vilkki, J., Ruane, J., Aasland, M., Rønningen, K., and S. Lien. 2001. Quantitative trait loci affection clinical mastitis and somatic cell count in dairy cattle. Mammalian Genome. 12: 837-842.

Knafler, G.J., Clark, J.A., Boersma, P.D., and J.L. Bouzat. 2012. MHC diversity and mate choice in the magellanic penguin, Spheniscus magellanicus. Journal of Heredity. 103: 759-768.

Knapp, L. 2005. The ABCs of MHC. Evolutionary Anthropology. 14: 28-37.

Knight, J.C. 2013. Genomic modulators of the immune response. Trends in Genetics. 29: 74-83.

Kohler, H., Bayry, J., Nicoletti, A., and S.V. Kaveri. 2003. Natural autoantibodies as tools to predict the outcome of immune response? Scandanavian Journal of Immunology. 58: 285-289.

Koivula, M., Mäntysaari, E.A., Negussie, E., and T. Serenius. 2005. Genetic and phenotypic relationships among milk yield and somatic cell count before and after clinical mastitis. Journal of Dairy Science. 88:827-833.

Kolbehdari, D., Wang, Z., Grant, J.R., Murdoch, B., Prasad, A., Xiu, Z., Marques, E., Stothard, P., and S.S. Moore. 2009. A whole genome scan to map QTL for milk production traits and somatic cell score in Canadian Holstein bulls. Journal of Animal Breeding and Genetics. 126: 216-227.

Kühn, C., Bennewitz, J., Reinsch, N., Xu, N., Thomsen, H., Looft, C., Brockmann, G.A., Schwerin, M., Weimann, C., Hiendleder, S., Erhardt, G., Medjugorac, I., Förster, M., Brenig, B., Reinhardt, F., Reents, R., Russ, I., Averdunk, G., Blümel, J., and E. Kalm. 2003. Quantitative trait loci mapping of functional traits in the German Holstein cattle population. Journal of Dairy Science. 86: 360-368.

L

Lee, Y. and J.A. Nelder. 2006. Double hierarchical generalized linear models. Journal of the Royal Statistical Society. Series C: Applied Statistics. 55: 139-185.

Lund, M.S., Sahana, G., Andersson-Eklund, L., Hastings, N., Fernandez, A., Schulman, N., Thomsen, B., Viitala, S., Williams, J.L., Sabry, A., Viinalass, H., and J. Vilkki. 2007. Joint analysis of quantitative trait loci for clinical mastitis and somatic cell score on five chromosomes in three Nordic dairy cattle breeds. Journal of Dairy Science. 90: 5282-5290.

Lund, M.S., Guldbrandtsen, B., Buitenhuis, A.J., Thomsen, B., and C. Bendixen. 2008. Detection of quantitative trait loci in Danish Holstein cattle affecting clinical mastitis, somatic cell score, udder conformation traits, and assessment of associated effects on milk yield. Journal of Dairy Science. 91: 4028-4036.

Lutz H.U. and S. Miescher. 2008. Natural antibodies in health and disease: An overview of the first international workshop on natural antibodies in health and disease. Autoimmunity reviews. 7: 405-409.

M

Madsen, P., Shariatti, M.M., and J. Ødegård. 2008. Genetic analysis of somatic cell score in Danish Holsteins using a liability-normal mixture model. Journal of Dairy Science. 91: 4355-4364.

Mallard B.A., Burnside E.B., Burton J.H. and B.N. Wilkie. 1983. Variation in serum immunoglobulins in Canadian Holstein-Friesians. Journal of Dairy Science. 66: 862-866.

Malonia, S.K., Sinha, S., Lakshminarasimhan, P., Singh, K., Jalota-Badhwar, A., Rampalli, S., Kaul-Ghanekar, R., and S. Chattopadhyay. 2010. Gene regulation by SMAR1: Role in cellular homeostasis and cancer. Biochimica et Biophysica Acta - Reviews on Cancer. 1815: 1-12.

Mark, T., Fikse, W.F., Emanuelson, U., and J. Philipsson. 2002. International genetic evaluations of Holstein sires for milk somatic cell and clinical mastitis. Journal of Dairy Science. 85: 2384-2392.

Mark, T., and P. Sandøe. 2010. Genomic dairy cattle breeding: risks and opportunities for cow welfare. Animal Welfare. 19(S): 113-121.

Matter, M.S., and A.F. Ochsenbein. 2008. Natural antibodies target virus-antibody complexes to organized lymphoid tissue. Autoimmunity Reviews. 7: 480-486.

Matukumalli, L.K., Lawley, C.T., Schnabel, R.D., Taylor, J.F., Allan, M.F., Heaton, M.P., O'connell, J., More, S.S., Smith, T.P.L.,. Sonstegard, T.S., and C.P. Van Tassell. 2009. Development and characterization of a high density SNP genotyping assay for cattle. PLoS ONE. 4:e5350. doi:10.1371/journal.pone.0005350

McCarthy, S., Berry, D.P., Dillon, P., Rath, M., and B. Horan. 2007. Effect of strain of Holstein-Friesian and feed system on udder health and milking characteristics. Livestock Science. 107: 1–28.

Medzhitov, R., and Janeway, C.A. Jr. 1997. Innate immunity: impact on the adaptive immune response. Current Opinion in Immunology. 9: 4-9.

Meredith, B.K., Kearney, J.F., and D.P. Berry. 2010. Genome-wide associations for milk production and somatic cell count in Irish Holstein-Friesian cattle. Pdf 0737 in Proc. 9th World Conference on Genetics Applied to Livestock Production, Leipzig, Germany. Gesellschaft für Tierzuchtwissenschaften e.V., Giessen, Germany.

Meyer, K. 1992. Bias and sampling covariances of estimates of variance components due to maternal effects. Genetics Selection Evolution. 24: 487-509.

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

Mrode, R.A., and G.J.T. Swanson. 2003. Estimation of genetic parameters for somatic cell count in the first three lactations using random regression. Livestock Production Science. 79: 239-247.

Mulder, H.A., Bijma, P., and W.G. Hill. 2008. Selection for uniformity in livestock by exploiting genetic heterogeneity of residual variance. Genetics Selection Evolution. 40: 37-59.

Mulder, H.A., Rönnegård, L., Fikse, W.F., Veerkamp, R.F., and E. Strandberg. 2013. Estimation of genetic variance in macro- and micro-environmental sensitivity using double hierarchical generalized linear models. Genetics Selection Evolution. 45:23. doi:10.1186/1297-9686-45-23

N

Neeteson-van Nieuwenhoven, A., Knap, P., and S. Avendaño. 2012. The role of sustainable commercial pig and poultry breeding for food security. Animal Frontiers. 3: 52-57.

O

Ochsenbein, A.F., Fehr, T., Lutz, C., Suter, M., Brombacher, F., Hengartner, H., and R.M. Zinkernagel. 1999. Control of early viral and bacterial distribution and disease by natural antibodies. Science. 286: 2156-2159.

