From the Institute of Animal Breeding and Genetics
Justus-Liebig-University Gießen

GENETIC EPIDEMIOLOGY OF BOVINE INFECTIOUS DISEASES

Dissertation
to obtain the doctoral degree (Dr. agr.)
The Faculty of Agricultural Science,
Nutritional Science and Environmental Management of
Justus-Liebig-University Gießen, Germany

Presented by
Medhat Mahmoud
Born in Cairo, Egypt

Gießen, Sep-2018

1st Referee: Prof. Dr. Sven König

Institute of Animal Breeding and Genetics

Justus-Liebig-University Gießen, Germany

2nd Referee: Prof. Dr. Dirk Hinrichs

Department of Animal Breeding

University of Kassel

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 3

CHAPTER 4

LIST OF FIGURES

CHAPTER 1

CHAPTER 3

CHAPTER 4

asterisk. I show results that do not pass multiple-testing correction as smaller squares. All genetic correlations in this report can be found in tabular form in Table S1. **92**

Figure 2. Genetic correlations among the 14 cow resistance traits analyzed by G-REML. Blue, positive genetic correlation; red, negative genetic correlation. Larger squares correspond to more significant P values. Genetic correlations that are different from zero at a false discovery rate (FDR) of 1% are shown as full-sized squares. Genetic correlations that are significantly different from zero after Bonferroni correction for the 105 tests in this analysis are marked with a yellow asterisk. I show results that do not pass multiple-testing correction as smaller squares. All genetic correlations in this report can be found in tabular form in Table S2. **94**

Figure 3. Estimated genetic correlations of birthweight (BW) and average daily gain (ADG) with all resistance traits in calves. This plot compares the genetic correlation between BW and all calf resistance traits with the genetic correlation between ADG and all calf resistance traits obtained from G-REML. The horizontal axis indicates pairs of phenotypes (BW and ADG), and the vertical axis indicates genetic correlation. Error bars represent standard errors. '***'P ≤ 0.001; '**'P ≤ 0.01; '*'P ≤ 0.05; '·'P ≤ 0.1. **95**

Figure 4. Estimated genetic correlations of average milk yield (AMY) and fat to protein ratio (FTP) with all resistance traits in cows. This plot compares the genetic correlation between AMY and all calf resistance traits with the genetic correlation between FTP and all calf resistance traits obtained from G-REML. The horizontal axis indicates pairs of phenotypes (AMY and FTP), and the vertical axis indicates genetic correlation. Error bars represent standard errors. '***'P ≤ 0.001; '**'P ≤ 0.01; '*'P ≤ 0.05; '·'P ≤ 0.1. **96**

Figure 5. P-values for the selected annotated candidate genes in all resistance traits in calves and cows. Dark red color means very high p-value, dark blue color means very low (i.e. more significant) p-value. **98**

CHAPTER 5

SUMMARY

A major part of animal husbandry is keeping animals healthy and disease-free. With respect to infectious diseases, this largely relies on control of the environment and prompt treatment or culling of infected animals. Until recently, little attention has been paid to the improvement of the genetic robustness or resistance of livestock to infectious pathogens. One reason is the lack of information regarding the genetic contribution of variation in resistance (or susceptibility) to pathogens between animals, and whether such resistance to one pathogen is positively or negatively associated with resistance to other pathogens and to livestock productivity such as growth rate or milk production. A major hurdle in providing this information has been the lack of data collected from animals affected by pathogens and the difficulty in estimating associations between resistance to disease using data from different animals affected by different diseases. Thus, the overall aim and content of this thesis is a comprehensive genetic evaluation of infectious disease traits using all available epidemiological and genomic tools. This thesis is divided into four chapters.

CHAPTER 1 and 2 is an introduction covering the most important epidemiological and genomic concepts necessary for readers from a wide variety of disciplines, with a focus on infectious diseases and the application of genomic tools to disease resistance.

In **CHAPTER 3,** relationships between certain infectious diseases (diarrhea and respiratory disease) in calves and the production and disease traits of first-parity cows during the early lactation period are analysed. Only weak phenotypic associations between calf and cow traits were identified. Moreover, genetic correlations and those based on single nucleotide polymorphisms (SNP)-marker effects were close to zero. The results indicate that calf and cow diseases are different traits, implying that selection on calf traits is unrelated to genetic improvements in lactating cows.

To the best of our knowledge, I estimate for the first time in cattle (or any other species) (in **CHAPTER 4)** the contribution of genetics to the resistance to multiple infectious pathogens, and the associations among different pathogen-resistance traits and growth and milk production, using data from a large infectious disease surveillance program of German dairy cattle in conjunction with genomic information**.** I show that there are often positive associations between resistance to similar diseases, and also an example of a negative association (increased resistance to one disease increases susceptibility to another) . This information will guide the breeding of

more robust animals that require less antibiotics and other treatments to stay healthy, in addition to stimulating further work related to the elucidation of mechanisms that underpin these associations.

Finally, in **CHAPTER 5**, a general discussion focusing on the most pertinent results and observations from the previous chapters is given. In the context of host resistance genetics, this thesis questions the traditional hypothesis of disease resistance selection based on multiple pathogen infections in cattle (genetic correlation), since every pathogen resistance trait can have either a positive or negative correlation with other pathogens or with productivity and performance traits (real pleiotropy).

Definitions and abbreviations (adapted from Veterinary Epidemiology (Thrusfield, 2007))

Outbreak = "An occurrence of disease in an agricultural or breeding establishment, including all buildings and adjoining premises, where animals are present".

Prevalence = "The number of instances of disease or related attributes (infection or presence of antibodies) in a known population, at a designated time, without distinction between old and new cases".

Incidence = "The number of new cases that occur in a known population over a specified period".

Mortality = "Analogous to incidence measures where the relevant outcome is death associated with, rather than new cases of, a specific disease".

Survival = "The probability of animals with a specific disease remaining alive for a specified length of time".

Resistance = "Refers to mechanisms that restrict the reproduction rate of pathogens within a host, for instance, by blocking pathogen entry or limiting pathogen replication".

Tolerance = "Refers to the ability of a host to limit the detrimental impact that pathogens can inflict on host performance (growth and milk production), without affecting pathogen burden".

cGDS = general calf disease status

cDIA = calf diarrhea

cRD = calf respiratory disease

flGDS = first-lactation general disease status

flRD = first lactation respiratory disease

flDIA = first-lactation diarrhea

flFF = first-lactation female fertility disorders

flCLAW = first-lactation claw disorders

flMAST = first-lactation mastitis

flMET = first-lactation metabolic disorders

Staph. Aureus = Staphylococcus aureus

Staph. Haemolyticus = Staphylococcus haemolyticus

Strep. Agalactiae Streptococcus agalactiae

Strep. Dysgalactiae = Streptococcus dysgalactiae

Strep. Uberis = Streptococcus uberis

CHAPTER 1

General introduction

Preface

Livestock breeders are faced with the challenge of maintaining and improving the health and welfare of farm animals whilst simultaneously being pressured by society to reduce antibiotic and drug treatments in animals. However, despite the commencement of breeding livestock for increased disease resistance, there is very limited information regarding the consequences of breeding for resistance to single pathogens on the resistance to other pathogens or on production traits. To optimise breeding programmes, estimates of the genetic correlations between these traits are required, which is what this thesis provides for the first time in Holstein dairy cattle, the single most important livestock breed worldwide. To the best of our knowledge, such large, carefully controlled and diagnosed recordings of multiple pathogens within the same basic population (of livestock or otherwise) have not been performed; thus, such analyses have not been possible. Large-scale genomic studies of individual pathogens in humans are beginning to appear; however, I am not aware of any studies showing pleiotropy between diseases and I expect that my thesis will stimulate interest in this approach.

When considering the genetic evaluation of infectious disease resistance, it is necessary to be clear and consistent in the definitions and main concepts of epidemiology, to ensure that readers from different disciplines have a common level of understanding.

Epidemiological measurements

In the most recent edition of *A Dictionary of Epidemiology* (Sixth Edition, 2014), Porta redefined the definition of epidemiology as "the study of the occurrence and distribution of health-related events, states, and processes in specified populations, including the study of the determinants influencing such processes, and the application of this knowledge to control relevant health problems". A literal translation of the word "epi-demi-ology", based on its ancient Greek roots επι- (epi) = upon, δημο- (demo) = people, and λογο (logo-) = discoursing, which is "the study of that which is upon the people".

Uses of epidemiology

There are five general objectives of epidemiology: (1) determination of the origin of a disease with a known cause; (2) investigation and control of a disease for which the cause is either unknown or poorly understood; (3) acquisition of information regarding the ecology and natural history of a disease; (4) planning, monitoring, and assessment of disease control programs; and (5) assessment of the economic effects of a disease, and analysis of the costs and economic benefits of alternative control programs (Porta, 2014).

Infectious disease measurement & quantitation

Counting the affected animals in a population is essential to the investigation of any disease and its quantitative description. Furthermore, it is usually desirable to describe the who, what, when, and where of the disease, and to relate the number of diseased animals to the population size at risk of developing the disease, allowing assessment of disease burden. For instance, the report of ten cases of paratuberculosis infection on a farm does not indicate the true extent of the problem unless the report is considered in terms of the total number of animals on the farm, and how many animals are capable of transmitting the infectious pathogen (Krieger, 2001).

Determinants of infection diseases

Prior to the mid-twentieth century (the microbial revolution), the investigation of infectious diseases began to dominate, and epidemiologists developed a traditional model. The determinants were classified into those associated with the host, the agent, and the environment. These three groups of factors are called the *epidemiological triad* (Figure 1) (Rockett, 1999; Krieger, 2001) in which the individual components can interact in a variety of ways resulting in various states of health of an individual or population.

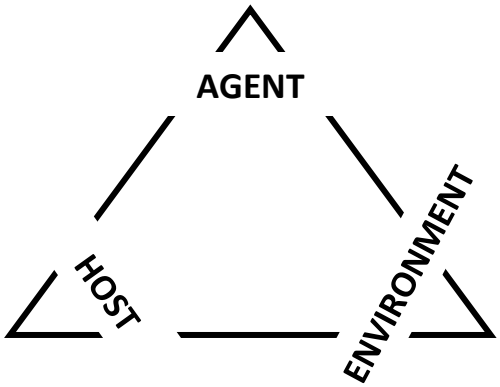


Figure 1. The epidemiologic triad, adapted from Rockett (1999)

Agent determinants

"Infectious agent" is a term that originally referred to a living entity, or microorganism (bacterium, virus, fungus, parasite), capable of causing disease. However, this term also includes certain chemical components, for instance, L-tryptophan contaminant, and physical injuring forces. All kinds of agents vary in their ability to infect or induce a disease in animals and plants (Mills et al., 1995), and this ability mainly depends upon the level of host acquired immunity or inherent susceptibility. "Infectivity", "pathogenicity", and "virulence" are the main terms used to describe and quantitate the ability of any agent to infect or induce disease to a host. *Infectivity* is the ratio of the number of infected individuals to the number of exposed individuals, *pathogenicity* is the ratio of the number of clinically apparent disease cases to the number of infected individuals, and *virulence* is the ratio of the number of severe or fatal cases to the number of clinically apparent cases (Last, 2001). Several biological agents (pathogens) will be covered and quantitated in this thesis.

Bacterial pathogen: Bacteria can cause a disease by three different mechanisms: The production of toxins in foods and water, spreading via parasites, or the causation of direct host inflammation. Among the bacterial pathogens found on dairy farms, I selected those most frequently existing in the German herds that are under investigation: Salmonella, Escherichia coli, *Staphylococcus aureus* (Staph. Aureus), *Staphylococcus haemolyticus* (Staph. Haemolyticus), *Streptococcus agalactiae* (Strep. Agalactiae), *Streptococcus dysgalactiae* (Strep. Dysgalactiae), *Streptococcus uberis* (Strep. Uberis), *Clostridium perfringens* (Clost. Perfringens), and *Mycobacterium avium paratuberculosis* (Mycobac. Paratuberculosis) (Figure 2).

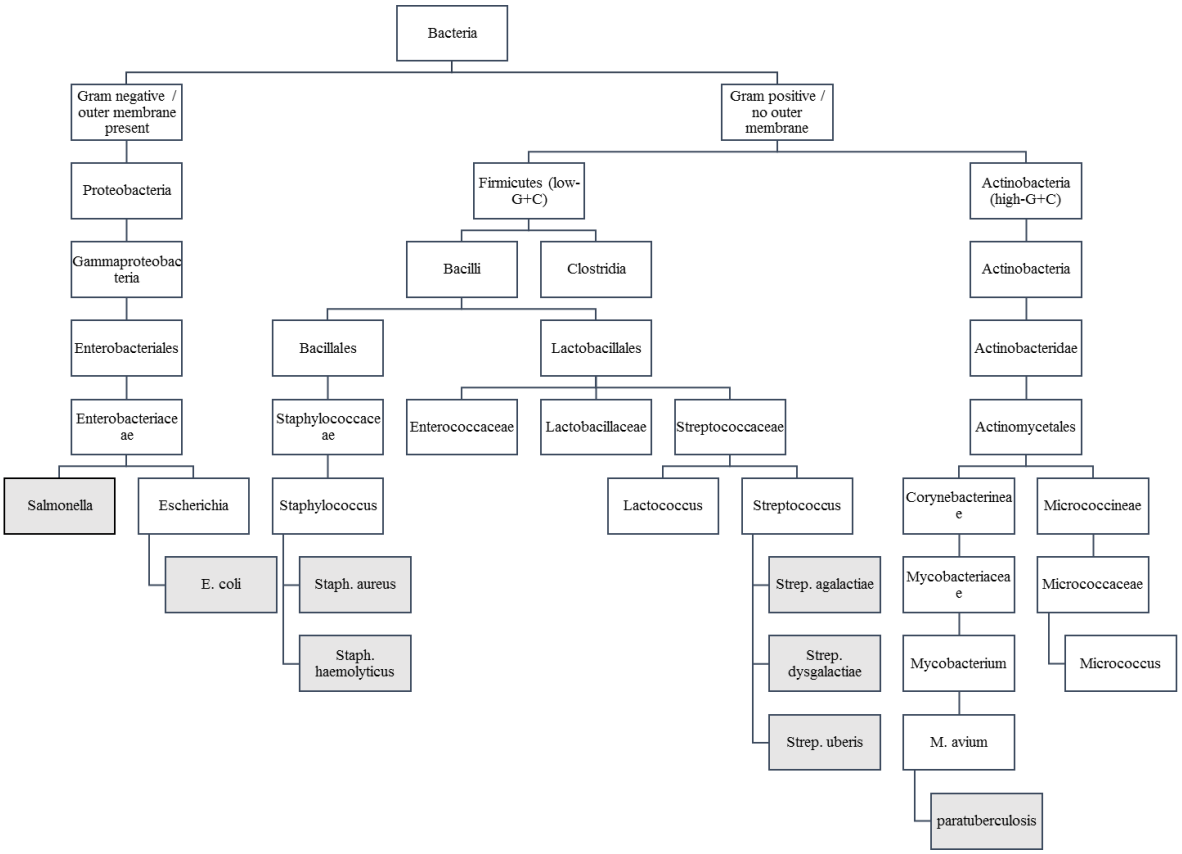


Figure 2. The phylogenetic tree of the selected bacterial pathogens.
"Low G+C": Their DNA typically has fewer G and C DNA bases than A and T bases as compared to other bacteria.
"High G+C": Their DNA typically has higher G and C DNA bases than A and T bases as compared to other bacteria.

Viral Pathogens. A highly contagious agent that has been classified by scientists to be between a living and non-living organism. Viruses cannot replicate on their own but can replicate very fast in a living cell to induce disease in the host. In my study I selected the following strains: *Rotavirus, Bovine Respiratory Syncytial Virus and Bovine herpesvirus 1* (Figure 3).

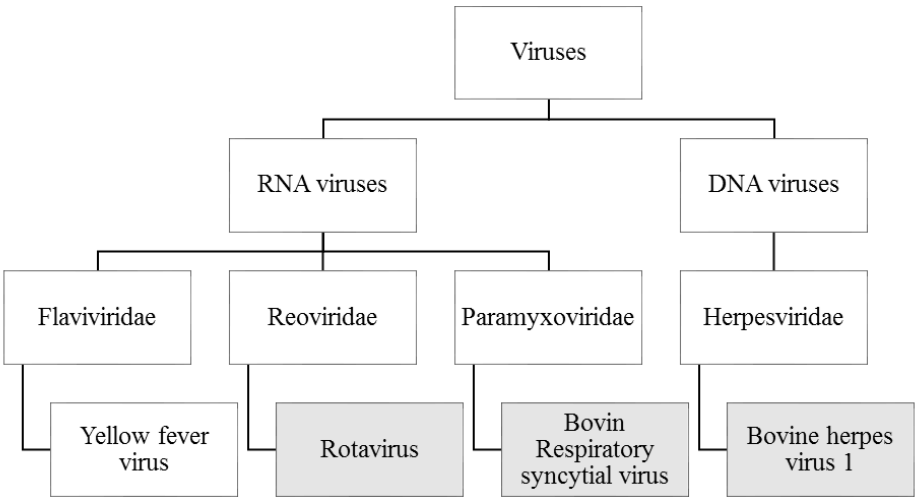


Figure 3. The phylogenetic tree of the selected viral pathogens.

Fungal pathogen. A biological agent that are not particularly dangerous; however, certain types can cause harmful diseases. In this study I selected only one strain of fungi: *Trichophyton* (Figure 4).

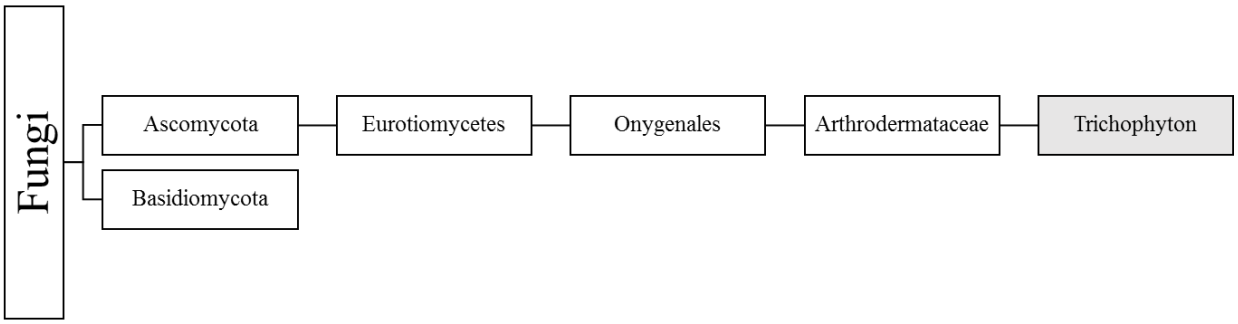


Figure 4. The phylogenetic tree of the selected fungal pathogens.

Parasitic pathogen. A biological agent that lives on/in a host animal and obtains its food from or at the expense of its host. In addition to Myiasis, the following parasites cause very harmful diseases in cattle: *Cryptosporidium, Coccidia, Bovicola bovis. Dictyocaulus viviparous and Chorioptic scabies* (Figure 5).

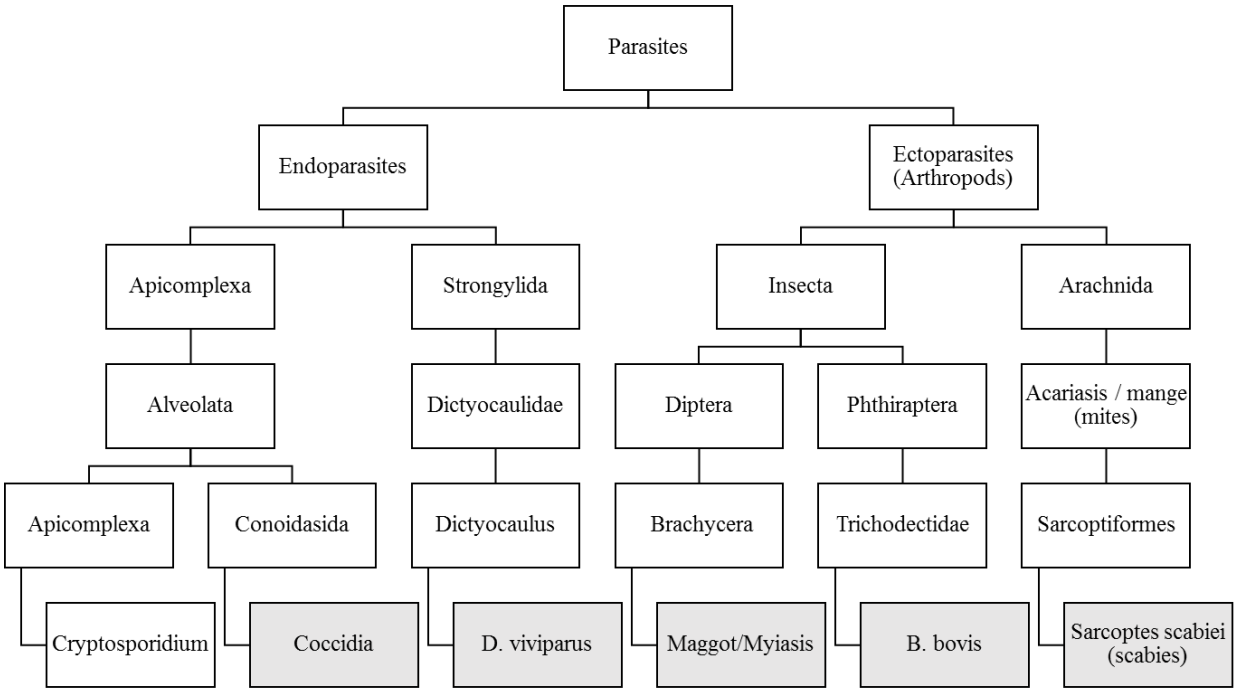


Figure 5. The phylogenetic tree of the selected parasitic pathogens.

Environment determinants

Environmental determinants are external factors that affect host immunity and agent infectivity, and include physical factors such as weather, climate, and location; biological factors such as insects that transmit certain agents; and socioeconomic factors such as overcrowding, sanitation, general husbandry, and management practices (Kirkwood, 2003). Since the climate is the same in Northeast Germany, particular attention will be paid to climate in the form of temporal distribution.

Temporal distribution of infectious diseases.

The random or systematic fluctuation in disease occurrence in a given population (Appendix - A) can be classified into four major distributions (endemic, epidemic, pandemic, and sporadic).

***Endemic*:** The constant presence or usual frequency of an infection in a given population. An example is infection with *Dictyocaulus viviparus*, which has stayed constant over the last 4 years (Figure 6).

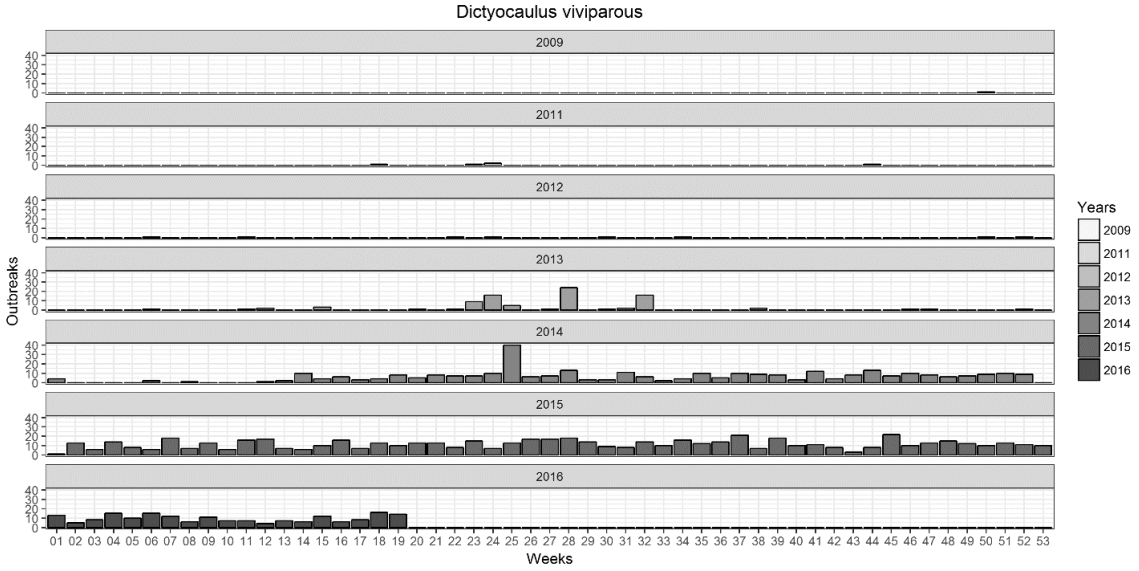


Figure 6. Example of endemic trends in infectious disease occurrence (infection of Dictyocaulus viviparous).

.

***Epidemic*:** A sudden unexpected increase in the occurrence of outbreaks of a disease in a given population. In my study, infection of Mycobac. Paratuberculosis during the 35th week of 2015 is a good example of an epidemic infection; when the number of outbreaks exceeds the expected (endemic) level (Figure 7).

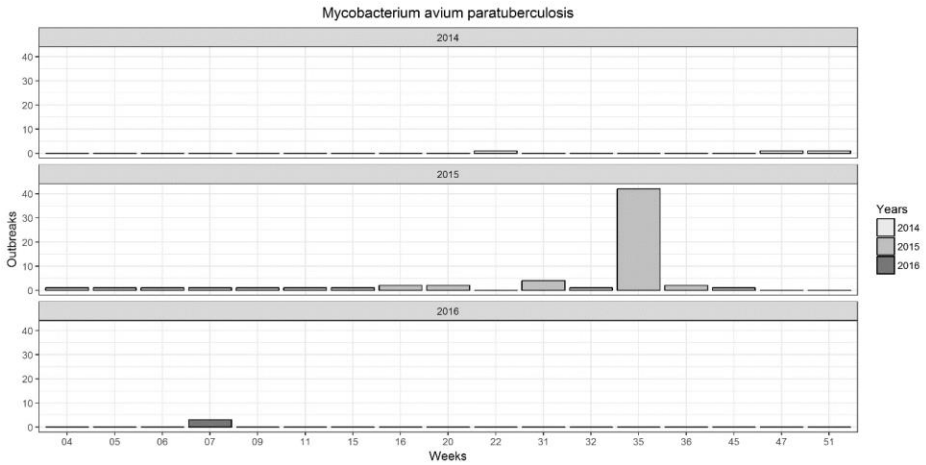


Figure 7. Example of epidemic trends in infectious disease occurrence (infection of Mycobac. Paratuberculosis).

***Pandemic*:** A widespread epidemic that usually affects a large proportion of the population. Infection with bovine herpesvirus 1 is an example of this distribution (Figure 8).

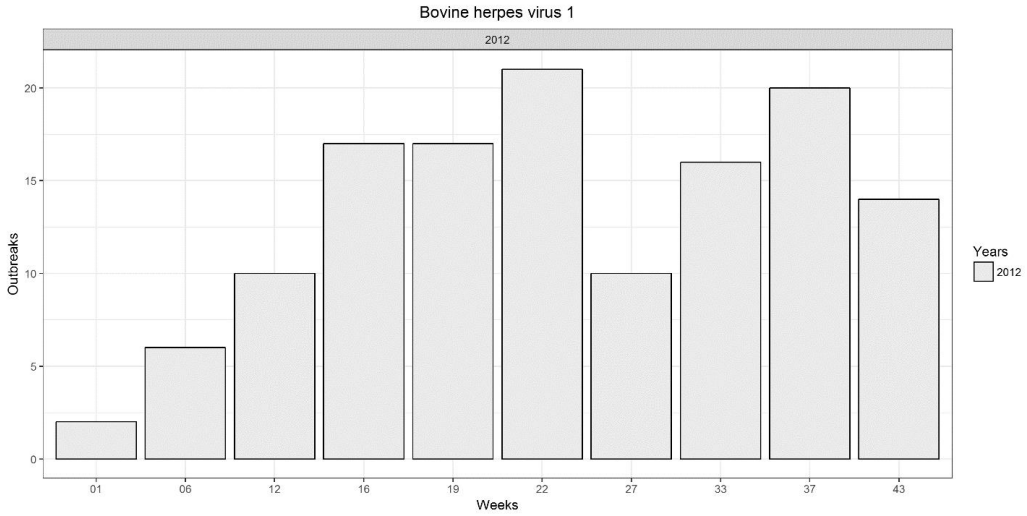


Figure 8. Example of pandemic trends in infectious disease occurrence (infection of Bovine herpesvirus 1).

***Sporadic*:** Infection with the *Bovicola bovis* insect can help to illustrate this type of distribution (Figure 9), where infection occurs irregularly and indiscriminately.

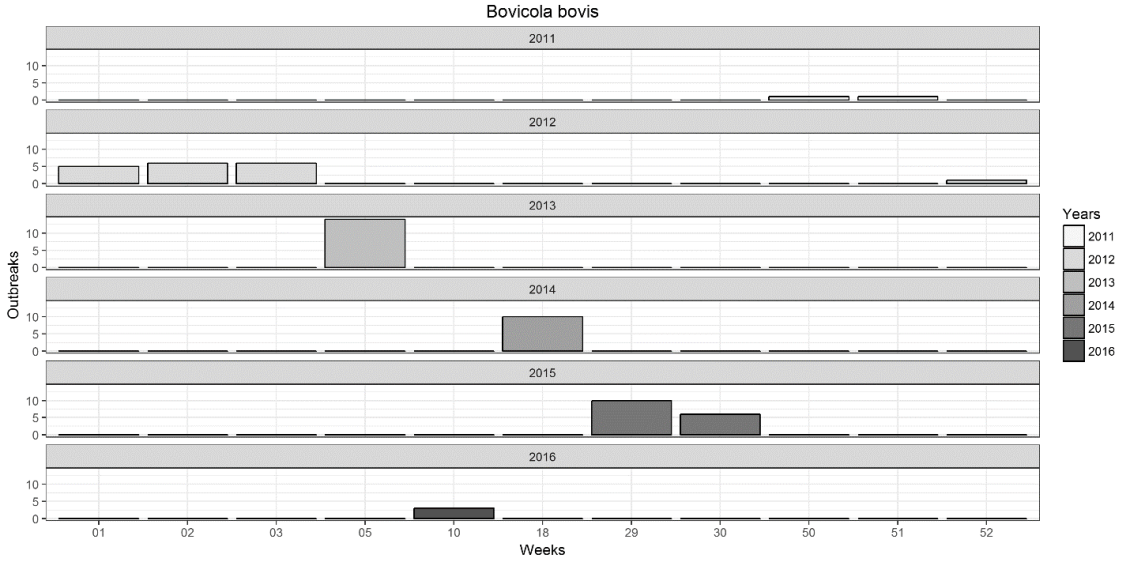


Figure 9. Example of sporadic trends in infectious disease occurrence (infection of Bovicola bovis).

Host determinants

There are three main host characteristics that control the distribution of infectious disease in any population: age, sex, and genotype of the host.

***Host age*:** Host age shows a distinct association with the frequency or occurrence of many infectious disease. For instance, many viral and bacterial infections are more likely to occur (sometimes with a higher mortality rate) in young as compared with old animals. In contrast, many fungal and parasitic infestations induce milder responses in young as compared with old animals (Thrusfield, 2007). To determine the impact of host age on pathogen infectivity, I illustrate the number of cases of each pathogen during an animal's life (Appendix - B). Thus, I have classified all pathogens into three groups:

1- Calf diseases: those recorded in the first five months of animal life (Figure 10A)
2- Cow diseases: those recorded after the first calving (Figure 10B)
3- Calf-cow diseases: those recorded at both ages (Figure 10C).

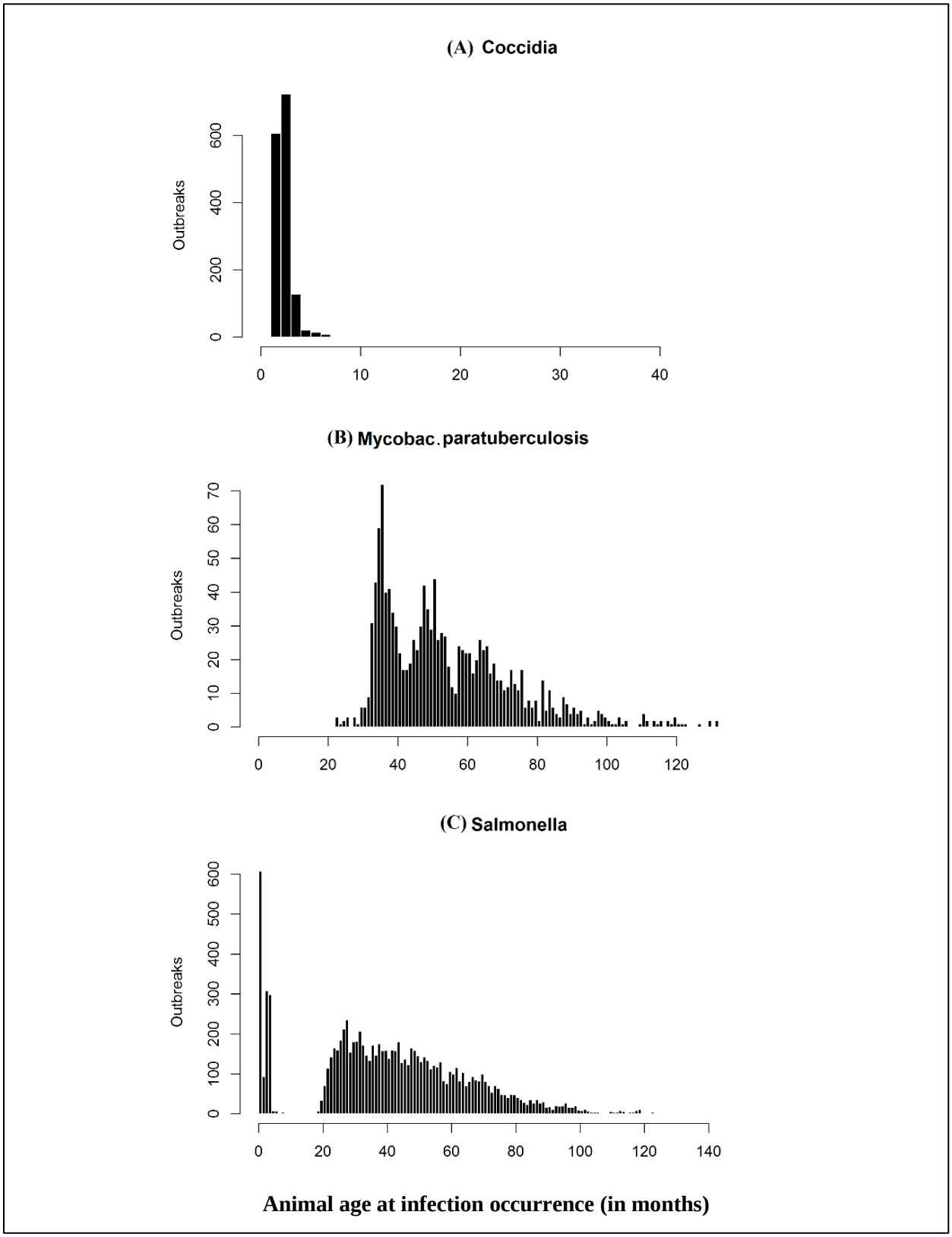


Figure 10. Impact of host age on infectious disease occurrence, A- Coccidia (Calf diseases), B- Mycobac. Paratuberculosis (Cow diseases) and C- Salmonella (Calf-Cow diseases).

***Host sex*:** Sex differences with respect to hormone type, production type, and genetics of the host can cause many differences in disease occurrence. For instance, female hormones such as oestrogen have been found to have an effect on the development of diabetes mellitus in dogs (Marmor et al., 1982). Moreover, sex chromosomes (X and Y) can induce certain genes associated with resistance to infectious diseases, by being either sex-linked, sex-limited, or sex-influenced (Patterson and Medway, 1966).

***Host genetic*s:** According to Nicholas (2010), genetically controlled disease traits belong to one of three categories: 1) chromosomal abnormality diseases, 2) Mendelian diseases (controlled by a few inherited genes), and 3) quantitative diseases (controlled by many genes). The first two categories represent the non-infectious diseases that are genetically disordered. Infectious diseases are mostly defined as quantitative traits, where many genes are in control either by resistance or susceptibility to pathogens. Thus, the mode of inheritance of resistance or susceptibility to infectious diseases will be discussed in more detail in **Chapters 2 and 3.**

Economic loss and prevention methods

The spread of infectious disease in farm animals results in a large economic loss for both developing and developed countries. Many pathogens (mostly bacterial and viral) have a high impact on the safety of animal products, public perception of livestock production industries, and animal welfare. Moreover, new infectious disease threats continue to emerge due to the impacts of climate change and globalisation (migration and product exchange) (Foresight Project, 2006). Hence, the management and control of infectious diseases in animals and humans must be a high priority of ongoing research.

To date, the current strategies for the prevention and cure of infectious diseases are culling diseased animals and vaccination and selection of resistant animals (**host**); biological control of a pathogen (**agent**) by the introduction of a natural enemy or predator); or biosecurity and sanitation (**environment**). Recent developments in high-throughput genotyping/sequencing platforms allow for greater understanding of the genetic architecture of host responses to infection, by identifying genomic loci contributing to host resistance and susceptibility, quantitating the magnitude and direction of single nucleotide polymorphism (SNP) effects, and elucidating the interaction between genes. Although the use of SNP genotyping technology is a promising approach, it must be

critically evaluated prior to application to genomic estimating breeding values (G-EBV) for use in conventional selection schemes. Some of these concepts are discussed below.

Due to the complexity of resistance to infectious disease, using host genetics to improve animal health and resistance will be a valuable approach for certain pathogens, while for others it may be of low priority compared with other pathogen-control strategies, if even appropriate. Therefore, genetic evaluation of all available resistance to infectious pathogen traits and the identification of certain markers underlining these traits are required. Furthermore, special attention should be given to estimating the genetic correlation among these traits and the pleiotropy between resistance traits and productivity and performance traits.

Genetics of disease resistance

Current breeds of cattle and other livestock species have inherited their genome from the historical wild breeds (progenitors). Despite the clear phenotypic variation across livestock breeds and within species, only small differences among such animals have been revealed at the DNA level through molecular studies. Host genetic variation in disease resistance is one of these similarities between current and progenitor breeds, and is one of the reasons that the co-evolution process between the host and its invading pathogen is crucial for the survival of both entities (Khibnik and Kondrashov, 1997)

More than 50 diseases in all major livestock species have been evaluated to quantitate the genetic variation involved in host resistance (Bishop, 2005). Three important factors are assumed to be the main sources of genetic variation in host resistance: (1) the fluctuation in selection pressure across time and environments, especially with respect to disease resistance; (2) during certain historical epidemics, not all the population succeeded in becoming completely resistant to the infection; and (3) artificial selection of other productivity traits affected the genetic variance of host resistance through pleiotropy between productivity and resistance. These genetic variations cover all types of pathogens and genetic architecture of host resistance (Bishop, 2010), and also confirm that there is a host genetic variation in resistance to almost all diseases; however, further investigations are required.

Genomics and disease resistance

Following the discovery of the structure of DNA in 1953 (the beginning of molecular revolution), geneticists began to develop novel genotyping technologies that led to the identification of genetic markers (microsatellites). After the sequencing of the whole genome in 2001, the complete genome sequences of several livestock species became available, subsequently aiding in the detection of SNPs throughout the entire genome. The molecular revolution resulted in several advantages:

1- The ability to identify the causative mutations responsible for host resistance against the most dangerous pathogens.
2- The ability to determine the exact sequence of certain genes underlying the genetic variation between hosts and agents.
3- The ability to detect the product of each gene (protein) and their transcriptional location (tissue) and time (age).

These advantages presenting a proper tools with which to dissect and understand host genetic variation in infectious disease resistance.

Pleiotropy versus genetic correlation

At the single locus level, pleiotropy is a biological phenomenon in which a single genomic variant (gene or marker) influences more than one phenotypic trait. Genetic correlation, however, is the statistical interpretation of this phenomenon and is used to quantitate the level of influence one variant has on multiple traits. For instance, if one variant affects multiple traits in the same direction (positive or negative effect), this generates a genetic correlation; however in principle, there could exist a balance of positive and negative effects on two traits such that no genetic correlation exists. In animal breeding, pleiotropy has been known for decades, since the selection for milk yield has certain effects on other traits. In 1957, Farthing and Legates reported one of the first negative genetic correlations between milk yield and fat percentage in milk. In **Chapters 2** and **3** I report several further genetic correlations.

Biological pathway analysis

Biological pathway analysis is an approach employed to evaluate the association between a select set of genes (biological pathways) and a trait of interest, and can be used to assess the cumulative genetic effects across multiple genes. Typically, biological pathway analysis consists

of two main steps: (1) to annotate SNPs to their corresponding genes (identifying genes linked to these SNPs) and assign these genes to their corresponding biological pathways using a public database such as KEGG; (2) to test the association between a biological pathway (set of genes with assigned SNPs) and the phenotype using statistics (*p*-values) obtained in a genome-wide association study (GWAS).

Scope of the thesis

The scope of this thesis includes several important topics related to infectious disease resistance traits; from the epidemiological status of infectious disease in Northeast Germany to the genetic evaluation of these traits using pedigree, genome-wide SNPs, and gene network analysis, including the correlations among these diseases and their impact on performance and productivity in German Holstein cattle. The objectives of this thesis can be summarised in the following five points:

1- All necessary definitions and epidemiological descriptions of all infectious disease occurrence in Northeast Germany, including the federal states of Mecklenburg-West Pomerania and Berlin-Brandenburg, are stated in **CHAPTER 1**.

2- Estimation of the genetic relationship between calf diseases and cow production and health traits is the main aim of **CHAPTER 3,** on phenotypic (using linear and generalised linear mixed models), quantitative genetic (using the pedigree relationship matrix), and genomic (using GWAS summary statistics) scales.

3- There are two main objectives in **CHAPTER 4**: a) Implementation of the G-REML method to estimate the genetic parameters for all resistance traits and their genetic correlations with other productivity and performance traits; b) Estimation of pleiotropy among all resistance traits using post-GWAS function analysis on gene and gene-sets (biological pathways).

4- The final **CHAPTER 5** discusses the outcomes of all previous chapters and summarises the simple comparison between pleiotropy and genetic correlation, indicating the best phenotypes for each genetic analysis.

REFERENCES

Bishop, S.C. 2005. Disease resistance: Genetics. In: Pond, W.G. and Bell, A.W. (eds) Encyclopedia of Animal Science. Marcel Dekker, Inc., New York, pp. 288–290.

Bishop S.C. Axford R.F.E. Nicholas F.W. Owen J.B. 2010. Breeding for disease resistance in farm animals. 3rd edn CABI

Farthing, B.R., and J.E. Legates. 1957. Genetic Covariation between Milk Yield and Fat Percentage in Dairy Cattle1. J. Dairy Sci. 40:639–646.

Foresight Project (2006) Infectious diseases: Preparing for the future. Executive Summary. Office of Science and Innovation, London.

Khibnik, A.I. and Kondrashov, A.S. 1997. Three mechanisms of Red Queen dynamics. Proceedings of the Royal Society of London Series B – Biological Sciences 264, 1049–1056.

Kirkwood, J.K. 2003. Welfare, husbandry and veterinary care of wild animals in captivity; changes in attitudes, progress in knowledge and techniques. In: International Zoo Yearbook, 38. Eds Olney, P.JS. and Fiskin, F.A, pp. 124-130. The Zoological Society of London, London

Krieger N. 2001. A glossary for social epidemiology. J Epidemiol Community Health 55:693–700.

Last JM, editor. Dictionary of epidemiology 2001. 4th ed. New York: Oxford University Press; p. 61.