Ochsenbein A.F. and R.M. Zinkernagel. 2000. Natural antibodies and complement link innate and acquired immunity. Immunology Today. 21: 624-630.

Oltenacu, P.A., and D.M. Broom. 2010. The impact of genetic selection for increased milk yield on the welfare of dairy cows. Animal Welfare. 19(S): 39-49.

P

Pastoret, P.P. 2007. Challenges and issues of early life vaccination in animals and humans. Journal of Comparative Pathology. 137: S2-3.

Petersson, K.J., Gustafsson, H., Strandberg, E., and B. Berglund. 2006. Atypical progesterone profiles and fertility in Swedish dairy cows. Journal of Dairy Science. 89: 2529-2538.

Ploegaert, T.C.W. 2010a. Parameters for natural resistance in bovine milk (PhD-thesis). Pg 68-80 Wageningen University, Wageningen, The Netherlands. ISBN 978-90-8585-827-0.

Ploegaert, T.C.W. 2010b. Relation of natural antibodies with risk for mastitis and high somatic cell count in Dutch Holstein Friesian cows. In: Parameters for natural resistance in bovine milk (PhD thesis), pp. 82-92. Wageningen University, Wageningen, The Netherlands. ISBN: 978-908585-827-0.

Ploegaert, T.C.W., Wijga, S., Tijhaar, E., Van der Poel, J.J., Lam, T.J.G.M., Savelkoul, H.F.J., Parmentier, H.K., and J.A.M. van Arendonk. 2010. Genetic variation of natural antibodies in milk of Dutch Holstein-Friesian cows. Journal of Dairy Science. 93: 5467-5473.

Ploegaert, T.C.W., Tijhaar, E., Lam, T.J.G.M., Taverne-Thiele, A., Van der Poel, J.J., Van Arendonk, J.A.M., Savelkoul, H.F.J. and H.K. Parmentier. 2011. Natural antibodies in bovine milk and blood plasma: Variability among cows, repeatability within cows, and relation between milk and plasma titers. Veterinary Immunology and Immunopathology. 144: 88-94.

Potts, W.K., Manning, C.J., and E.K. Wakeland. 1994. The role of infectious disease, inbreeding and mating preferences in maintaining MHC genetic diversity: an experimental test. Philosophical Transactions: Biological Sciences. 346: 369-378.

Pozsonyi, É., György, B., Berki, T., Bánlaki, Z., Buzás, E., Rajczy, K., Hossó, A., Prohászka, Z., Szilágyi, Á., Cervenak, L., and G. Füst. 2009. HLA-association of serum levels of natural antibodies. Molecular Immunology. 46: 1416-1423.

Pravieux, J.J., Poulet, H., Charreyre, C., and V. Juillard. 2007. Protection of newborn animals through maternal immunization. Journal of Comparative Pathology. 137: S32-34.

Pryce, J.E., Bolormaa, S., Chamberlain, A.J., Bowman, P.J., Savin, K., Goddard, M.E., 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. Journal of Dairy Science. 93: 3331-3345.

R

Rainard, P, and C. Riollet. 2006. Innate immunity of the bovine mammary gland. Veterinary Research. 37: 369-400.

Remmelink, G., Blanken, K., Van Middelkoop, J., Ouweltjes, W. and H. Wemmenhove. 2012. Handboek Melkveehouderij 2012. Wageningen UR Livestock Research, Wageningen, The Netherlands.

Reusch, T.B., Haberli, M.A., Aeschlimann, P.B., and M. Milinski. 2001. Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. Nature 414: 300–302.

Riggio, V., Portolano, B., Bovenhuis, H., and S.C. Bishop. 2010. Genetic parameters for somatic cell score according to udder infection status in Valle del Belice dairy sheep and impact of imperfect diagnosis of infection. Genetics Selection Evolution. 42: 30.

Rönnegård, L., Felleki, J., Fikse, F., Mulder, H.A., and E. Strandberg. 2010. Genetic heterogeneity of residual variance – estimation of variance components using double hierarchical generalized linear models. Genetics Selection Evolution. 42: 8-17. doi: 10.1186/1297-9686-42-8

Rönnegård, L., Felleki, J., Fikse, F., Mulder, H.A., and E. Strandberg. 2013. Variance component and breeding value estimation for genetic heterogeneity of residual variance in Swedish Holstein dairy cattle. Journal of Dairy Science. 96: 2627-2636.

Rodo, J., Gonçalves, L.A., Demengeot, J., Coutinho, A., and C. Penha-Conçalves. 2006. MHC class II molecules control murine B cell responsiveness to lipopolysaccharide stimulation. Journal of Immunology. 177: 4620-4626.

Rupp, R., and D. Boichard. 1999. Genetic parameters for clinical mastitis, somatic cell score, production, udder type traits, and milking ease in first lactation Holsteins. Journal of Dairy Science. 82: 2198-2204.

Rupp R. and D. Boichard. 2003. Genetics of resistance to mastitis in dairy cattle. Veterinary research. 34: 671-688.

Rupp, R., Hernandez, A., and B.A. Mallard. 2007. Association of bovine leukocyte antigen (BoLA) DRB3.2 with immune response, mastitis, and production and type traits in Canadian Holsteins. Journal of Dairy Science. 90: 1029-1038.

Rupp, R., and D. Boichard. 2003. Genetics of resistance to mastitis in dairy cattle. Veterinary Research. 34: 671-688.

S

Schmidt, W.N., Mathahs, M.M., and Z. Zhu. 2012. Heme and HO-1 inhibition of HCV, HBV, and HIV. Frontiers in Pharmacology 3, Article 129. doi: 10.3389/fphar.2012.00129.

Schopen, G.C.B., Visker, M.H.P.W., Koks, P.D., Mullaart, E., Van Arendonk, J.A.M., and H. Bovenhuis. 2011. Whole-genome association study for milk protein composition in dairy cattle. Journal of Dairy Science. 94: 3148-3158.

Schroeder H.W. and L. Cavacini . 2010. Structure and function of immunoglobulins. Journal of Allergy and Clinical Immunology. 125: S41-52.

Schulman, N.F., Viitala, S.M., De Koning, D.J., Virta, J., Mäki-Tanilla, A., and J.H. Vilkki. 2004. Quantitative trait loci for health traits in Finnish Ayrshire cattle. Journal of Dairy Science. 87: 443-449.

Siwek M., Buitenhuis B., Cornelissen S., Nieuwland M., Knol E.F., Crooijmans R., Groenen M., Parmentier H., and J. Van der Poel. 2006. Detection of QTL for innate: Non-specific antibody levels binding LPS and LTA in two independent populations of laying hens. Developmental and Comparative Immunology. 30: 659-666.

Seegers, H., Fourichon, C., and F. Beaudeau. 2003. Production effects related to mastitis and mastitis economics in dairy cattle herds. Veterinary Research. 34: 475-491.

Snyderman, R., and E.J. Goetzl. 1981. Molecular and cellular mechanisms of leukocyte chemotaxis. Science. 213: 830-837.

Sodeland, M., Kent, M.P., Olsen, H.G., Opsal, M.A., Svendsen, M., Sehested, E., Hayes, B.J., and S. Lien. 2011. Quantitative trait loci for clinical mastitis on chromosomes 2, 6, 14 and 20 in Norwegian Red cattle. Animal Genetics. 42: 457-569.

Sordillo, L.M., and K.L. Streicher. 2002. Mammary gland immunity and mastitis susceptibility. Journal of Mammary Gland Biology and Neoplasia. 7: 135-146.

Star, L., Frankena, K., Kemp, B., Nieuwland, M.G.B., and H.K. Parmentier. 2007. Natural humoral competence and survival in layers. Poultry Science. 86: 1090-1099.

Stelwagen, K., Carpenter, E., Haigh, B., Hodgkinson, A., and T.T. Wheeler. 2009. Immune components of bovine colostrum and milk. Journal of Animal Science. 87(Suppl. 1): 3-9.

Story, J.D., and R. Tibshirani. 2003. Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences of the United States of America. 100: 9440-9445.

Sumimoto, H., Hata, K., Mizuki, K., Ito, T., Kage, Y., Sakaki, Y., Fukumaki, Y., Nakamura, M., and K. Takeshige. 1996. Assembly and activation of the phagocyte NADPH oxidase. Journal of Biological Chemistry. 271: 22152-22158.

Sun Y., Parmentier H.K., Frankena K., and J.J. Van der Poel. 2011. Natural antibody isotypes as predictors of survival in laying hens. Poultry Science. 90: 2263-2274.

T

Takeshima, S., and Y. Aida. 2006. Structure, function and disease susceptibility of the bovine major histocompatibility complex. Animal Science Journal 77:138-150.

Takeshima, S., Matsumoto, Y., Chen, J., Yoshida, T., Mukoyama, H., and Y. Aida. 2008. Evidence for cattle major histocompatibility complex (BoLA) class II DQA1 gene heterozygote advantage against clinical mastitis caused by Streptococci and Escherichia species. Tissue Antigens. 72: 525-531.

Tal-Stein, R., Fontanesi, L., Dolezal, M., Scotti, E., Bagnato, A., Russo, V., Canavesi, F., Friedmann, A., Soller, M., and E. Lipkin. 2010. A genome scan for quantitative trait loci affection milk somatic cell score in Israeli and Italian Holstein cows by means of selective DNA pooling with single- and multiple-marker mapping. Journal of Dairy Science. 93: 4913-4927.

Tao, W., and B. Mallard. 2007. Differentially expressed genes associated with Staphylococcus aureus mastitis of Canadian Holstein cows. Veterinary Immunology and Immunopathology. 120: 201-211.

Thompson-Crispi, K.A., Sewalem, A., Miglior, F., and B.A. Mallard. 2012. Genetic parameters of adaptive immune response traits in Canadian Holsteins. Journal of Dairy Science. 95: 401-409.

Tumbarello, D.A., Waxsel, B.J., Arden, S.D., Bright, N.A., Kendrick-Jones, J., and F. Buss. 2012. Autophagy receptors link myosin VI to autophagosomes to mediate Tom1-dependent autophagosome maturation and fusion with the lysosome. Nature cell biology. 14: 1024-1047.

U

Urioste, J.I., Franzén, J., and E. Strandberg. 2010. Phenotypic and genetic characterization of novel somatic cell count traits from weekly or monthly observations. Journal of Dairy Science. 93: 5930-5941.

Uthaisangsook, S., Day, N.K., Bahna, S.L., Good, R.A., and S. Haraguchi. 2002. Innate immunity and its role against infections. Annals of Allergy, Asthma and Immunology. 88: 253-264.

V

Vaarst, M., and C. Enevoldsen. 1997. Patterns of clinical mastitis manifestations in Danish organic dairy herds. Journal of Dairy Research. 64: 23-37.

Van Knegsel, A.T.M., De Vries Reilingh, G., Meulenberg, S., Van den Brand, H., Dijkstra, J., Kemp, B., and H.K. Parmentier. 2007. Natural antibodies related to energy balance in early lactation dairy cows. Journal of Dairy Science. 90: 5490-5498.

Veerkamp, R.F., Simm G., and J.D. Oldham. 1994. Effects of interaction between genotype and feeding system on milk production, feed intake, efficiency and body tissue mobilization in dairy cows. Livestock Production Science. 39: 229-241.

Veerkamp, R.F., Oldenbroek, J.K., Van Der Gaast, H.J., 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. Journal of Dairy Science. 83: 577–583.

Vollmers, H.P., and S. Brändlein. 2006. Natural IgM antibodies: from parias to parvenus. Histology and Histopathology. 21: 1355-1366.

Vollmers H.P. and S. Brändlein. 2009. Natural antibodies and cancer. New Biotechnology. 25: 294-298.

W

White, P.C. 1989. Molecular genetics of the class III region of the HLA complex. In: Immunobiology of HLA. Immunogenetics and Histocompatibility. B. Dupont (ed.), Springer-Verlag, New York, United States of America. Pages 62-69.

White, S.N., Mousel, M.R., Herrmann-Hoesing, L.M., Reynolds, J.O., Leymaster, K.A., Neibergs, H.L., Lewis, G.S., and D.P. Knowles. 2012. Genome-wide association identifies multiple genomic regions associated with susceptibility to and control of ovine lentivirus. PLOS ONE 7. doi: 10.1371/journal.pone.0047829.

Wijga, S., Parmentier, H.K., Nieuwland, M.G.B., and H. Bovenhuis. 2009. Genetic parameters for levels of natural antibodies in chicken lines divergently selected for specific antibody response. Poultry Science. 88: 1805-1810.

Wijga, S., Bovenhuis, H., Bastiaansen, J.W.M., Van Arendonk, J.A.M., Ploegaert, T.C.W., Tijhaar, E., and J.J. van der Poel. 2013. Genetic parameters for natural antibody isotype titers in milk of Dutch Holstein-Friesians. Animal Genetics. doi: 10.1111/age.12038

Wilmink J.B.M. 1987. Adjustment of test-day milk, fat and protein yield for age, season and days-in-milk. Livestock Production Science. 16: 335-348.

Windig, J.J., Ouweltjes, W., Ten Napel, J., De Jong, G., Veerkamp, R.F., and Y. de Haas. 2010. Combining somatic cell count traits for optimal selection against mastitis. Journal of Dairy Science. 93: 1690-1701.

Windig, J.J., Urioste, J.I. and E. Strandberg. 2013. Integration of epidemiology into the genetic analysis of mastitis in Swedish Holstein. Journal of Dairy Science. 96: 2617-2626.

Wolf, J.B., Brodie III, E.D., Cheverud, J.M., Moore, A.J., and M.J. Wade. 1998. Evolutionary consequences of indirect genetic effects. Trends in Ecology and Evolution. 13: 64-69.

Woof J.M., and M.A. Kerr. 2006. The function of immunoglobulin A in immunity. Journal of Pathology. 208: 270-282.

Wrighton, K.H., Lin, X., and X. Feng. 2012. Critical regulation of TGF signaling by Hsp90. Proceedings of the National Academy of Sciences of the United States of America. 105: 9244–9249.