Marmor, M., Willeberg, P., Glickman, L.T., Priester, W.A., Cypess, R.H. and Hurvitz, A.L. 1982. Epizootiologic patterns of diabetes mellitus in dogs. American Journal of Veterinary Research, 43, 465-471

Mills, D., Kunoli, H., Keen, N.T. and Mayama, S. (Eds) 1995. Molecular Aspects of Pathogenicity and Resistance. APS Press, St Paul.4

Nicholas, F.W. 2003. Introduction to Veterinary Genetics, 2nd edn. Blackwell Publishing, Oxford.

Patterson, D.F. and Medway, W. 1966. Hereditary diseases of the dog. Journal of the American Veterinary Medical Association, 149, 1 741-1 754

Porta, M.S., S. Greenland, M. Hernán, I. dos S. Silva, J.M. Last, and International Epidemiological Association eds. 2014. A Dictionary of Epidemiology. Six-edition. Oxford University Press, Oxford.

Rockett IRH. 1999. Population and health: An introduction to epidemiology. 2nd ed. Washington, DC: Population Bulletin.

Sinnecker, H. 1976. General Epidemiology (translated by N. Walker). John Wiley and Sons, London

Thrusfield, M. 2007. Veterinary Epidemiology. Wiley.

Nicholas, F.W. 2010. Introduction to Veterinary Genetics, 2nd edn. Blackwell Publishing, Oxford

CHAPTER 2

Genetic resistance to diseases and parasites

History and fundamentals

The control of diseases and parasites is one of the major problems encountered in the efficient production of livestock. Evidence suggests that some animals may be genetically resistant or genetically susceptible to disease and to parasites. In the long run it is not the proportion of individuals that die that are of the greatest economic importance but those that are so affected that they make slow and inefficient gains or are of lowered fertility and producing ability. Disease and parasitic infestations in livestock cause millions of dollars of losses to livestock producers every year in all parts of the world. This has been true since animals were first domesticated. Selection for superior performance in livestock also includes a certain amount of automatic natural selection for genetic resistance to disease and parasites because those that perform best must be healthy and free from infections or infestations. To date man has done very little deliberate selection for genetic resistance to disease and parasites although nature has always selected in this direction.

Much more selection for disease resistance by man has been done in plants than in animals. The development of strains of plants resistant to certain diseases has been successful. Two good examples are the development of varieties of wheat that are resistant to stem rust (Mago, et al., 2005; Rutkoski et al., 2011) —a disease that ravaged wheat in the spring wheat areas a few years ago. More recently genetic susceptibility to Southern Corn Blight (Bruns et al., 2017) caused considerable loss to farmers in the Corn Belt through proper breeding methods, however, this disease was brought under control in a year or two. New plant diseases of economic importance appear from time to time, but plant geneticists are able, as a general rule, to quickly develop new varieties and strains which are genetically resistant to these diseases.

Problems of genetic control of disease in animals

Procedures for developing genetically resistant strains or varieties of plants are seldom applicable to animals, and there are several reasons why this is true. Plants are much more fertile than animals, and individual plants are much cheaper to produce. For example, one ear of corn can produce enough seed to produce many plants in a single year. One animal, on the other hand, produces only one or only a few progenies each year. A new generation of plants can be produced every year and sometimes almost twice each year if seed produced in temperate zones are transported and planted in tropical or semi-tropical climates. Livestock have a much longer generation interval than plants. True, it is possible to produce a new generation of swine each year, but in cattle and horses the generation interval is four or more years.

Plants may be exposed to a disease, and those found to be resistant may be used for reproductive purposes. Exposure of animals to a disease to identify those genetically resistant would involve large economic losses and would be impractical for livestock producers and veterinarians because of high death losses and possible reservoirs of a disease in those that survive. Those that survive might be a source of infection to nearby herds and flocks. Genetic resistance to disease in livestock is often due to many genes (polygenes) and not to a single pair of genes. In such cases the development of a strain of animals resistant to a disease would be a long, slow process because the animals produced would have to be exposed to the disease to determine if they were resistant before selections could be made.

Probably the most important problem encountered in developing strains of genetically resistant animals is that genetic resistance to disease appears to be specific and not general. In other words, one might eventually develop a strain of animals genetically resistant to one disease but they would be susceptible to another. It is obvious, then, that developing strains of animals that are genetically resistant to disease involves many problems and may not be economically feasible.

Infectious and non-infectious diseases

The terms genetic resistance and susceptibility to disease imply that some individuals when exposed to a disease become ill whereas others do not. For practical purposes I may consider both resistance and susceptibility together in discussing the relationship of genetics to disease. If I was trying to develop resistant lines or strains of animals I would be trying to increase the proportion of resistant animals in a population while decreasing the proportion of those that would be susceptible.

Disease may be defined as any condition of animals in which there is a deviation from a normal state of health: an interference with the vital functions of the body (Campbell et al., 1975). Two general kinds of disease are known: infectious and non-infectious. Infectious diseases are caused by pathogenic organisms such as bacteria, multicellular parasites, fungi, protozoa, and viruses. Non-infectious diseases may result from mechanical ailments such as wounds and rupture; digestive disturbances such as bloat: chemical diseases caused by the intake of such chemicals as arsenic and nitrates: the ingestion of poisonous plants; and genetic disorders such as hemophilia and diabetes. Usually when we think of genetic resistance to disease we are considering only

infectious diseases, but non-infectious diseases also can be of economic importance to man and the livestock industry.

Evidence for genetic resistance to infectious diseases

Some individuals in a large population possess genetic resistance to most, if not all, infectious diseases. Much of this evidence is from general observation because little research has been done to clarify this point in the case of many diseases. Two or more species of animals often differ in their degree of resistance or susceptibility to certain diseases. Species differences are genetic differences. Cattle do not appear to be susceptible to certain diseases affecting swine such as hog cholera; man is apparently immune to many diseases affecting livestock; and vice versa. Certain diseases, however, such as brucellosis affect man, cattle, swine and other animals although the organism causing the disease in one species may be a different strain than the one causing the disease in another zebu cattle (*Bos Indicus*) appear to be resistant to some diseases to which european cattle (*Bos Taurus*) are susceptible (Francis, 1966).

Individuals within the same species also vary in their resistance or susceptibility to an infectious disease. Before vaccines for polio were developed (Hannik and Cohen, 1978.) not everyone contracted the disease although probably most were exposed to it. When a virulent disease strikes a herd or flock seldom do all individuals show symptoms of the disease. Those who do Contract the disease show a of variation: from those that show only slight symptoms to those that die Of course, the question arises as to whether or not all individuals in a herd or flock have the same degree of exposure to the disease-causing organism, but even the degree of exposure can, in some instances, have a genetic base. For example, the thickness of the skin or hair or the presence of disease repelling or destroying secretions on the surface of the body, or within the body, might prevent infections in some animals. Even these body defenses may have a genetic base.

A good example of genetic resistance and susceptibility to disease has been described for the Australian rabbit (Ingersoll, 1964). In 1859, twenty-four wild rabbits (*Oryctoloqus cuniculus*) were brought from Europe to Australia, thirteen of which were released that year on an estate near Geelong on the southern coast of Victoria. Three years later they were looked upon as pests and six years later 20,000 rabbits were killed on this estate, and it was estimated that 10,000 remained alive. By 1928, the rabbits had spread over nearly two-thirds of the Australian continent, and by

1953 it was estimated that between 500 million and 1 billion adult rabbits were present in an area of over one million square miles. It was estimated that if no rabbits were in Australia, the land could support an additional 100 million sheep, or the equivalent in cattle, and the increased income to the Australian economy from wool and beef would be about 898 million dollars per year. The rapid growth of the rabbit population was attributed to a soil ideal for burrowing, a plentiful supply of food and the absence of several important natural enemies.

In 1950 and 1951 rabbits in several areas of Australia were inoculated with a virus known as infectious myxomatosis which in laboratory tests in Europe had proved to cause a high mortality rate in rabbits. The virus also had a high degree of host specificity and would not affect other valuable species of animals. The disease spread rapidly and it was estimated that 80 to 90 percent of the rabbit population was destroyed. By 1953, it was thought that the rabbit problem in Australia had been virtually eliminated. In sub sequent years, however, there was an increase in the rabbit populations much of which was attributed to resistance to the disease. Research also indicated that genetic changes in the virus itself may have resulted in strains developing which were less virulent.

In combating *leucosis deliberate* selection by man for increased resistance to disease was practiced. Leucosis (lymphomatosis) is a disease of poultry that became one of the poultryman's worst problems. It caused 30 to 50 percent of the mortality among pullets during their first year of life. Some estimates are that losses from this disease ran as high as 75 million dollars annually. The disease is characterized by an accumulation of lymphocytes in the nerves, the visceral organs. and sometimes the iris of the eyes. it is now thought that the disease is caused by two or more different viruses.

Cornell University (Hutt, 1958) designed a selection experiment wherein two strains of White Leghorns were bred for resistance to the disease and a third for increased susceptibility. Selection in both directions was effective cases of lymphomatosis gradually decreased in the resistant lines so that its occurrence was almost negligible. Egg production also increased in these 1ines Selection for increased susceptibility made the strain so susceptible that a few years more than half of the pullets died of the disease. The development of a vaccine against this disease helped save the problem in controlling this disease, and nullified, to a certain extent, all of the work that had been done to select for genetic resistance to this disease.

Body defenses against disease

The body has many defenses against infectious diseases beginning with the skin outside the body and including several mechanisms within the body. Many of these defense mechanisms are under genetic control as shown by breed and line differences in susceptibility to infections as well as certain reported genetic defects which make affected individuals more susceptible to infections.

The skin. The skin is one of the first lines of defense against bacterial infections. It plays largely a mechanical role. Bacteria usually cannot penetrate the intact hide or skin although they sometimes invade the sweat glands and cause local infections. Cuts or tears in the hide or skin which penetrate to the deeper layers of tissues are Sources of bacterial entry and infections. Even very young children know that when they cut their finger they should rush to their mother to have the cut treated with an antiseptic and wrapped in a cloth bandage to keep out the dirt and speed the healing of the cut. The thickness of the hide and hair in animals as well as the presence of skin secretions helps prevent the entry of bacteria and may repel certain disease-carrying insects. Zebu cattle are popular in tropical in semi-tropical areas because they are heat tolerant. They are also resistant to many tick or insect carried infectious organisms. They are superior to the British breeds of cattle in their ability to resist such diseases. Some animals such as the horse have the ability to shake their skin Whish serves to frighten insects and possibly avoid insect bites.

Secretion of the body. The alimentary canal, the nasal cavity, and the upper lung passages are lined with a thick, slimy secretion called mucus which traps and prevents bacteria from migrating because of its stickiness. The linings of these body passages also have cell layers which prevent bacterial entrance into the underlying tissue in a way similar to the skin on the surface of the body. It is known that swarms of bacteria are normally present in the small intestine and colon, but they usually do not penetrate to the underlying tissues in large numbers. One kind of scours in baby pigs has been reported to be due to an accumulation of large numbers of E. coil in tie intestines (Babakhani et al., 1993). A report from England suggests that some of these bacteria possess a sticky substance on their surface which attaches them to the intestinal wall where they multiply in large numbers and produce toxins causing scours in baby pigs. Injections of the sticky material on the bacteria into sons causes them to build antibodies against the sticky material so that the bacteria can't stick to the intestinal walls and therefore do not accumulate in large numbers

and produce toxins. The treated sow will transmit the antibodies she has produced to her pigs when they are born, through colostrum. preventing this type of scours in the pigs. This same report also stated that some pigs naturally do not allow the E. coli to stick to the intestinal wall and their ability to prevent this is heritable.

The flow of secretions in the body surfaces, in the interior canals of the body, and within the glands nay wash away bacteria and prevent their entrance into the body. For example, tears continually wash the eye free of bacteria laden dust particles. Possible secretions of this kind on the skin and Surface of the body may possess bactericidal or repellent properties (secretory IgA-immunoglobulins). Breed differences in cattle in their susceptibility to pink eye are well known, but what these genetic differences might be hasn't been fully determined through research. The gastric juice of the stomach has a high acid content which destroys Some bacteria in the digestive tract or severely retards their growth. Ciliated cells in the upper lung passages of mammals beat in such a direction that the mucus coating of the cells moves slowly upward and the bacteria the mucus are moved up out of the passages into the pharynx Where coughing often removes them. The nasal passages also possess ciliated cells which act in a similar manner Whether or not there are genetic variations in these mechanisms of defense is not known.

Phagocytois. Once disease organisms penetrate the body tissues the neutrophils (granulocytes) and macrophages (monocytes) of the peripheral blood s and destroy them by the process known as phngocytosis. In this process they engulf the bacteria or foreign particles and begin digesting them. Neutrophils and macrophages both process an abundance of lysosomes which are filled with proteolytic enzymes especially fitted for digesting these Objects. Phagocytotic cells also contain bactericidal agents which kill bacteria before they can multiply and destroy the phagocyte itself. Some genetic defects are known which have major detrimental effects upon the neutrophils so that they do not function normally in both man and animals. Among these are the Chediak-Higashi syndrome and cyclic neutropenia (Kritzler, 1964).

The Chediak—Higashi syndrome in humans is characterized by partial albinism and the presence of giant particles in cells which normally produce cytoplasmic particles, especially in the leukocytes, or white blood cells. This syndrome is usually associated with frequent recurrent severe infections in either gastro-intestinal tract, the skin, or the respiratory tract. Frequent fevers of unknown origin and ulceration of the oral mucosa also occur. The relationship between the large

granules and increased infection is unknown, but it has been postulated that many of the giant granules are lysosomes, and their large size may be due to the improper function and distribution of the lysosomal enzymes which function in the phagocytosis of bacteria in the peripheral blood stream. Perhaps these lysosomes do not rupture to release the enzymes necessary for the phagocytosis and destruction of bacteria. The syndrome is inherited as an autosomal recessive (Kritzler, 1964) and has also been reported in mink cattle, bison, whales, and mice (Witkop, 1975).

Cyclic neutropenia refers to the disappearance of neutrophils in the peripheral blood due to an arrest of their production in the bone marrow. The Condition has been described in both dogs and man. The neutrophils disappear from the peripheral blood stream at intervals of 8 to 12 days in days and in intervals of about 21 days in children. Cyclic neutropenia appears to be associated with a recessive grey coat color in collie dogs (Ford, 1969). It does not appear to be associated with a lack of pigment production in humans. both humans and dogs the periods of neutropenia are accompanied by local or Systemic infections and fever. In collies, death almost always occurs before maturity. The exact cause of death is unknown but it probably involves accompanying complications of the disappear of the neutrophils.

The antigen-antibody response. This is an important mechanism in the body to guard against disease brought into play when foreign substances such as bacteria eater the body. The immune reaction is quite complicated although white blood cells known as lymphocytes play an important role in this body defense mechanism against disease. I will mention only two major genetic defects that have been reported which interfere with the immune reaction. These are agammaglobulinemia in humans and an immunodeficiency in Arabian foals (Gitlin, 1957 and Mcguire, 1974). Agammaglobulinemia is sex-linked recessive trait affecting only boys. The disease is characterized by the failure to produce sufficient antibodies to resist infections and death usually occurs at an early age unless the condition is recognized and proper treatment is given (Gitlin, 1957). The immuno-deficitncy in Arabian foals has been identified with a syndrome characterized by lymphopenia, immunoglobulin deficiency, thymus hypoplasia, absence of spleen and lymph node lymphocytes, and an increased susceptibility to infection. Death usually occurs within a month or two after birth because of pneumonia or other infections (Mcguire, 1974). It appears to be inherited as an autosomal recessive. A deficiency of an immunoglobulin class (IgG2) has been reported in cattle and is accompanied by an increased susceptibility to infection (Nansen,

1972). A lethal syndrome in certain calves of the Black Pied Danish cattle breed has also been reported and is accompanied by an extreme hypoplasia of the thymus. lymph nodes, and the spleen (Brummerstedt, 1971). It appears to be inherited as an autosomal recessive.

Interferon. This is a protein substance produced by cells, particularly those of the spleen. Interferon interferes with virus reproduction probably by producing antiviral polypeptides which inhibit the normal translation of mRNA. It may also prevent the virus from gaining entrance into the cells which would also inhibit virus reproduction. Since interferon is a protein in nature and the function of a gene is to send the code for protein structure to the ribosomes in the cytoplasm, its production is under genetic control and there should be genetic variations among individuals in their resistance or susceptibility to Virus infections. To date no specific examples of genetic variations in interferon production have been reported although certainly individuals are known to vary in their resistance to viral diseases. Factors other than interferon production could be involved, however.

Iron. Recent studies have shown that the ability of the host to reduce the amount of iron in the blood plasma in response to bacterial invasion is important in the resistance or susceptibility to disease (Weinberg, 1943). The ability of the host to withhold iron from microbes is called nutritional immunity. Since similar quantities of iron are required for the growth in numbers of invading microbes. It is not known if there are genetic variations among animals in their ability to withhold iron from microbes, but perhaps this could be another area for study of genetic resistance ability to disease.

Genetic resistance to non-infectious disease

Several different forms of non-infectious diseases appear to have a genetic basis. It is not the disease which is inherited, as a general rule, but an inherent weakness in the tissues of the body which allows the disease to produce observable symptoms. Mechanical ailments such as hernias are known to have a genetic base in several species (Warrent, et al., 1931). A hernia (or rupture) refers to the protrusion of internal body organs or tissues through some opening surrounding a cavity. A common hernia in both males and females is one in which the intestines protrude through an opening around the umbilicus into the skin. This is called an umbilical hernia. Scrotal hernias found in males are due to the protrusion of the intestines through the inguinal canal into the scrotum. A hereditary form of umbilical hernia has been reported in Holstein-Freisian cattle

(Warrent, et al., 1931). It occurred mostly in males and was inherited as an autosomal recessive. Umbilical hernia has also been noted in Poland China swine and probably was inherited as an epistatic trait because it appeared in a high percentage of the offspring produced by crossing two different inbred lines. Neither line had a high incidence of this defect when bred pure. Scrotal hernia in swine has been reported to be due to two recessive genes (Warwick, 1943).

Diabetes mellitus is a disease in which there is an impairment of carbohydrate metabolism accompanied by an increased metabolism of proteins and fats. Diabetes in humans appears to be a growing health problem in the United States (Maugh, 1975). There are two major forms of the disease: juvenile onset diabetes and maturity onset diabetes. Either form can strike individuals at any age. The juvenile onset form begins abruptly and symptoms occur almost immediately. This is the most difficult form of diabetes to control and almost always requires the administration of insulin. Maturity onset diabetes begins slowly and usually produces milder symptoms. This form can usually be controlled by dietary regulation or by the administration of oral drugs which stimulate the release of insulin. Recent evidence suggests that viruses play a role in initiating juvenile onset diabetes, but a genetic predisposition to diabetes is also important in the onset of the disease. Genetic susceptibility may be the most important factor in the maturity onset form. If viruses are an important cause of diabetes, it may be possible to develop vaccines against as it. Forms of diabetes also occur in animals, but they have not been as closely studied as in humans. Some of the evidence that viruses are implicated in the onset of diabetes has been obtained in studies with laboratory animals. If viruses are implicated in the occurrence of juvenile onset diabetes, this sheds a new light on this particular disease. Previously it was thought that this disease was due only to an inborn error of metabolism.

In regions of Africa where malarial infections are prevalent in humans, mortality from such infections is consistently lower in individuals heterozygous for the sickle cell gene (genotype Ha-Hs). Mortality is high in individuals which are homozygous for the sickle cell gene (Hs-Hs) because of anemia and in individuals not carrying the gene (Ha-Ha) because of malaria. Thus, the heterozygous individuals in these areas survive in larger proportions than where malaria is not prevalent. The sickle cell gene, therefore, is involved in some way in resistance to malarial infections.

Substances in the diet related to disease

Malformations, or defects, may occur in animals because of the consumption of certain substances in their diet. It is not known if there is a genetic resistance to such substances, though the trouble involved does vary in different individuals. Farmers in the western U.S. observed for many years that about one percent of their ewes would deliver defective lambs with malformations of the face and head in which the eye was often located in the center of the face, lending the name cyclopia to lending the name cyclopia to this deformity. Research has proven that this defect was caused by pregnant ewes consuming amounts of the plant *veraturum californicum* on about the 14^{th} day of pregnancy (Mulvilhill, 1972). This plant grew in only in the high-altitude ranges grazed by affected flocks. An epidemic in Kentucky, in which over 900 pigs on nine farms were born with stiff and deformed joints, appeared to be caused by some unidentified substances in burley tobacco stalks eaten by all mothers of the defective pigs during pregnancy.

Hydrocephalus in a strain of laboratory rats at the University of Missouri, Biochemistry Department has been reported (Newberne, 1958). Apparently, this condition was aggravated by a deficiency of vitamin B12, or folic acid, in the diet of the mothers during pregnancy. The incidence of this defect seemed to appear more frequently in this particular strain of rats than some but no genetic basis was reported. The occurrence of defects due to toxins in the ration of farm animals, or a deficiency of certain substances, may be important from several standpoints. They may cause considerable economic losses to livestock producers; they may confuse the issue of whether or not they are of genetic or non-genetic origin; and they may also be used as a model to learn more about the causes of such defects in humans and animals.

Genetic resistance to certain drugs may be of economic importance. For example, the use of warfarin, an anticoagulant rodenticide has been effective in controlling the rat population. It now appears, however has be rats in Europe and the U.S. are genetically resistant to these substances and subsequently they are no longer effective in killing such animals (Jackson and kaurene, 1972). The growth of molds on foodstuffs, resulting in unpleasant flavors or causing other undesirable changes in products, has been observed for many years. More recently it has been found that some molds can manufacture chemical substances that ate poisonous or produce toxic symptoms when men or animals eat food containing these products. These chemical Substances are generally referred to ns mycotoxins. They may remain in food long after the molds that produce

them die, and some of them may even resist usual conditions encountered in cooking or food processing. Mycotoxins may also be present in food that does not shown the presence of molds. The records show that hundreds of humans and livestock have been poisoned by these substances.

Aflatoxins, one specific group of mycotoxins, have been extensively studied. Some aflatoxins are known to interact with the DNA of the chromosome and to interfere with the synthesis of RNA, affecting the transcription of genetic information in animal cells, animal tissues, and icroorga11h Aflatoxins have been shown to be very potent cancer producing substances in several species or animals. Many diseases such as ergot poisoning in animals appear to be due to mycotoxins (Wogan, 1972). Although there is no definite proof as yet, mycotoxins may be involved in fescue foot in cattle. Since varying degrees of severity of this disease are observed in animals, resistance or susceptibility to these substances may be heritable. However, this has not been determined experimentally.

Indicators of genetic resistance of susceptibility to disease

The development of strains of animals genetically resistant to disease exposing them to a disease and then locating those which are genetically resistant is an impractical procedure because of death losses and poor performance in the susceptible individuals. It may be possible. However, to find traits in animals which may indicate that they will be resistant to a disease. Thus, selection for resistance to disease could be based on these measurable indicators without exposure to the disease. One of the best examples of an indicator of resistance to a disease pigment around the eye in Herefords (Bonsma, 1949). Hereford cattle produced in climates with intense sunlight are subject to eye infections and even true cancer of the eyes or the eye lids. Some of these cancers become so severe that they cause the death of the affected individual. Selection of breeding stock with pigment on the eye lids or around the eye drastically reduces the eye trouble, perhaps because it reduces the exposure of the tissue of the eye or lids to the rays of the sun. Something more than this may be involved, however, because white Charolais or Charbray do not have the high incidence of eye problems as Herefords in the same regions.

In all probability other indicators of genetic resistance to disease will be found through future research. Areas of research which could be fruitful from this standpoint are the analysis of body, secretions, antibodies of the blood, and the phagocytic ability of leucocytes, to determine if differences exist in these traits between resistant and susceptible animals. It was mentioned

previously that partial albinism in humans and some animals, as shown in the Chediak-Higashi syndrome as well as others, are associated with an increased incidence of bacterial infections. Thus, albinism is sometimes an indicator of increased susceptibility to disease. Another point of interest involving coat color is that albino and black mice differ in their response to alcohol. After receiving a standard dose or alcohol, pure black mice slept an average of 65 minutes, whereas albinos slept 175 minutes. Black heterozygotes slept an average of 81 minutes (Witkop, 1975). Individuals with the Hermansky-Pudlak syndrome are extremely sensitive to aspirin and aspirin-like drugs which makes the mild bleeding defect they possess much more severe than usual because of a platelet defect (Witkop, 1975).

Resistance to internal and external parasites

Evidence suggests that some animals are genetically resistant to internal as well as external parasites. Romney sheep England have been found to be more resistant to *Trichostrongyle* worms than breeds developed in areas where exposure to these parasites was less severe. Similar results were obtained in California when the resistant Romney sheep were compared with sheep of other breeds (Gregory, 1937). Texas selection experiments also indicated that there was a genetic resistance to stomach worms (*Haemonchus Contortus*) in both sheep and goats (Warwick, 1943). Close observation shows that there is great individual variation in the number of flies and lice infesting cattle. It is not known if genetic resistance or susceptibility is involved but it may be. Certainly, this suggests that research in this area is needed. Zebu cattle are much more resistant to ticks than British breeds of cattle (Hutt, 1958). Therefore, they are less susceptible to tick-borne diseases, because of the reduction of the possibility of tick-borne pathogens entering their body.

Conclusion

Much evidence shows that animals are genetically resistant or susceptible to many infectious and non-infectious diseases as well as to Internal and extern parasites. Limited efforts have been made to develop strains or breeds that possess genetic resistance, however. The main reason more work of this nature has not been pursued is that certain problems have been encountered in developing such lines. For example, genetic resistance in many cases is determined by many pairs of genes (Quantitative traits), genetic resistance is usually specific and animals have to be exposed to a disease to find those that are resistant.

Developing disease resistant strains would be more practical if the resistant individuals could be identified without exposing them to the disease, this can be done by using the modern selection techniques like Genomic selection (Meuwissen et al., 2001). Genetic correlation and some other indicators of genetic resistance (such as red pigment around the eyes coinciding with a reduction in the incidence of cancer eye in Hereford cattle) would make selection for resistance more rapid and successful. Much more work needs to be done in this area. The normal Occurrence of resistance to disease and parasites is probably more important than is realized because the ability to resist disease is over looked until the disease strikes and then we notice only the affected animals.

REFERENCES

Babakhani F. K. , G. A. Bradley, L. A. Joens. 1993. Infection and Immunity. 61 (8) 3466-3475

Bonsma j. C. 1949. Breeding cattle for increased adaptability to, tropical and subtropical environments. Jags. 39: 204.

Brummerstedt e., t. Flagstad a. Basse, and e. Andresen. 1971. The effut of zinc on calves with hereditary Thymus Hypoplasia (Lethal Trait A 46). APMS Section A, 79: 686.

Bruns, H.A., 2017. Southern corn leaf blight: a story worth retelling. Agronomy Journal, 109(4), pp.1218-1224.

Campbell j. R., and F. Lasley. 1975. The science of animals that serve mankind 2nd edition.

Ford l. 1969. Hereditary aspects of human and canine cyclic neutropenia. JH. 60: 293.

Francis, J., 1966. Resistance of Zebu and other Cattle to Tick Infestation and Babesiosis with Special Reference to Australia: An Historical Review. Br Vet J. 122(7):301-307

Gitlin d., and C. E. Janeway. 1957. Agammaglobulinemia. Sa.197: 93.

Gregory, p. W. 1937. Resistance to worms in rommey sheep. Pasap, 316.

Hannik, C. A. and H, Cohen, 1978. New approach to the production of concentrated and purified inactivated polio and rabies tissue culture vaccines. Developments in biological standardization, 41, pp.159-168.

Hutt, f. E. 1958. Genetic resistance to disease in domestic animals. Cornell university press.

Ingersoll j. M. 1964. The Australian rabbit. As, 52: 265.

Jackson, W. B., and kaurene. 1972. Resistance of wild norway rats in north carolina to warfarin rodenticide. S, 176: 1343.

Kritzler, R. A. 1964. Chediak-higashi syndrome cytologic and serum lipid observations in a case and family. Ajm, 36: 583.

Mago, R., Bariana, H.S., Dundas, I.S., Spielmeyer, W., Lawrence, G.J., Pryor, A.J. and Ellis, J.G., 2005. Development of PCR markers for the selection of wheat stem rust resistance genes Sr24 and Sr26 in diverse wheat germplasm. Theoretical and Applied Genetics, 111(3), pp.496-504.

Maugh, T. H. 1975. Diabetes: epidemiology suggests a viral connection. S. 188: 347.

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

Mcguire, T. C., M. J. Poppie, and k. L. Banks. 1974. Immunodeficiency: a fatal genetic disease in arabian foals. Javma 164: 70.

Mulvilhill J. F. 1972.congenital and genetic diseases in domestic animals. S, 176, 132.

Nansen, P. 1972. Selective immunoglobulin deficiency in cattle and susceptibility to infection. Apms, 80:49.

Newberne, P. M., and b. L. O'dell. 1958. Histopathology of hydrocephalus resulting from a deficiency of vitamin b12. Psebm, 97: 62.

Rutkoski, J.E., Heffner, E.L. and Sorrells, M.E., 2011. Genomic selection for durable stem rust resistance in wheat. Euphytica, 179(1), pp.161-173.

Warren, T. R., and F. W. Atkeson. 1931. Inheritance of hernia in a family of holstein-friesian cattle. JH, 22: 345.

Warwick, B. L., R. O. Bery, R. D. Turk and C. O. Morgan. 1949. Selection of sheep and goats for resistance to stomach worms, haemonchus contortus. Jas, 8: 609.

Warwick, E. J., A. B. Chapman, and burr ross. 1943. Some anomalies in pigs. JH, 34: 349.

Weinberg, E. P. 1974. Iron and susceptibility to infectious disease. S, 184:952.

Witkop, C. J., JR. 1975. Albinism. Nh, 84:48.

Wogan, G. H. 1972. Moulds and mycotoxins—a direct threat to public health. FTAF, 28: 6a.

CHAPTER 3

Phenotypic, genetic and SNP-marker associations between calf diseases and subsequent performances and disease occurrences of first-lactation German Holstein cows

M. Mahmoud, T. Yin, K. Brügemann and S. König

Institute of Animal Breeding and Genetics
Justus-Liebig-University Gießen, 35390 Gießen, Germany

J. Dairy Sci. 100:2017–2031
https://doi.org/10.3168/jds.2016-11767

ABSTRACT

A total of 31,396 females born in the years 2010 to 2013 from 43 large-scale Holstein-Friesian herds were phenotyped for calf and cow disease traits using a veterinarian diagnosis key. Calf diseases were the general disease status (cGDS), calf diarrhea (cDIA) and calf respiratory disease (cRD) recorded from birth to the age of two months. Incidences were 0.48 for cGDS, 0.28 for cRD, and 0.21 for cDIA. Cow disease trait recording focussed on the early period directly after calving in first parity including the interval from 10 d before calving up to lactation day 200. Also for cows, at least one entry for the respective disease implied a score = 1 (sick), and otherwise a score = 0 (healthy). Corresponding cow diseases were the first-lactation general disease status (flGDS), first-lactation diarrhea (flDIA), and first-lactation respiratory disease (flRD). Additional cow disease categories included mastitis (flMAST), claw disorders (flCLAW), female fertility disorders (flFF) and metabolic disorders (flMET). A further cow trait category considered first-lactation test-day production traits from official test-days 1 and 2 after calving. The genotype dataset included 41,256 SNP from 9,388 females with phenotypes. Linear and generalized linear mixed models with a logit link-function were applied to Gaussian and categorical cow traits, respectively, considering the calf disease as a fixed effect. Most of the calf diseases were not significantly associated with the occurrence of any cow disease. By trend, increasing risks for the occurrence of cow diseases were observed for healthy calves, indicating mechanisms of disease resistance with aging. Also by trend, occurrence of calf diseases was associated with decreasing milk, protein and fat yield. Univariate linear and threshold animal models were used to estimate heritabilities and breeding values (EBV) for all calf and cow traits. Heritabilities for cGDS and cRD were 0.06, and 0.07 for cDIA. Genetic correlations among all traits were estimated using linear-linear animal models in a series of bivariate runs. The genetic correlation between cDIA and cRD was 0.29. Apart from the genetic correlation between flRD with cGDS (-0.38), EBV-correlations and genetic correlations between calf diseases with all cow traits were close to zero. Genome wide association studies (GWAS) were applied to estimate SNP-effects for cRD and cDIA, and for the corresponding traits observed in cows (flRD and flDIA). Different significant SNP-markers contributed to cDIA and flDIA, or to cRD and flRD. The average correlation coefficient between cRD and flRD considering SNP-effects from all chromosomes was 0.01, and

-0.04 between cDIA and flDIA. In conclusion, calf diseases were inappropriate early predictors for genetic improvements of early lactation cow traits in parity one.
Key words: calf and cow diseases, genetic parameters, genetic and genomic associations

INTRODUCTION

As reviewed by Egger-Danner et al. (2015), a variety of recent quantitative-genetic studies focussed on the estimation of genetic (co)variance components for cow health traits, with a focus on udder health, claw disorders, metabolism, and female fertility. Especially in the phase of a naturally negative energy balance during the first third of lactation, disease incidences were quite large (Gernand et al., 2012). For most of the disease traits recorded during the early lactation period, low heritabilities and antagonistic genetic relationships with productivity, were identified. Nevertheless, in deterministic predictions (König et al., 2005), sustainable selection response for cow health after calving was only generated when aiming on direct selection strategies instead of using health indicators, e.g., functional conformation traits.

In the era of genomic selection, large cow training sets combining health phenotypes with high-throughput genomic SNP marker data, allow alternative health trait selection strategies (Buch et al., 2012). A second major advantage of genomic selection is due to the substantial reduction of generation intervals (Schaeffer, 2006). In such a generation interval perspective, traits measured at an early stage of life, e.g. calf traits, might be valuable early predictors for subsequent cow health or cow productivity. Pronounced genetic correlations between calf and cow traits furthermore allow efficient selection strategies among young females, with associated effects on farm economy, i.e., to avoid expensive raising up of all female offspring.

The most important disease problems observed in calves in the past decades in North-America (Curtis et al., 1988) as well as in central Europe (Perez et al., 1990) included respiratory infections and different kinds of diarrhea. Despite optimizations of the farm management including calf feeding strategies and improvements of calf husbandry systems, both disease categories still have the highest incidence rates. The importance of both diseases was shown based on a random herd sampling in the Norwegian Red population (Gulliksen et al., 2009a,b), on a farmer survey in Austria (Klein-Jöbstl et al., 2015), or when referring to official calf mortality statistics in the US and Asia (Cho and Yoon, 2014). Especially in the Holstein breed

(e.g., Becher et al., 2004), moderate to quite large incidences were identified for respiratory infections and diarrhea.

Presence of diarrhea and / or respiratory diseases in calves affected the performance and productivity of the same animal later in life. In most cases, occurrence of a calf disease increased the probability for the occurrence of heifer health disorders (e.g. Sivula et al., 1996). Warnick et al. (1994) also indicated that calf respiratory disease was phenotypically associated with an increased occurrence of dystocia at the first calving date. Rossini (2004) studied the phenotypic associations between calf respiratory and digestive diseases with age at first calving and first-lactation production traits. Non-significant effects were identified for 305-d milk and fat yield, but protein yield moderately decreased by 0.05 kg/d. Also age at first calving was higher for calves with multiple occurrences of respiratory disease. In contrast, Mousa et al. (2015) identified negative correlations between calf scours and age at first calving. Bünger et al. (1979) reported reduced dry-feed intake in lactating dairy cows due to infections during the calfhood stage. Reduced feed intake causes a negative energy balance with associated health problems (e.g. Collard et al., 2000). In a similar context, Beam et al. (2015) focussed on the effects of calf growth rates and levels of starter feed intake, and they identified positive association between those measurements during calfhood with performance traits and body weight in mature cows.

Phenotypic relationships revealed that calf diseases mostly appeared in clusters. This means that, e.g., a calf being susceptible for digestive infections, also showed symptoms for a respiratory disease. Rossini (2004) reported that occurrence of calfhood digestive disease was associated with a 2-fold increase of the probability for occurrence of calf respiratory disease. Also Lundborg et al. (2003) confirmed the positive relationships between calf diarrhea and calf respiratory disease. In a study conducted in New York State (Henderson et al., 2011), the residual correlation between calf respiratory and bloat disease was positive, that between respiratory and umbilical disease slightly negative.

So far, only a few studies estimated quantitative genetic parameters for diseases in dairy or dual-purpose cattle calves. In the study by Henderson et al. (2011), heritablities for bloat, respiratory and umbilical diseases were 0.040, 0.095, and 0.139, respectively). Fuerst-Waltl et al. (2010) estimated genetic parameters for diseases in Austrian dual-purpose Fleckvieh heifer calves: the heriabilty was 0.027 for the liability to diarrhea, and 0.039 for the liability of respiratory diseases. Based on the Norwegian health data recording system and using records from 250,212

Norwegian Red calves, the heritability for respiratory disease on the underlying liability scale was 0.05 (Heringstad et al., 2008). Given the low heritability combined with a low incidence rate (only 0.7% of the calves had a veterinary treatment for respiratory disease within the first 180 d of age), Heringstad et al. (2008) saw limited possibilities for genetic improvements. Nevertheless, the same authors pointed to the accurate genetic predictions of sires for respiratory disease resistance. With regard to genetic associations among calf traits, genetic correlations partly substantially differed from phenotypic correlations. For example, in the study by Henderson et al. (2011), the phenotypic correlation between calf bloat disease with respiratory disease was close to zero, but the genetic correlation between the same traits was quite high (0.62).

For functional traits with low heritabilities, and reflecting a possible "missing heritability" phenomenon, Lee et al. (2011) recommended genome-wide association studies to detect SNP-markers contributing to the phenotypic trait expression. Pimentel et al. (2011) suggested utilization of specific SNP-effects to improve selection response for antagonistically related production and fertility traits simultaneously. Up to now, genome-wide association studies (**GWAS**) focused on cow health, i.e., addressing classical health disorders (e.g. van der Spek et al., 2015), infectious diseases (e.g. Bermingham et al., 2014) or immune responses (Thompson-Crispi et al., 2014). To our knowledge and with regard to GWAS in calves, the size of SNP-marker effects was only estimated for calf birth weight (Cole et al., 2014).

The objective of the present study was to infer relationships between calf diseases and cow production and health from the early lactation period in first parity on i) phenotypic, ii) quantitative-genetic and iii) genomic scales. The objective implies for i) the application of linear and generalized linear mixed model analysis for cow traits considering the calf disease as a fixed effect, ii) the estimation of genetic correlations between calf and cow traits using the pedigree relationship matrix and animal models, and iii) the determination of SNP-marker effects for calf and cow diseases based on GWAS.

MATERIALS AND METHODS

Data

***Phenotypes*.** Health data recording was implemented in the framework of the dairy cattle contract herd system of northeast Germany, including the federal states of Mecklenburg-Westpommerania and Berlin-Brandenburg. Data from a subset of contracting testing herds in this region were previously used in the udder health study by Martin et al. (2013). Dairy cattle farmers used electronical recording systems based on the diagnosis key from Feucker and Staufenbiel (2003). This diagnosis key was also considered when developing the ICAR guidelines for recording, evaluation and genetic improvement of health traits (Stock et al., 2013). An explanation for this hierarchical disease entry system is given by Gernand et al. (2012) for the major cow disease categories claw disorders, mastitis, female fertility, and metabolism. I considered in total 31,396 female calves from 43 contract herds born in the years 2010, 2011, 2012, and 2013. All calves had a calving date from 2012 to 2015, and at least one official test-day record, because only for cows with official test-day records (= the data basis for official genetic evaluations) the "disease history" and remaining data (calving dates, pedigree, etc.) was available. The female calves born in the 43 herds are considered as possible replacements, implying the raising up of females in the same farms until selection decisions are made in the first third of the first lactation. Ages at first calving of the cows ranged from 20 to 39 months.

The three overall types of traits included a) calf disease traits, b) first-lactation disease traits, and c) first-lactation production traits. Disease traits were defined as a binary trait without considering repeated measurements. For the recording of calf diseases, I defined a time window from birth up to the age of two months, and for cow diseases, the interval from 10 d before calving up to lactation day 200 (210 d period). At least one entry for the respective disease implied a score = 1 (sick), and otherwise I assigned a score = 0 (healthy). The DIM at test-day 1 ranged from 5 to 50 days, with an average value of 20.78 days. Hence, the minimal time period for a cow to fall ill included 15 days (10 days before up to 5 days after calving). The calf disease traits included the categories a1) general calf disease status (**cGDS**), a2) calf diarrhea (**cDIA**), and a3) calf respiratory disease (**cRD**). For cGDS, the calf received the score = 0 when absolutely no disease entry was assigned in the disease database. All forms of calf scours, and in additional the recorded phenotype "diarrhea" resulting from digestive imbalances, were considered when defining cDIA. In detail,

and according to the diagnosis key, the following diseases contributed to cDIA: diseases of the intestine, rotavirus infection, *Escherichia coli* infection, salmonellosis, coccidiosis, and cryptosporidiosis. With regard to cRD, the following respiratory diseases were considered: diseases of nasal cavity and paranasal sinuses, diseases of the bronchial tubes and lungs, and the bovine respiratory syncytial virus (**BRSV**).

Accordingly, the cow disease traits within the 210 d time window included the categories b1) first-lactation general disease status (**flGDS**), b2) first-lactation diarrhea (**flDIA**), and b3) first-lactation respiratory disease (**flRD**). In addition to the traits included in cDIA, flDIA furthermore considered diarrhea symptoms due to diseases of the reticulum and the abomasum. Following the disease categorization by Gernand et al. (2012), further cow disease categories were b4) first-lactation mastitis (**flMAST**) including clinical and subclinical mastitis, b5) first-lactation claw disorders (**flCLAW**) including purulent and non-purulent claw disorders, b6) first-lactation female fertility disorders (**flFF**) including the sub-categories "uterus" (puerperal disorders, retained placenta, endometritis) and "ovary" (silent estrus, ovarian cysts, corpus luteum persistent), and b7) first-lactation metabolic disorders (f**lMET**) including ketosis, fatty liver syndrome, adiposity, and excessive loss of weight. The incidence rates for diseases of female calves (= diseased female calves in relation to 31,396 female calves) and of cows (= diseased cows in relation to 31,396 cows) for the defined time periods are given in Table 1.

With regard to the first-lactation production trait category c), I focussed on the measurements from the official test-days 1 and 2 after calving including c1) milk yield (**MY_1, MY_2**), c2) fat percentage (**FP_1**, **FP_2**), c3) protein percentage (**PP_1**, **PP_2**), c4) fat yield (**FY_1**, **FY_2**), c5) protein yield (**PY_1**, **PY_2**), and c6) the fat to protein ratio (**FPR_1**, **FPR_2**) as an indicator for energy balance. For the calf disease – cow disease trait, and for the calf disease – cow test-day 1 trait association studies, all datasets included 31,396 animals. Because of dairy cow disposal after test-day 1, the dataset for test-day 2 observations included 30,360 cows. The DIM at test-day 2 ranged from 27 to 105 days, with an average value of 52.92 days.

Table 1. Incidences for the following disease traits recorded in calves and in first-lactation German Holstein cows: cGDS = general calf disease status, cRD = calf respiratory disease, cDIA = calf diarrhea, flGDS = first-lactation general disease status, flRD = first lactation respiratory disease, flDIA = first-lactation diarrhea, flFF = first-lactation female fertility disorders, flCLAW = first-lactation claw disorders, flMAST = first-lactation mastitis, flMET = first-lactation metabolic disorders.