Z

Zelenay, S., Moraes Fontes, M.F., Fesel, C., Demengeot, J., and A. Coutinho. 2007. Physiopathology of natural auto-antibodies: The case for regulation. Journal of Autoimmunity 29: 229-235.

Zhang, Q., Boichard, D., Hoeschele, I., Ernst, C., Eggen, A., Murkve, B., Pfister-Genskow, M., Witte, L.A., Grignola, F.E., Uimari, P., Thaller, G., and M.D. Bishop. 1998. Mapping quantitative trait loci for milk production and

health of dairy cattle in a large outbred pedigree. Genetics. 149: 1959-1973.

Zschenker, O., Illies, T., and D. Ameisy. 2006. Overexpression of Lysosomal Acid Lipase and Other Proteins in Atherosclerosis. J. Biochem. 140:23–38.

Summary

Summary

Reduction of the impact of livestock diseases becomes increasingly important, not only for economic and environmental reasons, but also from a human and animal welfare perspective. One way to reduce the impact of disease is through incorporation of disease resistance into livestock breeding schemes. Livestock breeding has the potential to decrease the impact of disease on the long-term. Besides production traits, breeding schemes increasingly include traits related to welfare, such as disease resistance. Mastitis constitutes an important infectious disease in dairy cattle, where somatic cell score (SCS) is frequently used in dairy cattle breeding schemes as an indicator trait for mastitis. Still, there is a lot to be learned about the interpretation of genetic parameters and the biology of disease resistance and SCS. This PhD thesis aimed to obtain additional insight in disease resistance and SCS through the following objectives: 1) quantify the impact of genetics on innate immunity, represented by natural antibodies (NAb), through estimation of heritabilities and genetic correlations; 2) identify the genomic regions involved in SCS and NAb levels; and 3) quantify the impact of genetics on environmental sensitivity for SCS. Chapter 1 introduces the functional traits NAb levels and SCS, reviews the structure and function of NAb, discusses the relation of SCS with mastitis and introduces concepts of genome-wide association studies and environmental sensitivity. Chapter 2 covers the estimation of heritabilities for NAb levels in milk. In addition, genetic correlations between different NAb isotype levels are estimated. In Chapter 3 a genome-wide association study is performed on the NAb levels for which genetic parameters were estimated in chapter 2. Chapter 4 includes a genome wide association study on SCS, where both lactation-average SCS and the trait standard deviation in lactation-average SCS are analyzed. Chapter 5 studies the genetic variation in environmental sensitivity for SCS by means of a double hierarchical generalized linear model. Finally, chapter 6 discusses some of the biology that is behind NAb and SCS.

In chapter 2 genetic parameters for NAb isotypes immunoglobulin (Ig) A, IgG1 and IgM levels binding the bacterial antigens lipopolysaccharide, lipoteichoic acid, peptidoglycan and the model antigen keyhole limpet hemocyanin were estimated in Dutch Holstein-Friesian cows. Further, this study included total NAb levels binding the antigens mentioned above, making no isotype distinction. The study showed that NAb isotype levels are heritable, with heritabilities ranging from 0.06 to 0.55, and that these heritabilities were generally higher than heritabilities for total NAb levels. Genetic correlations were nearly all positive and ranged from -0.23 to 0.99. Strong genetic correlations were found between IgA and IgM. Genetic

correlations were substantially weaker when they involved IgG1, indicating that IgA and IgM have a common genetic basis, but that the genetic basis for IgG1 differs from that for IgA or IgM. Results from this study indicate that NAb isotype levels show the potential for effective genetic selection. Further, NAb isotypes may provide a better characterization of different elements of the immune response or immune competence. As such, NAb isotypes may enable more effective decisions when breeding programs start to include innate immune parameters.

Chapter 3 aimed to provide more insight in the largely unknown genomic basis of NAb levels in milk using the antibody phenotypes for which genetic parameters were estimated in chapter 2. Additionally, genotype information consisting of about 50,000 markers was available for these cows. Each individual marker was tested for detection of variation in NAb levels. Results show that information can be gained when total NAb levels are further subdivided into isotype levels, suggesting that analysis of detailed phenotypes can provide further insight in the genetic control of traits. Further, this study identified genomic regions associated with NAb levels on all bovine autosomes. A region on BTA23 was consistently found significantly associated with genetic variation in isotype IgM levels across the different binding specificities. The bovine major histocompatibility complex (MHC) is located near this region, making this a region of candidate gene(s) involved in NAb expression in dairy cows both from a functional and positional perspective.

Chapter 4 aimed to identify genomic regions associated with lactation-average SCS and the standard deviation of test-day SCS. It is one of the first studies to combine detailed phenotypic and genotypic cow data from research dairy herds located in different countries. The combined dataset contained up to 52 individual test-days per lactation and thereby aimed to capture temporary increases in SCS associated with infection, mainly by means of the standard deviation of test-day SCS. Phenotypic data consisted of 46,882 test-day records on 1,484 cows and genotypic data consisted of 37,590 markers. The association between each individual marker and the phenotypic data was estimated. The analyses showed that lactation-average SCS was significantly associated with a marker on chromosome 4 and a marker on chromosome 18. Likewise the standard deviation of test-day SCS was associated with this marker on chromosome 18. In addition, the standard deviation of test-day SCS was significantly associated with a marker on chromosome 6. Relatively few associations were found, suggesting that lactation-average SCS and the standard deviation of test-day SCS are controlled by multiple loci, each with a relatively small effect, distributed across the genome.

Chapter 5 studied environmental sensitivity for SCS. Sensitivity to respond to environmental factors can have a genetic origin. Environmental factors can be

divided into known and unknown environmental factors, referred to as macro- and micro environment, respectively. Macro-environmental sensitivity can be expressed as genetic variation in the slope of a reaction norm, whereas micro-environmental sensitivity can be expressed as differences in residual variance that have a genetic origin. The aim of the study was to estimate macro- and micro-environmental sensitivity for SCS in Holstein dairy cattle, where the macro-environment was defined as the herd-year average somatic cell count. Variance components were estimated with a double hierarchical generalized linear model extended with a reaction norm using the ASReml software. Both macro- and micro-environmental sensitivity were found for SCS and these two sensitivities were found to be positively correlated. Estimated variance components and genetic correlations obtained with the macro or micro model resembled those obtained with the macro-micro model, which suggests that use of macro or micro models does not lead to biased estimates resulting from micro models partly picking up macro-environmental sensitivity or vice versa. Knowledge on both forms of sensitivity, however, can aid in optimization of selection as correlations between the additive genetic variance in intercept, slope and environmental variance were all away from unity. We conclude that selection for reduced environmental sensitivity has the potential to reduce variability in animal performance due to known and unknown environmental factors and herewith increase predictability of performance across and within environments.