Trait category	Healthy	Diseased	Incidence (in %)
Calf health traits			
cGDS	16,429	14,967	47.67
cDIA	24,670	6,726	21.42
cRD	22,491	8,905	28.36
First-lactation health traits			
flGDS	9,499	21,897	69.74
flDIA	30,035	1,361	4.33
flRD	31,123	273	0.87
flMAST	24,880	6,516	20.75
flCLAW	24,520	6,876	21.90
flFF	19,106	12,290	39.15
flMET	30,890	506	1.61

Genotypes. From the 31,396 animals with phenotypes, 9,401 individuals were genotyped. Genotyping was performed with the *Illumina Bovine 50K SNP BeadChip* V2 (applied to 3,192 animals), and with the *Illumina Bovine Eurogenomics 10K low-density chip* (applied to 6,209 animals). The low-density genotypes were imputed by the project partner Vereinigte Informationssysteme Tierhalung (Verden, Germany) to the 50K panel applying the algorithm by Segelke et al. (2012), resulting in 45,613 SNP for all animals. Post-imputation SNP quality

checks identified 13 animals with almost identical SNP genotypes (larger than 95% of congruency). These animals were eliminated from the ongoing analyses. In total, 362 SNP with minor allele frequency lower than 0.01, and 3,995 SNP showing a significant ($P < 10^{-5}$) deviation from Hardy-Weinberg equilibrium, were discarded. All SNP had a genotype call rate greater than 95%. The final dataset after imputation and SNP-data editing included 41,256 SNP from 9,388 animals. When considering the genotyped animals, the disease incidence was 29.0% for cRD and 25.5% for cDIA, and 1.13% and 3.77% for flRD and flDIA, respectively.

Statistical models

Phenotypic impact of calf diseases on first-lactation cow traits. Generalized linear mixed models (**GLMM**) with a logit link function were applied to test the effect of binary calf diseases on first-lactation binary cow diseases, and to estimate least squares means for the probability of disease occurrence. For this purpose, the Glimmix procedure in SAS (Version 9.4, SAS Institute Inc., Cary, NC, USA) was used. The statistical model [1] was defined as follows:

$$\text{logit}(\pi_{rstuv}) = \log[\pi_{rstuv} / (1-\pi_{rstuv})] = \varphi + \text{Herd}_r + \text{CYear}_s + \text{CSeason}_t + \text{AFC}_u + \text{CD}_v \quad \textbf{[1]}$$

where

π_{rstuv}	=	probability of occurrence for the first-lactation cow disease
φ	=	overall mean effect
Herd_r	=	fixed herd effect (43 herds)
CYear_s	=	fixed effect of calving year (four years from 2012 – 2015)
CSeason_t	=	fixed effect of calving season (three months per season)
AFC_u	=	fixed effect of age at first calving (three classes)
CD_v	=	fixed effect of the calf disease (healthy or sick)

Significance of fixed effects was determined based on Wald type III tests (König et al. 2005).

Test-day production traits followed a Gaussian distribution. Hence, linear mixed models (lm-function in R, R Core Team, 2013) were applied to investigate the impact of calf diseases on first-lactation production traits, and to calculate least squares means. The statistical model [2] was:

$$y_{rstuv} = \varphi + \text{Herd}_r + \text{CYear}_s + \text{CSeason}_t + \text{AFC}_u + \text{CD}_v + e_{rstuv} \quad \textbf{[2]}$$

where y_{rstuv} represents the observations for the test-day traits, and e_{rstuv} is the random error term. Remaining effects are the same as specified for model [1].

Genetic associations among calf diseases and their relationships with first-lactation cow traits. According to the three overall trait categories, a) calf diseases, b) first-lactation diseases, and c) first-lactation production traits, genetic analyses considered three different univariate statistical models. For binary calf diseases, the GLMM (model [3]) was:

$$\text{logit}\,(\pi_{rstu}) = \log[\pi_{rstu} / (1-\pi_{rstu})] = \varphi + \text{Herd}_r + \text{BYear}_s + \text{BMonth}_t + \text{AG}_u \quad \textbf{[3]}$$

where

π_{rstuv}	=	probability of occurrence for the calf disease
φ	=	overall mean effect
Herd_r	=	fixed herd effect (43 herds)
BYear_s	=	fixed effect of birth year (four years from 2010 – 2013)
BMonth_t	=	fixed effect of birth month (12 months)
AG_u	=	random additive-genetic effect

For binary first-lactation cow diseases, the GLMM (model [4]) was:

$$\text{logit}\,(\pi_{rstuv}) = \log[\pi_{rstuv} / (1-\pi_{rstuv})] = \varphi + \text{Herd}_r + \text{CYear}_s + \text{CSeason}_t + \text{AFC}_u + \text{AG}_v \quad \textbf{[4]}$$

with effects as specified in models [1] and [3].

For Gaussian distributed first-lactation test-day traits, the linear model [5] was:

$$y_{rstuv} = \varphi + \text{Herd}_r + \text{CYear}_s + \text{CSeason}_t + \text{AFC}_u + \text{AG}_v + e_{rstuv} \quad \textbf{[5]}$$

with effects as in models [2] and [3].

Heritabilities and genetic variances were obtained from univariate animal models. Regarding the logit models [3] and [4], calculation of heritabilities followed the procedure as described by Southey et al. (2003), i.e.:

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + (\frac{\pi^2}{3})}$$

The genetic (co)variance structure for models [3], [4] and [5] was:

$$var\begin{bmatrix}\boldsymbol{u}\\ \boldsymbol{e}\end{bmatrix} = \begin{pmatrix}\sigma_a^2 A_u & 0\\ 0 & \sigma_e^2 I_n\end{pmatrix}$$

where **u** = vector additive genetic effects; **e** = vector of random residual effects; σ_a^2 and σ_e^2 are the variances for additive genetic and residual effects, respectively; $\mathbf{A_u}$ is an additive genetic

(co)variance matrix for cows and sires (animal model); and $\mathbf{I_n}$ is an identity matrix for n observations.

For the estimation of variance components, heritabilities and genetic values, the three univariate models [3], [4] and [5] were applied, using the AI-REML algorithm as implemented in the DMU software package (Madsen and Jensen, 2013). Genetic associations among calf diseases, and between calf diseases and first-lactation cow traits, were inferred using correlations between estimated breeding values (**EBV**) considering the most influential 242 sires with more than 30 daughters. For the 1,115 sires, the average no. of daughters with phenotypes per sire was 28.16, and 284 females had an unknown sire. Additionally, I calculated genetic correlations among all calf traits, and for all calf-cow trait combinations, using a series of bivariate linear-linear models. Such a strategy implies utilization of models []3] and [4] as linear models instead of GLMM definitions with a logit link function. According to Vinson and Kliewer (1976), genetic correlations from linear-linear and from linear-threshold models are expected to be the same.

Genome wide association study. Estimation of SNP-effects for the two calf traits (cRD and cDIA) and for the corresponding traits observed in first lactating cows (flRD and flDIA) was carried out using a linear mixed model GWAS approach, and applying the software package GCTA (Yang et al., 2011). In matrix notation, the following single SNP – threshold model was defined [6]:

$$\mathbf{l = Xb + Wa + Zu + e \ [6]}$$

where **l** referred to the vector of unobserved liabilities; **b** was the vector of the fixed effects (herd, birth year and birth month for cRD and cDIA; herd, calving year, calving season and age at first calving for flRD and flDIA), **X** was an incidence matrix for the fixed effects; **a** was the vector for the additive allele substitution effects of the candidate SNP to be tested for the association; **W** was the design matrix for SNP genotypes coded as 0, 1 or 2; **u** was the vector for additive polygenic effects (random effect captured by the genomic relationship matrix) with the corresponding incidence matrix **Z**; and **e** is the vector of random residual effects. The genomic relationship matrix considered all SNP except the chromosome on which the respective candidate SNP for testing is located (software package LOCO, Yang et al., 2014). As suggested by Yang et al. (2014), exclusion of the respective chromosomes was done in order to avoid double-counting. *P*-values $\leq 5 \times 10^{-5}$ were used to identify significant associations between single SNP with calf and cow

diseases. Such a threshold was used in previous GWAS in dairy cows for production and functional traits based on 50K SNP-chip genotype data (e.g., Minozzi et al., Dadousis et al., 2016).

RESULTS AND DISCUSSION

Calf-cow trait associations on the phenotypic scale

Incidences of the calf diseases by birth month (= no. of diseased calves in relation to all calves born in the same month, and considering the following time period of two month) are shown in Figure 1. Incidences were larger for calves born in the late autumn, early spring and winter seasons compared to the summer season. Calf disease incidence peaks were observed for calves born in January and February 2010, 2011 and 2013, and in April 2012. An insight into historical meteorological data (Brügemann et al., 2013) indicated extremely high ambient temperature x humidity indices larger than 60 for these winter months. High air moisture in winter calf indoor housing systems was identified as a major determinant for the occurrence of cRD (e.g. Gulliksen et al., 2009b). A further increase of pathogen virulence in indoor production systems in autumn and winter seasons was due to the combination of high humidity with high concentrations of dust (e.g., Cobb et al., 2014). Interestingly, during the "peak incidence periods" for cRD, a simultaneous increase of cDIA was observed. Such associations between these most important calf diseases were also carried out in physiological experiments (Howard et al., 1989).

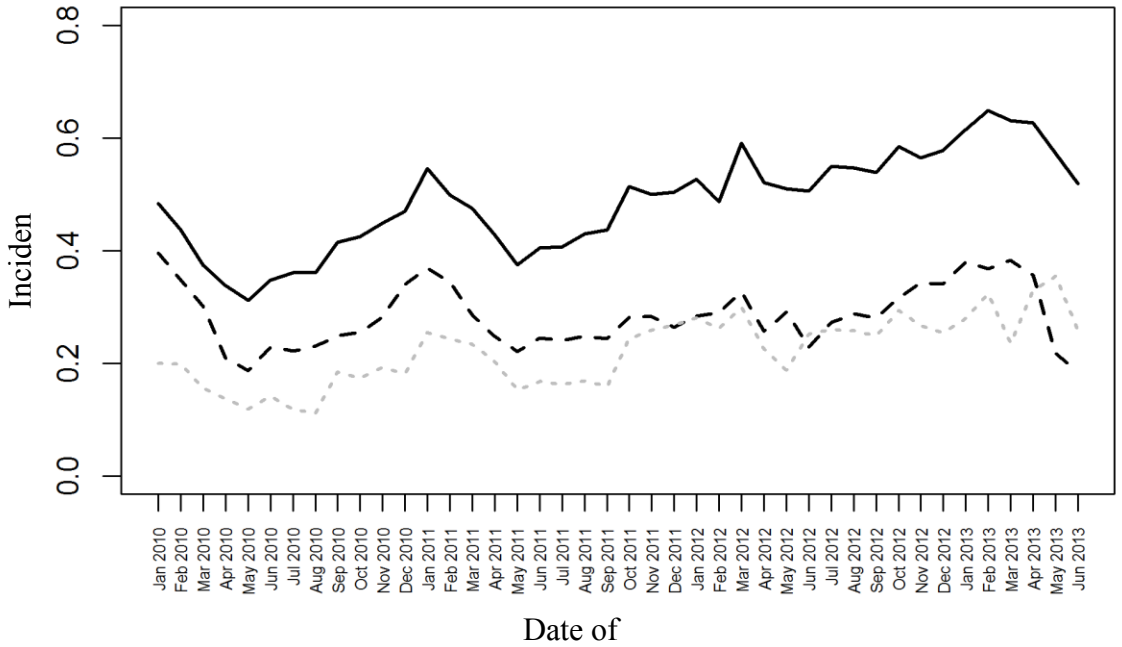


Figure 1. Calf disease incidences (solid black line = general calf disease status, dashed black line = calf respiratory disease, dotted grey line = calf diarrhea) by date of birth.

***Impact of calf diseases on cow diseases*.** Most of the calf diseases were not significantly associated with the occurrence of any cow disease ($P > 0.05$) in the 210 d period around calving (Table 2, results from model [1]). For some calf-cow disease trait associations, I identified an increasing risk for the occurrence of cow diseases for healthy calves. This effect even was highly significant ($P \leq 0.001$) when studying the impact of cGDS and cDIA on flGDS, and was significant ($P \leq 0.01$) for the impact of both calf traits on flMAST. A comprehensive literature review by Van der Fels-Klerx et al. (2002) addressed the associations between respiratory diseases early in life with mortality and reproduction up to first calving. In this regard, most researchers identified detrimental cRD-effects, e.g., an increasing death risk by factor 3.4 between 90 and 900 days of age for heifers with pneumonia treatments during the first three month of their life (Waltner-Toews et al., 1986). Also Svensson et al. (2006) reported increased morbidity risks during the rearing period for calves with diarrhea or respiratory problems. Few studies focussed on the effects of calf diseases on the occurrence of diseases in milking cows. Hultgren and Svensson (2009) identified a higher risk for clinical mastitis for females with severe diarrhea at the age between three and seven months compared to healthy calves. However, they could not explain reduced clinical mastitis incidence rates for mild cases of calf diarrhea. Assumptions address acquired immunity (Hara et al., 2012), or the development of disease resistances with age. Colditz et al. (1996) reported decreasing susceptibility in adult animals, and they also referred to an improved immunity status. Adams and Templeton (1998) stretched this concept of disease resistance, involving immune and non-immune mechanisms. Those explanations might support the non-significant associations between same calf and cow disease categories, i.e. the non-significant effects of cDIA on flDIA, and of cRD on flRD (Table 2). König et al. (2005) identified decreasing incidences of the claw disorder *dermatitis digitalis* with increasing parities, and in addition to the development of disease resistance, they assumed effects of selection. In contrast, Santin et al. (2008) and Faber et al. (2002) found a higher prevalence for *cryptosporidium* and *Eimeria* infections in in young females compared to adult cows.

Table 2. Differences of least squares means (Dif. LSM) for the probability of the occurrence of a first-lactation cow disease (flGDS = first-lactation general disease status, flRD = first lactation respiratory disease, flDIA = first-lactation diarrhea, flFF = first-lactation female fertility disorders, flCLAW = first-lactation claw disorders, flMAST = first-lactation mastitis, flMET = first-lactation metabolic disorders) with regard to the comparisons of solutions for healthy calves minus solutions for diseased calves (cGDS = general calf disease status, cDIA = calf diarrhea, cRD = calf respiratory disease).

	Calf disease								
	cGDS			cDIA			cRD		
Cow disease	Dif. LSM	P-value	Odds ratio	Dif. LSM	P-value	Odds ratio	Dif. LSM	P-value	Odds ratio
flGDS	0.045	***	1.21	0.044	***	1.20	0.014	n.s.	1.06
flDIA	0.001	n.s.	1.13	0.001	n.s.	1.07	0.001	*	1.25
flRD	0.000	n.s.	1.39	0.000	n.s.	1.33	0.000	n.s.	1.12
flMAST	0.007	**	1.11	0.012	***	1.18	0.001	n.s.	1.01
flCLAW	0.000	n.s.	1.01	0.001	n.s.	1.03	0.002	n.s.	1.02
flFF	0.001	n.s.	1.06	0.001	*	1.10	0.001	n.s.	1.02
flMET	0.000	n.s.	1.12	0.000	n.s.	1.01	0.000	n.s.	1.05

Significance level: *** = $P \leq 0.001$; ** = $P \leq 0.01$; * = $P \leq 0.05$; n.s. = not significant

***Impact of calf diseases on cow production traits*.** Occurrence of calf diseases was associated with decreased milk yield, protein yield and fat yield at the first two official test-days after calving in parity one. Corresponding least squares means from model [2] are listed in Table 3.3 for cGDS, in Table 4 for cRD, and in Table 5 for cDIA. However, most of the least square means differences for the comparison "healthy versus diseased calf" were not significantly different ($P > 0.05$). Most obvious effects were observed for milk yield, e.g. a highly significant advantage ($P < 0.001$) of healthy calves (calf disease cRD) on MY_2 (Table 4). Due to the naturally antagonistic relationships between milk volume and the content traits fat percentage and protein percentage, only marginal differences were identified for both calve groups with regard to FY_1, FY_2, PY_1, and PY_2. Past calf disease-cow productivity association studies (e.g., Hatch et al., 1974; Britney at al., 1984) identified no significant difference in milk production between groups created according the calfhood health status. In contrast, in more recent studies, milk production was higher for healthy versus cDIA calves (Svensson and Hultgren, 2008). Modern dairy cows have been selected on increasing productivity since decades. Al-Kanaan (2016) hypothesised effects of intensive milk-yield selection on covariances between functional and production traits over a 30-year time.

For FPR (health/energy efficiency indicator), non-significant ($P > 0.05$) differences of least squares means were observed, with different trends for different calf disease traits (Table 3). A large FPR indicates metabolic disorders such as ketosis. Accordingly, the calf disease status was not related with flMET recorded in first parity cows (Table 2). The relationship between FPR and the probability for the occurrence of health disorders is not linear. Bergk and Swalve (2011) identified an intermediate optimum for FPR, whereas extremely high as well as quite low FPR increased the risk for involuntary cow cullings early in lactation.

Table 3. Least squares means (LSM) for test-day production traits of first-lactation cows in dependency of their general calf disease status (healthy or diseased calf)

Production trait[1]	Healthy calf		Diseased calf		Difference of LSM	P-value[2]
	LSM	S.E.	LSM	S.E.		
MY_1	29.126	0.304	29.034	0.306	0.091	NS
MY_2	32.938	0.259	32.698	0.261	0.241	**
FP_1	4.404	0.035	4.418	0.036	-0.014	NS
FP_2	3.774	0.025	3.783	0.026	-0.009	NS
PP_1	3.231	0.015	3.232	0.015	-0.001	NS
PP_2	3.073	0.011	3.078	0.011	-0.005	NS
FY_1	1.264	0.013	1.258	0.013	0.006	NS
FY_2	1.237	0.011	1.226	0.011	0.011	**
PY_1	0.928	0.009	0.926	0.009	0.003	NS
PY_2	1.008	0.008	0.999	0.008	0.009	**
FPR_1	1.371	0.011	1.367	0.011	0.004	NS
FPR_2	1.234	0.008	1.233	0.009	0.000	NS

[1]MY = milk yield, FP = fat percentage, PP = protein percentage, FY = fat yield, PY = protein yield, FPR = fat to protein-ratio; *_1 indicates the observation from the first test-day after calving, and *_2 indicates the observation from the second test-day after calving

[2]Significance level: *** = $P \leq 0.001$; ** = $P \leq 0.01$; * = $P \leq 0.05$; n.s. = not significant

Table 4. Least squares means (LSM) for test-day production traits of first-lactation cows in dependency of their calf respiratory disease status (healthy or diseased calf)

Production trait[1]	Healthy calf		Diseased calf		Difference of LSM	P-value[2]
	LSM	S.E.	LSM	S.E.		
MY_1	29.159	0.302	28.913	0.309	0.246	*
MY_2	32.924	0.258	32.630	0.263	0.295	***
FP_1	4.405	0.035	4.415	0.036	-0.010	NS
FP_2	3.766	0.025	3.784	0.026	-0.018	*
PP_1	3.231	0.015	3.233	0.015	-0.002	NS
PP_2	3.073	0.011	3.077	0.011	-0.004	NS
FY_1	1.265	0.013	1.254	0.013	0.011	**
FY_2	1.237	0.011	1.222	0.011	0.015	***
PY_1	0.929	0.009	0.922	0.009	0.007	**
PY_2	1.007	0.008	0.997	0.008	0.010	***
FPR_1	1.371	0.010	1.367	0.011	0.003	NS
FPR_2	1.235	0.008	1.231	0.009	0.004	NS

[1]MY = milk yield, FP = fat percentage, PP = protein percentage, FY = fat yield, PY = protein yield, FPR = fat to protein-ratio; *_1 indicates the observation from the first test-day after calving, and *_2 indicates the observation from the second test-day after calving

[2]Significance level: *** = $P \leq 0.001$; ** = $P \leq 0.01$; * = $P \leq 0.05$; n.s. = not significant

Table 5. Least squares means (LSM) for test-day production traits of first-lactation cows in dependency of their calf diarrhea disease status (healthy or diseased calf)

Production Trait[1]	Healthy calf		Diseased calf		Difference of LSM	P-value[2]
	LSM	S.E.	LSM	S.E.		
MY_1	29.105	0.302	29.010	0.313	0.095	NS
MY_2	32.873	0.257	32.679	0.267	0.195	*
FP_1	4.410	0.035	4.423	0.036	-0.013	NS
FP_2	3.775	0.025	3.798	0.026	-0.023	**
PP_1	3.231	0.015	3.234	0.016	-0.002	NS
PP_2	3.075	0.011	3.080	0.011	-0.005	NS
FY_1	1.261	0.013	1.263	0.014	-0.002	NS
FY_2	1.233	0.011	1.233	0.011	0.000	NS
PY_1	0.928	0.009	0.926	0.009	0.002	NS
PY_2	1.005	0.008	1.001	0.008	0.004	NS
FPR_1	1.369	0.010	1.372	0.011	-0.003	NS
FPR_2	1.232	0.008	1.239	0.009	-0.006	*

[1]MY = milk yield, FP = fat percentage, PP = protein percentage, FY = fat yield, PY = protein yield, FPR = fat to protein-ratio; *_1 indicates the observation from the first test-day after calving, and *_2 indicates the observation from the second test-day after calving

[2]Significance level: *** = $P \leq 0.001$; ** = $P \leq 0.01$; * = $P \leq 0.05$; n.s. = not significant

Calf-cow trait associations on the quantitative-genetic scale

Genetic parameters for calf diseases. I applied univariate threshold animal models, and bivariate linear-linear models. All models converged. Due to the large dataset combined with a high average no. of offspring per sire, S.E. for heritabilities and genetic correlations were quite small. Theoretically, results from sire threshold models are more reliable compared to results from animal threshold model because of the "extreme data category problem", resulting in possible convergence problems (e.g. Hoeschele and Tier, 1995).

Heritabilities for calf diseases were 0.06 for cGDS, 0.07 for cRD, and 0.06 for cDIA (Table 6). Almost identical heritabilities for cGDS and cRD, and for cGDS and cDIA, were expected, because cRD and cDIA represent subsets of cGDS. Additionally, the high EBV and genetic correlations between the overall disease trait cGDS and subset diseases reflect such overlaps in trait definitions. Those high correlations between an unspecified disease trait definition cGDS compared to detailed recorded single diseases cRD or cDIA justifies a quite simple disease recording system ("producer data") instead the utilization of sophisticated diagnosis keys on a veterinarian basis. In most cases, correlations among EBV were "less extreme" in comparison to genetic correlations. For EBV accuracies of one, EBV and genetic correlations are assumed to be equal. In the present study, EBV correlations are always an underestimation of genetic correlations, because prediction accuracies (r_{TI}) were smaller than one (Calo et al., 1973). For sires with daughter records and following the principles of selection index theory for the low heritability traits cDIA and cRD, $r_{TI} = (2nh^2)/(4 + (n-1)h^2)$ = 0.76, with *n* being 40 daughter records per sire. The correlation between EBVs for cDIA and EBVs for cRD was low (0.13), and only moderate for the genetic association between both traits (r_g = 0.29). Hence, different genes seem to contribute to the different calf disease phenotypes, implying limited selection response for cRD when selection is based on cDIA, and vice versa. The phenotypic correlation between cRD and cDIA was of similar magnitude (0.18). Previous studies (e.g. Henderson et al., 2011) reported substantial differences when comparing genetic, phenotypic and residual correlations among calve diseases.

Table 6. Heritabilities (diagonal) with corresponding S.E., genetic correlations with corresponding S.E. (above diagonal), and correlations among estimated breeding values considering 242 sires with greater equal than 30 daughters (below diagonal) for the calf disease traits general disease status (cGDS), diarrhea (cDIA), and respiratory disease (cRD)

	cGDS	**cDIA**	**cRD**
cGDS	0.06 ± 0.01	0.63 ± 0.07	0.78 ± 0.05
cDIA	0.48	0.06 ± 0.01	0.29 ± 0.09
cRD	0.71	0.13	0.07 ± 0.01

Genetic relationships between calf and cow diseases. Heritabilites for cow diseases in the early lactation period in first parity were in a narrow range between 0.001 (flRD and flMET) and 0.11 (flDIA) (Table 7). Low heritability estimates and low additive-genetic variances for cow diseases in the period directly after calving were also identified by König et al. (2005) for claw disorders when applying robust single trait animal models, or by Gernand et al. (2012) for clinical mastitis, claw disorders and metabolic disorders based on repeated measurement analyses. Such small heritabilities for disease traits during the quite sensitive period around calving suggest modifications of feeding, husbandry or management strategies to improve the cow health status.

Correlations between EBV for calf diseases with EBV for first-lactation cow diseases were close to zero (Table 7). The highest positive EBV correlation (0.07) was between cGDS and flCLAW, and the most pronounced negative value was -0.09 between cDIA and flFF. Also most of the genetic correlations between calf and cow diseases were in a range close to zero, and confirming EBV correlations. Only the genetic correlations between flRD with cGDS (-0.38) and between flMET with cDIA (-0.21) substantially differed from the corresponding EBV-correlations (-0.05, and 0.00, respectively). A negative genetic calf-cow disease trait correlation (as found for most of the trait combinations) indicates that genetic resistance against a calf disease increases genetic susceptibility to a cow disease. Such findings on the genetic scale are in agreement with phenotypic results (Table 2), i.e. indicating by trend decreasing probabilities for the occurrence of a cow disease when the same female was infected during calfhood. Such concepts of resistance and susceptibility on both phenotypic and genetic scales were discussed by König (2012).

However, most of the phenotypic calf-cow disease associations were statistically not significant (*P > 0.05*), and correlations among EBV were not significantly different from zero. Those results imply that calf diseases are unsuitable early predictors with regard to the application of indirect selection strategies during calfhood, aiming on genetic improvements of a milking cows' health status in early lactation in first parity.

Table 7. Heritabilities (h^2) with corresponding S.E. for first-lactation cow disease traits, correlations between estimated breeding values for cow disease traits with calf disease traits considering 242 sires with greater equal than 30 daughters ($r_{EBV\text{-}sire}$), and genetic correlations (r_g) between cow and calf disease traits with corresponding S.E.

		Calf disease[2]					
		cGDS		cDIA		cRD	
Cow disease[1]	h^2	r_g	$r_{EBV\text{-}sire}$	r_g	$r_{EBV\text{-}sire}$	r_g	$r_{EBV\text{-}sire}$
flGDS	0.07	-0.10	-0.09	-0.18	-0.01	-0.03	-0.07
	±0.01	±0.10		± 0.10		± 0.10	
flDIA	0.11	-0.08	-0.08	0.00	-0.08	-0.18	-0.03
	±0.02	±0.10		± 0.10		± 0.10	
flRD	0.00	-0.38	-0.05	-0.02	0.05	-0.05	-0.02
	±0.02	±0.14		± 0.17		± 0.12	
flMAST	0.07	-0.09	-0.02	-0.10	-0.05	0.22	0.03
	±0.01	±0.10		± 0.10		± 0.10	
flCLAW	0.06	0.03	0.07	0.11	0.04	-0.01	0.02
	±0.01	±0.10		±0.10		±0.10	
flFF	0.05	0.00	-0.07	-0.05	-0.09	-0.08	-0.02
	±0.01	±0.10		±0.11		±0.10	
flMET	0.00	-0.03	0.00	-0.21	0.00	-0.02	0.00
	±0.00	±0.17		±0.17		±0.16	

[1]flGDS = first-lactation general disease status, flDIA = first-lactation diarrhea, flRD = first-lactation respiratory disease, flMAST = first-lactation mastitis, flCLAW = first-lactation claw disorders, flFF = first-lactation female fertility disorders, flMET = first-lactation metabolic disorders

[2] cGDS = general calf disease status, cDIA = calf diarrhea, cRD = calf respiratory disease

Genetic relationships between calf diseases and cow production traits. Heritabilities for test-day production traits (Table 8) reflect the well-known pattern of parameter estimates (e.g. Brügemann et al., 2011): heritabilities for all traits were throughout higher at the official test-day 2 compared to corresponding results at test-day 1, and across traits, again highest heritabilities were estimated for the content traits FP and PP (0.27 and 0.36, respectively). Smaller heritabilities at test-day 1 were due to larger residual variances compared to estimates later in lactation. For the health trait indicator FPR, heritabilities (0.14 for FPR_1 and 0.19 for FPR_2) were substantially larger compared to the corresponding health trait flMET (0.001, Table 7), or in comparison to corresponding single diseases such as ketosis diagnoses based on producer data (e.g. Zwald et al., 2004).

The FPR recorded early in lactation might be a valuable indicator for cow health or cow disposals (Bergk and Swalve, 2011), but in analogy with remaining test-day traits, only weak EBV relationships and genetic correlations with calf disease traits were identified (Table 8). However, negative EBV correlations as well as negative genetic correlations were found between FPR_1 and FPR_2 with all calf disease traits. Such results confirm the negative correlations between calf diseases cDIA and cRD with the cow diseases flFF and flMET, because especially extremely high FPR indicate metabolic disorders (code = 1) early in lactation. All genetic correlations between MY_1, MY_2, FY_1, FY_2, PY_1, and PY_2 with the calf disease traits were slightly positive. Positive genetic correlations between productivity and functional traits (genetic antagonism) suggest that breeding on, e.g., increasing milk yield, simultaneously increases the risk for the occurrence of a disease.

Nevertheless, due to the "very close to zero EBV and genetic correlations" between a calf disease with all first-lactation test-day production traits, breeders cannot use those calf traits as early predictors to select, e.g., high-yielding milking cows for the first lactation period in parity one. From a dairy cattle breeding perspective, it is imperative to improve selection response on the cow-dam pathway of selection, because of the increasing replacement rates combined with quite long generation intervals. Weigel et al. (2012) defined the potential for improvements of on-farm selection strategies when genotyping female calves, but it remains questionable if commercial farms will invest in genotyping activities. As an alternative, it might be worthwhile to identify phenotypic indicators in potential selection candidates early in life, as achieved by Biermann et al. (2015) for meat quality traits in pigs.

Table 8. Heritabilities (h^2) with corresponding S.E. for first-lactation test-day production traits, correlations between estimated breeding values for test-day production traits with calf disease traits considering 242 sires with greater equal than 30 offspring ($r_{EBV\text{-}sire}$), and genetic correlations (r_g) between test-day production traits and calf disease traits with corresponding S.E.

[1]MY = milk yield, FP = fat percentage, PP = protein percentage, FY = fat yield, PY = protein yield,

Production Trait[1]		Calf disease[2]					
		cGDS		cDIA		cRD	
	h^2	r_g	$r_{EBV\text{-}sire}$	r_g	$r_{EBV\text{-}sire}$	r_g	$r_{EBV\text{-}sire}$
MY_1	0.16 ± 0.01	0.04 ± 0.08	0.05	0.06 ± 0.08	0.03	0.04 ± 0.08	0.02
MY_2	0.25 ± 0.02	0.02 ± 0.07	0.01	0.08 ± 0.08	0.03	0.02 ± 0.07	-0.03
FP_1	0.13 ± 0.01	0.04 ± 0.07	-0.03	-0.15 ± 0.09	-0.06	0.03 ± 0.08	-0.06
FP_2	0.27 ± 0.02	-0.02 ± 0.07	-0.04	-0.14 ± 0.07	-0.02	-0.04 ± 0.07	-0.06
PP_1	0.15 ± 0.01	0.03 ± 0.07	0.03	-0.03 ± 0.08	-0.12	0.05 ± 0.08	0.03
PP_2	0.36 ± 0.02	-0.02 ± 0.07	0.09	0.01 ± 0.07	-0.03	0.02 ± 0.08	0.08
FY_1	0.14 ± 0.01	0.08 ± 0.08	0.02	0.06 ± 0.09	-0.01	0.07 ± 0.08	-0.01
FY_2	0.19 ± 0.02	0.02 ± 0.08	-0.02	0.06 ± 0.08	0.03	0.05 ± 0.08	-0.08
PY_1	0.14 ± 0.01	0.08 ± 0.08	0.08	0.08 ± 0.09	-0.03	0.08 ± 0.08	0.05
PY_2	0.19 ± 0.01	0.04 ± 0.08	0.07	0.08 ± 0.08	0.02	0.04 ± 0.08	0.01
FPR_1	0.14 ± 0.01	-0.02 ± 0.08	-0.05	-0.17 ± 0.08	-0.03	-0.01 ± 0.08	-0.08
FPR_2	0.19 ± 0.02	-0.04 ± 0.07	-0.10	-0.14 ± 0.08	0.00	-0.01 ± 0.07	-0.12

FPR = fat to protein-ratio; *_1 indicates the observation from the first test-day after calving, and *_2 indicates the observation from the second test-day after calving

[2] cGDS = general calf disease status, cDIA = calf diarrhea, cRD = calf respiratory disease

SNP-marker associations between calve and cow traits

Different SNP-markers from different chromosomes contributed to either the occurrence of a calf or a cow disease. With regard to the Manhattan plots for respiratory diseases, significant SNP-effects were identified on chromosomes 8 and 19 for cRD (Figure 2A), but on chromosome 7 and 8 for flRD (Figure 2B). On chromosome 8, I identified SNP within a segment of 8 cM significantly associated with both traits cRD and flRD, but the exact location and the direction of marker effects differed between calves and cows. The correlation based on SNP-effects between the both traits cRD and flRD was only 0.05 considering SNP on chromosome 7, and was marginally larger for SNP on chromosome 8 (0.14) (Figure 3). The average correlation coefficient between cRD and flRD considering SNP-effects from all chromosomes was 0.01, supporting the "zero-EBV-correlations" between both traits in the quantitative-genetic analysis (Table 7). Pimentel et al. (2011) combined 50K SNP data with results from gene expression profiles, and they estimated SNP-effects located in relevant chromosomal regions. Despite the naturally antagonistic relationship between productivity and fertility, they identified some significant SNP simultaneously improving both trait categories. In this context, Swalve (2014) suggested to select only those bull dams and bull sires carrying the favourable alleles, in order to achieve "a gradual change of genetic antagonisms".

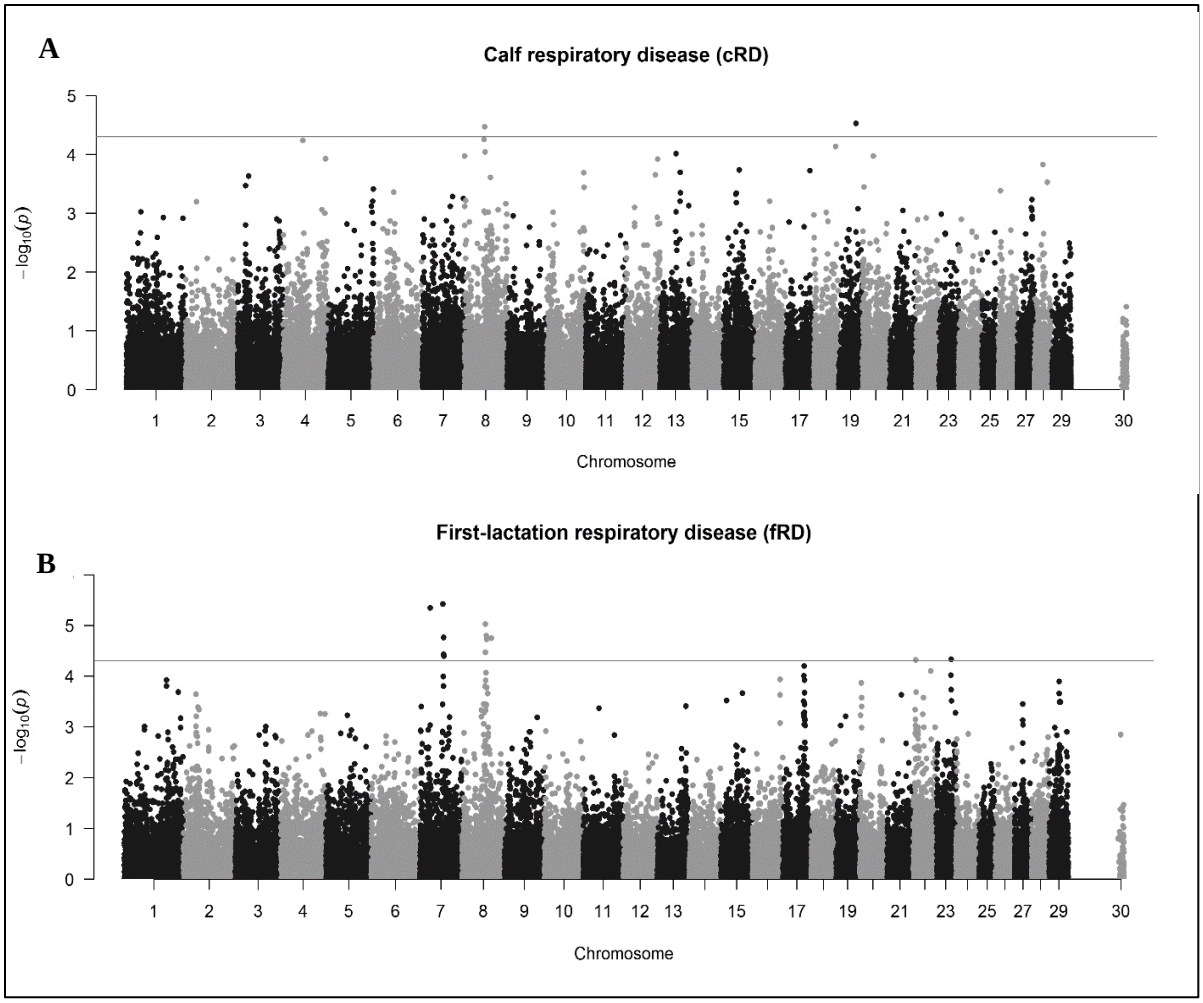


Figure 2. Manhattan plots for SNP-allele substitution effects for (**A**) cRD = calf respiratory disease and (**B**) flRD = first-lactation respiratory disease. The horizontal line indicates the genome-wide significance threshold value for the 50k genotypes (P-value = 5 x 10^{-5}).

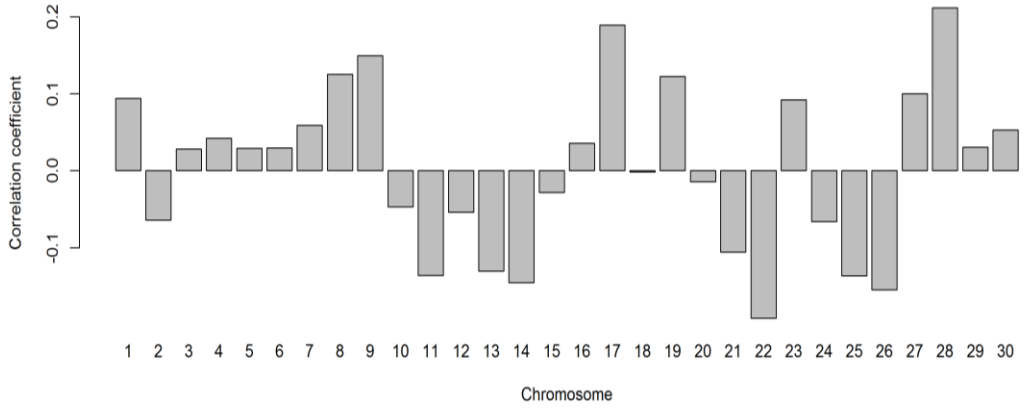


Figure 3. Correlation coefficient between marker effects on different chromosomes between cRD = calf respiratory disease and flRD = first-lactation respiratory disease.

However, for the same denomination of a calf and a cow disease, only a limited no. of SNP affected both traits significantly. This is especially valid for diarrhea. For cDIA, I identified nine significant SNP located on chromosomes 1, 2, 5, 10, 11, 14, and 24 (Figure 4a), but only two SNP on chromosomes 4 and 10 significantly contributed to flDIA (Figure 4b). The highest correlation between cDIA and flDIA based on SNP effects was identified on chromosome 23 (0.17), and a pronounced antagonistic relationship exhibited chromosome 29 (-0.29) (Figure 5). The average correlation coefficient between cDIA and flDIA considering SNP-effects from all chromosomes was -0.04.

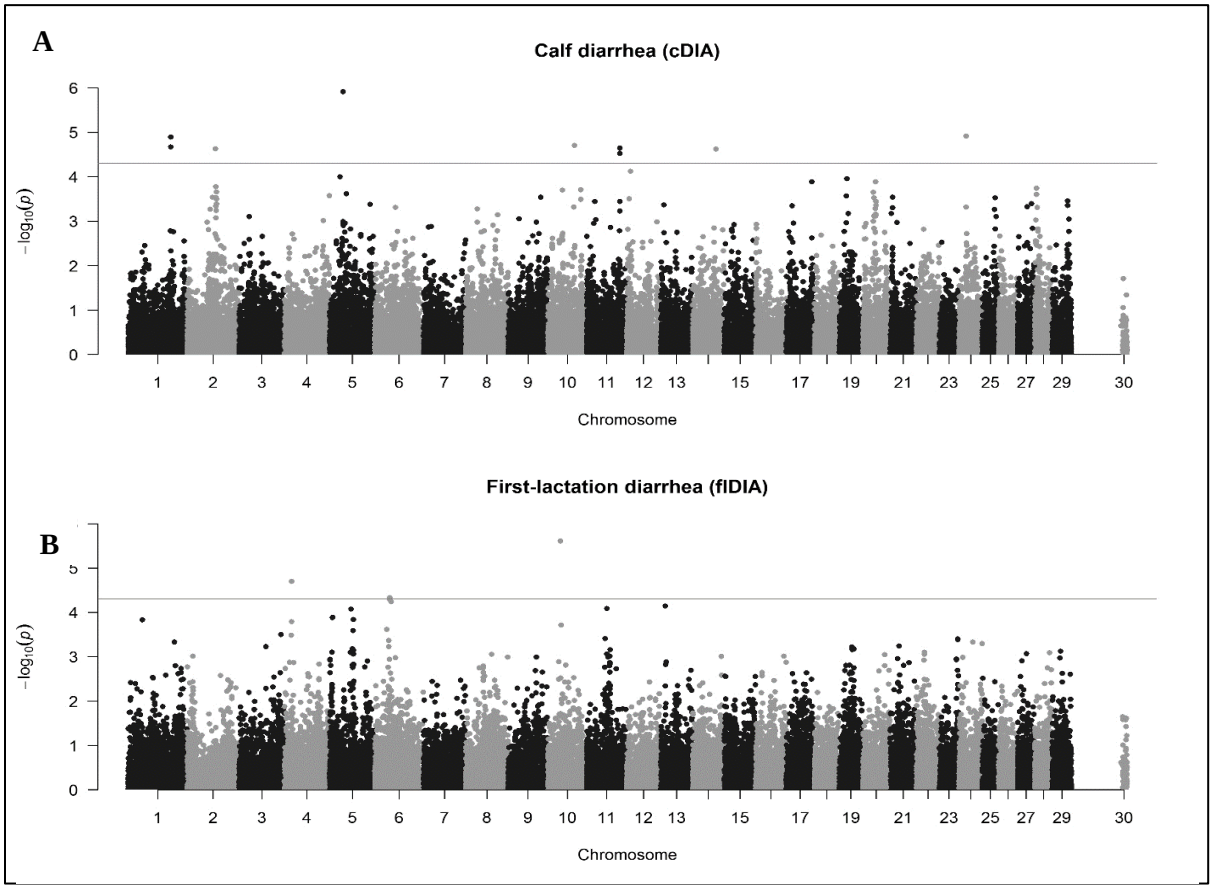


Figure 4. Manhattan plots for SNP-allele substitution effects for (A) cDIA = calf diarrhea and (B) flDIA = first-lactation diarrhea. The horizontal line indicates the genome-wide significance threshold value for the 50k genotypes (P-value =5 x 10-5).

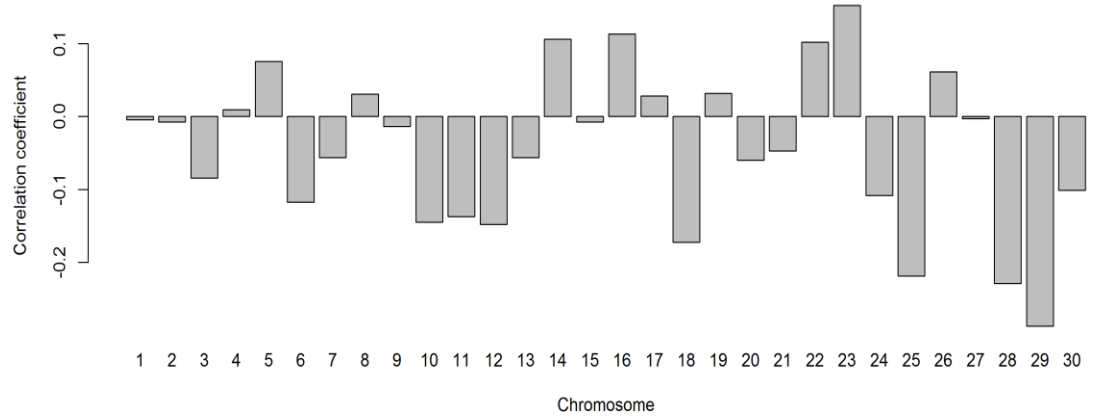


Figure 05. Correlation coefficient between marker effects on different chromosomes between cDIA = calf diarrhea and flDIA = first-lactation diarrhea.

Results from GWAS can be used to identify functional mutations. The most significantly associated marker for cRD on chromosome 19 at 46.75 cM was located near to a QTL detected for *"Bovine respiratory disease"* in cattle (chromosome 19, 46.75cM) (Keele et al., 2015). For flRD, the significant marker at chromosome 23 was located close to another QTL for *"Bovine respiratory disease"* (Keele et al., 2015). Three markers were significantly associated with flDIA. The most significantly associated marker on chromosome 10 at 34.38 cM was in close distance to a QTL detected for *"Bovine viral diarrhea virus susceptibility"* in cattle (Casas et al., 2015).

CONCLUSIONS

For the calf diseases cGDS, cDIA and cRD, incicences were quite high (0.48, 0.28 and 0.21, respectively), but rather low additive-genetic variances and heritabilities were identified. Heritabilities were 0.06 for cGDS and cDIA, and 0.07 for cRD. The EBV correlation between cDIA and cRD was 0.13, and the genetic correlation between both traits was 0.29. Apart from the genetic correlation between flRD with cGDS (-0.38), EBV-correlations as well as genetic correlations between calf diseases and cow traits from early lactation in parity one, were close to zero. Those results imply that both calf diseases cDIA and cRD, as well as calf and cow disease, are different traits. Hence, calf traits are inappropriate indicators or inappropriate early predictors to genetically improve cow health or cow productivity directly after calving in first parity. Correlations based on SNP-marker effects from all chromosomes were also close to zero between same definitions for calf and cow diseases. GWAS identified different significant SNP-markers contributing to cDIA and flDIA, or to cRD and flRD. Furthermore, on the phenotypic scale, weak relationships between calf diseases and cow traits were identified. By trend, infected calves had lower probabilities for the occurrence of a cow disease, pointing to mechanisms of genetic resistance after infections.

ACKNOWLEDGEMENT

The authors gratefully acknowledge funding from the German Federal Ministry of Education and Research (BMBF) and from the Förderverein Bioökonomieforschung e.V. (FBV) / German Holstein Association (DHV) for the collaborative project "KMU-innovativ-10: Kuh-L – cow calibration groups for the implementation of selection strategies based on high-density genotyping in dairy cattle", grant no. 031A416C.

REFERENCES

Adams, L. G., and J. W. Templeton. 1998. Genetic resistance to bacterial diseases of animals. Rev. Sci. Tech. 17:200–219.

Al-Kanaan, A. 2016. Heat stress response for physiological traits in dairy and dual purpose cattle populations on phenotypic and genetic scales. PhD thesis, Faculty of Organic Agriculture, University of Kassel.

Beam, D., K. J. Stalder, A. J. Heinrichs, and C. D. Dechow. 2015. Association of calf growth and feed intake traits with mature cow milk yield and body weight. Animal Industry Report. 661:30.

Becher, K. A., I. D. Robertson, D. M. Fraser, D. G. Palmer, and R. C. A. Thompson. 2004. Molecular epidemiology of Giardia and Cryptosporidium infections in dairy calves originating from three sources in Western Australia. Vet. Parasitol. 123:1–9.

Bergk, N., and H. H. Swalve. 2011. Fat-to-protein-ratio in early lactation as an indicator of herdlife for first lactation dairy cows. Züchtungskunde. 83:89–103.

Bermingham, M. L., S. C. Bishop, J. A. Woolliams, R. Pong-Wong, A. R. Allen, S. H. McBride, J. J. Ryder, D. M. Wright, R. A. Skuce, S. W. McDowell, and E. J. Glass. 2014. Genome-wide association study identifies novel loci associated with resistance to bovine tuberculosis. Heredity 112:543–551.

Biermann, A. D. M., T. Yin, U. U. König von Borstel, K. Rübesam, B. Kuhn, and S. König. 2015. From phenotyping towards breeding strategies: using in vivo indicator traits and genetic markers to improve meat quality in an endangered pig breed. Animal. 9:919–927.

Britney, J. B., S. W. Martin, J. B. Stone, and R. A. Curtis. 1984. Analysis of early calfhood health status and subsequent dairy herd survivorship and productivity. Prev. Vet. Med. 3:45–52.

Brügemann, K., E. Gernand, U. U. von Borstel, and S. König. 2011. Genetic analyses of protein yield in dairy cows applying random regression models with time-dependent and temperature x humidity-dependent covariates. J. Dairy Sci. 94:4129–4139.

Brügemann, K., E. Gernand, U. U. von Borstel, and S. König. 2013. Application of random regression models to infer the genetic background and phenotypic trajectory of binary conception rate by alterations of temperature × humidity indices. Livestock Science. 157:389–396.

Buch, L. H., M. Kargo, P. Berg, J. Lassen, and A. C. Sørensen. 2012. The value of cows in reference populations for genomic selection of new functional traits. Animal. 6:880–886.

Bünger, U., P. Schmoldt, J. Pongé, H. Rotermund, P. Kaphengst, T. Motsch, W. Brade, and W. Kleiner. 1979. [Effect of diseases on the nutrient intake by suckling calves]. Arch. Exp. Veterinarmed. 33:347–365.

Calo, L. L., R. E. McDowell, L. D. VanVleck, and P. D. Miller. 1973. Genetic aspects of beef production among Holstein-Friesians pedigree selected for milk production. J. Anim. Sci. 37:676–682.

Casas, E., B. E. Hessman, J. W. Keele, and J. F. Ridpath. 2015. A genome-wide association study for the incidence of persistent bovine viral diarrhea virus infection in cattle. Anim. Genet. 46:8–15.

Cho, Y.-I., and K.-J. Yoon. 2014. An overview of calf diarrhea - infectious etiology, diagnosis, and intervention. J. Vet. Sci. 15(1): 1–17.

Cobb, C. J., B.S. Obeidat, M. D. Sellers, A. R. Pepper-Yowell, and M. A. Ballou. 2014. Group housing of Holstein calves in a poor indoor environment increases respiratory disease but does not influence performance or leukocyte responses. J. Dairy Sci. 97:3099–3109.

Colditz, I. G., D. L. Watson, G. D. Gray, and S. J. Eady. 1996. Some relationships between age, immune responsiveness and resistance to parasites in ruminants. Int. J. Parasitol. 26:869–877.

Cole, J. B., B. Waurich, M. Wensch-Dorendorf, D. M. Bickhart, and H. H. Swalve. 2014. A genome-wide association study of calf birth weight in Holstein cattle using single nucleotide polymorphisms and phenotypes predicted from auxiliary traits. J. Dairy Sci. 97:3156–3172.

Collard, B. L., P. J. Boettcher, J. C. M. Dekkers, D. Petitclerc, and L. R. Schaeffer. 2000. Relationships between energy balance and health traits of dairy cattle in early lactation. J. Dairy Sci. 83:2683–2690.

Curtis, C. R., H. N. Erb, and M. E. White. 1988. Descriptive epidemiology of calfhood morbidity and mortality in New York Holstein herds. Prev. Vet. Med. 5:293–307.

Dadousis, C., S. Biffani, C. Cipolat-Gotet, E.L. Nicolazzi, A. Rossoni, E. Santus, G. Bittante, and A. Cecchinato. 2016. Genome-wide association of coagulation properties, curd firmness modeling, protein percentage, and acidity in milk from Brown Swiss cows. J. Dairy Sci.

Egger-Danner, C., J. B. Cole, J. E. Pryce, N. Gengler, B. Heringstad, A. Bradley, and K. F. Stock. 2015. Invited review: overview of new traits and phenotyping strategies in dairy cattle with a focus on functional traits. Animal 9:191–207.

Faber, J. E., D. Kollmann, A. Heise, C. Bauer, K. Failing, H. J. Bürger, and H. Zahner. 2002. Eimeria infections in cows in the periparturient phase and their calves: oocyst excretion and levels of specific serum and colostrum antibodies. Vet. Parasitol. 104:1–17.

Feucker, W., and R. Staufenbiel. 2003. Zentraler Diagnoseschlüssel. http://www.portal-rind.de/index.php?module=Downloads&func=prep_hand_out&lid=17. Accessed 16.03.2010.

Fuerst-Waltl, B., A. Koeck, C. Fuerst, and C. Egger-Danner. 2010. Genetic analysis of diarrhea and respiratory diseases in Austrian Fleckvieh heifer calves. CD-ROM Communication 4–138/Article No. 304 in Proc. 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Germany, August 1-6, 2010.

Gernand, E., P. Rehbein, U. U. von Borstel, and S. König. 2012. Incidences of and genetic parameters for mastitis, claw disorders, and common health traits recorded in dairy cattle contract herds. J. Dairy Sci. 95:2144–2156.

Gulliksen, S. M., K. I. Lie, and O. Osterås. 2009a. Calf health monitoring in Norwegian dairy herds. J. Dairy Sci. 92 :1660–1669.

Gulliksen, S. M., E. Jor, K. I. Lie, I. S. Hamnes, T. Løken, J. Akerstedt, and O. Osterås. 2009b. Enteropathogens and risk factors for diarrhea in Norwegian dairy calves. J. Dairy Sci. 92:5057–5066.

Hatch, R. D., B. O. Brodie, S. L. Spahr, and B. G. Erwin. 1974. Calf pneumonia-enteritis; a herd survey. Vet. Med. Small. Anim. Clin. 69:1290–1294.

Hara S., U. S. Diesterbeck, S. König, S., and C.-P. Czerny. 2012. Transcriptional analysis of equine λ-light chains in the horse breeds Rhenish-German Coldblood and Hanoverian Warmblood. Vet. Immunol. and Immunopath. 145 (1-2):50-65.

Henderson, L., F. Miglior, A. Sewalem, J. Wormuth, D. Kelton, A. Robinson, and K. E. Leslie. 2011. Short communication: genetic parameters for measures of calf health in a population of Holstein calves in New York State. J. Dairy Sci. 94:6181–6187.

Heringstad, B., Y. M. Chang, D. Gianola, and O. Østerås. 2008. Short Communication: Genetic analysis of respiratory disease in Norwegian Red calves. J. Dairy Sci. 91:367–370.

Hoeschele, I., and B. Tier. 1995. Estimation of variance components of threshold characters by marginal posterior modes and means via Gibbs sampling. Genet. Sel. Evol. 27:519-540.

Howard, C. J., M. C. Clarke, and J. Brownlie. 1989. Protection against respiratory infection with bovine virus diarrhoea virus by passively acquired antibody. Vet. Microbiol. 19:195–203.

Hultgren, J., and C. Svensson. 2009. Lifetime risk and cost of clinical mastitis in dairy cows in relation to heifer rearing conditions in southwest Sweden. J. Dairy Sci. 92:3274–3280.

Keele, J. W., L. A. Kuehn, T. G. McDaneld, R. G. Tait, S.A. Jones, T. P. L. Smith, S. D. Shackelford, D. A. King, T. L. Wheeler, A. K. Lindholm-Perry, and A. K. McNeel. 2015. Genomewide association study of lung lesions in cattle using sample pooling. J. Anim. Sci. 93:956–964.

Klein‐Jöbstl, D., T. Arnholdt, F. Sturmlechner, M. Iwersen, and M. Drillich. 2015. Results of an online questionnaire to survey calf management practices on dairy cattle breeding farms in Austria and to estimate differences in disease incidences depending on farm structure and management practices. Acta Vet. Scand. 57:44.

König, S., A. R. Sharifi, H. Wentrot, D. Landmann, M. Eise, and H. Simianer. 2005. Genetic parameters of claw and foot disorders estimated with logistic models. J. Dairy Sci. 88:3316–3325.

Lee, S. H., N. R. Wray, M. E. Goddard, and P. M. Visscher. 2011. Estimating missing heritability for disease from genome-wide association studies. Am. J. Hum. Genet. 88:294–305.

Lundborg, G. K., P. A. Oltenacu, D. O. Maizon, E. C. Svensson, and P. G. A. Liberg. 2003. Dam-related effects on heart girth at birth, morbidity and growth rate from birth to 90 days of age in Swedish dairy calves. Prev. Vet. Med. 60:175–190.

Madsen P., Jensen J. 2013. DMU – a package for analysing multivariate mixed models. Version 6, release 5.2. Aarhus University, Foulum, Denmark. Available from http://dmu.agrsci.dk.

Martin, G., M. Wensch-Dorendorf, R. Schafberg, and H. H. Swalve. 2013. A comparison of udder health trait definitions in German Holstein dairy cattle based on mastitis treatment data. Livest Sci. 151:219–227.

Minozzi, G., E.L. Nicolazzi, A. Stella, S. Biffani, R. Negrini, B. Lazzari, P. Ajmone-Marsan, and J.L. Williams. 2013. Genome wide analysis of fertility and production traits in Italian Holstein cattle. PLoS One. 8:e80219.

Mousa, M. S., A. J. Seykora, and M. A. Abdellatif. 2015. Genetic and phenotypic correlations of age at first calving with some calf-hood fitness and health in Holstein heifers. Egypt. J. Anim. Prod. 52:37–43.

Perez, E., J. P. T. M. Noordhuizen, L. A. Van Wuijkhuise, and E. N. Stassen. 1990. Management factors related to calf morbidity and mortality rates. Livest. Prod. Sci. 25:79.

Pimentel, E. C. G., S. Bauersachs, M. Tietze, H. Simianer, J. Tetens, G. Thaller, F. Reinhardt, E. Wolf, and S. König. 2011. Exploration of relationships between production and fertility traits in dairy cattle via association studies of SNPs within candidate genes derived by expression profiling. Anim. Genet. 42:251–262.

R Core Team. 2013. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rossini, K. 2004. Effects of calfhood respiratory and digestive disease on calfhood morbidity and first lactation production and survival rates. MS Thesis. Virginia State Tech Univ., Blacksburg.

Santin, M., J. M. Trout, and R. Fayer. 2008. A longitudinal study of cryptosporidiosis in dairy cattle from birth to 2 years of age. Vet. Parasitol. 155:15–23.

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

Segelke, D., J. Chen, Z. Liu, F. Reinhardt, G. Thaller, and R. Reents. 2012. Reliability of genomic prediction for German Holsteins using imputed genotypes from low-density chips. J. Dairy Sci. 95:5403–5411.

Sivula, N. J., T. R. Ames, and W. E. Marsh. 1996. Management practices and risk factors for morbidity and mortality in Minnesota dairy heifer calves. Prev. Vet. Med. 27:173–182.

Southey, B. R., S. L. Rodriguez-Zas, and K. A. Leymaster. 2003. Discrete time survival analysis of lamb mortality in a terminal sire composite population. J. Anim. Sci. 81:1399–1405.

Stock, K. F., J. Cole, J. Pryce, N. Gengler, A. Bradley, L. Andrews, B. Heringstad and C. Egger-Danner. 2013. Standardization of health data – ICAR guidelines including health key. In: Challenges and benefits of health data recording in the context of food chain quality, management and breeding, Egger-Danner, C., O.K. Hansen, K. Stock, J.E. Pryce, J. Cole, N. Gengler & B. Heringstad (editors), ICAR Technical Series No. 17, ICAR, Rome, Italy, p. 75-81.

Svensson, C., and J. Hultgren. 2008. Associations between housing, management, and morbidity during rearing and subsequent first-lactation milk production of dairy cows in southwest Sweden. J. Dairy Sci. 91:1510–1518.

Svensson, C., J. Hultgren, and P. A. Oltenacu. 2006. Morbidity in 3-7-month-old dairy calves in south-western Sweden and risk factors for diarrhea and respiratory disease. Prev. Vet. Med. 74:162–179.

Swalve, H.H. 2014. Use of high density marker data– the role of individual functional mutations. In: Proc. 10th Wld. Congr. Genet. Appl. Livest. Prod., Vancouver, Canada, 17 Aug.https://www.asas.org/docs/default-source/wcgalp-proceedings-oral/302_paper_10260_manuscript_1726_0.pdf?sfvrsn=2

Thompson-Crispi, K. A., M. Sargolzaei, R. Ventura, M. Abo-Ismail, F. Miglior, F. S. Schenkel, and B. A. Mallard. 2014. A genome-wide association study of immune response traits in Canadian Holstein cattle. BMC Genomics 15:559.

Van der Fels-Klerx, H. J., S. W. Martin, M. Nielen, and R. B. M. Huirne. 2002. Effects on productivity and risk factors of bovine respiratory disease in dairy heifers; a review for the Netherlands. Neth. J. Agric. Sci. 50:27–45.

van der Spek, D., J. A. M. van Arendonk, and H. Bovenhuis. 2015. Genome-wide association study for claw disorders and trimming status in dairy cattle. J. Dairy Sci. 98:1286–1295.

Vinson, W. E., and R. W. Kliewer. 1976. Overall classification as a selection criterion for improving categorically scored components of type in Holstein. J. Dairy Sci. 59:2104-2114.

Waltner-Toews, D., S. W. Martin, and A. H. Meek. 1986. The effect of early calfhood health status on survivorship and age at first calving. Can. J. Vet. Res. 50:314–317.

Warnick, L. D., H. N. Erb, and M. E. White. 1994. The association of calfhood morbidity with first calving age and dystocia in New York Holstein herds. Kenya Vet. 18:177.

Weigel, K. A., P. C. Hoffman, W. Herring, and T. J. Lawlor Jr. 2012. Potential gains in lifetime net merit from genomic testing of cows, heifers, and calves on commercial dairy farms. J. Dairy Sci. 95:2215–2225.

Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher. 2011. GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 88:76–82.

Yang, J., N. A. Zaitlen, M. E. Goddard, P. M. Visscher, and A. L. Price. 2014. Advantages and pitfalls in the application of mixed-model association methods. Nat Genet. 46:100–106.

Zwald, N. R., K. A. Weigel, Y. M. Chang, R. D. Welper, and J. S. Clay. 2004. Genetic selection for health traits using producer-recorded data. I. Incidence rates, heritability estimates, and sire breeding values. J. Dairy Sci. 87:4287–4294.

CHAPTER 4

Genome-wide pleiotropy and shared biological pathways for resistance to bovine pathogens

M Mahmoud[1,2], Y Zeng[2], M Shirali[3], T Yin[1], K Brügemann[1], S König[1], CS Haley[2,4]

[1]Institute of Animal Breeding and Genetics, Justus-Liebig-University Gießen, Gießen, Germany

[2]MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK

[3]Centre for Clinical Brain Sciences, University of Edinburgh, Kennedy Tower, Royal Edinburgh Hospital, Edinburgh, UK;

[4]Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Midlothian, UK.

PLoS ONE 13(4): e0194374.

https://doi.org/10.1371/journal.pone.0194374

ABSTRACT

Host genetic architecture is a major factor in resistance to pathogens and parasites. The collection and analysis of sufficient data on both disease resistance and host genetics has, however, been a major obstacle to dissection the genetics of resistance to single or multiple pathogens. A severe challenge in the estimation of heritabilities and genetic correlations from pedigree-based studies has been the confounding effects of the common environment shared among relatives which are difficult to model in pedigree analyses, especially for health traits with low incidence rates. To circumvent this problem I used genome-wide single-nucleotide polymorphism data and implemented the Genomic-Restricted Maximum Likelihood (G-REML) method to estimate the heritabilities and genetic correlations for resistance to 23 different infectious pathogens in calves and cows in populations undergoing natural pathogen challenge. Furthermore, I conducted gene-based analysis and generalized gene-set analysis to understand the biological background of resistance to infectious diseases. The results showed relatively higher heritabilities of resistance in calves than in cows and significant pleiotropy (both positive and negative) among some calf and cow resistance traits. I also found significant pleiotropy between resistance and performance in both calves and cows. Finally, I confirmed the role of the B-lymphocyte pathway as one of the most important biological pathways associated with resistance to all pathogens. These results both illustrate the potential power of these approaches to illuminate the genetics of pathogen resistance in cattle and provide foundational information for future genomic selection aimed at improving the overall production fitness of cattle.

INTRODUCTION

Infection is the colonization of the host by at least one domain of pathogens such as viruses, bacteria, fungi, and/or parasites (parasite infestation). Each type of pathogen is different in its invasion and replication in the host tissue, namely "infectivity" and/or the capacity to pass from one individual to another, namely "transmissibility" (Garrity et al., 2001, Alberts et al., 2002). Although **Error! Reference source not found.**variation in resistance to individual infectious agents (pathogens) is associated with levels of immunisation, disease treatment policies, diet and other environmental factors, previous studies of the resistance to various pathogens in animals (Bishop and Woolliams, 2014) and humans (Frodsham and Hill, 2004) have also revealed a major role of host genetic factors in pathogen resistance and host survival.

Host genetic architecture in cattle has been found to be a major factor in resistance to multifactorial diseases and disorders such as infertility, metabolic disorder, claw disorder and mastitis (Gernand et al., 2012). Several studies have addressed the importance of the genetic contribution of host resistance/susceptibility to different domains, species, and subspecies of pathogens (Berry et al., 2011, Raszek et al., 2016). The most studied pathogen subspecies in the field of dairy cattle breeding (particularly in Europe) are *Mycobacterium avium* subspecies *Tuberculosis* (Allen et al., 2010, Finlay et al., 2012, Bermingham et al., 2014, Tsairidou et al., 2014, Richardson et al., 2016, Tsairidou et al., 2016, Kiser et al., 2017), and *Mycobacterium avium* subspecies *Paratuberculosis* (Settles et al., 2009, Zanella et al., 2011, Alpay et al., 2014, Zare et al., 2014). Another important pathogen affecting calves is *Salmonella typhimurium*. Wray and Sojka (Wray and Sojka, 1978) reported some phenotypic variation in resistance to Salmonella between Jersey calves and Friesian calves, which may be due to the genetic variation among breeds. Templeton et al., (Templeton et al., 1990) noted that calves of sires with a high resistance to Brucellosis also show a high resistance to Salmonella, suggesting a genetic contribution to the resistance to this pathogen. In studies of the genetics of resistance to viral pathogens, resistance has mostly been measured indirectly as the symptoms of infection in animals, rather than directly measured as susceptibility to the pathogen itself. For instance, in 2008, Heringstad et al., (Heringstad et al., 2008) estimated the heritability of susceptibility to respiratory diseases to be 0.05±0.018. Estimation of genetic variation underlying the resistance/susceptibility for parasitic infestation has been well studied in small ruminants (Bishop and Morris, 2007), and to a small

extent in cattle (Stear and Murray, 1994). In Canadian Holstein, the heritability for susceptibility to *Neospora caninum* was in a range between 0.08±0.02 and 0.12±0.04 Pan et al. (2004). In Dutch Holstein-Friesian, the heritability of eggs/larvae count in animal feces was estimated to be from 0.00±0.02 to 0.25±0.05 (Coppieters et al., 2009). The resistance to ectoparasites in cattle was studied and reviewed by (Davis, 1993) who concluded that the heritability of resistance to ticks was 0.31, and 0.21 for resistance to Buffalo flies.

A number of studies have also detected quantitative trait loci (QTLs) associated with susceptibility/resistance to infection diseases in farm animals. Lee et al., (Lee et al., 2015) found 11 QTLs on three chromosomes (BTA15, BTA17, and BTA22), significantly associated with susceptibility to Foot-and-mouth disease in Holstein cattle. Casas et al., (Casas et al., 2015) searched for markers directly associated with the susceptibility to infection with Bovine viral diarrhea virus in feedlot cattle, and found a significant association on Chromosome BTA14.

Despite estimates of the heritability of resistance for some pathogens, genetic correlations (pleiotropy) among resistance between these domains of pathogens are still unknown, as is the level of pleiotropy among resistance to different species and subspecies within each domain of these pathogens. Several reasons lie behind the absence of the genetic correlations among resistance/susceptibility to various pathogens. First, the most popular method used in animal breeding for estimating genetic correlation has been the pedigree-base restricted maximum likelihood approach (Pedigree-REML). This method tests genetic overlap of traits between related individuals within pedigrees. To quantify the genetic correlation between traits in a family-based study, I may need to measure the traits in individuals with pedigree relationships (Bulik-Sullivan et al., 2015a). Consequently, it will be challenging and costly to repeat measurements on all animals, in particular for these low-prevalence traits and/or for traits where slaughter of the animals to measure their resistance is needed (i.e. endo-parasite infestation). Moreover, some disease traits (i.e. Bovine respiratory diseases) may result in death of the animal (at a young age) before it is possible to measure other traits (such as milk production after first calving) for which I want to test the correlation.

The genomics era has provided a solution to the problem of estimating genetic correlations, by allowing the genetic correlation to be estimated using genomic variants (i.e. SNPs) instead of using pedigree information, providing more precise and accurate estimates for the narrow-sense

heritability (h^2_{SNP}) in case-control studies and for quantitative traits as well as for the coheritability between such traits, which in particular does not need the measurement of multiple traits per animal. So far, several different methods have been developed for estimating the (co)heritability using genomic data: the first method uses the significantly associated SNPs found in large GWAS studies to estimate the causal relationships between risk factors and disease. This method is efficient only in case of the traits with many significant SNPs, which is usually not the case in resistance/susceptibility traits. For complex traits (especially for those recorded in case control studies), it is recommended to use genome-wide data (array or sequence data) instead of using only significant markers to estimate genetic correlation (Solovieff et al., 2013), and this approach has been implemented in two published methods; Genomic-restricted maximum likelihood (G-REML) (Yang et al., 2010, Chen et al., 2014) and polygenic scores (Purcell et al., 2009, Dudbridge, 2013). One of the limitations for the application of the last two methods in complex traits is the availability of individual-level genotype data. Hence, Bulik-Sullivan and colleagues (Bulik-Sullivan et al., 2015a, Bulik-Sullivan et al., 2015b) developed LD Score regression using the GWAS summary statistics instead of the individual-level genotype data for estimating heritabilities and genetic correlations.

In this study, I use individual-level genotypes for ~ 20000 animals from cow calibration groups (NB Cow calibration groups: is a programme initiated in Germany combining information for novel traits with high-density genetic markers based on ~20,000 genotyped cows, to offer a new perspective on breeding for improved disease resistance (Swalve, 2015)). I applied the G-REML method to estimate the heritability of the 23 resistance traits and the genetic correlations between these traits, and the correlations between the resistance traits with calf performance and cow productivity. I also implemented a post-GWAS functional analysis to estimate the pleiotropy based on different scales; scale 1) pleiotropy among all 23 resistance traits based on gene analysis; scale 2) pleiotropy among all 23 traits based on gene-set analysis (biological pathways), to understand the biological background of the underlying resistance/susceptibility to infectious diseases.

MATERIALS AND METHODS

Materials

The resistance to infectious disease traits in Holstein cattle was measured through the framework of the dairy cattle test-herd system of northeast Germany, including the federal states of Mecklenburg-Westpommerania and Berlin-Brandenburg. Dairy cattle farmers and veterinarians used electronic recording systems, which were based on the diagnosis key as developed by Feucker and Staufenbiel (Feucker and Staufenbiel, 2003). This diagnosis key was also considered when developing the International Committee for Animal Recording (ICAR) (Stock et al., 2013).

Phenotypes. In calves and cows, four main domains of pathogens were used to classify 23 resistance traits: (a) bacterial pathogens, (b) viral pathogens, (c) fungal pathogens, and (d) parasitic pathogens. I distinguished between calf resistance traits and cow resistance traits (Mahmoud et al., 2017). For the traits recorded in calves, I defined a time window from birth to the age of 150 days. For the traits recorded in cows, the window was from 20 days before first calving up to 365 days after first calving (a 385-day period). At least one entry for the respective pathogen implied a score = 0 for infected (non-resistant); otherwise, score = 1 for non-infected (resistant). The infected animal with a given pathogen was the animal (calf/cow) that was recorded as infected with this pathogen during its calf/cow age. A non-infected (resistant) calf to a given pathogen was defined as a calf that was found to be healthy on a farm infected with the given pathogen and that was born after the first record of this infection on that farm; a resistant cow to a given pathogen is defined as a cow that was found to be healthy in a farm where the first record of infection with the given pathogen was at least 20 days before its first calving date. Note that in common with other studies of natural infection in livestock and other species, I use resistance to define animals that did not become infected in a herd that was undergoing a specific disease challenge. For a given pathogen, all herds that showed no occurrence of infection were excluded because I do not know whether these herds were challenged with the particular pathogen.

All veterinary diagnosis and infection pathogen recording were done according to the ICAR (S8 Table, also available online through: www.icar.org). The nine pathogen resistance traits that were recorded in calves were Salmonella, Escherichia coli, Bovine respiratory syn., Bovine herpesvirus 1, Trichophyton, Cryptosporidium, Coccidia, Myiasis and Bovicola bovis. The 14

pathogen resistance traits that were recorded in cows were Salmonella, Escherichia coli, *Staphylococcus aureus* (Staph. Aureus), *Staphylococcus haemolyticus* (Staph. Haemolyticus), *Streptococcus agalactiae* (Strep. Agalactiae), *Streptococcus dysgalactiae* (Strep. Dysgalactiae), *Streptococcus uberis* (Strep. Uberis), *Clostridium perfringens* (Clost. Perfringens), *Mycobacterium avium paratuberculosis* (Mycobac. Paratuberculosis), Rotavirus, Trichophyton, Dictyocaulus viviparous, Bovicola bovis and Chorioptic scabies. For testing the genetic correlation between the resistance and performance in calves I considered two performance traits in calves: birthweight (in kg) and average daily gain (in g/day) during the first 360 days of calf life. For testing the genetic correlation between the resistance and productivity in cows, I considered two productivity traits in cows: average milk yield (in kg) during the first lactation and fat to protein ratio (in %) during the first lactation.

Genotypes. Genotyping was performed using the Illumina Bovine 50K SNP-BeadChip V2 (Illumina Inc., San Diego, CA), and with the Illumina Bovine Eurogenomics 10K low-density chip. The low-density genotypes (10K) were imputed by Vereinigte Information system Tierhalung (Verden, Germany) to the 50K panel applying the algorithm by Segelke et al. (Segelke et al., 2012). In post-imputation SNP quality checks; animals with almost identical SNP genotypes (>95% congruency across all SNPs) were eliminated from the ongoing analyses; SNPs with minor allele frequency <0.01 and SNPs showing a significant ($P < 10^{-5}$) deviation from Hardy-Weinberg equilibrium were discarded. All SNPs had a genotype call rate greater than 95%.

Genotyping was only undertaken for infected and resistant animals from populations where I could be sure that all animals genotyped have been challenged with the relevant pathogen (i.e. that they come from an infected population). Hence, the incidences from these genotyped samples do not reflect the actual incidences in the German Holstein population, but are likely to be higher as I have excluded data from herds where there is no evidence for a disease challenge for a particular pathogen. Further epidemic research, it is required to study the true incidence at the population scale. For a full statistical descriptive of the 23 resistance traits and the four quantitative traits in calves and cows in the genotyped sample, see Table 1.

Statistical models

Correction for population stratification in genome-wide data. The most common method for dealing with population stratification is principal component analysis (PCA) (Price et al., 2006, Price et al., 2010). Fitting the leading principal components in the model can correct for stratification for analyses such as estimating the proportion of variance explained by genome-wide SNPs and for genome-wide association studies (GWAS). Here, I applied principal components analysis to the genome-wide SNP data to infer continuous axes of genetic variation. Hence, the new axes will reduce the data dimensions (eigenvectors), describing as much variability as possible (eigenvalues): $V^{-1}(c^t c)V = D$, where $\boldsymbol{V}$ was the matrix of eigenvectors which diagonalizes the covariance matrix $\boldsymbol{c^t c}$ (covariance matrix of genotyped data), $\boldsymbol{D}$ was the diagonal matrix of eigenvalues of $\boldsymbol{c^t c}$. Then, I adjusted the phenotypes by including the first five eigenvectors as covariates in the model when estimating the proportion of variance explained by all the SNPs, or in G-REML and GWAS (see model-1).

A univariate mixed linear model was used to estimate the phenotypic variance explained by all autosomal SNPs (h^2_{SNPs}) by applying the genomic-restricted maximum likelihood analysis (G-REML), and using the GCTA software (Yang et al., 2011). In matrix notation, the model was defined as:

$$\mathrm{y} = \mathbf{X}\mathrm{b} + \mathrm{g}_{\mathbf{G}} + \varepsilon \quad (1)$$

where **y** refered to the vector of the quantitative trait (for performance and productivity traits) or of unobserved liabilities (for resistance traits); **b** was the vector of the fixed effects (herd, birth year and birth month for calf traits; herd, calving year, calving season, age at first calving for cow traits and the first 5 PCs), **X** was an incidence matrix for the fixed effects; $\boldsymbol{g_G}$ was the vector of aggregated effects of all autosomal SNPs with $var(\boldsymbol{g_G}) = A_G\sigma^2_G$ and A_G was the genomic relationship matrix (GRM). The heritability explained by all autosomal SNPs (h^2_{SNPs}) was defined as $h^2_{SNPs} = \sigma^2_{SNPs}/\sigma^2_p$ where σ^2_p was the phenotypic variance.

A bivariate model was used to estimate the genetic correlation among resistance traits and between resistance and performance or productivity traits (Yang et al., 2011).

$$\begin{bmatrix} \mathrm{y}_1 = \mathbf{X}_1\mathrm{b}_1 + \mathrm{g}_{\mathbf{G1}} + \mathrm{e}_1 \ \ (\text{trait } 1) \\ \mathrm{y}_2 = \mathbf{X}_2\mathrm{b}_2 + \mathrm{g}_{\mathbf{G2}} + \mathrm{e}_2 \ \ (\text{trait } 2) \end{bmatrix} \quad (2)$$

The two equations are the same as in model 1, while the (co)variance matrix was:

$$\text{var}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \sigma^2_{G_1}\mathbf{A} + \sigma^2_{e_1}\mathbf{I} & \sigma_{G_1G_2}\mathbf{A} + \sigma_{e_1e_2}\mathbf{I} \\ \sigma_{G_1G_2}\mathbf{A} + \sigma_{e_1e_2}\mathbf{I} & \sigma^2_{G_2}\mathbf{A} + \sigma^2_{e_2}\mathbf{I} \end{bmatrix} \quad (3)$$

where **A** is GRM, **I** is the identity matrix, σ^2_G is the genetic variance, σ^2_e is the residual variance and $\sigma_{G_1G_2}$ is the genetic covariance between the two traits. I assumed that all environmental correlations between pathogens were zero, as less than 10% of animals shared diagnoses for any pair of pathogens.

Significance thresholds. To determine the significance of the estimated pleiotropy among calf and cow resistance traits, two methods are commonly used to determine the significance threshold for genome-wide analysis: the false discovery rate and Bonferroni correction. The False discovery rate (FDR) correction was introduced by Benjamini and Hochberg (Benjamini and Hochberg, 1995). The FDR method first ranks all p-values from the smallest to the largest, and then adjusts each p-value accordingly:

$$\text{FDR corrected} - \text{p} = \frac{\text{Number of tests}}{\text{p} - \text{vlaue ranking}} * \text{p} - \text{value}$$

For the pleiotropy among resistance traits, I used the FDR of 1%. Bonferroni correction bases on the number of independent tests performed in each scenario (36 tests in calf traits and 105 tests in cow traits). The Bonferroni threshold was used at α=1%, and calculated as follows:

$$\text{Bonferroni threshold} = -\log10(\frac{\alpha}{\text{Number of tests}})$$

Biological pathway analysis. Biological pathway analysis is an approach where the association between a select set of genes (biological pathways) and a trait of interest (the resistance to different pathogens) was tested. This analysis can be used to test the cumulative genetic effects across multiple genes within a pathway.

Multi-marker Analysis of GenoMic Annotation (MAGMA): MAGMA used in my pathway analysis according to Leeuw et al., (Leeuw et al., 2015) with the following three steps: First, an annotation step to map SNPs onto genes using the bovine gene location (UMD3.1), from ensemble-biomaRt (www.ensembl.org/biomart). Second: a gene-based analysis step to compute gene p-values, using MAGMA and the output of GWAS from GCTA (pre-estimation of summary statistics of GWAS for each trait, performed using the GCTA software (Yang et al., 2011). Third:

a gene-set (Biological pathways) analysis step, to compute biological pathway p-values, using MAGMA (with the "*competitive gene-set analysis"* function) with publicly available BIOCARTA, KEGG and REACTOME database. All analyses in MAGMA are structured as a linear regression model on gene-level data.

$$Z = \beta_0 + \beta_1 \mathbf{G}_1 + e \quad \text{(Model – 4)}$$

Where $\boldsymbol{Z}$ was the phenotype vector, Gene-sets $\boldsymbol{G_1}$ were binary indicator variables, coded with "1" for genes in the gene-set, and with "0" otherwise. e was the residual vector. The intercept $\boldsymbol{\beta_0}$ represents the mean, and $\boldsymbol{\beta_1}$ the association specific to the gene-set 1. One last step was to illustrate the most significant gene rich network related to most of the resistance traits in calf and cows. This was done using the web-based software GeneMANIA (Leeuw et al., 2015), and then to illustrate the most significant related genes to all resistance traits (both in the same Figure) to see how these genes are related to this pathway. The post-GWAS functional analyses were performed using the MAGMA software (Leeuw et al., 2015) and the GWAS output from GCTA software (Yang et al., 2011).

RESULTS

SNP heritabilities.

Using a univariate G-REML model, I estimated the SNP-based heritabilities ($\mathbf{h^2_{SNP}}$) for all resistance, performance, and productivity traits (Table 1). SNP heritabilities of resistance to bacterial pathogens ranged from 0.03±0.01 to 0.21±0.01 in calves, and from 0.02±0.01 to 0.13±0.02 in cows. Higher SNP heritabilities were estimated for the resistance to viral pathogens, ranging from 0.16±0.03 to 0.22±0.03 in calves, and 0.10±0.01 in cows. For resistance to trichophyton (the only fungal pathogen) in calves, the heritability estimate was 0.17±0.03, and in cows it was 0.04±0.01. In parasitic infestation, the SNP heritability estimates in calves ranged from 0.04±0.01 to 0.14±0.03, while in cow's estimates varied from 0.19±0.02 to 0.25±0.02. For calf performance traits, the SNP heritability explained 0.30±0.01 of the phenotypic variance in birthweight, while 0.08±0.02 of the phenotypic variance in average daily gain was explained. For cow productivity traits, SNPs explained 0.19±0.02 of the variance in average milk yield, and 0.25±0.02 of the variance for the fat to protein ratio trait.

Table 1. Number of genotyped samples and estimated SNP-based heritabilities (h^2_{SNP}) for resistance and performance traits in calves; and resistance and productivity traits in cows.

Calf resistance traits	Infected	Resistant	Incidence	Total	h^2_{SNP}±s.e.
Bacterial pathogens					
Salmonella	271	570	0.32	841	0.21±0.01
Escherichia coli	19	483	0.04	502	0.03±0.01
Viral pathogens					
Bovine respiratory syn.	143	375	0.28	518	0.16±0.03
Bovine herpes virus 1	113	162	0.41	275	0.22±0.03
Fungal pathogen					
Trichophyton	421	875	0.32	1296	0.17±0.03
Parasitic pathogens					
Cryptosporidium	238	749	0.24	987	0.14±0.03
Coccidia	270	770	0.26	1040	0.11±0.03
Myiasis	362	107	0.77	469	0.13±0.02
Bovicola bovis	20	445	0.04	465	0.04±0.01
Calf performance traits	**Mean±s.d.**				
Birthweight (in kg)	41.09±4.83			17976	0.30±0.01
Average daily gain (in kg)	0.77±0.17			7673	0.08±0.02
Cow resistance traits	**Infected**	**Resistant**	**Incidence**		
Bacterial pathogens					
Salmonella	103	794	0.11	897	0.11±0.02
Escherichia coli	87	905	0.09	992	0.08±0.02
Staph. Aureus	102	805	0.11	907	0.12±0.01
Staph. Haemolyticus	379	802	0.32	1181	0.08±0.02
Strep. Agalactiae	51	1177	0.04	1228	0.11±0.06
Strep. Dysgalactiae	18	797	0.02	815	0.02±0.01
Strep. Uberis	101	798	0.11	899	0.13±0.02
Clost. Perfringens	21	238	0.08	259	0.12±0.06
Mycobac. Paratuberculosis	30	781	0.04	811	0.10±0.05
Viral pathogen					
Rotavirus	237	1253	0.16	1490	0.10±0.01
Fungal pathogen					
Trichophyton	11	1564	0.01	1575	0.04±0.01
Parasitic pathogens					
Dictyocaulus viviparus	671	652	0.51	1323	0.06±0.01
Bovicola bovis	25	929	0.03	954	0.06±0.02
Chorioptic scabies	50	940	0.05	990	0.10±0.02
Cow productivity traits	**Mean±s.d.**				
Average milk yield (in kg)	30.83±5.51			9959	0.19±0.02
Fat to protein ratio (in %)	1.24±0.17			9959	0.25±0.02

The pleiotropy among calf resistance traits. A bivariate analysis was used to estimate the level of pleiotropy among calf resistance traits using genome-wide SNPs for all 36 pairwise combinations of the nine resistance traits, shown in Figure 1 and S1 Table. The genetic correlation was significantly different from zero (according to FDR <1% and according to the Bonferroni threshold) between the resistance to Salmonella pathogen and Trichophyton pathogen (0.55±0.07), between Salmonella and Cryptosporidium (0.98±0.01), between Bovine respiratory syn. and Coccidia (0.46±0.06), between Bovine herpesvirus 1 and Bovicola bovis (0.74±0.18), between Cryptosporidium and Coccidia (0.52±0.06), and between Myiasis and Bovicola bovis (-0.54±0.06).

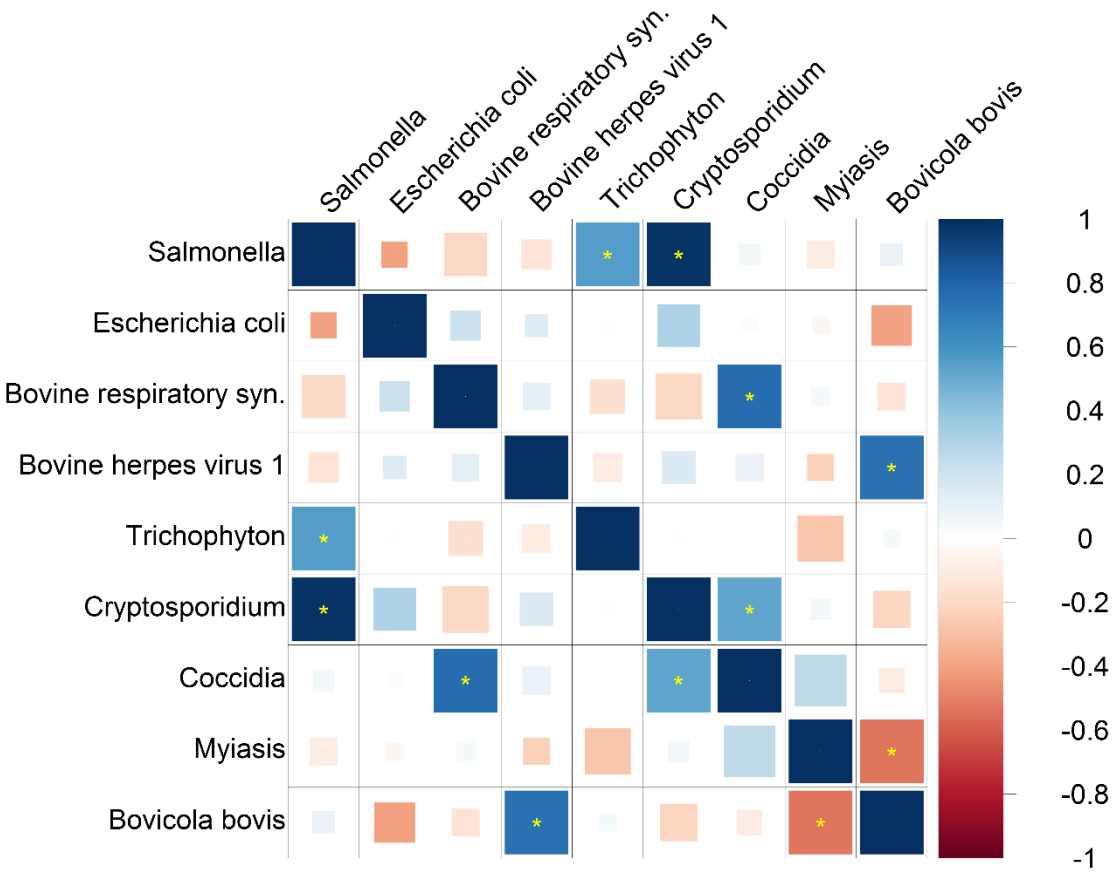


Figure 1. Genetic correlations among the 9 calf resistance traits analyzed by G-REML. Blue, positive genetic correlation; red, negative genetic correlation. Larger squares correspond to more significant P values. Genetic correlations that are different from zero at a false discovery rate (FDR) of 1% are shown as full-sized squares. Genetic correlations that are significantly different from zero after Bonferroni correction for the 36 tests in this analysis are marked with a yellow asterisk. I show results that do not pass multiple-testing correction as smaller squares. All genetic correlations in this report can be found in tabular form in S1 Table.