Chapter 6 discussed some of the biology behind NAb and SCS, as trait biology may affect genetic parameter estimation, identification of candidate genes and breeding programs. Knowledge on disease biology is of vital importance to fully understand the processes involved when selecting for increased disease resistance, as a better understanding enables a better prediction of the consequences. Chapter 6 contained three conclusions: firstly, analyses of cell types (detailed phenotypes) rather than SCS can provide further insight in the genetic control of SCS and mastitis. Secondly, no evidence was found for maternal genetic effects on NAb levels in milk. Maternal environmental effects, however, could play a role in NAb levels. Finally, genetic diversity in the MHC region is maintained by natural selection. Selective breeding and farm management practices may affect this genetic diversity, which could bring about negative effects on animal fitness, such as fertility problems. Selective breeding for specific MHC haplotypes may therefore impose a risk for negative effects on animal health.

Samenvatting

Samenvatting

Het verminderen van de impact van veeziekten wordt steeds belangrijker, niet alleen omwille van economie en milieu, maar ook omwille van menselijk en dierlijk welzijn. Een mogelijkheid om de impact van ziekte te verminderen is het opnemen van ziekteresistentie in fokprogramma's. Dierfokkerij biedt de mogelijkheid om de impact van dierziekten op de lange termijn te reduceren. Naast productiekenmerken bevatten fokprogramma's in toenemende mate kenmerken welke gerelateerd zijn aan welzijn, zoals ziekteresistentie. Mastitis is een belangrijke infectieziekte in melkvee, waar het celgetal (SCS) een kenmerk is dat veelvuldig in fokprogramma's wordt gebruikt als een indicator voor mastitis. Nog steeds is er veel te leren over de interpretatie van genetische parameters en de biologie achter ziekteresistentie en SCS. Deze PhD thesis had als doel aanvullend inzicht te verkrijgen in ziekteresistentie en SCS middels de volgende doelen: 1) kwantificeren van de impact van genetica op innate immuniteit door het schatten van erfelijkheidsgraden en genetische correlaties, waarbij natuurlijke antilichamen (NAb) als model kenmerk zijn genomen; 2) identificeren van regio's op het genoom welke een rol spelen in SCS en NAb niveaus; en 3) kwantificeren van de impact van genetica om milieugevoeligheid van SCS. Hoofdstuk 1 introduceert de kenmerken NAb niveaus en SCS, geeft een beknopt overzicht van de structuur en functie van NAb, bespreekt de relatie tussen SCS en mastitis en introduceert principes van genome-wide associatie studies en milieugevoeligheid. Hoofdstuk 2 betreft de schatting van erfelijkheidsgraden voor NAb niveaus in melk. Daarnaast zijn tevens genetische correlaties tussen verschillende NAb isotype niveaus geschat. In hoofdstuk 3 is een genome-wide associatie studie uitgevoerd voor de NAb niveaus waarvoor genetische parameters zijn geschat in hoofdstuk 2. Hoofdstuk 4 bevat een genome-wide associatie studie voor SCS, waar zowel lactatiegemiddeld SCS als het kenmerk standard deviatie voor test-dag SCS zijn geanalyseerd. Hoofdstuk 5 bestudeert de genetische variatie in milieugevoeligheid van SCS middels een double hierarchical generalized linear model. Tot slot bespreekt hoofdstuk 6 wat van de biologie rondom NAb en SCS.

In hoofdstuk 2 zijn genetische parameters geschat voor NAb isotypes immuunglobuline (Ig)A, IgG1 en IgM niveaus, welke in staat zijn de bacteriële antigenen lipopolysaccharide, lipoteichoic acid, peptidoglycaan en het model antigen keyhole limpet hemocyanin te binden. Deze NAb niveaus zijn gemeten in melk van Nederlandse Holstein-Friesian koeien. Naast isotypes waren in deze studie ook totale NAb niveaus opgenomen, welke geen onderscheid maken tussen isotypes. De studie liet zien dat NAb niveaus erfelijk zijn met erfelijkheidsgraden

variërend van 0.06 tot 0.55, en dat erfelijkheidsgraden voor NAb isotype niveaus over het algemeen hoger zijn dan erfelijkheidsgraden voor de totale NAb niveaus. Genetische correlaties waren vrijwel allen positief, variërend van -0.23 tot 0.99. Sterke genetische correlaties werden gevonden tussen IgA en IgM. Genetische correlaties waren aanzienlijk minder sterk wanneer de combinatie waarvoor de correlatie werd geschat IgG1 bevatte. Dit suggereert dat IgA en IgM een gezamelijke genetische oorsprong kennen, maar dat de genetische oorsprong voor IgG1 verschilt van dat voor IgA of IgM. Resultaten uit deze studie laten daarnaast zien dat NAb isotype niveaus potentie tot effectieve selectie hebben. Daarbij zouden NAb isotypes een betere karakterisering kunnen verschaffen met betrekking tot de verschillende elementen in een immuunrespons of immuun competentie. Zodoende zouden NAb isotypes effectievere beslissingen kunnen bewerkstelligen wanneer innate immuunparameters onderdeel worden fokprogramma's.

Hoofdstuk 3 had als doel meer inzicht te verschaffen in de grotendeels onbekende genomische basis van NAb niveaus in melk met behulp van de antilichaam fenotypes waarvoor genetische parameters zijn geschat in hoofdstuk 2. Daarbij was genotype informatie, bestaand uit ongeveer 50.000 markers, beschikbaar voor deze koeien. Iedere marker werd individueel getest op detectie van variatie in NAb niveaus. Resultaten laten zien dat extra informatie verkregen kan worden wanneer totale NAb niveaus verder worden onderverdeeld in isotype niveaus. Dit suggereert dat het analyseren van gedetailleerde fenotypes meer inzicht kan verschaffen in de genetische achtergrond van kenmerken. Daarnaast werden in deze studie genomische regio's gevonden op alle boviene autosomen welke geassocieerd werden met NAb niveaus. Een regio op chromosoom 23 was consistent geassocieerd met genetische variatie in isotype IgM niveaus ongeacht de bindingsspecificiteit van de antilichamen. Het boviene major histocompatibility complex (MHC) ligt dichtbij deze regio. Zowel vanuit een functioneel als een positioneel perspectief is dit een regio met kandidaat genen betrokken bij NAb expressie in melk koeien.

Hoofdstuk 4 had als doel genomische regio's te identificeren welke geassocieerd zijn met lactatiegemiddeld SCS en de standaarddeviatie van test-dag SCS. Het is een van de eerste studies welke gedetailleerde fenotypische en genotypische koe data, afkomstig van onderzoeksboerderijen in verschillende landen, combineert. De gecombineerde dataset bevatte tot 52 individuele test-dagen per lactatie met als doel tijdelijke stijgingen in SCS geassocieerd met infecties vast te leggen, met name middels de standaard deviatie van test-dag SCS. Fenotype data bestond uit 46.882 test-dag waarnemingen van 1.484 koeien en genotype data bestond uit 37.590

markers. De associatie tussen iedere individuele marker en de fenotype data werd geschat. De analyses lieten zien dat lactatiegemiddeld SCS significant geassocieerd was met een marker op chromosoom 4 en een marker op chromosoom 18. Ook de standaard deviatie van test-dag SCS was geassocieerd met deze marker op chromosoom 18. Verder was de standaarddeviatie van test-dag SCS significant geassocieerd met een marker op chromosoom 6. Relatief weinig associaties werden gevonden, wat suggereert dat lactatiegemiddeld SCS en de standaarddeviatie van test-dag SCS bepaald worden door meerdere genomische regio's ieder met een relatief klein effect, verdeeld over het genoom.