The pleiotropy among cow resistance traits. A second bivariate analysis was used to estimate the pleiotropy among cow resistance traits using genome-wide SNPs for all 105 pairwise combinations of the 14 resistance traits (all results are shown in Figure 2 and S2 Table). The estimated genetic correlation was significantly different from zero (according to FDR <1% and according to the Bonferroni threshold) between the resistance to Escherichia coli pathogen and Staph. Aureus pathogen (0.71±0.15), between Staph. Aureus and Strep. Uberis (0.56±0.10), between Staph. Haemolyticus and Strep. Dysgalactiae (1.00±0.34), between Staph. Haemolyticus and Chorioptic scabies (1.00±0.11), between Strep. Agalactiae and Strep. Dysgalactiae (1.00±0.28), between Strep. Agalactiae and Bovicola bovis (1.00±0.36), and between Rotavirus and Chorioptic scabies (-0.52±0.16). The estimated genetic correlation was significantly different from zero (with FDR <2%) between the resistance to Escherichia coli and Strep. Dysgalactiae (0.78±0.22), Strep. Agalactiae and Strep. Uberis (0.72±0.27), Strep. Agalactia and Trichophyton (0.96±0.33), and Rotavirus and Dictyocaulus viviparus (0.44±0.13).

The genetic correlation between calf and cow resistance traits was significantly different from zero for two different traits (S3 Table). The estimated genetic correlation between resistance to the Salmonella pathogen in calves and cows was -0.26±0.09 (p-value $\leq$ 0.001), and the genetic correlation between resistance to the Trichophyton pathogen in calves and cows was 0.18±0.12 (p-value $\leq$ 0.05).

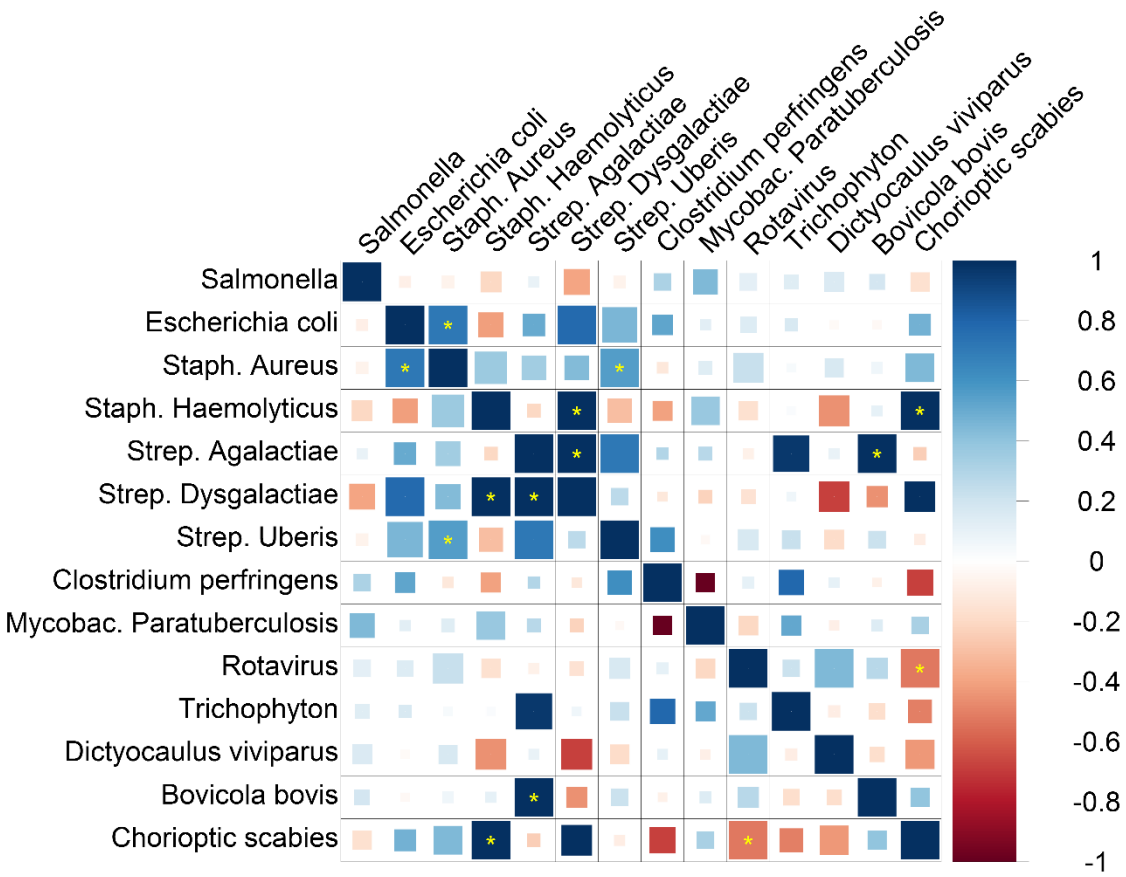


Figure 2. Genetic correlations among the 14 cow resistance traits analyzed by G-REML. Blue, positive genetic correlation; red, negative genetic correlation. Larger squares correspond to more significant P values. Genetic correlations that are different from zero at a false discovery rate (FDR) of 1% are shown as full-sized squares. Genetic correlations that are significantly different from zero after Bonferroni correction for the 105 tests in this analysis are marked with a yellow asterisk. I show results that do not pass multiple-testing correction as smaller squares. All genetic correlations in this report can be found in tabular form in S2 Table.

Pathogen resistance and calf performance. The genetic correlations of birthweight with the resistance to viral, fungal, and parasitic (except for Cryptosporidium) pathogens were highly significantly negative for resistance to Bovine herpes virus 1 ($P \leq 0.01$), and also significantly negative ($P \leq 0.05$) for resistance to the Bovine respiratory syn., Trichophyton, Coccidia and Bovicola bovis pathogens. There were no significant positive correlations between birthweight and resistance, although that with Myiasis approached significance ($P \leq 0.1$). The genetic correlation of average growth rate was very highly significantly ($P \leq 0.001$) negative with resistance to Salmonella, highly significantly negative with resistance to the Escherichia coli and Bovine respiratory syn. Pathogens, and negative and approaching significance ($P \leq 0.1$) with resistance to Bovine herpes virus 1 and Trichophyton pathogens. Positive correlations with average daily gain were only significant for Cryptosporidium ($P \leq 0.01$) and approaching significance for Coccidia ($P \leq 0.1$) (Figure 3). All correlations in calf performance traits had standard errors that ranged from 0.05 to 0.14 (S4 Table).

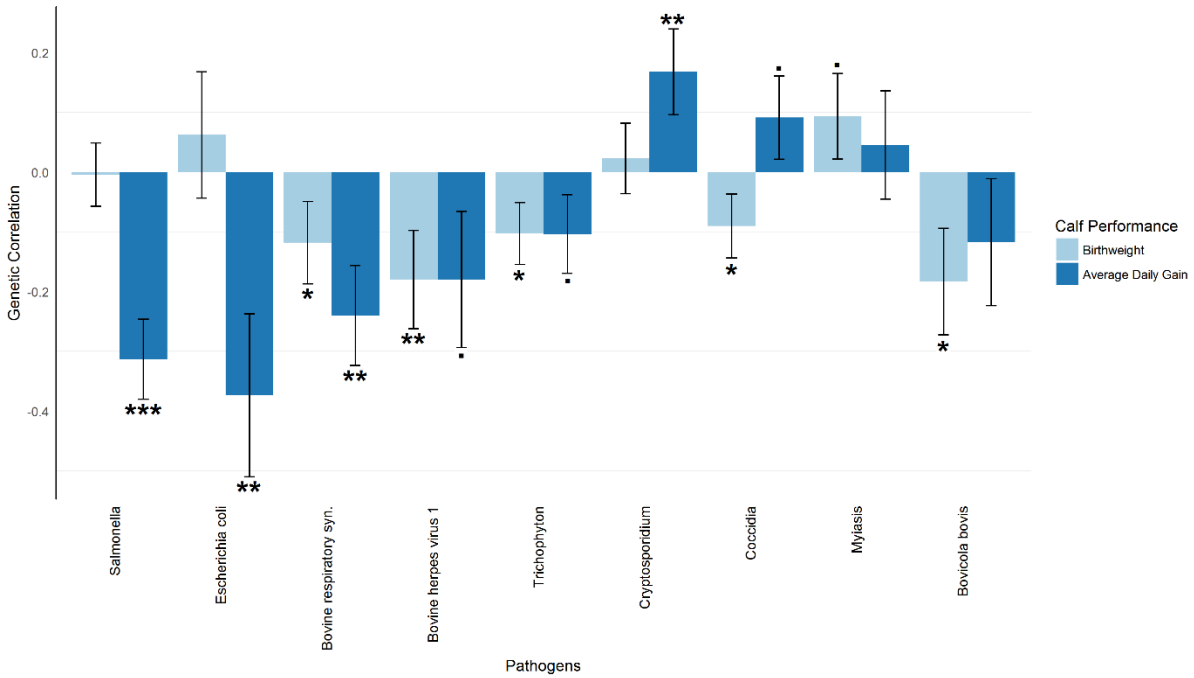


Figure 3. Estimated genetic correlations of birthweight (BW) and average daily gain (ADG) with all resistance traits in calves. This plot compares the genetic correlation between BW and all calf resistance traits with the genetic correlation between ADG and all calf resistance traits obtained from G-REML. The horizontal axis indicates pairs of phenotypes (BW and ADG), and the vertical axis indicates genetic correlation. Error bars represent standard errors. '***'$P \leq 0.001$; '**'$P \leq 0.01$; '*'$P \leq 0.05$; '·'$P \leq 0.1$.

Pathogen resistance and cow productivity. (Figure 4 and S5 Table). The resistance to bacterial pathogens were generally positively correlated with milk yield, the correlation being very highly significance positive ($P \leq 0.001$) with Staph. Aureus, highly significantly positive ($P \leq 0.01$) with Strep. Uberis and approaching significantly positive ($P \leq 0.1$) with Staph. Haemolyticus and Strep. Agalactiae. Milk yield was also significantly positively correlated with resistance to Rotavirus. Resistance to Trichophyton was approaching significance for a negative correlation with milk yield. Fat to protein ratio was positively correlated with resistance Staph. Aureus ($P \leq 0.01$), Rotavirus and Trichophyton ($P \leq 0.05$). Fat to protein ratio was significantly negatively correlated with Salmonella, Mycobac. Paratuberculosis and Chorioptic scabies ($P \leq 0.05$), and approaching significance for Staph. Haemolyticus ($P \leq 0.1$).

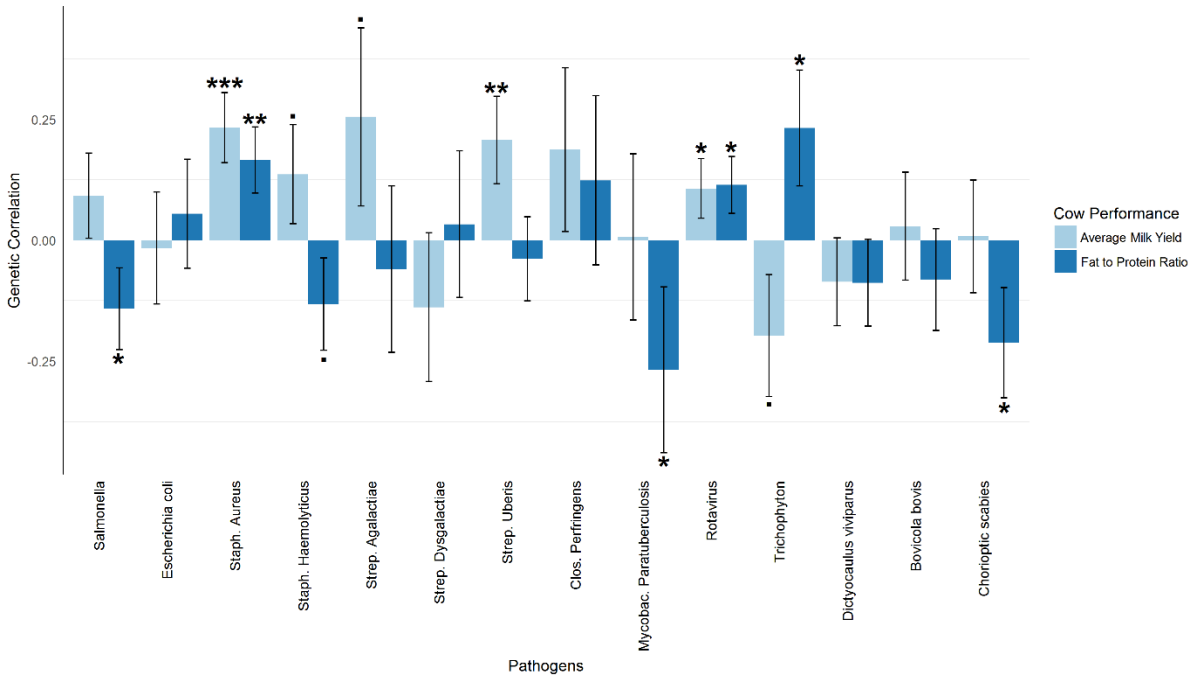


Figure 4. Estimated genetic correlations of average milk yield (AMY) and fat to protein ratio (FTP) with all resistance traits in cows. This plot compares the genetic correlation between AMY and all calf resistance traits with the genetic correlation between FTP and all calf resistance traits obtained from G-REML. The horizontal axis indicates pairs of phenotypes (AMY and FTP), and the vertical axis indicates genetic correlation. Error bars represent standard errors. *'***'P ≤ 0.001; '**'P ≤ 0.01; '*'P ≤ 0.05; '·'P ≤ 0.1.*

Biological pathway analysis. All SNPs (~50K SNPs) were annotated to the nearest gene where possible using a gene boundary extended by 20kb distance outside the transcription start site or transcription end site of the gene for all the 23 resistance traits, using MAGMA (version 1.06) and bovine gene location (UMD3.1) through ensemble-biomaRt (www.ensembl.org/biomart). This annotation pipeline resulted in a total of 16,094 genes ready for the next step of the analysis.

After testing the association of the 16,094 genes across all the 23 resistance traits, using p-values of summary statistics of GWAS from GCTA, I selected the top (most significant) 20 genes based on average p-values. The lowest average p-value for the correlation between the genes and all resistance traits was the average p-value for the RRM2B gene (average p-value = 0.28) (Figure 5, S6 and S7 Table). After estimating the p-values on the scale of genesets (Biological pathways), I selected the top 20 genesets based on average p-values (Figure 6, S8 and S9 Tables) out of 1083 pathways that were available from BIOCARTA, KEGG and REACTOME databases. I found that "*Reactome pre-notch transcription and translation*" and "*Biocarta B-lymphocyte pathway*" were the most associated pathways to all resistance traits according to their calculated p-values (the average p-values = 0.07 and 0.11 respectively) (Figure 6, S8 and S9 Tables). The final step was to illustrate the "*Biocarta B-lymphocyte pathway*" against the 19 genes that were associated with resistance at the gene-level. Figure 7 shows how the candidate genes (from gene analysis) were directly and indirectly connected to the "*Biocarta B-lymphocyte pathway*."

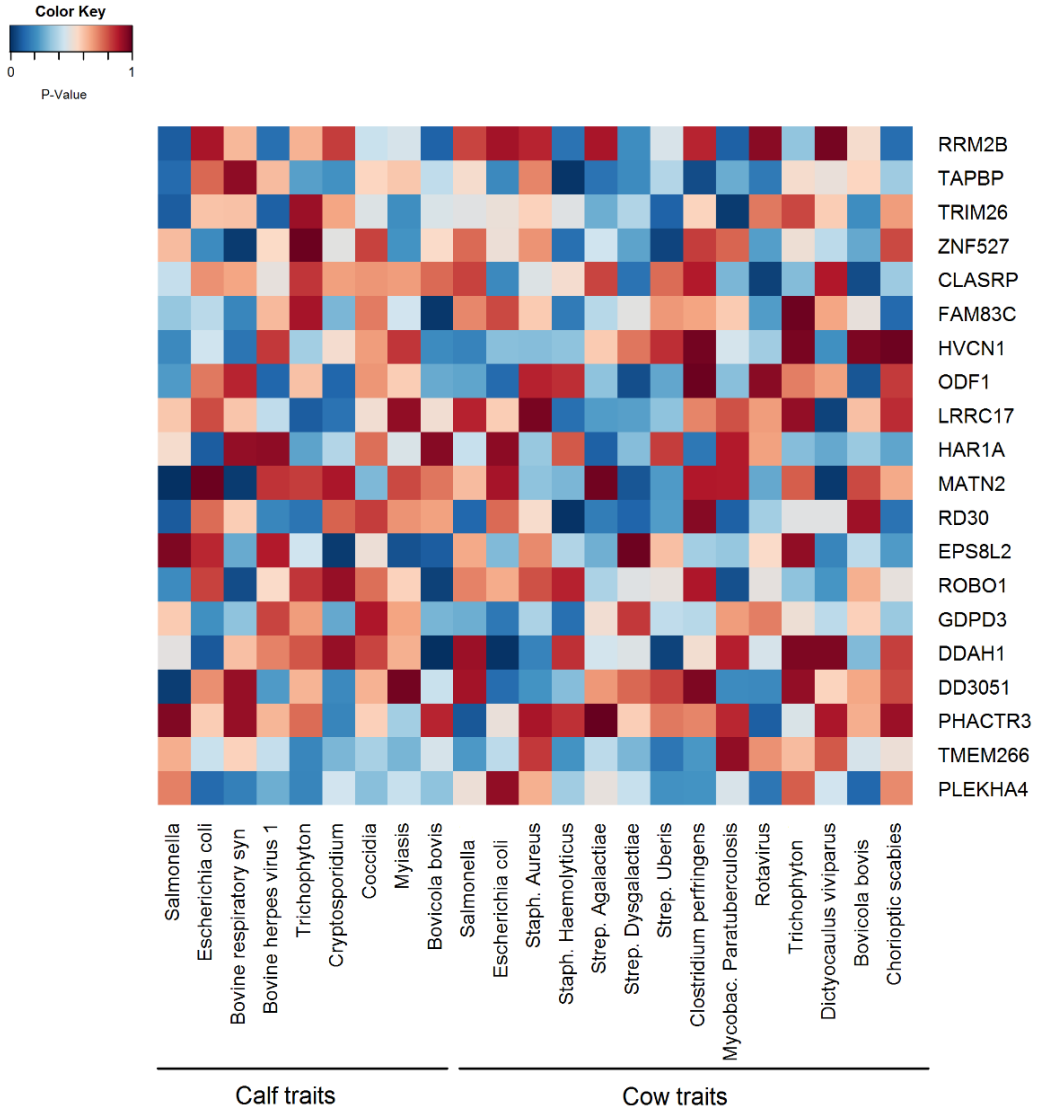


Figure 5. P-values for the selected annotated candidate genes in all resistance traits in calves and cows. Dark red color means very high p-value, dark blue color means very low (i.e. more significant) p-value.

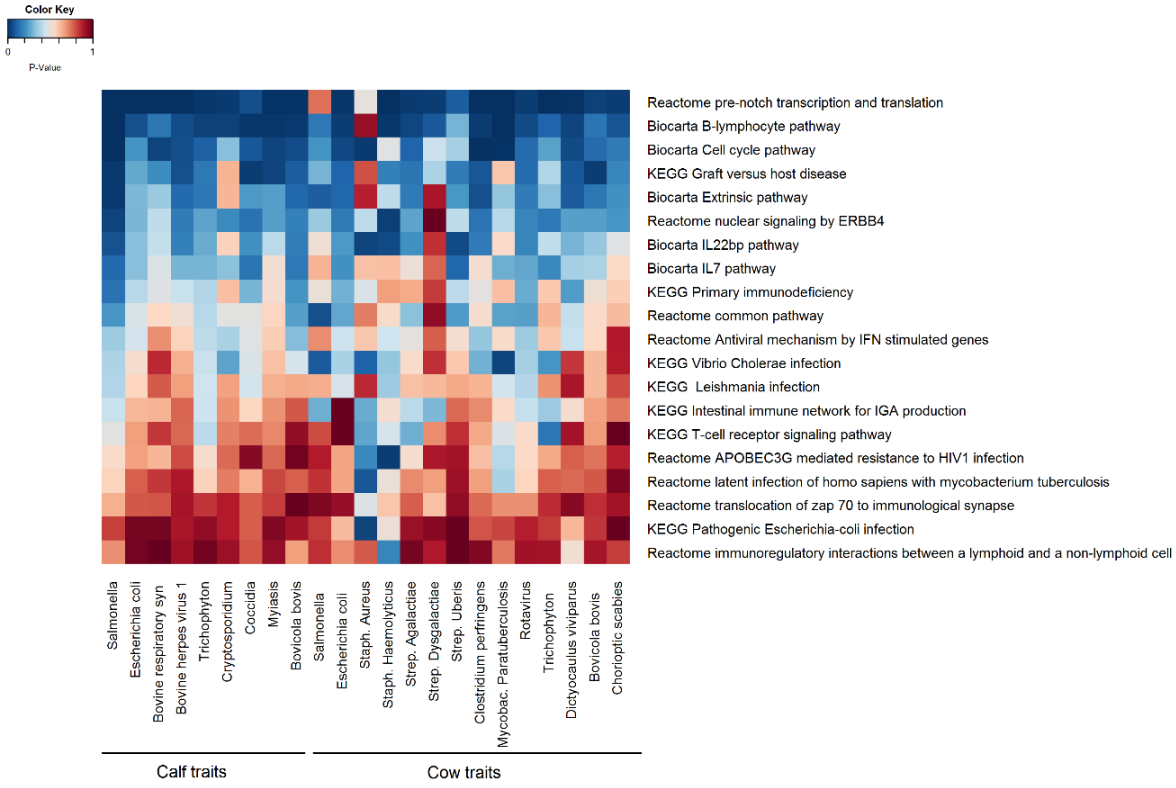


Figure 6. P-values for the selected 20 pathways tested across all resistance traits in calves and cows. Dark red color means very high p-value, dark blue color means very low (i.e. more significant) p-value.

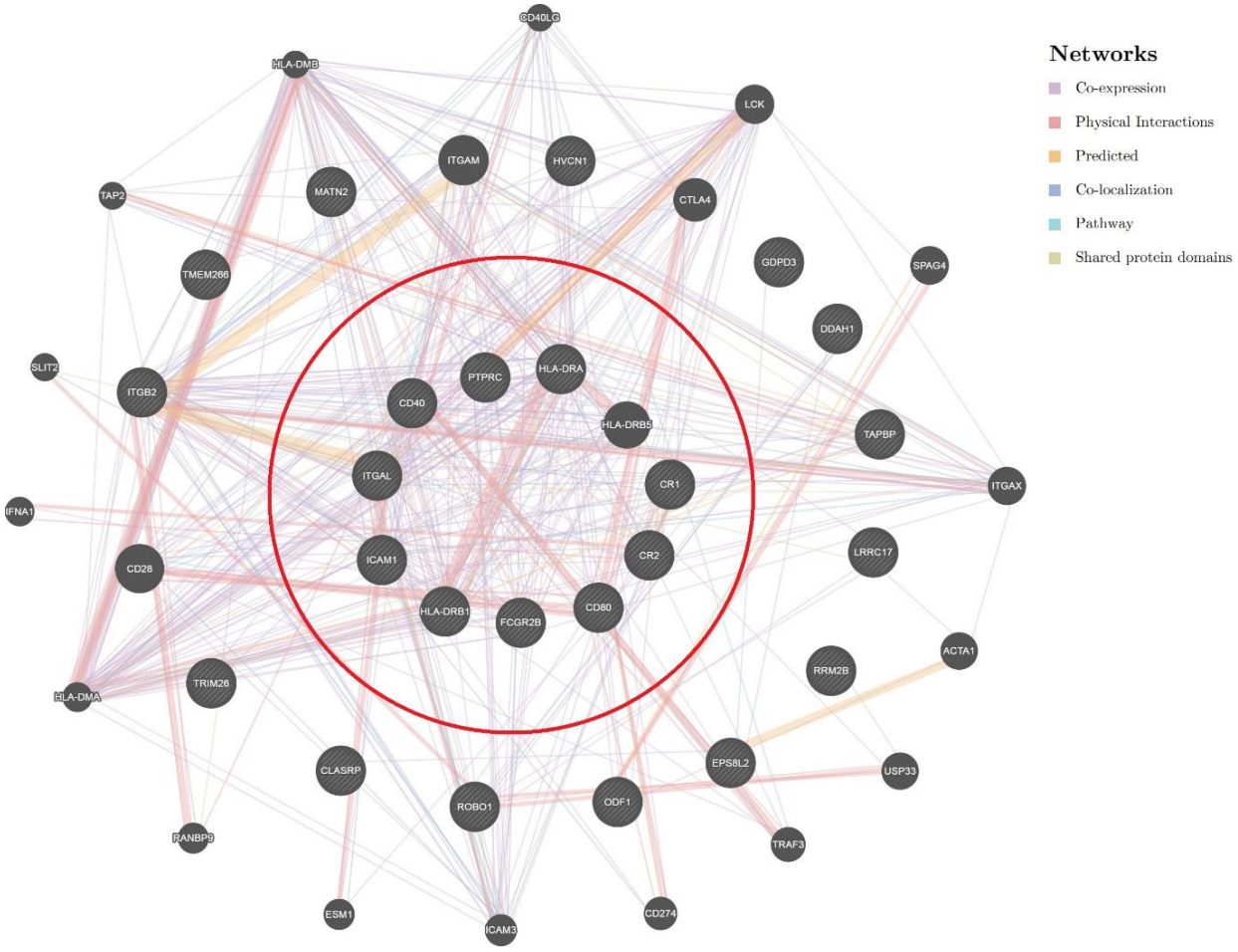


Figure 7. Network of gene-interactions the candidate genes (outside the red circle) and the "*Biocarta B-lymphocyte pathway*" (inside of the red circle).

DISCUSSION

We implemented G-REML using genome-wide SNP data to explore the genetic etiology of resistance to pathogens in cattle. This study thus provides the first genomic overview of the genetics of resistance across a range of pathogens together with the genetic correlations between pathogens and with production traits for Holstein cattle, the world's predominant dairy breed. I estimated the phenotypic variance explained by all SNPs (h^2_{SNPs}) for resistance, performance and productivity traits, and calculated the genetic correlation (r_g) between them, using genome-wide SNPs in calves and cows. The average heritability of resistance traits at a young age (calf traits) was larger than the average heritability of the same population at a later age (cow traits), reflecting the increase of the environmental effect potentially including increased pathogen exposure after first calving in cattle. The SNP heritabilities estimated using GREML in this study were of a similar scale to those for which estimates for individual pathogens have been previously published using pedigree and genome-wide data (Berry et al., 2011, Raszek et al., 2016).

Pleiotropy between resistance to pathogens and calf performance and cow productivity. There have only been very limited previous estimates of genetic correlations among resistance to different pathogens with which my estimates can be compared (e.g. (Passafaro et al., 2015)). I thus report a number of new findings that would be difficult to obtain from pedigree studies, including some very high estimates of positive pleiotropy between the resistance to bacterial, fungal and parasite pathogens. Negative pleiotropy was found between the resistance to most of the bacterial, viral and fungal pathogens with both performance traits that I analysed in calves (birthweight and average daily gain). Results supports the hypothesis that some resistance genes may negatively impact performance traits in young calves, perhaps reflecting a balance between energy expenditure on disease resistance and growth. However, clear positive pleiotropy found between resistance to bacterial, viral and fungal pathogens with average milk yield in cows supports the hypothesis that cows that are genetically less vulnerable to infections can produce more milk. This provides economic in addition to welfare justification for increasing focus of breeding objectives on these disease resistance traits.

Shared biological pathways. To date, few studies have implemented biological pathway analysis in animal health (Thompson-Crispi et al., 2014, Hamzić et al., 2015). This is the first study post-GWAS for resistance to most of all infection pathogens in dairy cattle. I found a group of 20 genes shared effects across all resistance traits and showed that "*Reactome pre-notch transcription and translation*" and "*Biocarta B-lymphocyte pathway*" are the most consistently associated pathways with resistance to different pathogens. The genetic correlations that I observed in this study show a clear pleiotropy (by the means of similar resistance mechanism against the 23 pathogens). The combination of accurate recording of multiple diseases and associated pathogens combined with associated genomic data is unusual and possibly so far unique in a mammalian species including humans. However, given the abundance of genomic data in humans combined with GWAS for a number of individual infectious diseases.

The scale of this dataset facilitated by the thorough electronic data collection within the dairy cattle test-herd system in north-eastern Germany has enabled us to obtain unique insight into the genetics of resistance across a range of pathogens. The corollary of these currently unique data is that the overall multivariate pattern will need to await further data collection from this or other similar programs before it is possible to replicate my results as a whole. Nonetheless, all significant correlations have a low standard error (S3, S4 and S5 Tables), and my estimates of heritabilities and pairwise genetic correlations are generally consistent with others in the literature where these are available (e.g. (Berry et al., 2011, Passafaro et al., 2015, Raszek et al., 2016, Mahmoud et al., 2017)).

The biological impact of the genetic pleiotropy on breeding and selection strategies. The estimated heritabilities and the pattern of genetic correlations between pathogens and with production traits provide valuable information allowing the further optimisation of cattle breeding programmes. For example, several low and non-significant genetic correlations were found among calf resistance traits as well as among cow resistance traits, questioning the traditional hypothesis of selection for mastitis resistance based on somatic cell count as a consequence of multiple pathogen infection in cattle (Heringstad et al., 2000) or in sheep (Rupp et al., 2009). Additionally, the highly significant negative correlation between resistance to the Salmonella pathogen in calves and cows implies that resistance traits in calves are not good indicators or early predictors for the resistance traits and genetic improvement of cow health after calving in first parity, supporting a

previous pedigree and GWA study (Mahmoud et al., 2017). The estimated zero genetic correlation between resistance to the Trichophyton pathogen in calves and cows indicates that, for this pathogen, calf and cow resistance, are genetically distinct. Thus, overall it is clear that to design an efficient breeding program, I need to take into account these results and utilise a programme combining selection in both cows and calves and consider how to most effectively collect and incorporate information on resistance and susceptibility to multiple pathogens. Fortunately, the increasing availability of genomic data in cattle combined with collection of data such as analysed here ultimately will facilitate genomic selection programmes that meet these objectives.

In summary, despite some limitations of available health traits and genotyped data, the use of G-REML method to estimate the genetic correlation among health traits and between health trait and performance and productivity traits promises to be a very valuable tool in the genetic improvement of animal health. Biological pathway analysis appears to be a very useful tool also, but at present I have had to use information for other species as have no biological pathways specifically tested and verified for cattle and the development of such databases would provide an invaluable resource for future research.

ACKNOWLEDGMENTS

We acknowledge the invaluable contributions of Pau Navarro, Carmen Amador and Charley Xia at the MRC Human Genetics Unit, University of Edinburgh and thank Andrea Wilson of The Roslin Institute, University of Edinburgh, for additional valuable comments on the manuscript. This research funded by the German Federal Ministry of Education and Research (BMBF), the Förderverein Bioökonomieforschung e.V. (FBV) / German Holstein Association (DHV) and through the collaborative project "KMU-innovativ-10: Kuh-L – cow calibration groups for the implementation of selection strategies based on high-density genotyping in dairy cattle" (grant no. 031A416C)

REFERENCES

Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. 2002. Introduction to Pathogens.

Allen, A. R., G. Minozzi, E. J. Glass, R. A. Skuce, S. W. J. McDowell, J. A. Woolliams, and S. C. Bishop. 2010. Bovine tuberculosis: the genetic basis of host susceptibility. Proceedings of the Royal Society B: Biological Sciences 277(1695):2737-2745.

Alpay, F., Y. Zare, M. H. Kamalludin, X. Huang, X. Shi, G. E. Shook, M. T. Collins, and B. W. Kirkpatrick. 2014. Genome-Wide Association Study of Susceptibility to Infection by Mycobacterium avium Subspecies paratuberculosis in Holstein Cattle. PLoS One 9(12):e111704.

Benjamini, Y. and Y. Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) 57(1):289-300.

Bermingham, M. L., S. C. Bishop, J. A. Woolliams, R. Pong-Wong, A. R. Allen, S. H. McBride, J. J. Ryder, D. M. Wright, R. A. Skuce, S. W. McDowell, and E. J. Glass. 2014. Genome-wide association study identifies novel loci associated with resistance to bovine tuberculosis. Heredity 112(5):543-551.

Berry, D. P., M. L. Bermingham, M. Good, and S. J. More. 2011. Genetics of animal health and disease in cattle. Ir Vet J 64(1):5.

Bishop, S. C. and C. A. Morris. 2007. Genetics of disease resistance in sheep and goats. Small Ruminant Res 70:48-59.

Bishop, S. C. and J. A. Woolliams. 2014. Genomics and disease resistance studies in livestock. Livest Sci 166(Supplement C):190-198.

Bulik-Sullivan, B., H. K. Finucane, V. Anttila, A. Gusev, F. R. Day, P.-R. Loh, C. ReproGen, C. Psychiatric Genomics, C. Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control, L. Duncan, J. R. B. Perry, N. Patterson, E. B. Robinson, M. J. Daly, A. L. Price, and B. M. Neale. 2015a. An atlas of genetic correlations across human diseases and traits. Nat Genet 47(11):1236-1241.

Bulik-Sullivan, B. K., P.-R. Loh, H. K. Finucane, S. Ripke, J. Yang, C. Schizophrenia Working Group of the Psychiatric Genomics, N. Patterson, M. J. Daly, A. L. Price, and B. M. Neale.

2015b. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 47(3):291-295.

Casas, E., B. E. Hessman, J. W. Keele, and J. F. Ridpath. 2015. A genome-wide association study for the incidence of persistent bovine viral diarrhea virus infection in cattle. Anim. Genet. 46(1):8-15.

Chen, G.-B., S. H. Lee, M.-J. A. Brion, G. W. Montgomery, N. R. Wray, G. L. Radford-Smith, P. M. Visscher, and I. B. D. G. C. International. 2014. Estimation and partitioning of (co)heritability of inflammatory bowel disease from GWAS and immunochip data. Hum. Mol. Genet. 23(17):4710-4720.

Coppieters, W., T. H. M. Mes, T. Druet, F. Farnir, N. Tamma, C. Schrooten, A. W. C. A. Cornelissen, M. Georges, and H. W. Ploeger. 2009. Mapping QTL influencing gastrointestinal nematode burden in Dutch Holstein-Friesian dairy cattle. BMC Genomics 10:96.

Davis, G. P. 1993. Genetic parameters for tropical beef cattle in northern Australia: a review. Aust. J. Agric. Res. 44(2):179-198.

Dudbridge, F. 2013. Power and Predictive Accuracy of Polygenic Risk Scores. PLoS Genetics 9(3):e1003348.

Feucker, W. and R. Staufenbiel. 2003. Zentraler Diagnoseschlüssel. http://www.portal-rind.de/index.php?module=Downloads&func=prep_hand_out&lid=17. Access 16.03.2010.

Finlay, E. K., D. P. Berry, B. Wickham, E. P. Gormley, and D. G. Bradley. 2012. A Genome Wide Association Scan of Bovine Tuberculosis Susceptibility in Holstein-Friesian Dairy Cattle. PLoS One 7(2):e30545.

Frodsham, A. J. and A. V. S. Hill. 2004. Genetics of infectious diseases. Hum. Mol. Genet. 13(suppl_2):R187-R194.

Garrity, G., D. R. Boone, and R. W. Castenholz. 2001. Bergey's Manual of Systematic Bacteriology: Volume One : The Archaea and the Deeply Branching and Phototrophic Bacteria: Archaea and the Deeply ... of Systematic Bacteriology. 2nd ed. 2001 edition ed. Springer, Edinburgh ; New York.

Gernand, E., P. Rehbein, U. U. von Borstel, and S. König. 2012. Incidences of and genetic parameters for mastitis, claw disorders, and common health traits recorded in dairy cattle contract herds. J. Dairy Sci. 95(4):2144-2156.

Hamzić, E., B. Buitenhuis, F. Hérault, R. Hawken, M. S. Abrahamsen, B. Servin, J.-M. Elsen, M.-H. Pinard - van der Laan, and B. Bed'Hom. 2015. Genome-wide association study and biological pathway analysis of the Eimeria maxima response in broilers. Genetics Selection Evolution 47:91.

Heringstad, B., Y. M. Chang, D. Gianola, and O. Østerås. 2008. Short Communication: Genetic Analysis of Respiratory Disease in Norwegian Red Calves. J. Dairy Sci. 91(1):367-370.

Heringstad, B., G. Klemetsdal, and J. Ruane. 2000. Selection for mastitis resistance in dairy cattle: a review with focus on the situation in the Nordic countries. Livestock Production Science 64(2):95-106.

Kiser, J. N., S. N. White, K. A. Johnson, J. L. Hoff, J. F. Taylor, and H. L. Neibergs. 2017. Identification of loci associated with susceptibility to Mycobacterium avium subspecies paratuberculosis (Map) tissue infection in cattle. Journal of Animal Science.

Lee, B.-Y., K.-N. Lee, T. Lee, J.-H. Park, S.-M. Kim, H.-S. Lee, D.-S. Chung, H.-S. Shim, H.-K. Lee, and H. Kim. 2015. Bovine Genome-wide Association Study for Genetic Elements to Resist the Infection of Foot-and-mouth Disease in the Field. Asian-australas. J. Anim. Sci. 28(2):166-170.

Leeuw, C. A. d., J. M. Mooij, T. Heskes, and D. Posthuma. 2015. MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLOS Computational Biology 11(4):e1004219.

Mahmoud, M., T. Yin, K. Brügemann, and S. König. 2017. Phenotypic, genetic, and single nucleotide polymorphism marker associations between calf diseases and subsequent performance and disease occurrences of first-lactation German Holstein cows. J. Dairy Sci. 100(3):2017-2031.

Pan, Y., G. B. Jansen, T. F. Duffield, S. Hietala, D. Kelton, C. Y. Lin, and A. S. Peregrine. 2004. Genetic susceptibility to Neospora caninum infection in Holstein cattle in Ontario. J. Dairy Sci. 87(11):3967-3975.

Passafaro, T. L., J. P. B. Carrera, L. L. d. Santos, F. S. S. Raidan, D. C. C. d. Santos, E. P. Cardoso, R. C. Leite, and F. L. B. Toral. 2015. Genetic analysis of resistance to ticks, gastrointestinal nematodes and Eimeria spp. in Nellore cattle. Veterinary Parasitology 210(3-4):224-234.

Price, A. L., N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, and D. Reich. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38(8):904-909.

Price, A. L., N. A. Zaitlen, D. Reich, and N. Patterson. 2010. New approaches to population stratification in genome-wide association studies. Nature Reviews Genetics 11(7):459-463.

Purcell, S. M., N. R. Wray, J. L. Stone, P. M. Visscher, M. C. O'Donovan, P. F. Sullivan, P. Sklar, D. M. Ruderfer, A. McQuillin, D. W. Morris, C. T. O'Dushlaine, A. Corvin, P. A. Holmans, M. C. O'Donovan, S. Macgregor, H. Gurling, D. H. R. Blackwood, N. J. Craddock, M. Gill, C. M. Hultman, G. K. Kirov, P. Lichtenstein, W. J. Muir, M. J. Owen, C. N. Pato, E. M. Scolnick, D. S. Clair, N. M. Williams, L. Georgieva, I. Nikolov, N. Norton, H. Williams, D. Toncheva, V. Milanova, E. F. Thelander, P. Sullivan, C. T. O'Dushlaine, E. Kenny, E. M. Quinn, K. Choudhury, S. Datta, J. Pimm, S. Thirumalai, V. Puri, R. Krasucki, J. Lawrence, D. Quested, N. Bass, C. Crombie, G. Fraser, S. L. Kuan, N. Walker, K. A. McGhee, B. Pickard, P. Malloy, A. W. Maclean, M. V. Beck, M. T. Pato, H. Medeiros, F. Middleton, C. Carvalho, C. Morley, A. Fanous, D. Conti, J. A. Knowles, C. P. Ferreira, A. Macedo, M. H. Azevedo, A. N. Kirby, M. A. R. Ferreira, M. J. Daly, K. Chambert, F. Kuruvilla, S. B. Gabriel, K. Ardlie, and J. L. Moran. 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460(7256):748-752.

Raszek, M. M., L. L. Guan, and G. S. Plastow. 2016. Use of Genomic Tools to Improve Cattle Health in the Context of Infectious Diseases. Frontiers in Genetics 7.

Richardson, I. W., D. P. Berry, H. L. Wiencko, I. M. Higgins, S. J. More, J. McClure, D. J. Lynn, and D. G. Bradley. 2016. A genome-wide association study for genetic susceptibility to Mycobacterium bovis infection in dairy cattle identifies a susceptibility QTL on chromosome 23. Genetics Selection Evolution 48:19.

Rupp, R., D. Bergonier, S. Dion, M. C. Hygonenq, M. R. Aurel, C. Robert-Granié, and G. Foucras. 2009. Response to somatic cell count-based selection for mastitis resistance in a divergent selection experiment in sheep. J. Dairy Sci. 92(3):1203-1219.

Segelke, D., J. Chen, Z. Liu, F. Reinhardt, G. Thaller, and R. Reents. 2012. Reliability of genomic prediction for German Holsteins using imputed genotypes from low-density chips. J Dairy Sci 95(9):5403-5411.

Settles, M., R. Zanella, S. D. McKay, R. D. Schnabel, J. F. Taylor, R. Whitlock, Y. Schukken, J. S. Van Kessel, J. M. Smith, and H. Neibergs. 2009. A whole genome association analysis identifies loci associated with Mycobacterium avium subsp. paratuberculosis infection status in US holstein cattle. Anim. Genet. 40(5):655-662.

Solovieff, N., C. Cotsapas, P. H. Lee, S. M. Purcell, and J. W. Smoller. 2013. Pleiotropy in complex traits: challenges and strategies. Nature Reviews Genetics 14(7):483-495.

Stear, M. J. and M. Murray. 1994. Genetic resistance to parasitic disease: particularly of resistance in ruminants to gastrointestinal nematodes. Veterinary Parasitology 54:161-176.

Stock, K. F., J. Cole, J. Pryce, N. Gengler, A. Bradley, L. Andrews, B. Heringstad, and C. Egger-Danner. 2013. Standardization of health data–ICAR guidelines including health key. ICAR Technical Series no. 17:75.

Swalve, H. H. 2015. German Projects for Genomic Based Breeding Programs in Holstein. in Interbull Technical Workshop and Industry Meeting. Walsrode and Verden/Aller, Germany.

Templeton, J. W., D. M. Estes, R. E. Price, R. Smith, III, and L. G. Adams. 1990. Immunogenetics of natural resistance to bovine brucellosis. Pages 396-399 in Proc. Proceedings of the 4th World Congress on Genetics applied to Livestock Production, Edinburgh 23-27 July 1990. XVI. Poultry, fish and horse genetics and breeding, growth and reproduction, immune response and disease resistance.

Thompson-Crispi, K. A., M. Sargolzaei, R. Ventura, M. Abo-Ismail, F. Miglior, F. Schenkel, and B. A. Mallard. 2014. A genome-wide association study of immune response traits in Canadian Holstein cattle. BMC genomics 15(1):559.

Tsairidou, S., S. Brotherstone, M. Coffey, S. C. Bishop, and J. A. Woolliams. 2016. Quantitative genetic analysis of the bTB diagnostic single intradermal comparative cervical test (SICCT). Genet Sel Evol 48.

Tsairidou, S., J. A. Woolliams, A. R. Allen, R. A. Skuce, S. H. McBride, D. M. Wright, M. L. Bermingham, R. Pong-Wong, O. Matika, S. W. J. McDowell, E. J. Glass, and S. C. Bishop. 2014. Genomic Prediction for Tuberculosis Resistance in Dairy Cattle. PLoS One 9(5):e96728.

Wray, C. and W. J. Sojka. 1978. Experimental Salmonella typhimurium infection in calves. Res. Vet. Sci. 25(2):139-143.

Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, M. E. Goddard, and P. M. Visscher. 2010. Common SNPs explain a large proportion of the heritability for human height. Nat Genet 42(7):565-569.

Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher. 2011. GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 88(1):76-82.

Zanella, R., M. L. Settles, S. D. McKay, R. Schnabel, J. Taylor, R. H. Whitlock, Y. Schukken, J. S. Van Kessel, J. M. Smith, and H. L. Neibergs. 2011. Identification of loci associated with tolerance to Johne's disease in Holstein cattle. Anim. Genet. 42(1):28-38.

Zare, Y., G. E. Shook, M. T. Collins, and B. W. Kirkpatrick. 2014. Genome-Wide Association Analysis and Genomic Prediction of Mycobacterium avium Subspecies paratuberculosis Infection in US Jersey Cattle. PLoS One 9(2):e88380.

CHAPTER 5

General Discussion

The main purpose of this thesis was to evaluate the possibility of controlling infectious diseases through the improvement of host genetic resistance using both traditional and modern methods in phenotypic, genetic, and genomic evaluation of infectious disease traits. The results have already been discussed and compared with other studies in each chapter. A statistical description of infectious diseases in Northeast Germany is given in CHAPTER 1, and temporal and age distributions are given in APPENDICES - A and – B, respectively. However, the knowledge gained from the individual chapters with respect to all the evaluations will be discussed in the following comprehensive points.

Until recently, hereditary resistance to disease was not a matter of great concern in the fields of animal production and animal health, veterinary activities too, were centered on specific immunization (vaccination) against infectious diseases rather than on the reinforcing of genetic resistance. While vaccination will remain an important method of controlling infectious disease in large production units, the protection of the animals against various non- infectious pathogenic agents still presents a largely unresolved problem. In CHAPTER 2 I discussed that some breeds or lines possess genetic resistance to a great many diseases, but this advantage has not yet been appropriately exploited in applied livestock work. Genetic resistance against swine atrophic rhinitis, bovine mastitis, ovine scrapie, avian leucosis, etc. has been demonstrated, but apart from the establishment of leucosis-resistant chicken lines, no attempts have as yet been made to obtain resistant populations in other species (CHAPTER 2). In recent years, however. investigations into natural resistance in animals have become increasingly important (CHAPTER 3 and 4).

Congenital resistance and immunity, as well as individual variations in physiological resistance to disease, have three main aspects: 1. congenital immunity or resistance to infection, 2. resistance to parasites and 3. resistance to other performance-limiting factors. Susceptibility or resistance differ greatly among species, breeds, lines, families and ages, as shown both by direct observation and by objective statistical evaluation in CHAPTER 3. Evidence has been obtained that lines resistant to certain infectious diseases can be established by selective breeding but this work is still largely in the experimental phase.

Natural selection for resistance to certain diseases may also occur, and this is very efficient with respect to the powers of resistance gained by the survivors, as exemplified by the development of resistance to myxomatosis infection in Australian rabbit (CHAPTER 2). Resistance depends on multiple interactions between pathogen, host, and environment (CHAPTER 1); these interactions, however, are realized as important limiting factors in artificial selection work (i.e. G*E interaction). The establishment of populations genetically resistant to particular infectious or

invasive diseases becomes a necessity when other approaches such as veterinary methods, hygiene measures, improvements in environmental conditions etc., fail to provide effective control.

The selection of animals for resistance to multiple pathogens is a difficult as shown in CHAPTER 3, but in many cases economically rewarding procedure, because it brings about a reduction of losses and results in a better economy of manpower and drug use. The resistance of the host organism to infectious or other pathogenic factors depends on the genetic polymorphism of the complement system. For example, highly significant correlation has been demonstrated between the serum immunoglobulin levels in new-born animals and their resistance to disease (Murphy et al., 2005). In certain cases, genetic resistance to disease during intrauterine life is clearly evident. Congenital resistance essentially means that pathogenic agents (bacteria or viruses) are unable to reproduce in the cells and tissues of the resistant host organism. The replication capacity of viruses, and the nature of the host's immune response, are both genetically determined (CHAPTER 2). A healthy animal also shows a certain degree of resistance to virus infections, but resistant individuals are fully protected against them. Conversely, chickens devoid of T-cells are very susceptible to infection by the Newcastle disease virus, and agammaglobulinemic chickens are extremely susceptible to it (CHAPTER 2). Hereditary susceptibility and resistance have been shown to play a major role in the epidemiology of many virus diseases of both humans and animals.

Improving animal nature-resistance to infectious diseases

As discussed in CHAPTER 1, among the four types of infectious disease outbreak occurrences, endemic infections are the most challenging for traditional disease control strategies. Thus, breeding to improve host resistance to infectious diseases is a complementary or perhaps alternative approach to controlling such diseases. The second important factor in selecting the resistance traits that fit the breeding goals is the genetic variance of these traits. As shown in CHAPTER 4, several resistance traits have enough genetic variance (narrow sense heritability) to improve the resistance to such pathogens ($h^2_{GCTA} = 0.22 \pm 0.03$ for *bovine herpes virus 1* pathogen). Moreover, genetic correction is also an important factor, especially when one resistance trait is genetically associated with another productivity or performance trait. One final factor that is essential to a good breeding scheme is the accurate measurement or recording of the phenotypes of the resistance or susceptibility traits. Most of the phenotypic disease traits are costly and logistically difficult; however, in the current genomic era, selection of animals based on their genotypes without the need

to expose them to infection or for them to be part of an epidemic is now possible (Bishop and Woolliams, 2014).

In general, stating the efficiency and availability of control strategies is difficult since they depend on specific pathogens and environment. For instance, certain infectious diseases such as paratuberculosis (in cattle), porcine reproductive and respiratory syndrome (PRRS) (in pigs), and nematode-related diseases (in sheep) cannot be eliminated by any 'non-genetic' control strategy, even vaccination (Bishop and Woolliams, 2014). The recent increase in public suspicion of vaccines and antibiotics in developed countries could also have an adverse impact on their use (Dube et al., 2013; Egger-Danner et al., 2015) due to consumer interest in healthy animal products (meat and milk free from antibiotic residuals). However, in developing countries, vaccine production technology may not be available and importing vaccines from other places may be unhelpful due to the variation in infectious pathogen strains around the world. Resistance to infectious diseases also is a complex property which depends on the virulence and dose of the pathogenic agent, and on environmental factors as well. The creation of resistant populations requires the selection of groups exposed to heavy infections, but both owners and veterinarians have objected to this approach although scores of experiments have verified its correctness. The course of an infectious disease within a population depends on the activity of the pathogenic agent, the resistance of the hosts to the infection, and the extent of elimination of extremely susceptible individuals from the population. Many authors have attempted to establish objective criteria for the assessment of natural resistance in animals. For example, Staykov et al. (2007) recommended the measurement of the bactericidal activity in rainbow trout serum, by the method of Pillemer modified by Abidov and Mirismailov (1979).

In cattle, extensive studies have shown that resistance to mastitis can be reinforced as a hereditary trait in families and populations of cattle. In fact, most cows carry the causal agent(s) of mastitis in an inactive state, but these agents become active and rapidly multiply only in predisposed individuals under appropriate environmental conditions (Hutt, 1958). Grooms (2006), found that the three main causes of calf losses: bovine viral diarrhea virus and leptospirosis are differed in their frequency of occurrence between sexes and breeds. In male and female calf groups reared under identical systems of management, significantly more bullocks than heifers are known to die from such diseases. *In poultry*, resistance to disease may be associated with certain blood group factors. Crittenden et al. (1970) demonstrated that embryos not possessing the antigen are highly susceptible. Some of those chicken strains in which genetic resistance to avian leucosis and Marek's disease has been reinforced by selective breeding are noted for excellent production characteristics. Another example: losses

from pullorum disease were high in the brown egg layers that are known to be more susceptible than white layers (Barrow et al., 2011). Resistance to infectious disease *in swine* was analyzed in Irish and British breeds. Gedymin et al. (1964) reported the hereditary transmission of resistance to tuberculosis in sow lines and boar lines. The incidence of tuberculosis differed significantly between resistant and susceptible families and lines. Lurie et al. (1951) referred to the selective breeding of *rabbits* for resistance to tuberculosis. The inheritance of congenital immunity to Brucella melitensis was demonstrated in *goats* by (León and Guerrero, 1962). The selective breeding of fish populations for resistance to disease has also been reported, i.e., resistance to certain infectious diseases was found to differ between fish populations reared under similar conditions (Murray and Peeler, 2005)

Modern technology and breeding for diseases resistance

Whole genome sequence-based selection. Various commercial bovine single nucleotide polymorphism chips (with densities ranging from 3k, 50k and 777k to whole‑genome sequence (WGS)) have been implemented in the several genome-wide prediction methods (Raymond et al., 2018; Meuwissen et al., 2013; Pryce et al., 2012). So far, results of several studies in livestock evaluation shown that the genomic estimated breeding values (GEBV) using WGS can be more accurate than the GEBV using other chip types (Iheshiulor et al., 2016; MacLeod et al., 2016; Brøndum et at., 2015; Druet et at, 2014). And since the SNP in WGS data may actually the mutation that causing the variant effect, then the estimation of SNPs effect could be performed once in the reference (training) population. The accuracy of a GEBV estimated for a moderately heritable trait using a multi-breed reference population was 0.70 as reported by Iheshiulor et al. (2016).

A genome-wide selection scheme for enhancing diseases resistance (Figure 2), for example, could be used. To optimizing the training population, a breeding company/institute need to sequence a considerable number of animals or bull that are closely related to the breeding population. Assume that the first estimates come from 100 sire families with 20 son each, and each son has 100 or more daughters in their proof for several production and disease resistance traits (infectious and non-infectious diseases). Daughter resistance and yield deviation are calculated which represent daughter average adjusted for all environmental effects and for merit of their dams. The sire and sons (2100 animals) need to be sequenced (Figure 2). After estimating the

SNPs effect from the training population, with accuracy of 0.70, roughly 5000 elite females are required as dams of young bulls (Figure 1). The entire female population does not need to be sequenced. Only those that would normally qualify as a dam of bulls needs to be sequenced. Let the number of cows sequenced be 2000 (pre-selected on the usual criteria for bull dams). Using the SNPs information from training population, 1000 cows are chosen from these 2000 cows as dams of the next generation of young bull (calves). Assume that 500 bull calves are born from the 1000 elite dams and each of the young bulls are sequences to obtain GEBV with accuracy more than 0.70. The breeder needs to select the top 20 bull to purchase. Then the top 5 bull are designed to mating the elite dams for the next generation (Figure 1).

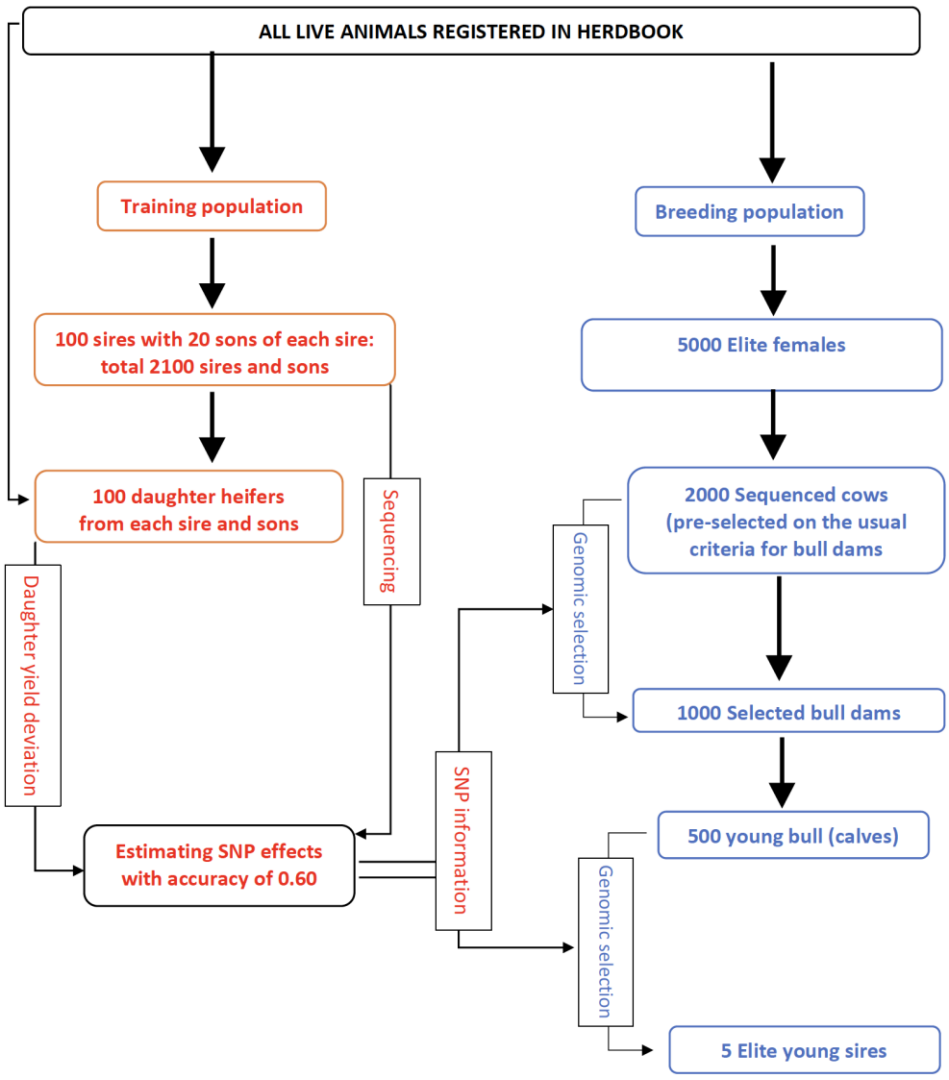


Figure 1. Example of genomic breeding program.

Embryonic stem cells-based selection. In dairy cattle, sires and dams can be selected based on their daughter health status or yield deviation at 5-7 years of age; or based on their genome after birth. However, in any breeding schemes that implementing either method, the generation length cannot be less than 2 years, where the cow reach maturity at approximately 15-18 months old and another 9 months of gestation. Embryonic stem cells (ESCs) technology introduced the visibility of producing a mature haploid male and female germ cell (sperms and eggs) from ESCs and unites them again with another pair of the opposite sex from different animal (in-vitro) to form a new zygote in just 3-4 months. Furthermore, using ESCs with embryo transfer technique can allows us to completely control of the genetic structure of the next generation, instead of distributing the elite bull semen in all participated farms in the breeding program. On 1981, embryonic stem cells (ESCs) were first isolated from mouse embryos (Evans and Kaufman, 1981; Martin, 1981). ESCs can spontaneously renewing and differentiating (*in vivo* and *in vitro*) into the derivatives of all three germ layers that can interact and forming all animal organs and tissues includes sperms and oocytes (Figure 2). During the last three decades, the derivation and propagation of ESCs in ruminants was unstable, however, Bogliotti et al. (2017) reported a stable derivation of bovine ESCs in a culture condition based on Wnt-pathway inhibition. A well-defined culture system has also been proposed to generate primordial germ cell-like cells (PGCLCs), and then spermatogonial stem cell-like cells (SSCLCs), from induced pluripotent stem cells (iPSCs) in cattle (Bogliotti et al. (2017). Therefore, assuming that a complete cycle (generation) to producing sperms and oocytes from ESCs is achievable in 3-4 months (Figure 2), the question is how to take advantage of this technology. How can genomic selection-based breeding schemes be modified (or replaced) so as to make faster genetic change? Further research is needed to look at these questions.

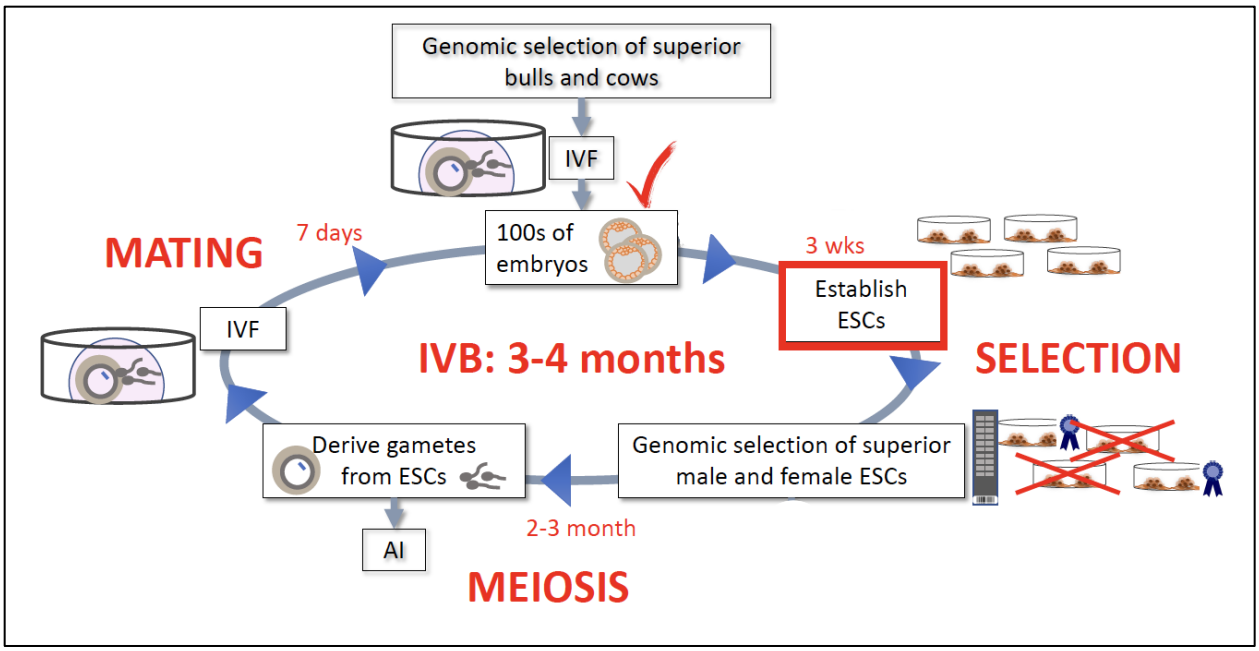


Figure 02. Embryonic stem cells-based selection.

The accuracy of genetic parameter estimates

This study represents the first effort in animal genetics to compare the estimate of heritability using genome-wide SNPs (genomic relationship matrix (GRM)) with the estimate of heritability using pedigree (kinship relationship matrix). Using the pedigree relationship in CHAPTER 3, the heritability estimates of average milk yield (0.20±0.01) and average fat-to-protein ratio (0.16±0.01) corresponded well with narrow-sense heritability estimates of average milk yield (0.19±0.02) and average fat-to-protein ratio (0.25±0.02) using GRM in CHAPTER 4, suggesting that there is no "*missing heritability*" (un-assayed) due to the low accuracy of pedigree, GRM, or algorithms used for estimation. In CHAPTER 3, the phenotypic and genetic association between calf and cow disease traits were not significantly different from zero, indicating that the early diagnosis of calf diseases cannot be used as a tool to improve the health traits of lactating cows. Furthermore, in CHAPTER 4, GRM estimated zero genetic correlation between resistance to *Trichophyton* in calves and cows, confirming the hypothesis that calf and cow resistance are two genetically different traits.

The relationship between some qualitative traits, i.e. the colour of skin, hair or feathers and the genetic susceptibility to certain diseases has long been known. For example, I discussed in CHAPTER 2, how white-faced Hereford cattle show a marked susceptibility to cancer of the eyelids on exposure to strong sunlight, while pigmented skin and hair around the eyes confer full protection against this condition (Bonsma, 1949). More examples in other literature: the so-called

white heifer disease (Rendel, 1952) is associated with sterility in Shorthorns cattle; grey sheep are noted for depressed vigor and digestive disorders (Craig, and Bitting. 1903); grey horses are frequently affected by melanoma (Valentine, 1995); Cyclic neutropenia, a condition frequently followed by death in the young, was identified long time ago as a recessive trait in grey Collie dogs (Lund 1967). Since pigmentation of the skin is a mendelian hereditary trait, selection for resistance to diseases that associated with such mendelian traits is relatively straightforward (Morris, 2007).

Genetic architecture of host resistance

A genetic contribution to the resistance or susceptibility of several infectious diseases has been well established by pedigree studies (Bishop and Morris, 2007; Bishop, 2010; Berry et al., 2011); however, the nature of their genetic architectures remains unknown. Genetic architecture refers to the number of genomic loci (SNP markers) that contribute to resistance, the minor allele frequency and effect magnitudes, and the interactions of alleles among genes. Understanding genetic architecture is the basis for progress in dissecting etiology, since it aids design of the appropriate study to successfully identify resistance variants.

We have reviewed how this thesis employed information regarding whole-genome SNPs to better understand the joint spectrum of allele frequencies and the magnitude and direction of the effect of variants of infectious disease resistance (APPENDIX - C). I focused particularly on resistance to the following pathogens in calves: Escherichia coli, Bovine respiratory syn., Cryptosporidium, and Coccidia; and resistance to the following pathogens in cows: Salmonella, Staph. Aureus, Staph. Haemolyticus, Strep. Agalactiae, Strep. Dysgalactiae Strep. Uberis, Mycobac. Paratuberculosis, Trichophyton, Bovicola bovis, and Chorioptic scabies. Four more production and performance traits were investigated: average milk yield, fat-to-protein ratio, birthweight, and average daily gain (in calves).

Figure 3 (Solovieff et al., 2013) shows that pleiotropy can be found in several forms. At a single loci level, a causal variant can be associated with multiple phenotypes (Figure 3a); At gene level, multiple variants in the same gene can be associated with different phenotypes (Figure 3b); At genomic level, multiple variants in the same region can be associated with different phenotypes (Figure 3c); A mediated pleiotropy, occurs when a genetic variant is associated with a given phenotype and this phenotype is correlated with another phenotype (Figure 3d); A pseudo pleiotropy is a phenotypical misclassification among two correlated traits (Figure 3e); A spurious pleiotropy can also appear as a result of a single marker that is correlated with multiple causative variants located in different genes with completely different functions (Figure 2f).

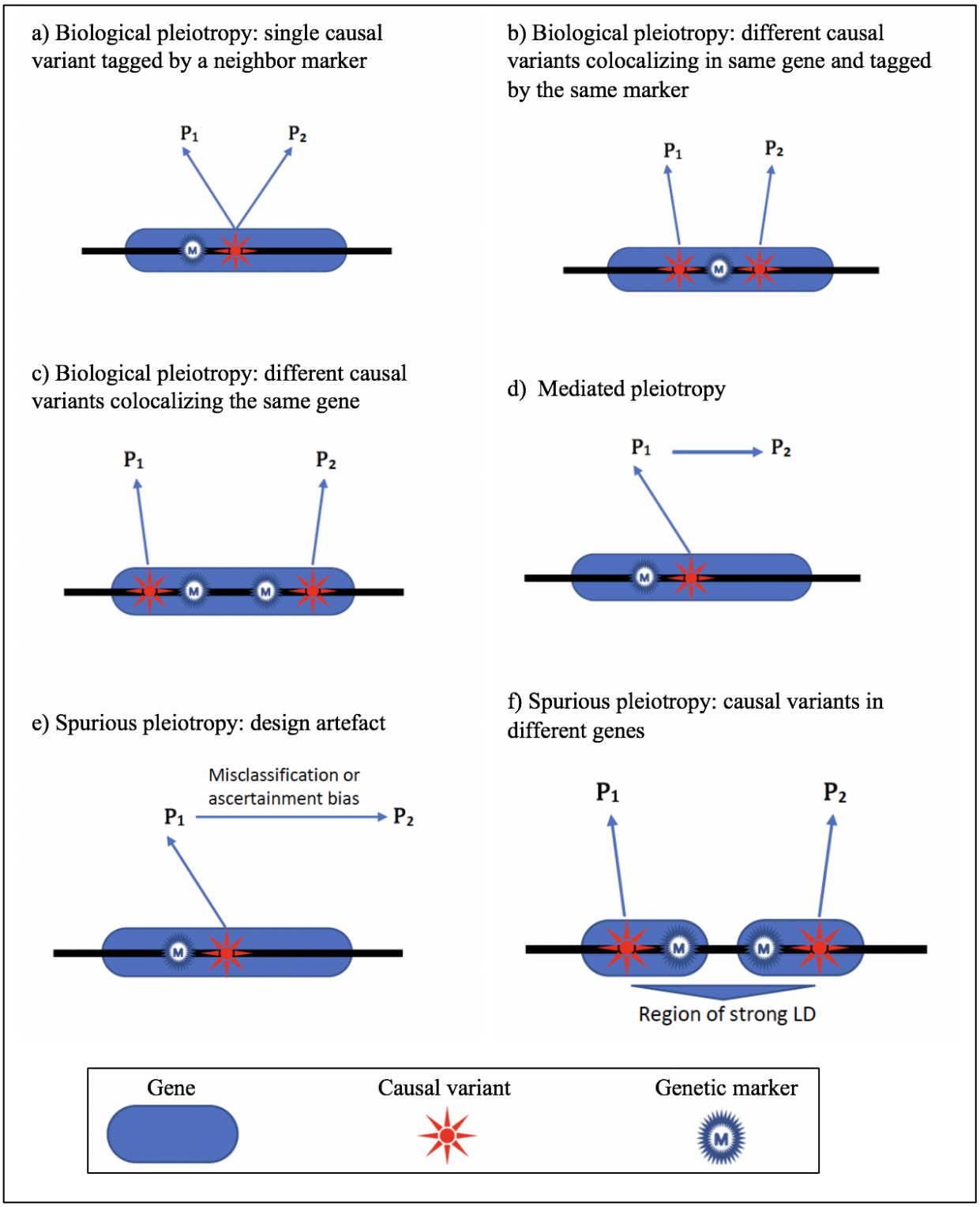


Figure 03. Types of pleiotropy, adapted from Solovieff et al., (2013)

Main conclusions

Single pathogen infectious disease resistance traits are the best phenotype to study the genetics of bovine infectious diseases. Additional conclusions of this thesis can be summarised as follows:

1- Prior to deciding any breeding goals to improve livestock health traits, an epidemiological study including temporal, age, and geographical distributions (if available) is required.

2- Drug treatments, antibiotics, vaccinations, and non-genetic prevention tools are extremely important for controlling infectious diseases in the short term; however, genetic improvement of host resistance to infectious diseases should be a complementary tool for long-term health improvement strategies.

3- Epidemic pattern, sufficient genetic variance, and a positive genetic correlation with other health, performance, and productivity traits are the most important characteristics of a good resistant trait for the genetic improvement of livestock health traits.

4- Most calf and cow diseases have genetically different traits, and selection for health traits should be carried out at both ages.

5- G-REML is one of the best (if not *the* best) methods for estimating genetic parameters for complex traits using genome-wide SNP data.

6- Biological pathway analysis appears to be a very useful tool for increasing the understanding of the biological background of any health trait.

References

Abidov, A. A., and M. I. Mirismailov. 1979. Influence of vaccination on some factors of natural resistance. Medizina, Tashkent: 28-74.

Barrow, P. A., and OC Freitas Neto. 2011.Pullorum disease and fowl typhoid—new thoughts on old diseases: a review. Avian pathology 40, no. 1: 1-13.

Bishop, S.C., and J.A. Woolliams. 2014. Genomics and disease resistance studies in livestock. Livest Sci 166:190–198. doi:10.1016/j.livsci.2014.04.034.

Berry, D.P., M.L. Bermingham, M. Good, and S.J. More. 2011. Genetics of animal health and disease in cattle. Ir Vet J 64:5. doi:10.1186/2046-0481-64-5.

Bishop S.C. Axford R.F.E. Nicholas F.W. Owen J.B. 2010. Breeding for disease resistance in farm animals. 3rd edn CABI

Bogliotti, Y. S., J. Wu, M. Vilarino, D. Okamura, D. A. Soto, C. Zhong, M. Sakurai, R. V. Sampaio, K. Suzuki, J. C. I. Belmonte and P. J. Ross. 2018. Efficient derivation of stable primed pluripotent embryonic stem cells from bovine blastocysts. Proceedings of the National Academy of Sciences, p.161.

Bonsma j. C. 1949. Breeding cattle for increased adaptability to, tropical and subtropical environments. Jags. 39: 204.

Brøndum, R. F., G. Su, L. Janss, G. Sahana, B. Guldbrandtsen, D. Boichard, and M.S. Lund. 2015. Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. Journal of dairy science, 98(6), pp.4107-4116.

Craig, R. A. and A. W. Bitting. 1903. Diseases of sheep (Vol. 87). Purdue University Agricultural Experiment Station.

Crittenden, L. B., W. E. Briles, and H. Ao Stone. 1970. Susceptibility to an avian leukosis-sarcoma virus: close association with an erythrocyte isoantigen. Science 169, no. 3952: 1324-1325.

Druet, T., I. M. Macleod and B. J. Hayes. 2014. Toward genomic prediction from whole-genome sequence data: impact of sequencing design on genotype imputation and accuracy of predictions. Heredity, 112(1), p.39.

Dube,E., Laberge,C., Guay,M., Bramadat,P., Roy,R., and Bettinger,J. 2013. Vaccine hesitancy: an overview. Hum. Vaccin. Immunother. 9, 1763–1773.doi: 10.4161/hv.24657

Egger-Danner,C., Cole,J.B., Pryce,J.E., Gengler,N., Heringstad,B., Bradley,A., et al. 2015. Invited review: overview of new traits and phenotyping strategies in dairy cattle with a focus on functional traits. Animal 9, 191–207.

Evans M. J., M. H. Kaufman. 1981: Establishment in culture of pluripotent cells from mouse embryo. Nature 292, 154–156.

Gedymin J., S. Alexandrowicz, W. Folejewski, M. Ratajszezak. 1964. Med. Weter. 20 (10), 577-640.

Grooms, D. L., 2006. Reproductive losses caused by bovine viral diarrhea virus and leptospirosis. Theriogenology, 66(3), pp.624-628.

Hutt, F.B., 1958. Genetic resistance to disease in domestic animals. Genetic resistance to disease in domestic animals.

Iheshiulor, O.O., J. A. Woolliams, X. Yu, R. Wellmann, and T. H. Meuwissen. 2016. Within-and across-breed genomic prediction using whole-genome sequence and single nucleotide polymorphism panels. Genetics Selection Evolution, 48(1), p.15.

León, A. P., and J. Guerrero. 1962. Inborn natural immunity to Brucella melitensis. II. Genetic factors of hereditary immunity to Brucella melitensis in goats. Revista de investigacion en salud publica 29, no. 2: 107.

Lurie, M. B., P. Zappasodi, A. M. Dannenberg, and I. B. Swartz. 1951. Constitutional factors in resistance to infection: The effect of cortisone on the pathogenesis of tuberculosis. Science 113, no. 2931: 234-237.

MacLeod, I.M., P. J. Bowman, C. J. Vander Jagt, M. Haile-Mariam, K. E. Kemper, A. J. Chamberlain, C. Schrooten, B. J. Hayes, and M. E. Goddard. 2016. Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. BMC genomics, 17(1), p.144.

Markov, G. G., I. Bradvarova, A. Mintcheva, P. Petrov, N. Shishkov, and R. G. Tsanev. 1969. Mycoplasma contamination of cell cultures: interference with 32P-labelling pattern of RNA." Experimental cell research 57, no. 2-3: 374-384.

Martin G. R. 1981: Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A 78, 7634–7638.

Meuwissen, T., B. Hayes, and M. Goddard. 2013. Accelerating improvement of livestock with genomic selection. Annu. Rev. Anim. Biosci., 1(1), pp.221-237.

Morris, C. A. 2007. A review of genetic resistance to disease in Bos taurus cattle. The Veterinary Journal, 174(3), pp.481-491.

Murphy, B. M., M. J. Drennan, F. P. O Mara and B. Earley, 2005. Cow serum and colostrum immunoglobulin (IgG~ 1) concentration of five suckler cow breed types and subsequent immune status of their calves. Irish journal of agricultural and food research, 44(2), p.205.

Murray, A. G., and E. J. Peeler. 2005. A framework for understanding the potential for emerging diseases in aquaculture. Preventive veterinary medicine 67, no. 2-3: 223-235.

Pryce, J.E. and H. D. Daetwyler, 2012. Designing dairy cattle breeding schemes under genomic selection: a review of international research. Animal Production Science, 52(3), pp.107-114.

Raymond, B., A. C. Bouwman, C. Schrooten, J. Houwing-Duistermaat, and R.F. Veerkamp. 2018. Utility of whole-genome sequence data for across-breed genomic prediction. Genetics Selection Evolution, 50(1), p.27.

Rendel, J. M., 1952. White heifer disease in a herd of dairy Shorthorns. Journal of Genetics, 51(1), pp.89-94.

Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M. and Smoller, J.W., 2013. Pleiotropy in complex traits: challenges and strategies. Nature Reviews Genetics, 14(7), p.483.

Staykov, Y., P. Spring, S. Denev, and J. Sweetman. 2007. Effect of a mannan oligosaccharide on the growth performance and immune status of rainbow trout (Oncorhynchus mykiss)." Aquaculture International 15, no. 2: 153-161.

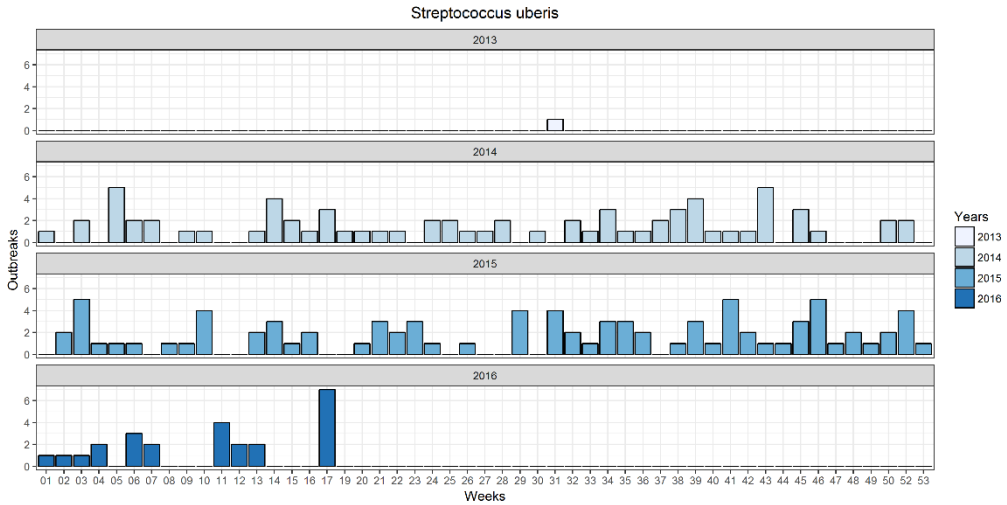
Valentine, B. A., 1995. Equine melanocytic tumors: a retrospective study of 53 horses (1988 to 1991). Journal of Veterinary Internal Medicine, 9(5), pp.291-297.

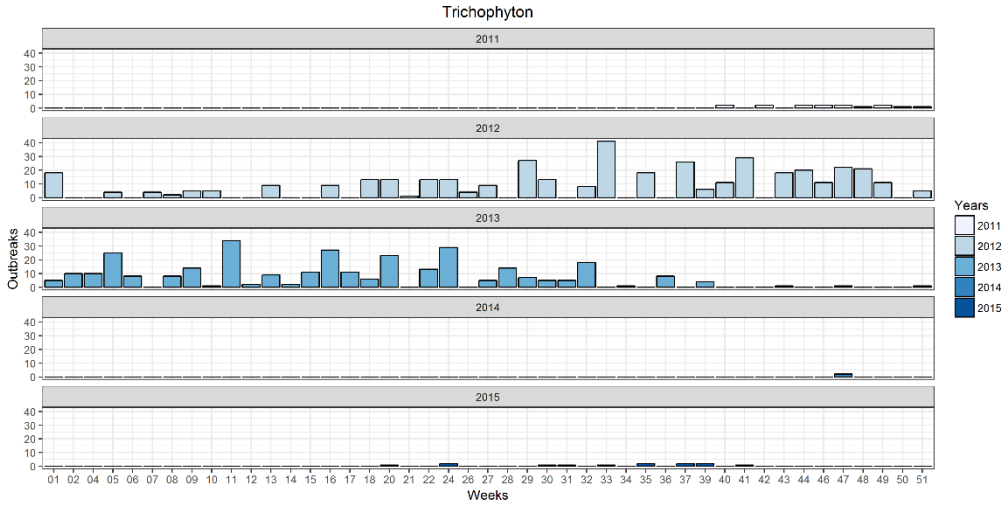
Lund, J. E., G. A. PADGETT and R. L., OTT. 1967. Cyclic neutropenia in grey collie dogs. Blood, 29(4), pp.452-461.

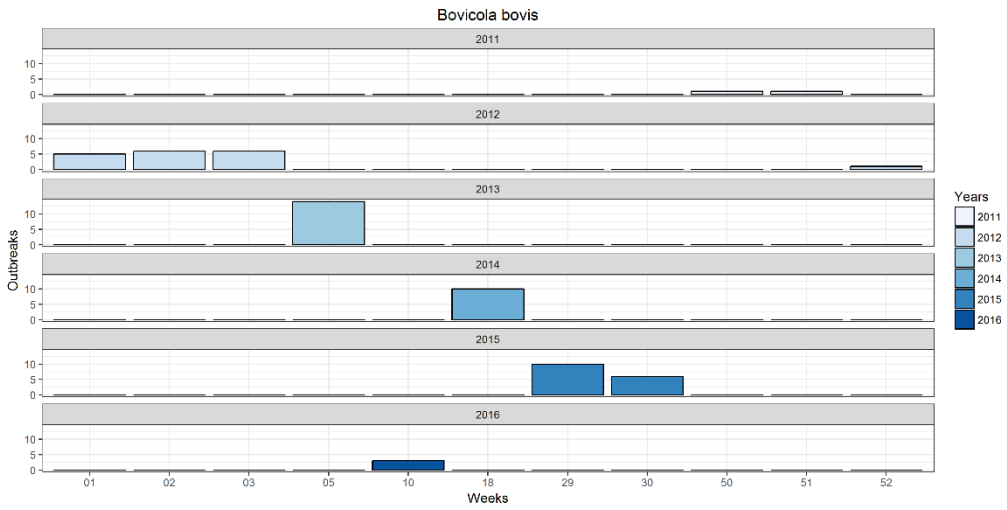
APPENDICES

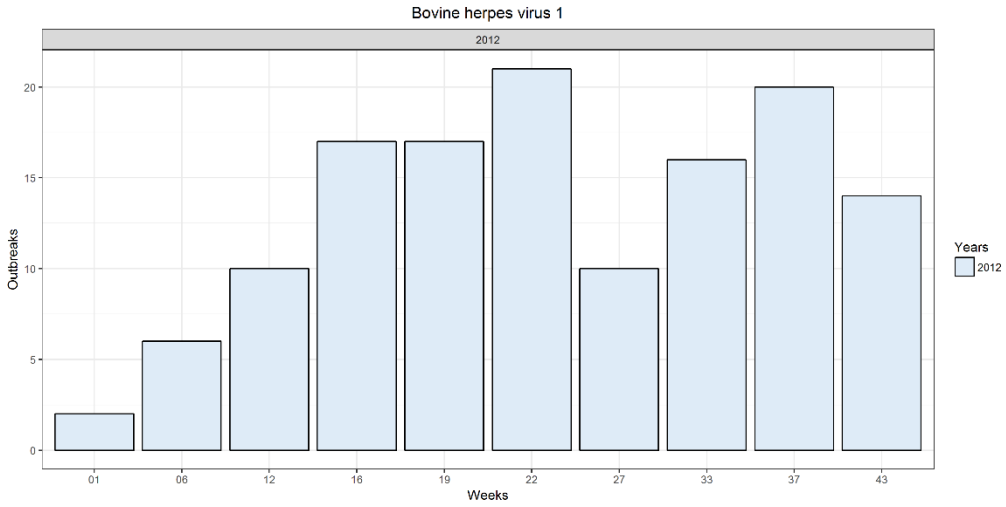
APPENDIX – A

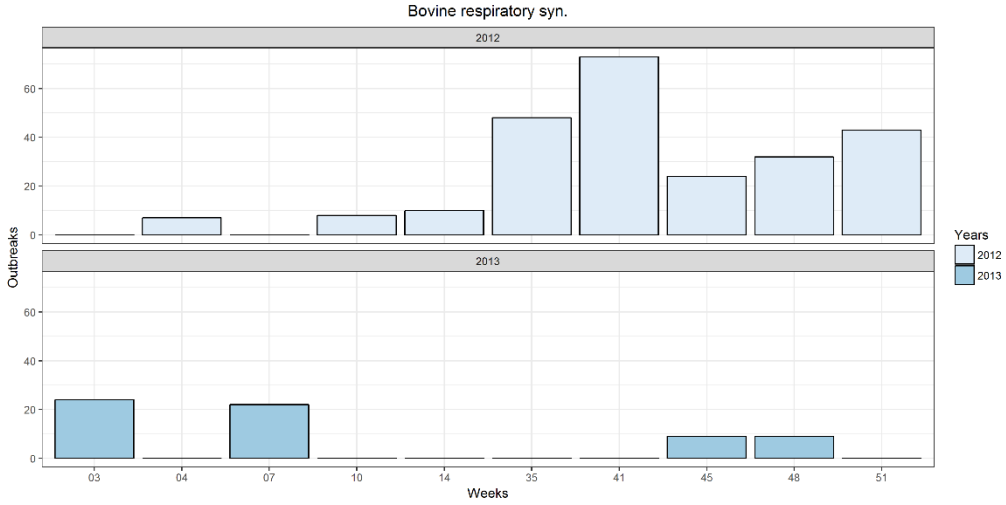
Temporal distribution (number of outbreaks per week) of infectious diseases in northeast of Germany from 2009 to 2016.