Hoofdstuk 5 bestudeerde de milieu gevoeligheid van SCS. Gevoeligheid tot reageren op milieu-invloeden kan een genetisch oorzaak hebben. Milieu-invloeden kunnen verdeeld worden in bekende en onbekende invloeden, ook wel macro- en micro milieu genoemd. Macro-milieu gevoeligheid kan omschreven worden middels de slope van een reactie norm, terwijl micro-milieu gevoeligheid omschreven kan worden middels verschillen in residuele variantie welke een genetische oorsprong hebben. Het doel van de studie was om macro- en micro-milieu gevoeligheid van SCS te schatten in Holstein melkvee, waar het macro-milieu gedefinieerd was als het boerderij-jaar gemiddelde celgetal. Variantie componenten werden geschat met een double hierarchical generalized linear model uitgebreid met een reactie norm, met behulp van de ASReml software. Zowel macro- en micro-milieugevoeligheid werd gevonden voor SCS en deze twee gevoeligheden waren positief gecorreleerd. Geschatte variantie componenten en genetische correlaties verkregen met het macro of micro model kwamen overeen met de schattingen verkregen met een macro-micro model, wat suggereert dat gebruik van macro of micro modellen niet leidt tot onjuiste schattingen doordat micro modellen gedeeltelijk macro-milieu gevoeligheid oppikken of andersom. Kennis van beide vormen van gevoeligheid, echter, kan nuttig zijn bij het optimaliseren van selectie omdat de additief genetische variantie in intercept, slope en milieu-variantie allen geen 1 waren. We concluderen dat selectie voor verminderde milieugevoeligheid potentie heeft variabiliteit in prestatie van dieren door bekende en onbekende milieu factoren te verminderen en hiermee voorspelbaarheid van prestatie binnen milieus en over milieus heen te vergroten.

Hoofdstuk 6 bespreekt wat van de biologie achter NAb en SCS, omdat biologie van kenmerken de schatting van genetische parameters, identificatie van kandidaat genen en fokprogramma's kan beïnvloeden. Kennis van biologie van ziekten is belangrijk om ziekteprocessen te doorgronden wanneer selectie is gericht op verhoogde ziekteresistentie, omdat een beter begrip een betere inschatting van de consequenties van selectie mogelijk maakt. Hoofdstuk 6 bevatte 3 conclusies: 1)

analyse van cel types (gedetailleerde fenotypes) in plaats van SCS kan mogelijk verder inzicht verschaffen in de genetische achtergrond van SCS en mastitis; 2) er zijn geen aanwijzingen gevonden voor een betrokkenheid van maternaal genetische effecten op NAb niveaus. Maternale milieu-effecten zouden echter wel een rol kunnen spelen in NAb niveaus in melk; 3) genetische diversiteit in de regio welke genen van het major histocompatibility complex bevat wordt in stand gehouden door natuurlijke selectie. Gerichte fokkerij en bedrijfsmanagement zou deze diversiteit kunnen beïnvloeden, wat negatieve effecten op fitness van dieren met zich mee kan brengen, bijvoorbeeld vruchtbaarheidsproblemen. Gericht fokken op specifieke MHC haplotypes kan daardoor een risico op negatieve effecten op diergezondheid met zich meebrengen.

Curriculum Vitae

Susan Wijga was born on November 3, 1984 in Oploo, the Netherlands. She was raised in Oploo and obtained her high school diploma in 2002 at Scholengemeenschap Stevensbeek in Stevensbeek. That same year she started her bachelor study 'Animal Sciences' in Leeuwarden where she specialized in 'Animal care and nutrition' and 'Laboratory animal management'. During this study she had internships at the animal laboratory at Utrecht University, University of Groningen and the Radboud University in Nijmegen. Her bachelor thesis focused on nutrient digestion in newly weaned piglets which she fulfilled at Schothorst Feed Research in Lelystad. In 2006 she obtained her bachelor degree and she started her master study 'Animal Sciences' later that year. She fulfilled two major theses; one at the Animal Breeding and Genomics Centre, Wageningen University, which focused on genetic parameters of natural antibody levels in chicken, and one at the Cell biology and Immunology group, Wageningen University, where she applied ELISA to obtain natural antibody levels in bovine milk. Subsequently, she estimated genetic parameters for these antibody levels. Her minor thesis was fulfilled at the Animal Breeding and Genomics Centre, Wageningen University, and this time genetic parameters for success as a guide dog were estimated. In 2009 she graduated and started her PhD research, which was part of the RobustMilk project. The results of this research are described in this thesis. From 2007 until 2013 she was also active as a member of the Committee on animal experiments of Wageningen University. Currently, Susan is working as a researcher at TOPIGS Research Center IPG in Beuningen.

Susan Wijga is op 3 november 1984 geboren in Oploo, Nederland. Ze is opgegroeid in Oploo en heeft in 2002 haar HAVO diploma behaald aan Scholengemeenschap Stevensbeek in Stevensbeek. Datzelfde jaar is ze begonnen aan haar bachelor studie 'Diermanagement' in Leeuwarden, waar ze de specialisaties 'Voeding en verzorging' en 'Proefdierbeheer' heeft gevolgd. Tijdens deze studie liep ze stage op dierlaboratoria van Utrecht Universiteit, Rijksuniversiteit Groningen en Radboud Universiteit Nijmegen. Haar bachelor thesis richtte zich op nutriëntenvertering in pasgespeende biggen, welke ze completeerde bij Schothorst Feed Research in Lelystad. In 2006 heeft ze haar bachelor diploma behaald en later dat jaar begon ze aan haar master studie 'Animal sciences'. Ze completeerde twee major theses; één bij het Animal Breeding and Genomics Centre, Wageningen Universiteit, welke zich richtte op het schatten van genetische parameters van natuurlijke antilichaam niveaus in kippen, en een aan de leerstoelgroep Celbiologie en Immunologie, Wageningen Universiteit, waar ze ELISA toepaste om natuurlijke antilichaam niveaus in koemelk te verkrijgen. Vervolgens werden voor deze antilichaam niveaus genetische parameters geschat. Haar minor thesis werd volbracht aan het Animal Breeding and Genomics Centre, Wageningen Universiteit, en deze keer werden genetische parameters voor succes als blindengeleidenhond geschat. In 2009 behaalde ze haar master diploma en begon ze het PhD onderzoek beschreven in deze thesis, welke deel uitmaakte van het RobustMilk project. Van 2007 tot 2013 was ze daarnaast actief als lid van de dierexperimentencommissie van Wageningen Universiteit. Momenteel is Susan als onderzoeker aangesteld bij TOPIGS Research Center IPG in Beuningen.