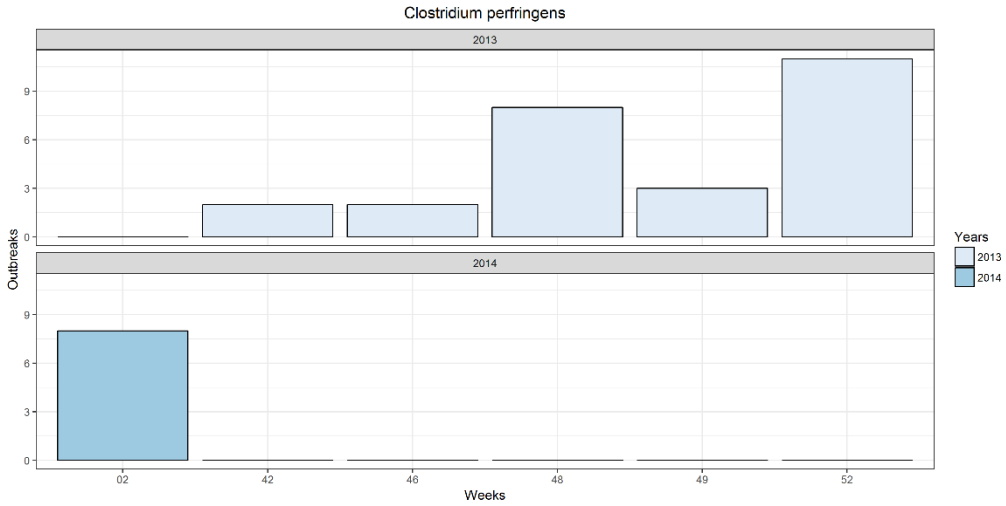


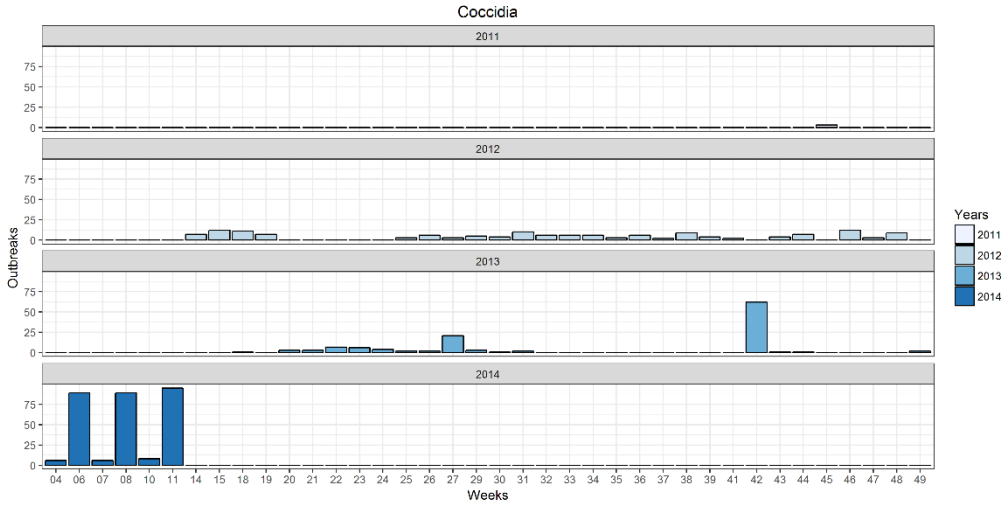


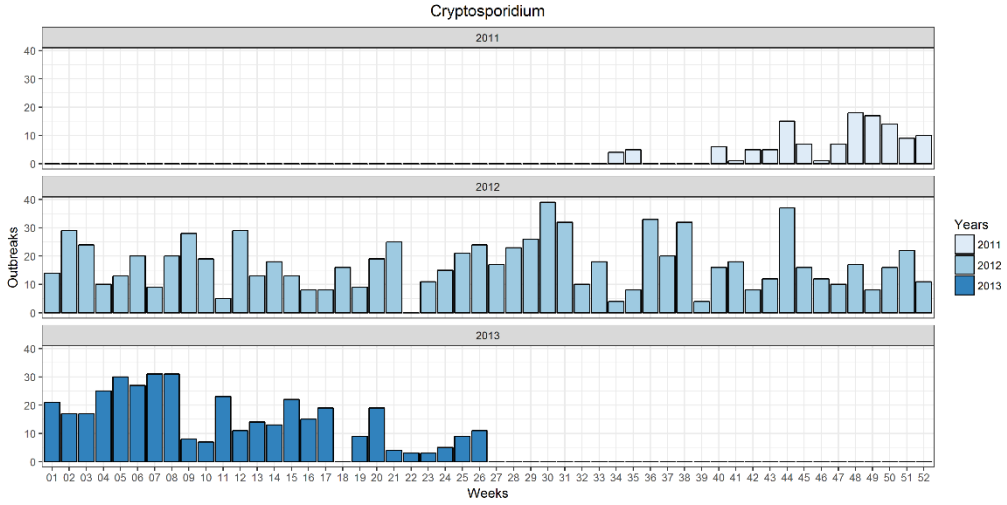


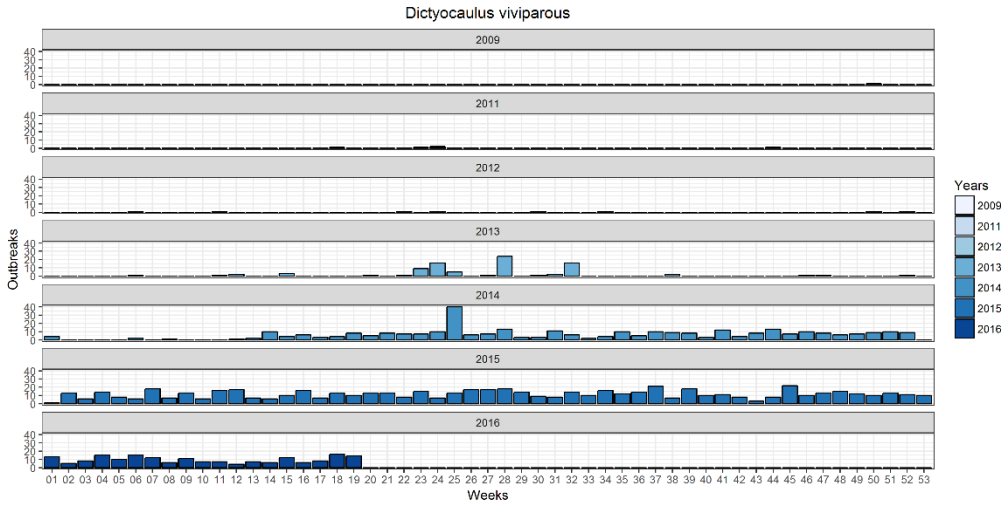

Bovine herpes virus 1
2012
Outbreaks
0
5
10
15
20
01
06
12
16
19
22
27
33
37
43
Weeks
Years
2012

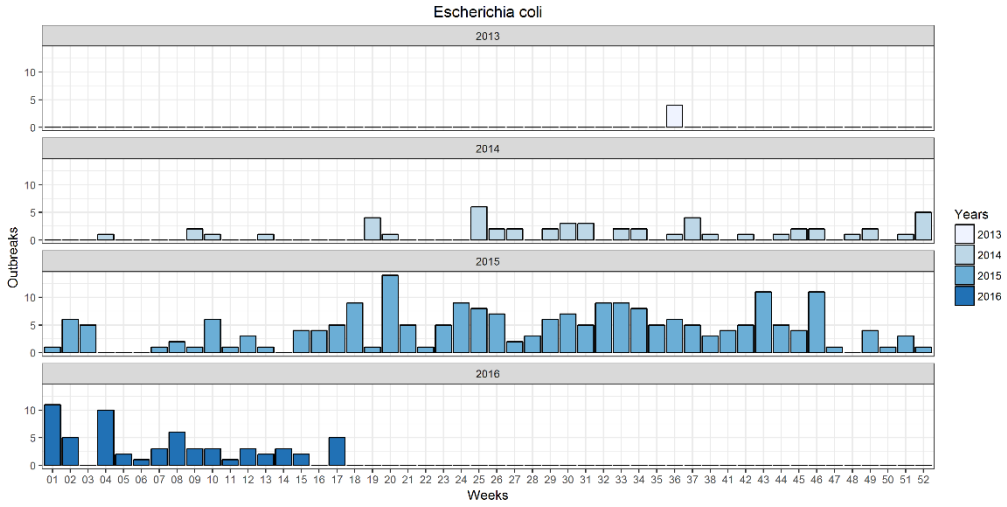

Bovine respiratory syn.
2012
2013
Outbreaks
0
20
40
60
03
04
07
10
14
35
41
45
48
51
Weeks
Years
2012
2013

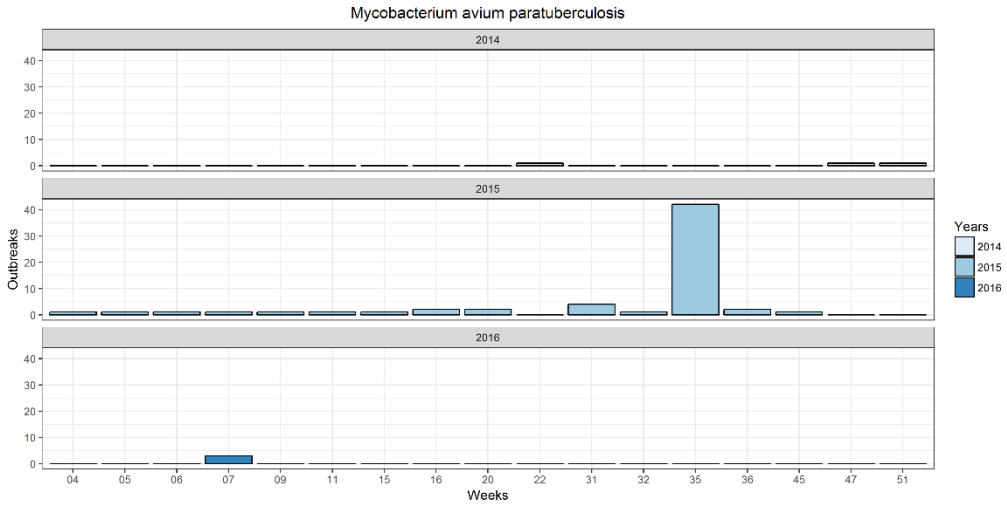

Clostridium perfringens
2013
2014
Outbreaks
0
3
6
9
02
42
46
48
49
52
Weeks
Years
2013
2014

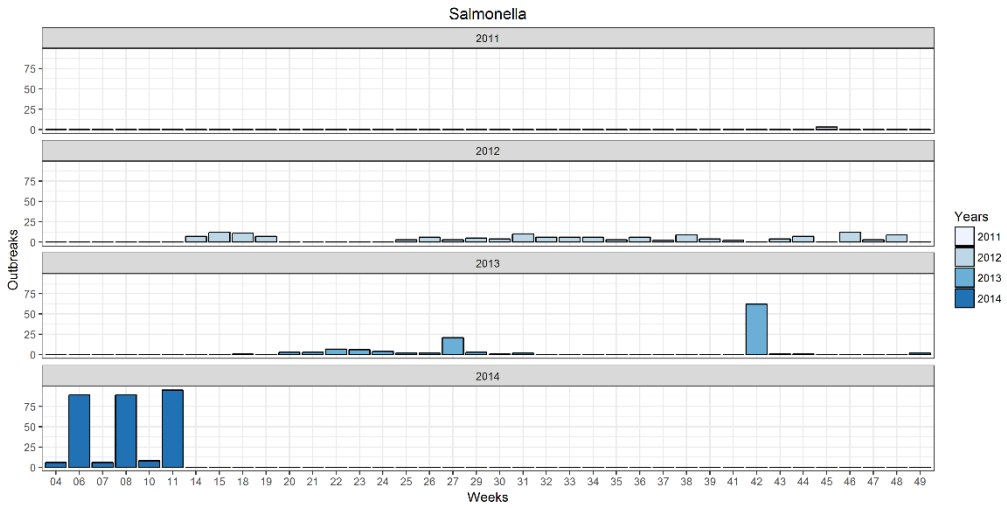

Coccidia
2011
2012
2013
2014
Outbreaks
Weeks
Years
2011
2012
2013
2014

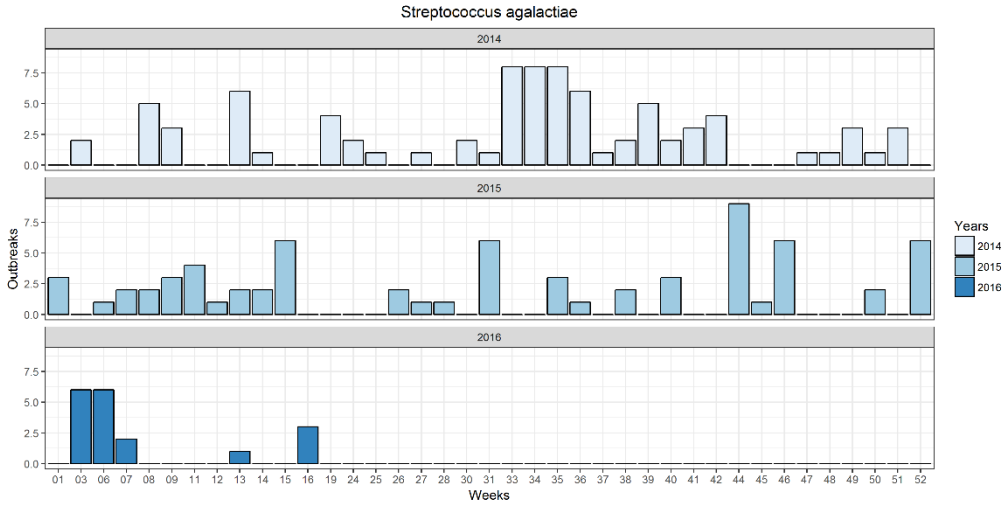

Cryptosporidium
2011
2012
2013
Outbreaks
Weeks
Years
2011
2012
2013

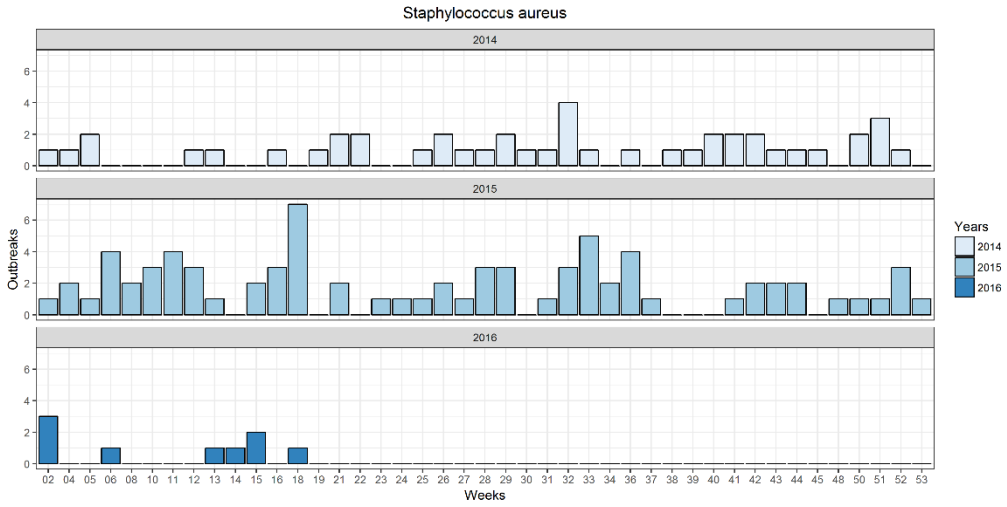

Dictyocaulus viviparous
2009
2011
2012
2013
2014
2015
2016
Outbreaks
Weeks
Years
2009
2011
2012
2013
2014
2015
2016

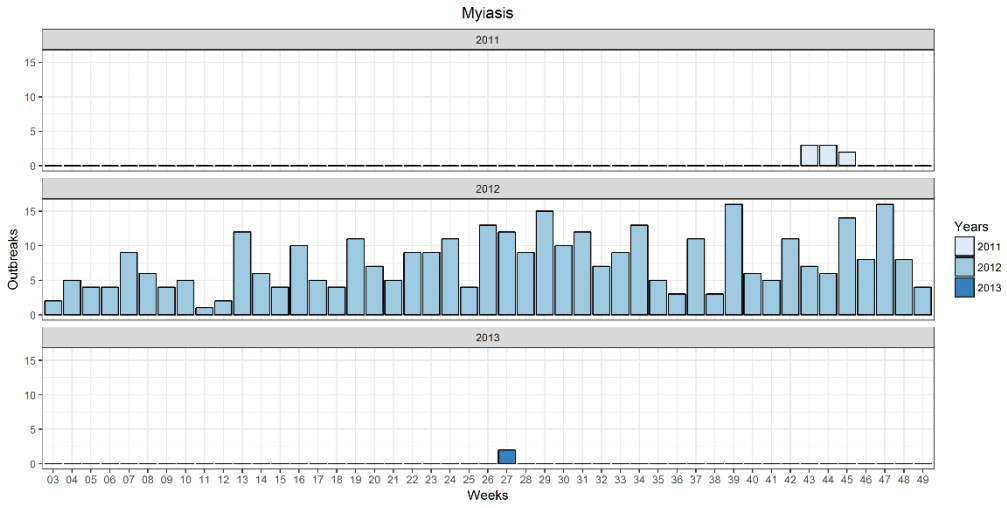

Escherichia coli
2013
2014
2015
2016
Outbreaks
Weeks
Years
2013
2014
2015
2016

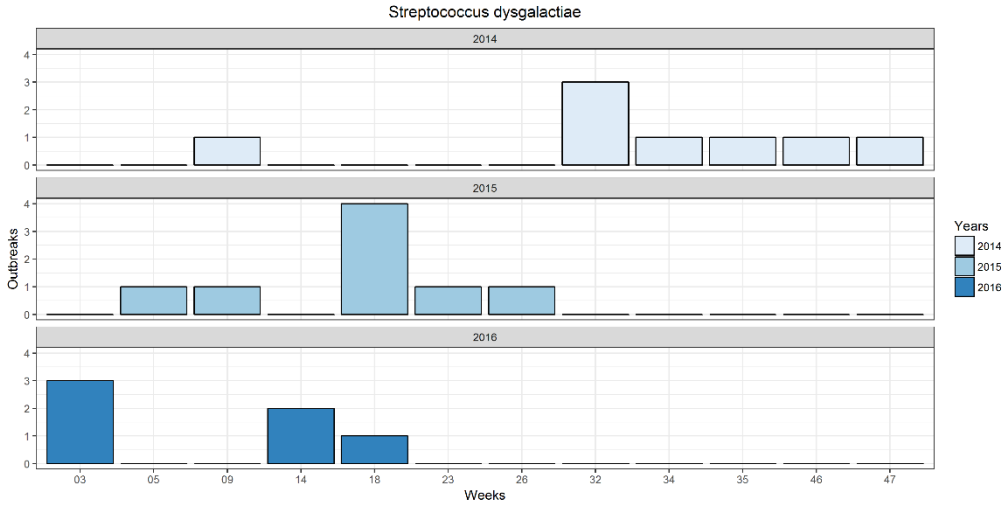

Mycobacterium avium paratuberculosis
2014
2015
2016
Outbreaks
Weeks
Years
2014
2015
2016

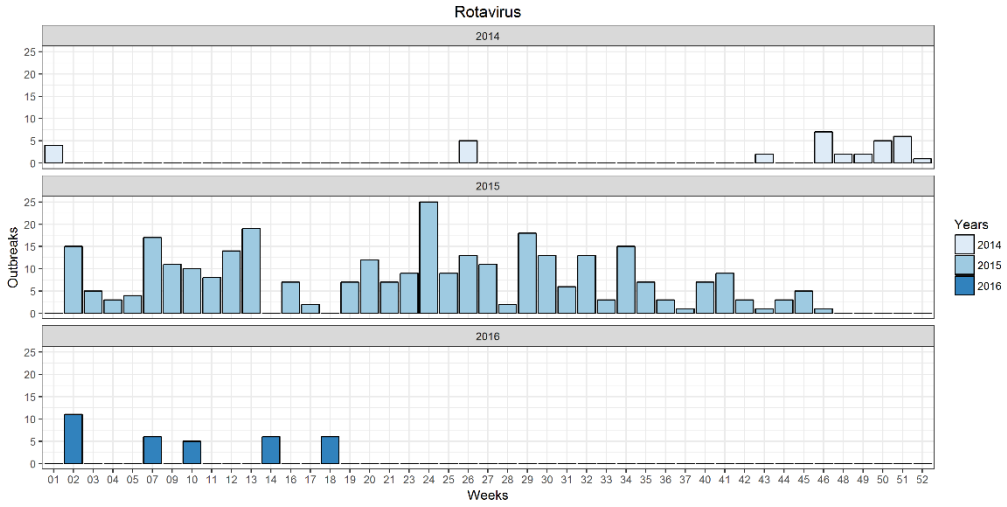

Salmonella
2011
2012
2013
2014
Outbreaks
Weeks
Years
2011
2012
2013
2014

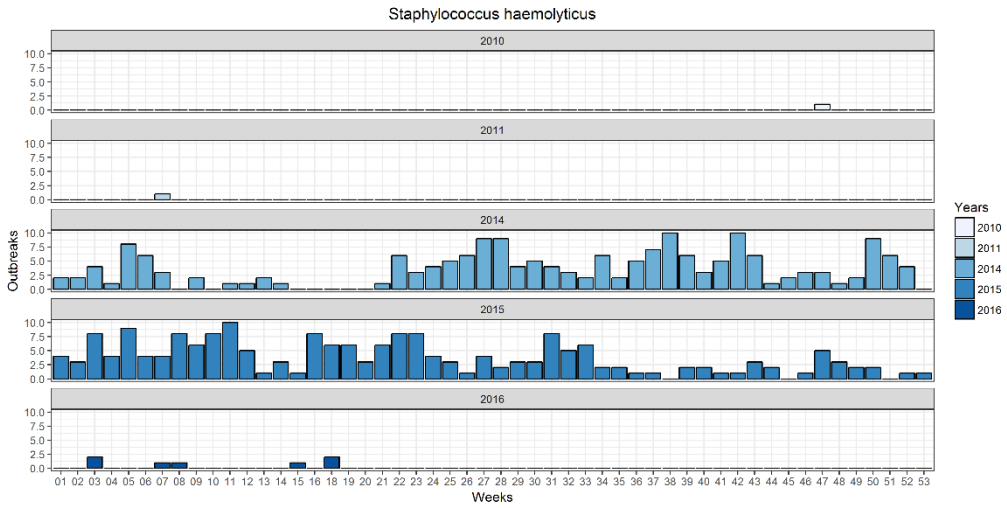

Streptococcus agalactiae
2014
2015
2016
Outbreaks
Weeks
Years
2014
2015
2016


Staphylococcus aureus
2014
2015
2016
Outbreaks
Weeks
Years
2014
2015
2016


Myiasis
2011
2012
2013
Outbreaks
Weeks
Years
2011
2012
2013

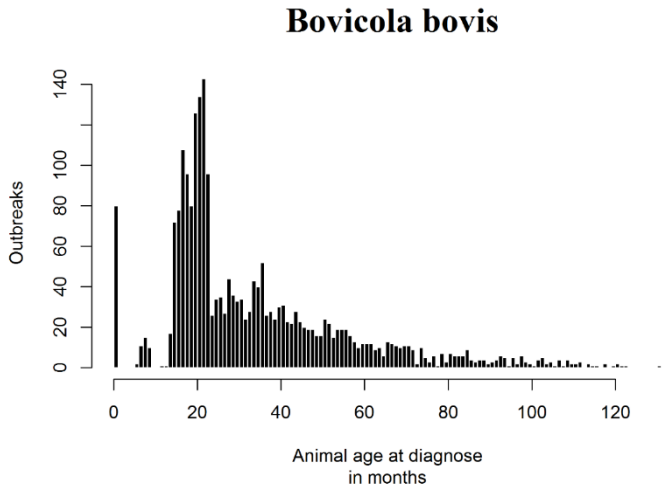

Streptococcus dysgalactiae
2014
2015
2016
Outbreaks
Weeks
Years
2014
2015
2016

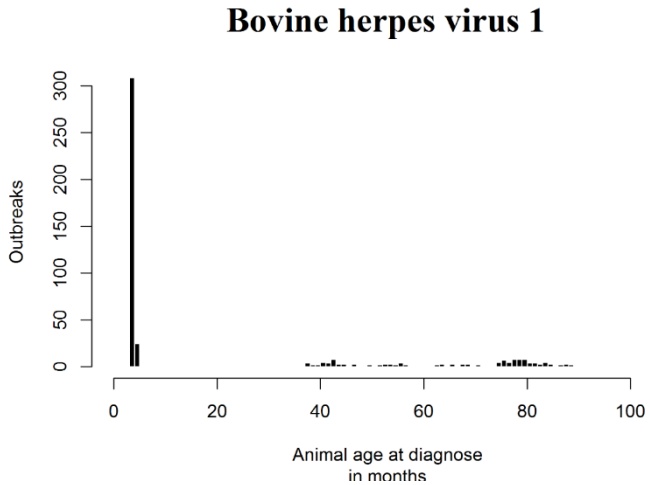

Rotavirus
2014
2015
2016
Outbreaks
Weeks
Years
2014
2015
2016

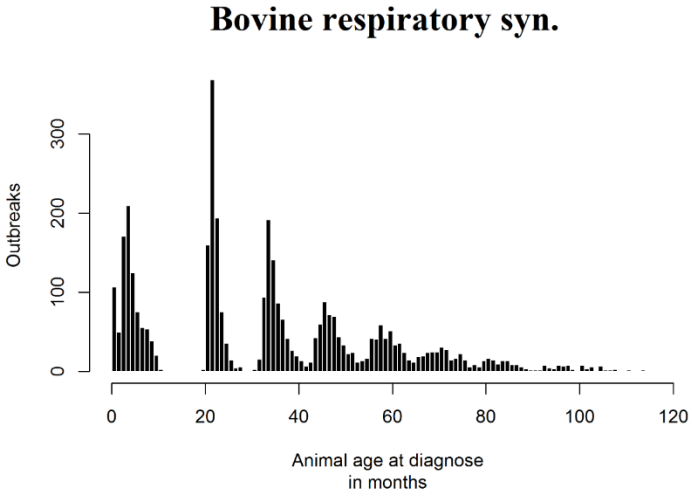

Staphylococcus haemolyticus
2010
2011
2014
2015
2016
Outbreaks
Weeks
Years
2010
2011
2014
2015
2016

APPENDIX – B

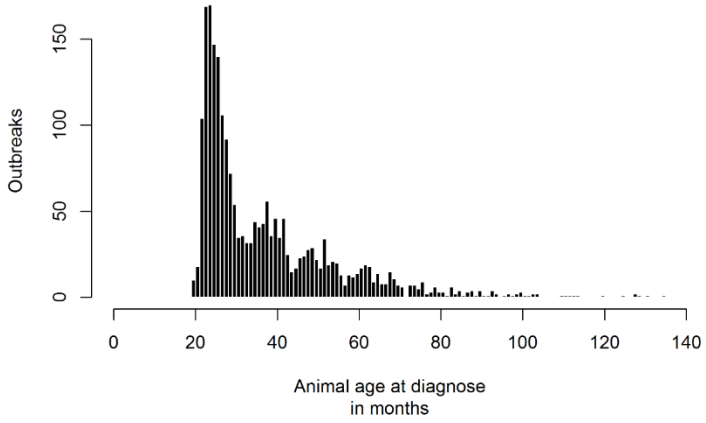
Age based distribution (number of outbreaks per age in month) of infectious diseases in northeast of Germany.



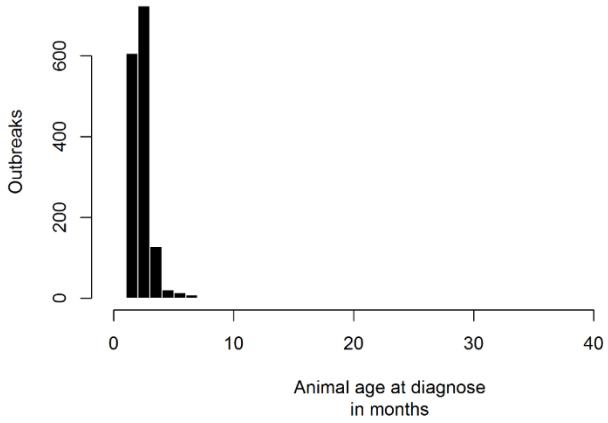




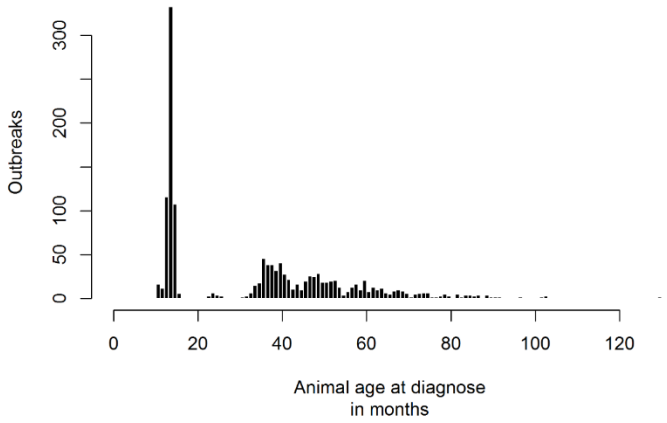
Clost. Perfringens

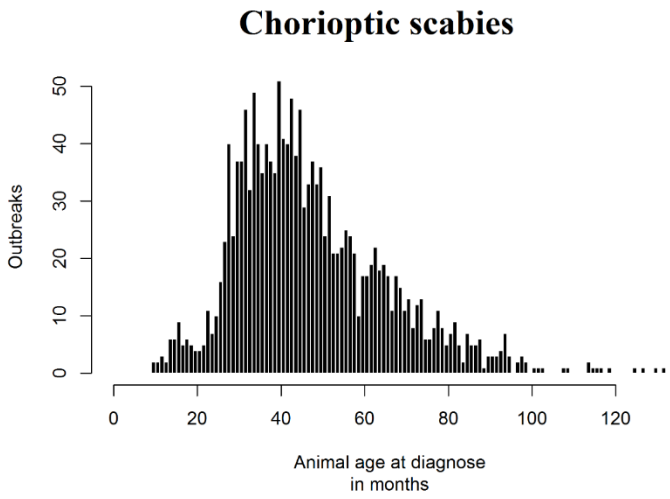


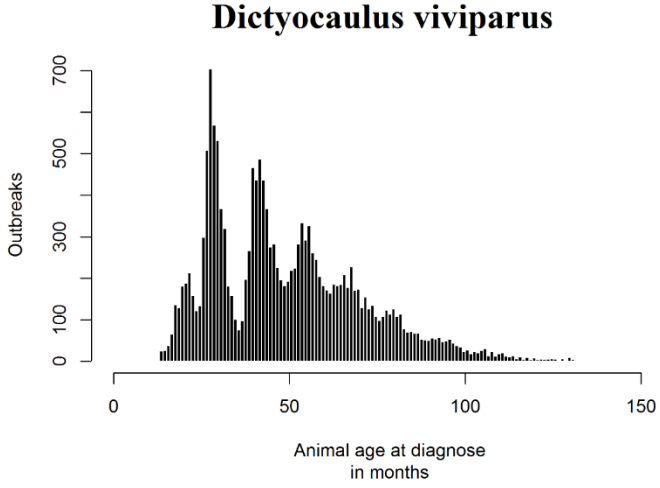
Coccidia

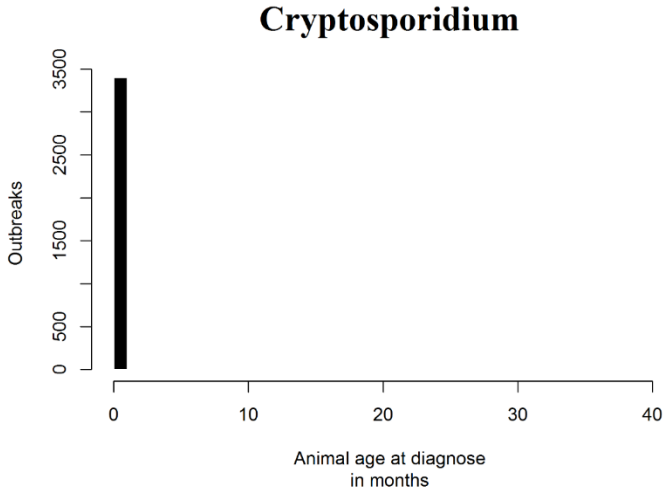


Bovine Virusdiarrhea

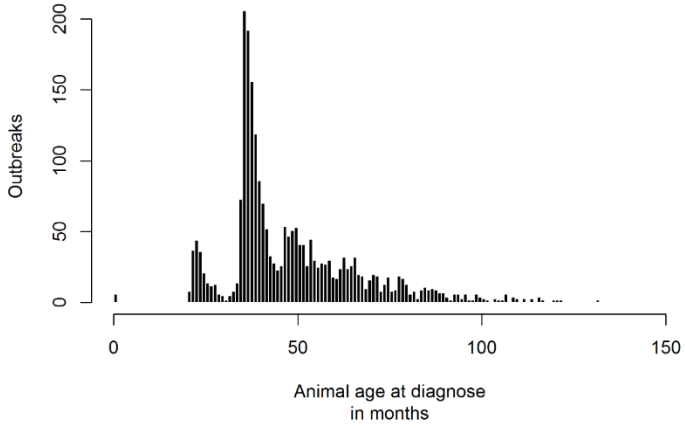



Chorioptic scabies
Outbreaks
0
10
20
30
40
50
0
20
40
60
80
100
120
Animal age at diagnose
in months

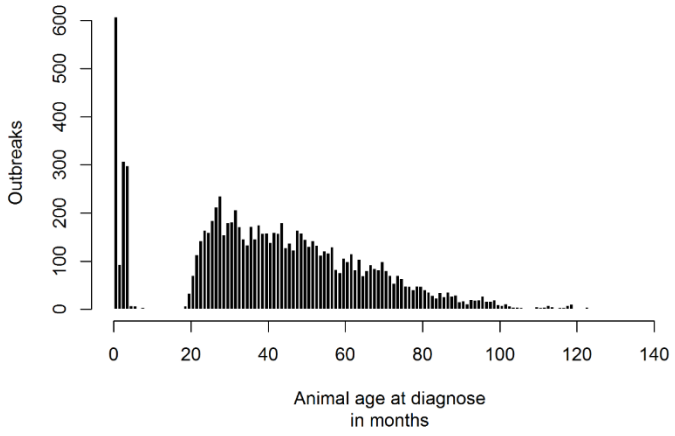

Dictyocaulus viviparus
Outbreaks
0
100
300
500
700
0
50
100
150
Animal age at diagnose
in months


Cryptosporidium
Outbreaks
0
500
1500
2500
3500
0
10
20
30
40
Animal age at diagnose
in months

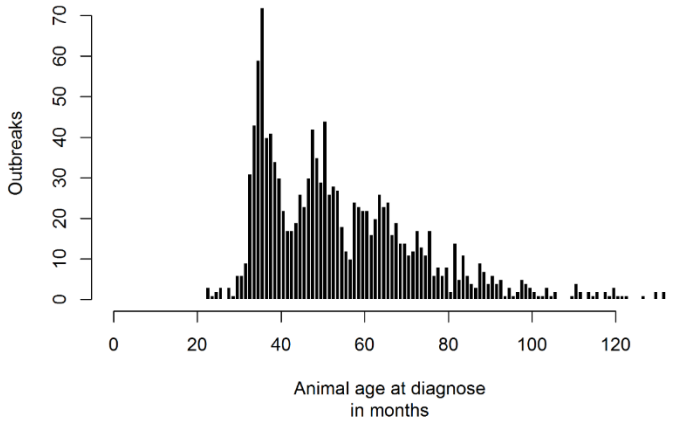
Rotavirus



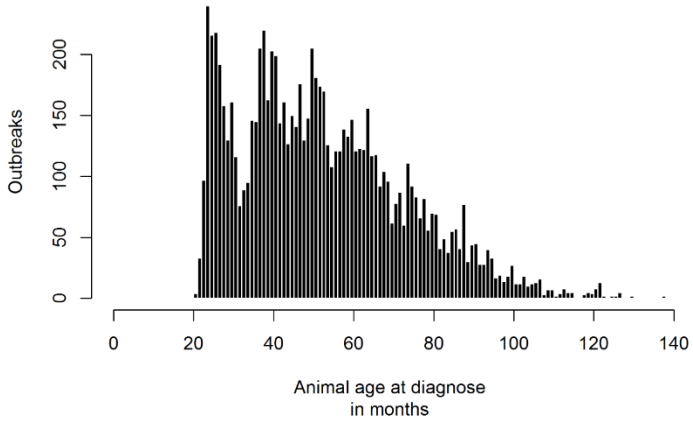
Salmonella



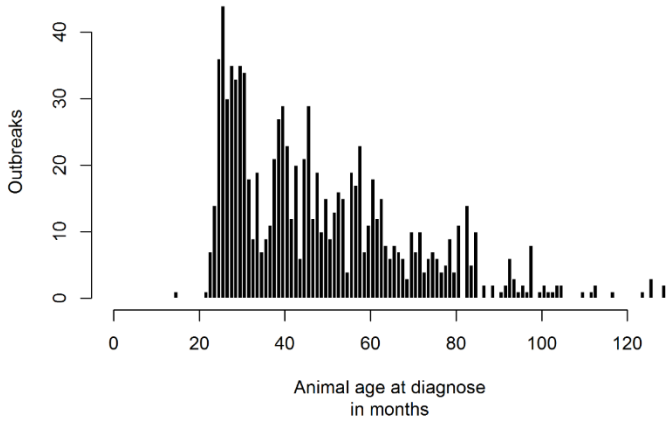
Mycobac. Paratuberculosis



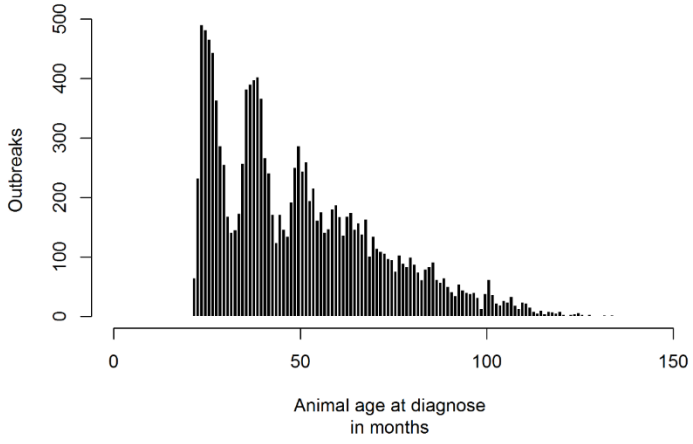
Staph. Aureus

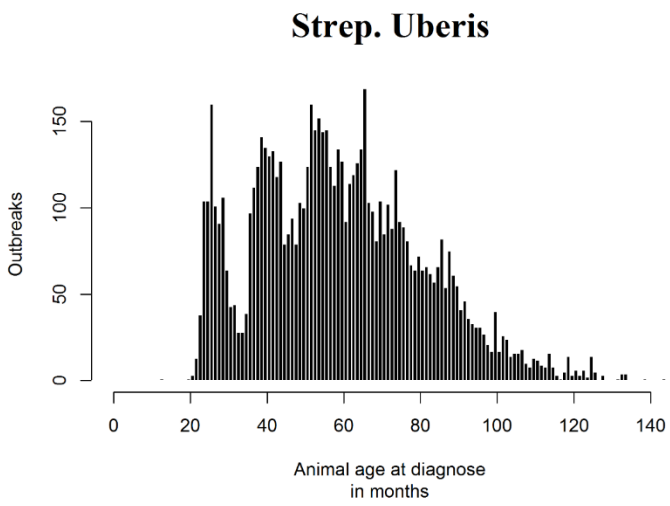


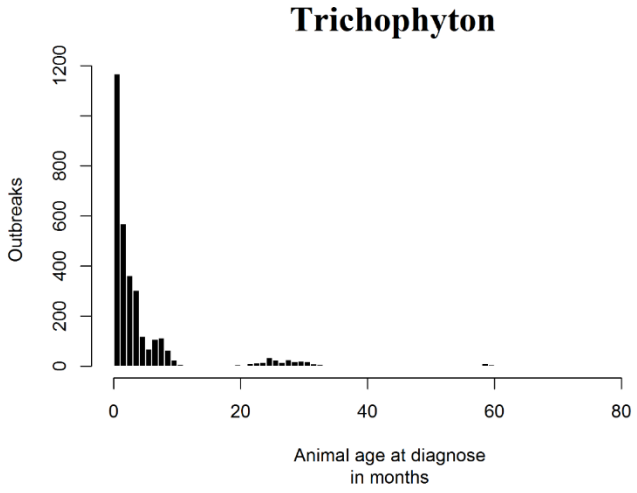
Strep. Dysgalactiae



Staph. Haemolyticus

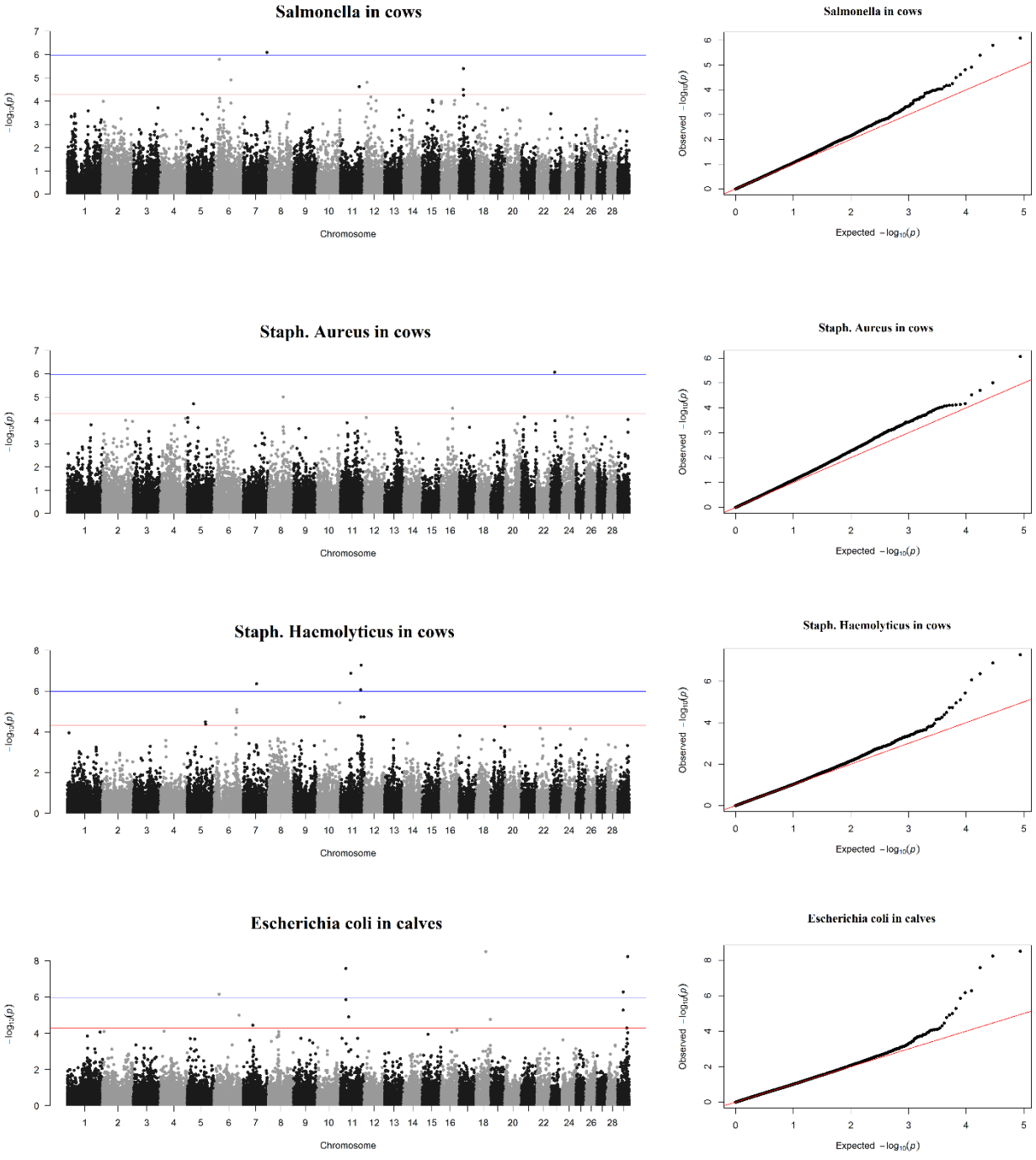


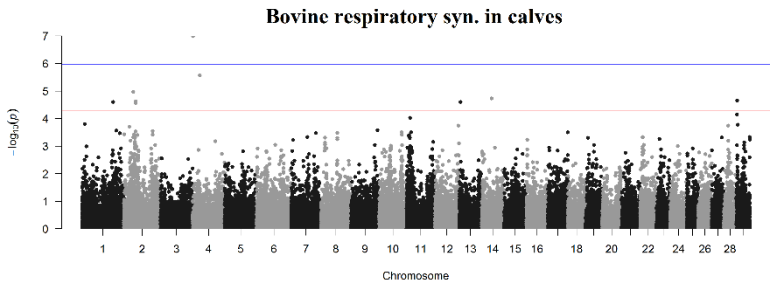

Strep. Uberis
Outbreaks
0
50
100
150
0
20
40
60
80
100
120
140
Animal age at diagnose
in months

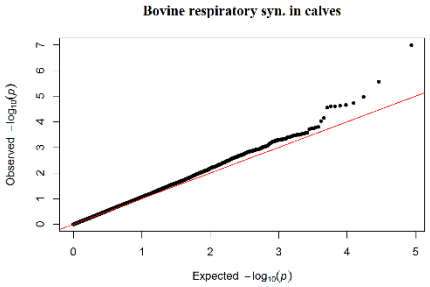

Trichophyton
Outbreaks
0
200
400
600
800
1200
0
20
40
60
80
Animal age at diagnose
in months

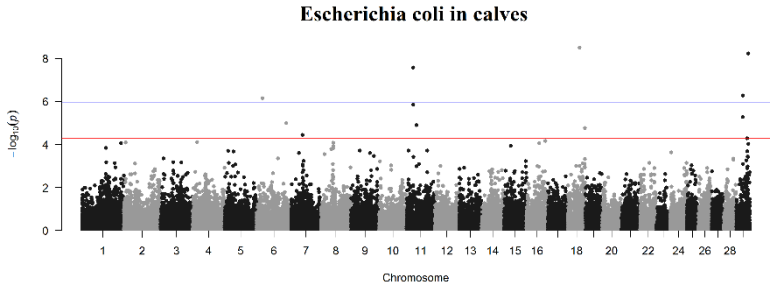
APPENDIX – C

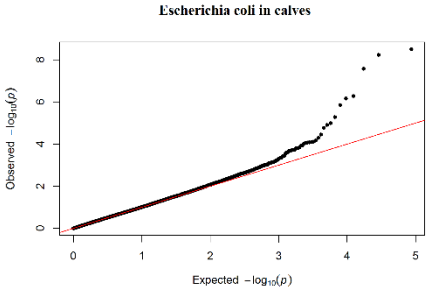
Manhattan plots for SNP allele substitution effects for resistance to infectious diseases in calves and cows. The red line indicates the genome-wide significance threshold value for the 50K (~50,000) genotypes (P-value = 5×10^{-5}). The blue line indicates the genome-wide significance threshold value based on Bonferroni threshold at α = 5% (P-value = $.05/5\times10^{4}$).