Publications

Peer reviewed publications

Wijga, S., Parmentier, H.K., Nieuwland, M.G.B., and H. Bovenhuis. 2009. Genetic parameters for levels of natural antibodies in chicken lines divergently selected for specific antibody response. Poultry Science. 88: 1805-1810.

Wijga, S., Russenberger, J., Van den Berg, S.M., Havlena, B., and H.C.M. Heuven. 2009. Genetic parameters for guiding ability in purpose bred dogs. Animal Genetics. 40: 1004-1005.

Ploegaert, T.C.W., Wijga, S., Tijhaar, E., Van der Poel, J.J., Lam, T.J.G.M., Savelkoul, H.F.J., Parmentier, H.K., and J.A.M. van Arendonk. 2010. Genetic variation of natural antibodies in milk of Dutch Holstein-Friesian cows. Journal of Dairy Science. 93: 5467-5473.

Wijga, S., Bastiaansen, J.W.M., Wall, E., Strandberg, E., De Haas, Y., Giblin, L., and H. Bovenhuis. 2012. Genomic associations with somatic cell score in first-lactation Holstein cows. Journal of Dairy Science. 95: 899-908.

Berry, D.P., Bastiaansen, J.W.M., Veerkamp, R.F., Wijga, S., Wall, E., Berglund, B., 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: 1206-1215.

Wijga, S., Bovenhuis, H., Bastiaansen, J.W.M., Van Arendonk, J.A.M., Ploegaert, T.C.W., Tijhaar, E., and J.J. van der Poel. 2013. Genetic parameters for natural antibody isotype titers in milk of Dutch Holstein-Friesians. Animal Genetics. doi: 10.1111/age.12038

Manuscripts in preparation

Wijga, S., Bastiaansen, J.W.M., Van der Poel, J.J., Parmentier, H.K., Van Arendonk, J.A.M., and H. Bovenhuis. Genome-wide associations for natural antibodies in milk of Holstein-Friesians.

Wijga, S., Mulder, H.A., and E. Strandberg. Macro- and micro-environmental sensitivity for test-day milk yield and somatic cell score in Swedish Holstein dairy cattle.

Conference proceedings

Wijga, S., Bastiaansen, J.W.M., Wall, E., Strandberg, E., De Haas, Y., Berry, D.P., and H. Bovenhuis. A Whole Genome Association Study to Detect Loci Associated with Somatic Cell Score in Dairy Cattle. 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany. 1-6 August 2010.

Wijga, S., Bastiaansen, J.W.M., Wall, E., Strandberg, E., De Haas, Y., Giblin, L., and H. Bovenhuis. Genomic regions associated with somatic cell score in dairy cattle. International conference on udder health and communication, Utrecht, the Netherlands. 25 – 27 October 2011.

Wijga, S., Bastiaansen, J.W.M., Van Arendonk, J.A.M., Ploegaert, T.C.W., and Van der Poel, J.J. Genomic regions associated with natural antibody titres in milk of Dutch Holstein-Friesian cows. 62nd Annual Meeting of the European Association for Animal Production, Stavanger, Norway. 26 August – 2 September.

Wijga, S., Bastiaansen, J.W.M., Van Arendonk, J.A.M., Ploegaert, T.C.W., and J.J. Van der Poel. Genomic regions affecting innate health of Dutch dairy cows. 63rd Annual meeting of the European Association for Animal Production, Bratislava, Slovakia. 27-31 August 2012.

Training and education

Training and education

The basic package (3 ECTS)

WIAS introduction course	2009
Ethics and philosophy of animal science	2010

Scientific exposure (12 ECTS)

International conferences

9th WCGALP, Leipzig, Germany	2010
62nd EAAP, Stavanger, Norway	2011
International conference on udder health and communication, Utrecht, the Netherlands	2011
63rd EAAP, Bratislava, Slovakia	2012

Seminars and workshops

Friends or fiends: consequences of social interactions for artificial breeding programs and evolution, Wageningen, the Netherlands	2009
Genetics and immunology of insect bite hypersensitivity in horses, Wageningen, the Netherlands	2009
WIAS science day, Wageningen, the Netherlands	2009-2013
Genomics and animal breeding, Wageningen, the Netherlands	2011
Healthy as a (sport) horse, Wageningen, the Netherlands	2011
New opportunities for conservation genetics with genome wide information, Wageningen, the Netherlands	2012

In-depth studies (13 ECTS)

Summer institute in statistical genetics, Liège, Belgium	2009
WIAS advanced statistics course: design of experiments	2009
Genomic selection in animal breeding, Enaforsholm, Sweden	2010
Quantitative genetics with a focus on 'Selection Theory', Wageningen, the Netherlands	2011
Genomic selection in livestock, Wageningen, the Netherlands	2011
Advanced methods and algorithms in animal breeding with focus on genomic selection, Wageningen, the Netherlands	2012

Professional skills support courses (4 ECTS)

Techniques for writing and presenting a scientific paper	2010
PhD competence assessment	2010
Mobilizing your –scientific- network	2012
Career perspectives	2012

Research skills training (3 ECTS)

Preparing own PhD research proposal	2009

Didactic skills training (7 ECTS)

Supervising practicals and excursions

BSc course 'Animal Breeding and Genetics, Wageningen, the Netherlands	2010-2011
MSc course 'Proefdierkunde', Wageningen, the Netherlands	2011-2013
Review RMC proposals, Wageningen, the Netherlands	2011-2012

Supervising theses

MSc major thesis	2012
MSc major thesis	2012

Management skills training (6 ECTS)

Membership of boards and committees

Member dier experimenten commissie (DEC) Wageningen University	2009-2013

Training and education total   **48 ECTS**

Dankwoord

Dankwoord

Tot ongeveer 3 maanden voor mijn MSc afstuderen vroeg ik me af: 'wie bij zijn volle verstand kiest toch voor een PhD traject?'. Deze vraag is me ook het afgelopen jaar veelvuldig (lees dagelijks) door het hoofd geschoten en nu ik dit zo schrijf kan ik maar tot één conclusie komen: ik dus....

Ik kan dit dankwoord maar op één manier beginnen en dat is met het bedanken van de degenen die het in de eerste plaats mogelijk hebben gemaakt dat ik dit proefschrift heb kunnen schrijven: mijn ouders. Pap en mam, jullie hebben het altijd belangrijk gevonden dat Tom en ik zouden studeren, dat we iets zouden doen wat we leuk vonden en dat we zouden presteren naar ons beste kunnen. Volgens mij zijn we hier beiden in geslaagd, maar dit hadden we nooit kunnen doen zonder jullie onvoorwaardelijke liefde en steun. Al die vakanties die jullie 'rundum Hause' hebben doorgebracht omdat het vakantiegeld bestemd was voor collegegeld, de tijd en het geld dat jullie geïnvesteerd hebben om mij een fijn thuis te kunnen geven op mijn studentenkamers in Leeuwarden en Renkum, alle autoritjes die jullie hebben gemaakt om mij (en Tom) te komen helpen wanneer dat nodig was, zelfs naar Leeuwarden rijden om bijvoorbeeld een kapotte koelkast te vervangen was niet te veel moeite. We grappen wel eens dat ik duidelijk niet vooraan in de rij stond toen de lengte en het figuur werden uitgedeeld, maar ik denk dat dat komt omdat ik vooraan in een andere rij stond: de rij waar de beste ouders werden uitgedeeld. Mam, dank voor alle keren dat je de zorg voor mijn meiden op je hebt genomen en sorry voor alle dingen die we dit jaar niet hebben kunnen doen omdat dit proefschrift telkens weer in de weg stond. Ik weet dat jullie het wel weten, maar toch maar even voor de zekerheid en voor alle anderen die dit lezen: ik hou waanzinnig veel van jullie.

Het schrijven van dit proefschrift was net zo min mogelijk geweest zonder de hulp van mijn promotor en mijn co-promotoren. Johan, dank voor het in mij gestelde vertrouwen toen je me hebt aangesteld op dit project. Dat ik een proefschrift zou schrijven had ik zelf niet voor mogelijk gehouden, maar jij zag dit anders. Dank voor deze unieke kans, ondanks het zware afgelopen jaar ben ik bij dat ik deze kans gegrepen heb. In het begin van mijn project was je op de achtergrond aanwezig en beschikbaar wanneer nodig, maar met name het afgelopen jaar is het contact wat intensiever geworden rondom de afronding van het project. Dank voor je tijd die je in je drukke agenda hebt vrijgemaakt om te helpen en voor je scherpzinnige en eerlijke feedback op het door mij geleverde werk.

Henk, dank voor de vrijheid en het vertrouwen. Ik werk graag zelfstandig en vraag om input wanneer ik vind dat ik hier behoefte aan heb. Jij vond dit allemaal prima en dat heb ik zeer gewaardeerd. Hoewel jij in het begin het idee had dat ik eigenlijk niet op deze feedback zat te wachten, heb ik er altijd veel waarde aan gehecht en het stilvallen en de gereserveerdheid bij het krijgen feedback had niet zozeer te maken met dat ik moeite had met kritiek, maar meer met dat ik alles wat je zei goed wilde onthouden om hiervan te leren en mezelf te verbeteren.

John, je bent in mijn ogen altijd de rust zelve en ogenschijnlijk laat je druk en deadlines nooit de overhand krijgen. Dit heb ik als erg prettig ervaren en dit heeft voor mij een waardevolle bijdrage geleverd aan de werksfeer en aan de kwaliteit van het werk. Je opmerkingen waren altijd duidelijk en terecht en ik heb ze nagenoeg allemaal ter harte genomen. Je bent erg hulpvaardig en betrokken, ook al kom je hierdoor niet altijd toe aan je eigen werkzaamheden.

Henk, John en Johan, het was een plezier onder jullie supervisie te mogen promoveren en ik wens jullie alle goeds in jullie toekomst.

Thanks to all my colleagues at ABG. There's many of you, so I'm not going to mention all of you by name. Thanks to all of you that helped me with my work one way or another, or that gave PhD life that little bit extra during coffee breaks. Two of you I'd like to mention by name. Han, we worked together on macro-micro environmental sensitivity, a difficult topic that I tried to tackle in a relatively short timeframe. Thank you for all your patience in explaining the theory, and explaining it again, …and again. I appreciate the time investment you made and that you kept my deadlines in mind as this was the final study of this thesis. I wish you all the best in your future, both in your professional and personal life.

Alex, thank you for all your help with the computer and the difficulties that came with it. Although I'm sure I interrupted you many times, you always made time immediately and I want you to know that I appreciated it. Some problems were difficult to tackle, but fortunately most had a relatively easy answer (for you… ☺) You made life quite a bit easier. I wish you all the best.

Dank aan mijn mede DEC leden, met wie ik de afgelopen 6 jaar iedere maand vergaderd heb. Ik heb er altijd enorm veel plezier in gehad om met jullie te discussiëren over de ingediende proefplannen. Graag had ik nog een tijdje gebleven, maar helaas komt er een tijd om afscheid te nemen. Ook voor jullie alle goeds in jullie toekomst.

I'd also like to express my gratitude towards my friends. To all my friends, new and old, thank you for your support and your interest in my project and progression.

Eva, uit het oog is zeker niet uit het hart. Weet dat ik altijd maar 1 telefoontje van je verwijderd ben.

Panya, although you're far away in Norway, it feels as if you're very close by as you're always just one e-mail away. Thank you for your help and advice during the PhD and finishing this thesis. Also thank you for your postcards, please don't stop sending them ☺

Dianne, ik blijf het apart vinden hoezeer wij op elkaar lijken. Bedankt voor je luisterend oor en alle gezelligheid. Ook al zien we elkaar niet meer dagelijks, ik hoop dat we toch nog regelmatig samen een kopje thee zullen doen.

Katrijn and Ewa, you silly girls. Thank you so much for all the great 'girls nights' and the kidnappings. Great memories ☺

Mathijs, ik hoop dat we onze etentjes kunnen voortzetten ondanks dat we (tijdelijk?) met 1 persoon minder zijn.

Ze kunnen niet lezen en ze snappen er niets van wanneer het tegen ze wordt gezegd, maar een van de grootste bronnen van vreugde (en ook verdriet) tijdens mijn PhD waren mijn fretten Jasmin en Julliët en mijn kat Isabel. Dank voor de troostende lik en het bezorgen van een schaterlach wanneer nodig. Julliët, jou moeten laten gaan was de aller moeilijkste beslissing die ik heb moeten maken, het heeft me laten inzien hoe sterk ik kan zijn als dat nodig is. Als dat me er niet onder krijgt dan zal het afronden van deze PhD dat zeker niet doen.

Mijn lieve familie, waarmee ik het toch wel erg getroffen heb.

Tom, ook al zien we elkaar niet zoveel, we weten toch wel wat we aan elkaar hebben. Emiel en Mario, mijn tweede ouders ☺ Jullie waren er toen ik nog een huilbaby was, tijdens alle belangrijke momenten in mijn leven, ook nu en ik weet zeker ook in de toekomst. Weet dat ik heel veel van jullie hou. Gerard, Harriëtte, Jaap, Kristin, Astrid en Marc. Dank voor jullie steun, knuffels en belangstelling. Astrid, bedankt voor je bijzondere bijdrage aan mijn proefschrift. Jouw tekeningen op de kaft van dit proefschrift vind ik erg bijzonder, en dit maakt het dan toch een soort van familieproject ☺

En tot slot Richard, maat, partner ☺ Dank voor je geduld deze afgelopen maanden, ik weet dat ik je regelmatig op de proef heb gesteld (net als jij mij ;-)). Bedankt voor de zorg voor mijn (onze) meisjes wanneer ik even niet in staat was voor ze te zorgen. Als onze relatie een road-trip, kamperen, een verbouwing, fretten en een

PhD project overleeft dan zegt dat mij dat wij samen een sterk team zijn dat wel tegen een stootje kan. Ik hou van je.

Liefs, Susan

Colophon

Colophon