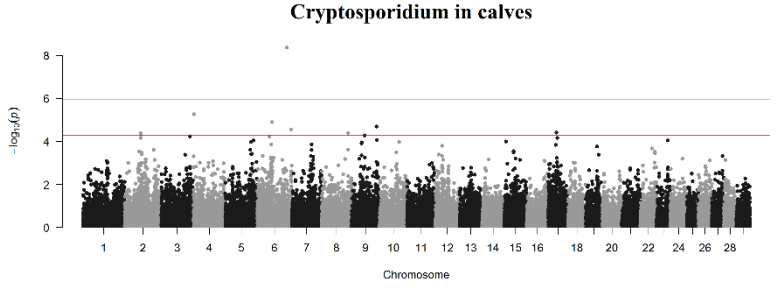


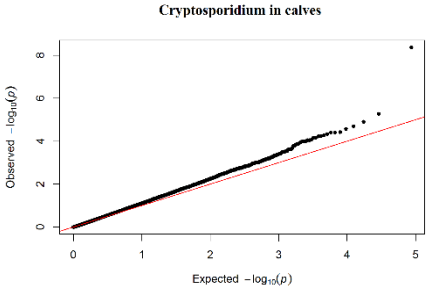

Bovine respiratory syn. in calves
$-\log_{10}(p)$
Chromosome

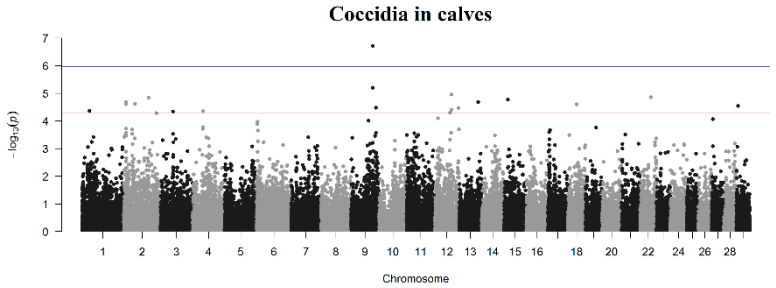

Bovine respiratory syn. in calves
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

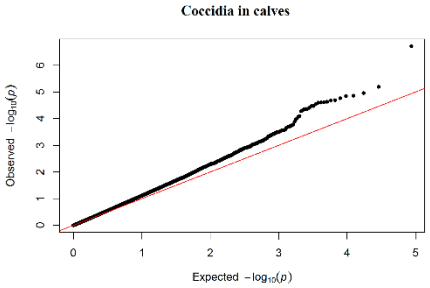

Escherichia coli in calves
$-\log_{10}(p)$
Chromosome

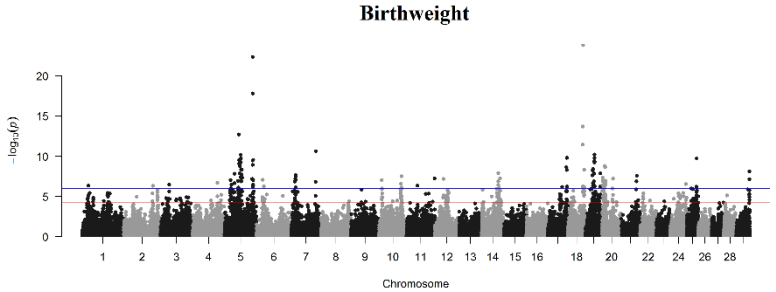

Escherichia coli in calves
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

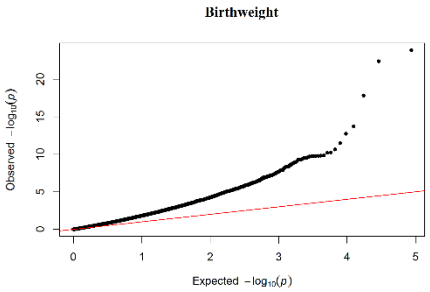

Cryptosporidium in calves
$-\log_{10}(p)$
Chromosome

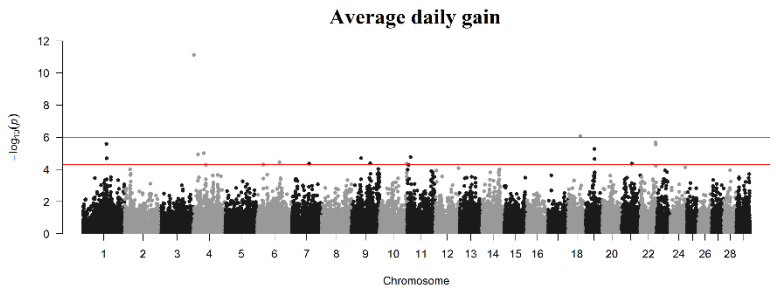

Cryptosporidium in calves
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

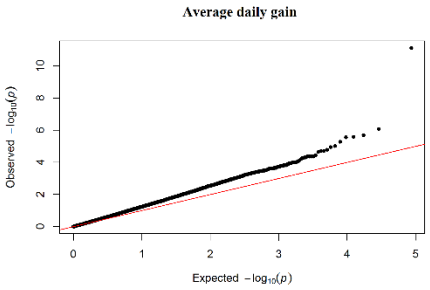

Coccidia in calves
$-\log_{10}(p)$
Chromosome

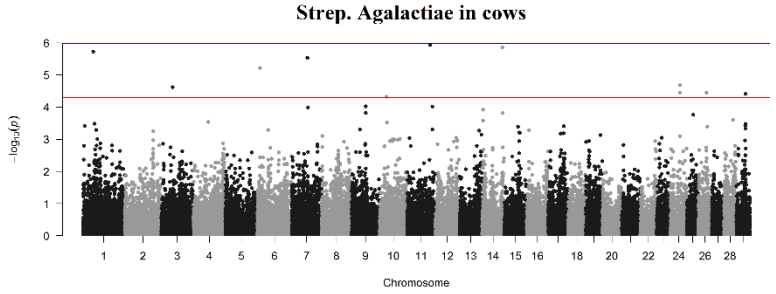

Coccidia in calves
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

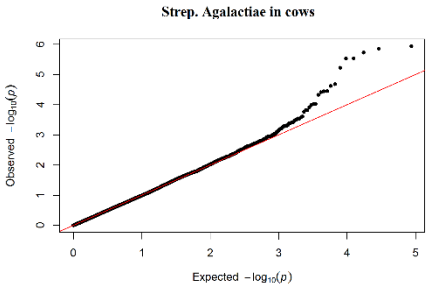

Birthweight
$-\log_{10}(p)$
Chromosome

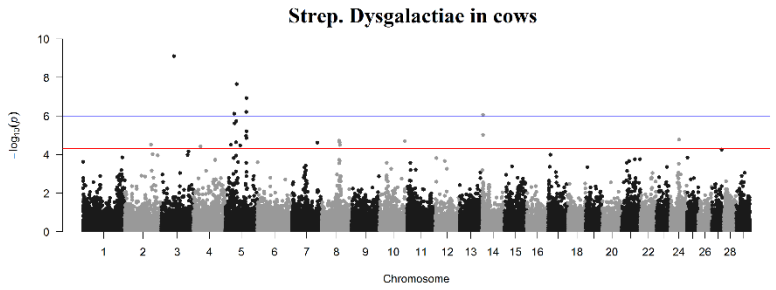

Birthweight
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

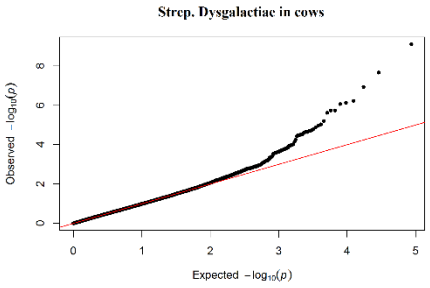

Average daily gain
$-\log_{10}(p)$
Chromosome

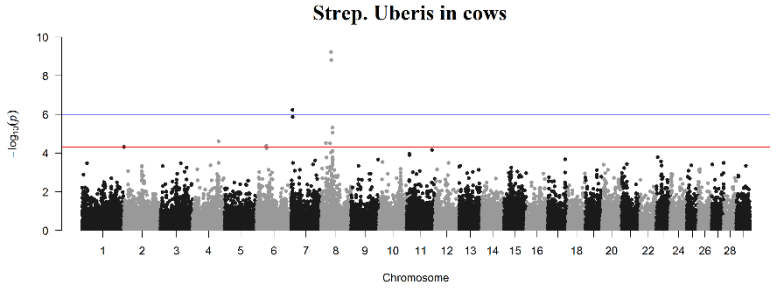

Average daily gain
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

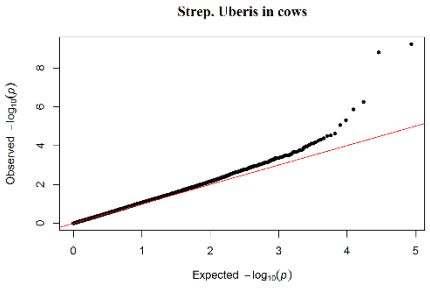

Strep. Agalactiae in cows
$-\log_{10}(p)$
Chromosome

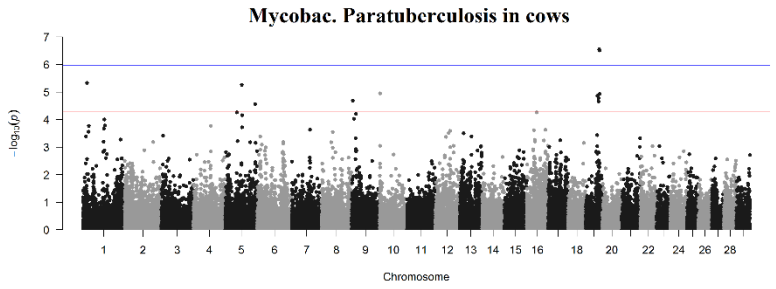

Strep. Agalactiae in cows
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

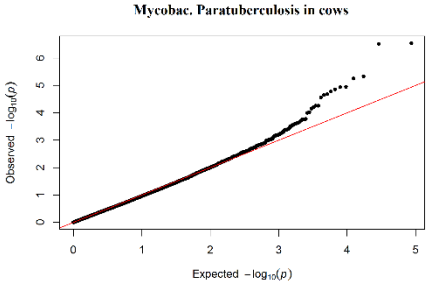

Strep. Dysgalactiae in cows
$-\log_{10}(p)$
Chromosome

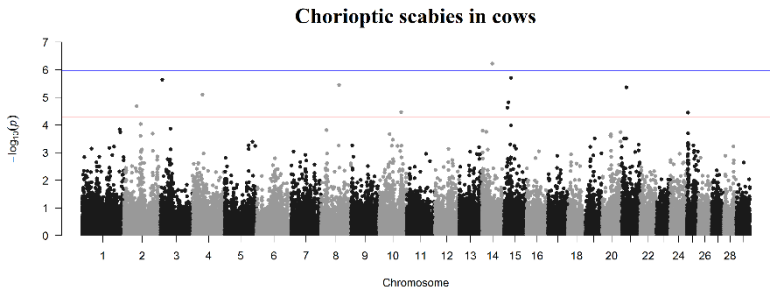

Strep. Dysgalactiae in cows
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

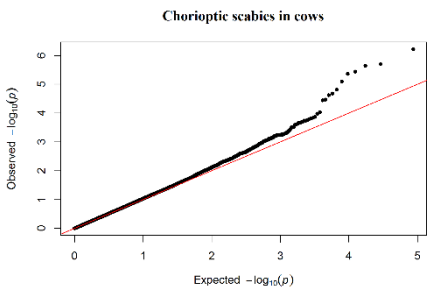

Strep. Uberis in cows
$-\log_{10}(p)$
Chromosome

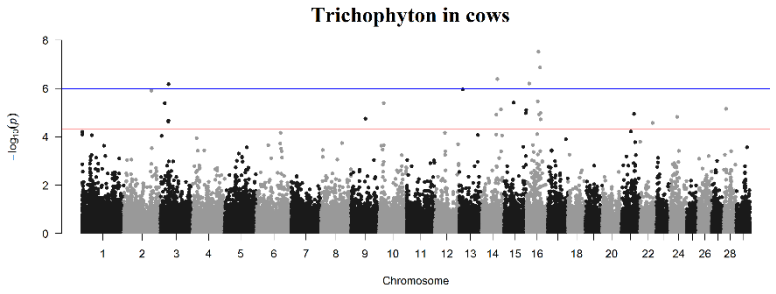

Strep. Uberis in cows
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

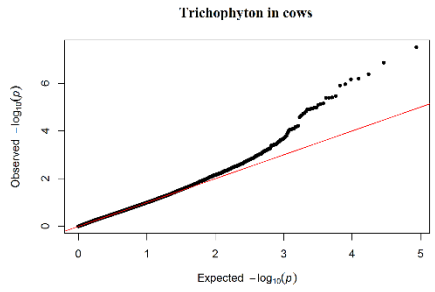

Mycobac. Paratuberculosis in cows
$-\log_{10}(p)$
Chromosome

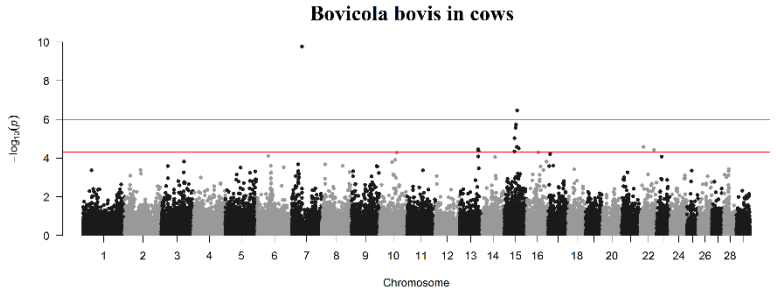

Mycobac. Paratuberculosis in cows
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

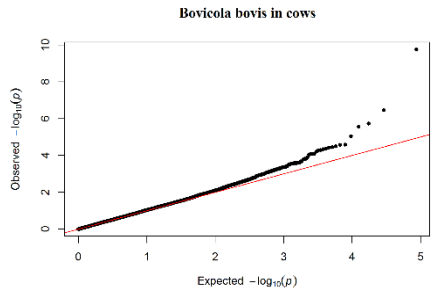

Chorioptic scabies in cows
Chromosome

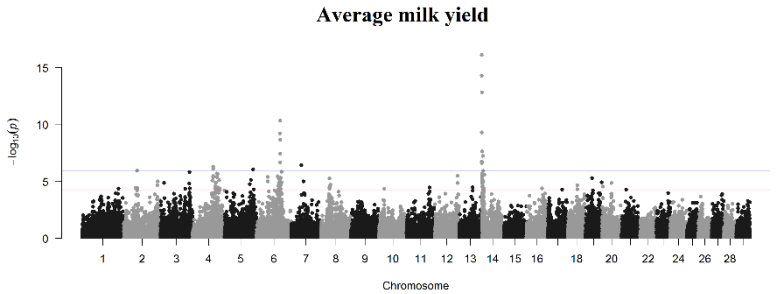

Chorioptic scabies in cows
Expected

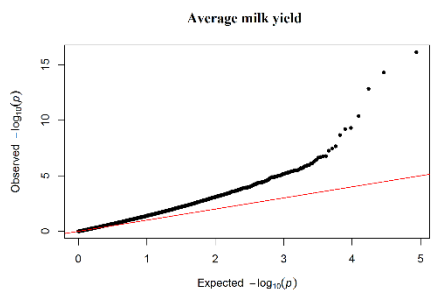

Trichophyton in cows
Chromosome

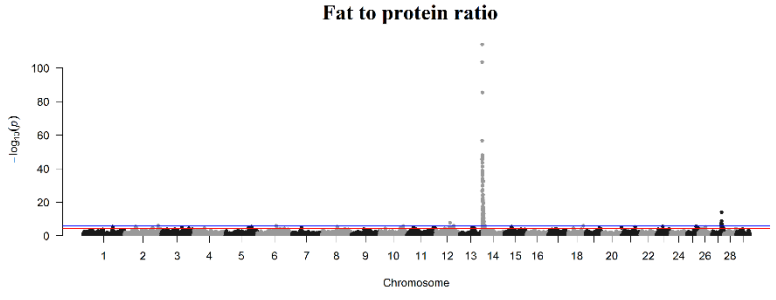

Trichophyton in cows
Expected

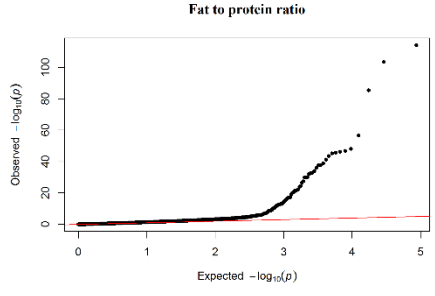

Bovicola bovis in cows
Chromosome


Bovicola bovis in cows
Expected


Average milk yield
Chromosome


Average milk yield
Expected


Fat to protein ratio
Chromosome


Fat to protein ratio
Expected

SUPPORTING INFORMATION

S1 Table. Genetic correlations among the 9 calf traits analyzed by G-REML

	Salmonella	Escherichia coli	Bovine respiratory syn.	Bovine herpes virus 1	Trichophyton	Cryptosporidium	Coccidia	Myiasis	Bovicola bovis
Salmonella	1.0000	-0.4052	-0.2032	-0.1409	0.5550	0.9850	0.0517	-0.1083	0.0828
Escherichia coli	-0.4052	1.0000	0.2175	0.1471	0.0005	0.3019	0.0144	-0.0427	-0.4028
Bovine respiratory syn.	-0.2032	0.2175	1.0000	0.1194	-0.1686	-0.2006	0.7602	0.0475	-0.1446
Bovine herpes virus 1	-0.1409	0.1471	0.1194	1.0000	-0.1052	0.1533	0.0806	-0.2394	0.7408
Trichophyton	0.5550	0.0005	-0.1686	-0.1052	1.0000	0.0014	0.0000	-0.2752	0.0476
Cryptosporidium	0.9850	0.3019	-0.2006	0.1533	0.0014	1.0000	0.5200	0.0596	-0.2131
Coccidia	0.0517	0.0144	0.7602	0.0806	0.0000	0.5200	1.0000	0.2589	-0.1076
Myiasis	-0.1083	-0.0427	0.0475	-0.2394	-0.2752	0.0596	0.2589	1.0000	-0.5380
Bovicola bovis	0.0828	-0.4028	-0.1446	0.7408	0.0476	-0.2131	-0.1076	-0.5380	1.0000

S1_pv Table. p-values for genetic correlations among the 9 calf traits analyzed by G-REML

	Salmonella	Escherichia coli	Bovine respiratory syn.	Bovine herpes virus 1	Trichophyton	Cryptosporidium	Coccidia	Myiasis	Bovicola bovis
Salmonella	1.00000	0.21680	0.04224	0.15480	0.00001	0.00001	0.31460	0.19040	0.29320
Escherichia coli	0.21680	1.00000	0.15490	0.27390	0.49890	0.04522	0.46440	0.41450	0.05614
Bovine respiratory syn.	0.04224	0.15490	1.00000	0.19590	0.10320	0.02887	0.00001	0.36840	0.19130
Bovine herpes virus 1	0.15480	0.27390	0.19590	1.00000	0.17600	0.11390	0.18110	0.20560	0.00004
Trichophyton	0.00001	0.49890	0.10320	0.17600	1.00000	0.49410	0.99000	0.03011	0.39370
Cryptosporidium	0.00001	0.04522	0.02887	0.11390	0.49410	1.00000	0.00001	0.32190	0.07943
Coccidia	0.31460	0.46440	0.00001	0.18110	0.99000	0.00001	1.00000	0.01507	0.23680
Myiasis	0.19040	0.41450	0.36840	0.20560	0.03011	0.32190	0.01507	1.00000	0.00001
Bovicola bovis	0.29320	0.05614	0.19130	0.00004	0.39370	0.07943	0.23680	0.00001	1.00000

S1_se Table. Standard error for genetic correlations among the 9 calf traits analyzed by G-REML

	Salmonella	Escherichia coli	Bovine respiratory syn.	Bovine herpes virus 1	Trichophyton	Cryptosporidium	Coccidia	Myiasis	Bovicola bovis
Salmonella	0.00000	0.13754	0.11903	0.12632	0.06941	0.00572	0.08731	0.11903	0.14357
Escherichia coli	0.13754	0.00000	0.16625	0.25121	0.16138	0.16678	0.16436	0.20233	0.24341
Bovine respiratory syn.	0.11903	0.16625	0.00000	0.16402	0.10824	0.10217	0.06155	0.14148	0.16584
Bovine herpes virus 1	0.12632	0.25121	0.16402	0.00000	0.12216	0.14252	0.13137	0.17570	0.18802
Trichophyton	0.06941	0.16138	0.10824	0.12216	0.00000	0.09149	0.08701	0.09269	0.11737
Cryptosporidium	0.00572	0.16678	0.10217	0.14252	0.09149	0.00000	0.06161	0.12826	0.15017
Coccidia	0.08731	0.16436	0.06155	0.13137	0.08701	0.06161	0.00000	0.11646	0.14477
Myiasis	0.11903	0.20233	0.14148	0.17570	0.09269	0.12826	0.11646	0.00000	0.05856
Bovicola bovis	0.14357	0.24341	0.16584	0.18802	0.11737	0.15017	0.14477	0.05856	0.00000

S2 Table. Genetic correlations among the 14 cow traits analyzed by G-REML

	Salmonella	Escherichia coli	Staph. Aureus	Staph. Haemolyticus	Strep. Agalactiae	Strep. Dysgalactiae	Strep. Uberis	Clostridium perfringens	Mycobac. Paratuberculosis	Rotavirus	Trichophyton	Dictyocaulus viviparus	Bovicola bovis	Chorioptic scabies
Salmonella	1.0000	-0.0895	-0.0678	-0.2029	0.0955	-0.3980	-0.0692	0.3153	0.4459	0.1194	0.1371	0.1567	0.1807	-0.1616
Escherichia coli	-0.0895	1.0000	0.7132	-0.4137	0.5041	0.7786	0.4534	0.5291	0.1255	0.1438	0.1632	-0.0268	-0.0343	0.4756
Staph. Aureus	-0.0678	0.7132	1.0000	0.3684	0.3433	0.4347	0.5576	-0.1223	0.1393	0.2268	0.0349	0.1680	0.0697	0.4410
Staph. Haemolyticus	-0.2029	-0.4137	0.3684	1.0000	-0.2001	1.0000	-0.3083	-0.4047	0.3709	-0.1605	0.0207	-0.4536	0.1068	1.0000
Strep. Agalactiae	0.0955	0.5041	0.3433	-0.2001	1.0000	1.0000	0.7177	0.2922	0.2774	-0.0748	0.9606	0.0962	1.0000	-0.2483
Strep. Dysgalactiae	-0.3980	0.7786	0.4347	1.0000	1.0000	1.0000	0.2673	-0.1290	-0.2207	-0.1517	0.0639	-0.6848	-0.4593	1.0000
Strep. Uberis	-0.0692	0.4534	0.5576	-0.3083	0.7177	0.2673	1.0000	0.6117	-0.0303	0.1629	0.2284	-0.1900	0.2136	-0.1000
Clostridium perfringens	0.3153	0.5291	-0.1223	-0.4047	0.2922	-0.1290	0.6117	1.0000	-1.0000	0.1086	0.7843	0.1079	-0.0719	-0.6852
Mycobac. Paratuberculosis	0.4459	0.1255	0.1393	0.3709	0.2774	-0.2207	-0.0303	-1.0000	1.0000	-0.2072	0.5153	-0.0841	0.1416	0.3233
Rotavirus	0.1194	0.1438	0.2268	-0.1605	-0.0748	-0.1517	0.1629	0.1086	-0.2072	1.0000	0.2176	0.4406	0.2743	-0.5243
Trichophyton	0.1371	0.1632	0.0349	0.0207	0.9606	0.0639	0.2284	0.7843	0.5153	0.2176	1.0000	-0.0968	-0.1728	-0.5088
Dictyocaulus viviparus	0.1567	-0.0268	0.1680	-0.4536	0.0962	-0.6848	-0.1900	0.1079	-0.0841	0.4406	-0.0968	1.0000	-0.1749	-0.4380
Bovicola bovis	0.1807	-0.0343	0.0697	0.1068	1.0000	-0.4593	0.2136	-0.0719	0.1416	0.2743	-0.1728	-0.1749	1.0000	0.3915
Chorioptic scabies	-0.1616	0.4756	0.4410	1.0000	-0.2483	1.0000	-0.1000	-0.6852	0.3233	-0.5243	-0.5088	-0.4380	0.3915	1.0000

S2_pv Table. p-values for Genetic correlations among the 14 cow traits analyzed by G-REML

	Salmonella	Escherichia coli	Staph. Aureus	Staph. Haemolyticus	Strep. Agalactiae	Strep. Dysgalactiae	Strep. Uberis	Clostridium perfringens	Mycobac. Paratuberculosis	Rotavirus	Trichophyton	Dictyocaulus viviparus	Bovicola bovis	Chorioptic scabies
Salmonella	1.00000000	0.31430000	0.29330000	0.07823000	0.35710000	0.02794000	0.29560000	0.14190000	0.03540000	0.13970000	0.21350000	0.09169000	0.18450000	0.10450000
Escherichia coli	0.31430000	1.00000000	0.00000005	0.03466000	0.05676000	0.00061520	0.00281900	0.08365000	0.34860000	0.16930000	0.26560000	0.44520000	0.44010000	0.06456000
Staph. Aureus	0.29330000	0.00000005	1.00000000	0.00624700	0.03711000	0.03579000	0.00000000	0.33920000	0.24900000	0.01085000	0.44470000	0.10730000	0.32980000	0.01531000
Staph. Haemolyticus	0.07823000	0.03466000	0.00624700	1.00000000	0.24810000	0.00003137	0.04074000	0.09084000	0.01573000	0.10240000	0.46420000	0.00917000	0.34900000	0.00003676
Strep. Agalactiae	0.35710000	0.05676000	0.03711000	0.24810000	1.00000000	0.00041290	0.00074810	0.29940000	0.23650000	0.35880000	0.00205000	0.36370000	0.00003951	0.25950000
Strep. Dysgalactiae	0.02794000	0.00061520	0.03579000	0.00003137	0.00041290	1.00000000	0.14320000	0.39400000	0.24490000	0.21850000	0.42010000	0.00903500	0.07495000	0.00956900
Strep. Uberis	0.29560000	0.00281900	0.00000000	0.04074000	0.00074810	0.14320000	1.00000000	0.03635000	0.45220000	0.07647000	0.11200000	0.09711000	0.13650000	0.34610000
Clostridium perfringens	0.14190000	0.08365000	0.33920000	0.09084000	0.29940000	0.39400000	0.03635000	1.00000000	0.10410000	0.31400000	0.03212000	0.36480000	0.42610000	0.02569000
Mycobac. Paratuberculosis	0.03540000	0.34860000	0.24900000	0.01573000	0.23650000	0.24490000	0.45220000	0.10410000	1.00000000	0.09529000	0.08959000	0.38260000	0.32240000	0.13820000
Rotavirus	0.13970000	0.16930000	0.01085000	0.10240000	0.35880000	0.21850000	0.07647000	0.31400000	0.09529000	1.00000000	0.14330000	0.00056590	0.07414000	0.00023900
Trichophyton	0.21350000	0.26560000	0.44470000	0.46420000	0.00205000	0.42010000	0.11200000	0.03212000	0.08959000	0.14330000	1.00000000	0.30560000	0.16310000	0.04369000
Dictyocaulus viviparus	0.09169000	0.44520000	0.10730000	0.00917000	0.36370000	0.00903500	0.09711000	0.36480000	0.38260000	0.00056590	0.30560000	1.00000000	0.19630000	0.01344000
Bovicola bovis	0.18450000	0.44010000	0.32980000	0.34900000	0.00003951	0.07495000	0.13650000	0.42610000	0.32240000	0.07414000	0.16310000	0.19630000	1.00000000	0.10460000
Chorioptic scabies	0.10450000	0.06456000	0.01531000	0.00003676	0.25950000	0.00956900	0.34610000	0.02569000	0.13820000	0.00023900	0.04369000	0.01344000	0.10460000	1.00000000

S2_se Table. Standard error for Genetic correlations among the 14 cow traits analyzed by G-REML

	Salmonella	Escherichia coli	Staph. Aureus	Staph. Haemolyticus	Strep. Agalactiae	Strep. Dysgalactiae	Strep. Uberis	Clostridium perfringens	Mycobac. Paratuberculosis	Rotavirus	Trichophyton	Dictyocaulus viviparus	Bovicola bovis	Chorioptic scabies
Salmonella	0.00000000	0.17930700	0.11988600	0.14774200	0.27210300	0.21486400	0.12722400	0.29564300	0.26783500	0.11155500	0.17214400	0.12188600	0.18843500	0.13253200
Escherichia coli	0.17930700	0.00000000	0.14602200	0.19935800	0.31877600	0.22788300	0.17160600	0.34726900	0.32653800	0.13805700	0.23405500	0.17957200	0.21914400	0.22444700
Staph. Aureus	0.11988600	0.14602200	0.00000000	0.13656300	0.20994900	0.22060800	0.09941600	0.27086000	0.21678000	0.08872100	0.17733900	0.12525900	0.15051200	0.15923100
Staph. Haemolyticus	0.14774200	0.19935800	0.13656300	0.00000000	0.28681500	0.33974000	0.15444800	0.32323200	0.20262800	0.12277100	0.21910600	0.16893200	0.20131700	0.11148700
Strep. Agalactiae	0.27210300	0.31877600	0.20994900	0.28681500	0.00000000	0.27823500	0.27038400	0.54080300	0.42513500	0.18668400	0.33503300	0.27803900	0.36029800	0.36842300
Strep. Dysgalactiae	0.21486400	0.22788300	0.22060800	0.33974000	0.27823500	0.00000000	0.21634400	0.47359200	0.36043000	0.19735700	0.29507300	0.29607600	0.28465100	0.18360000
Strep. Uberis	0.12722400	0.17160600	0.09941600	0.15444800	0.27038400	0.21634400	0.00000000	0.31306600	0.26219200	0.11367400	0.18235200	0.13404100	0.18037300	0.17623300
Clostridium perfringens	0.29564300	0.34726900	0.27086000	0.32323200	0.54080300	0.47359200	0.31306600	0.00000000	0.23754900	0.22824100	0.36920000	0.31051700	0.37034200	0.35772400
Mycobac. Paratuberculosis	0.26783500	0.32653800	0.21678000	0.20262800	0.42513500	0.36043000	0.26219200	0.23754900	0.00000000	0.18200600	0.39253900	0.28370600	0.32191700	0.31085300
Rotavirus	0.11155500	0.13805700	0.08872100	0.12277100	0.18668400	0.19735700	0.11367400	0.22824100	0.18200600	0.00000000	0.15635900	0.12962900	0.13398100	0.15565500
Trichophyton	0.17214400	0.23405500	0.17733900	0.21910600	0.33503300	0.29507300	0.18235200	0.36920000	0.39253900	0.15635900	0.00000000	0.18431600	0.19529100	0.25033400
Dictyocaulus viviparus	0.12188600	0.17957200	0.12525900	0.16893200	0.27803900	0.29607600	0.13404100	0.31051700	0.28370600	0.12962900	0.18431600	0.00000000	0.18621600	0.16662600
Bovicola bovis	0.18843500	0.21914400	0.15051200	0.20131700	0.36029800	0.28465100	0.18037300	0.37034200	0.32191700	0.13398100	0.19529100	0.18621600	0.00000000	0.24293800
Chorioptic scabies	0.13253200	0.22444700	0.15923100	0.11148700	0.36842300	0.18360000	0.17623300	0.35772400	0.31085300	0.15565500	0.25033400	0.16662600	0.24293800	0.00000000

S3 Table. Genetic correlations between resistance to the same pathogen in calves and cows analyzed by G-REML

	rg±se	p-value
	Salmonella in cow	
Salmonella in calf	`-0.26±0.09	0.00
	Trichophyton in cow	
Trichophyton in calf	0.18±0.12	0.05

S4 Table. Genetic correlations between all calf resistance traits and either birth weight or average daily gain in calves.

	Birthweight			Average Growth Rate		
	rg	se	p-value	rg	se	p-value
Salmonella	0.00	0.05	0.47	-0.31	0.07	0.00
Escherichia coli	0.06	0.11	0.28	-0.37	0.14	0.01
Bovine respiratory syn.	-0.12	0.07	0.04	-0.24	0.08	0.00
Bovine herpes virus 1	-0.18	0.08	0.01	-0.18	0.11	0.06
Trichophyton	-0.10	0.05	0.02	-0.10	0.07	0.06
Cryptosporidium	0.02	0.06	0.35	0.17	0.07	0.01
Coccidia	-0.09	0.05	0.05	0.09	0.07	0.09
Myiasis	0.09	0.07	0.10	0.05	0.09	0.32
Bovicola bovis	-0.18	0.09	0.02	-0.12	0.11	0.15

S5 Table. Genetic correlations between all cow resistance traits and either milk yield or fat to protein ratio.

	Avergae Milk Yield			Fat to protein ratio		
	rg	se	p-value	rg	se	p-value
Salmonella	0.09	0.09	0.15	-0.14	0.08	0.05
Escherichia coli	-0.02	0.12	0.45	0.05	0.11	0.32
Staph. Aureus	0.23	0.07	0.00	0.17	0.07	0.01
Staph. Haemolyticus	0.14	0.10	0.09	-0.13	0.10	0.09
Strep. Agalactiae	0.26	0.18	0.07	-0.06	0.17	0.37
Strep. Dysgalactiae	-0.14	0.15	0.19	0.03	0.15	0.42
Strep. Uberis	0.21	0.09	0.01	-0.04	0.09	0.33
Clos. Perfringens	0.19	0.17	0.14	0.12	0.17	0.25
Mycobac. Paratuberculosis	0.01	0.17	0.48	-0.27	0.17	0.05
Rotavirus	0.11	0.06	0.05	0.11	0.06	0.03
Trichophyton	-0.20	0.13	0.07	0.23	0.12	0.03
Dictyocaulus viviparus	-0.09	0.09	0.17	-0.09	0.09	0.18
Bovicola bovis	0.03	0.11	0.40	-0.08	0.10	0.23
Chorioptic scabies	0.01	0.12	0.47	-0.21	0.11	0.03

S6 Table. P-values for the candidate 20 genes and their average across all resistance traits in calves.

GENES	Salmonella	Escherichia coli	Bovine respiratory syn.	Bovine herpes virus 1	Trichophyton	Cryptosporidium	Coccidia	Myiasis	Bovicola bovis	average
RRM2B	0.26859	0.089055	0.11793	0.45461	0.65545	0.69188	0.22954	0.19906	0.14449	0.3167339
TAPBP	0.34604	0.35999	0.93837	0.90786	0.029303	0.22265	0.29926	0.056335	0.00059918	0.3511564
TRIM26	0.34604	0.35999	0.93837	0.90786	0.029303	0.22265	0.29926	0.056335	0.00059918	0.3511564
ZNF527	0.16508	0.53242	0.32451	0.088463	0.48215	0.56997	0.52815	0.29985	0.087357	0.3419944
CLASRP	0.10827	0.12103	0.035917	0.73979	0.050476	0.45912	0.15029	0.36995	0.073461	0.2342560
FAM83C	0.10827	0.12103	0.035917	0.73979	0.050476	0.45912	0.15029	0.36995	0.073461	0.2342560
HVCN1	0.10827	0.12103	0.035917	0.73979	0.050476	0.45912	0.15029	0.36995	0.073461	0.2342560
ODF1	0.42448	0.37305	0.19361	0.52158	0.14369	0.014857	0.26629	0.066781	0.00015851	0.2227218
LRRC17	0.38982	0.37378	0.37433	0.37081	0.23917	0.60628	0.47505	0.076145	0.12159	0.3363306
HAR1A	0.08169	0.64702	0.23642	0.50272	0.29816	0.0045598	0.0087639	0.2845	0.16401	0.2475382
MATN2	0.3508	0.33396	0.14885	0.37445	0.25037	0.21045	0.2271	0.016993	0.58999	0.2781070
RD30	0.69617	0.55464	0.59512	0.31293	0.2402	0.3301	0.23491	0.052439	0.097613	0.3460136
EPS8L2	0.15584	0.57725	0.022641	0.72137	0.52511	0.01161	0.14583	0.40157	0.15246	0.3015201
ROBO1	0.11744	0.76669	0.15639	0.062243	0.5067	0.24613	0.82772	0.10511	0.11853	0.3229948
GDPD3	0.18484	0.48047	0.79621	0.15742	0.47756	0.5732	0.28214	0.081451	0.077479	0.3456411
DDAH1	0.67103	0.49631	0.20517	0.17176	0.27111	0.29458	0.15727	0.011845	0.0069182	0.2539992
DD3051	0.16654	0.20024	0.13945	0.40631	0.32308	0.23311	0.17473	0.76091	0.44913	0.3170556
PHACTR3	0.30413	0.24537	0.075493	0.7684	0.004768	0.083934	0.018805	0.77932	0.42088	0.3001222
TMEM266	0.40222	0.64529	0.57915	0.43093	0.1413	0.45141	0.068665	0.069891	0.56321	0.3724518
PLEKHA4	0.72569	0.13163	0.18023	0.28799	0.19521	0.45111	0.32647	0.43158	0.33389	0.3404222

S7 Table. P-values for the candidate 20 genes and their average across all resistance traits in cows.

GENES	Salmonella	Escherichia coli	Staph. Aureus	Staph. Haemolyticus	Strep. Agalactiae	Strep. Dysgalactiae	Strep. Uberis	Clostridium perfringens	Mycobac. Paratuberculosis	Rotavirus	Trichophyton	Dictyocaulus viviparus	Bovicola bovis	Chorioptic scabies	average
RRM2B	0.088603	0.65044	0.090584	0.1721	0.51009	0.29216	0.21374	0.05784	0.36585	0.065931	0.35246	0.14935	0.28877	0.33055	0.2591763
TAPBP	0.088536	0.23732	0.08013	0.299	0.13555	0.083628	0.060067	0.82712	0.24188	0.60375	0.6281	0.19454	0.035759	0.51771	0.2880779
TRIM26	0.088536	0.23732	0.08013	0.299	0.13555	0.083628	0.060067	0.82712	0.24188	0.60375	0.6281	0.19454	0.035759	0.51771	0.2880779
ZNF527	0.034665	0.19564	0.069157	0.27791	0.12189	0.090902	0.47352	0.93747	0.025831	0.2855	0.69997	0.072147	0.39339	0.55666	0.3024751
CLASRP	0.1426	0.87455	0.015463	0.55133	0.68735	0.36743	0.20933	0.12267	0.37394	0.18302	0.268	0.90729	0.36193	0.15492	0.3728445
FAM83C	0.1426	0.87455	0.015463	0.55133	0.68735	0.36743	0.20933	0.12267	0.37394	0.18302	0.268	0.90729	0.36193	0.15492	0.3728445
HVCN1	0.1426	0.87455	0.015463	0.55133	0.68735	0.36743	0.20933	0.12267	0.37394	0.18302	0.268	0.90729	0.36193	0.15492	0.3728445
ODF1	0.14944	0.2371	0.19205	0.63285	0.19003	0.13575	0.47802	0.28069	0.35281	0.58153	0.97448	0.41147	0.58953	0.1575	0.3830893
LRRC17	0.634	0.054235	0.10895	0.31098	0.444	0.29367	0.555	0.10988	0.068993	0.24126	0.22006	0.31875	0.61126	0.37581	0.3104891
HAR1A	0.11279	0.69552	0.14435	0.38544	0.89407	0.2986	0.67386	0.094468	0.25953	0.55191	0.43409	0.27579	0.31421	0.071126	0.3718396
MATN2	0.20138	0.60791	0.061021	0.6434	0.091944	0.60523	0.82535	0.071212	0.67128	0.61463	0.19835	0.078816	0.099024	0.18836	0.3541362
RD30	0.89346	0.17346	0.054638	0.23843	0.52198	0.42038	0.78023	0.29538	0.14454	0.17505	0.43869	0.14619	0.2092	0.018139	0.3221262
EPS8L2	0.25485	0.88321	0.14045	0.45385	0.012855	0.34315	0.49644	0.36168	0.94882	0.14307	0.50613	0.06803	0.12294	0.22635	0.3544161
ROBO1	0.42795	0.23893	0.16762	0.094745	0.036399	0.17644	0.22639	0.80676	0.51074	0.48638	0.40913	0.17889	0.30174	0.73289	0.3425003
GDPD3	0.22949	0.5954	0.24166	0.36553	0.093575	0.035028	0.91888	0.064854	0.50815	0.087327	0.30901	0.4084	0.24631	0.53728	0.3314924
DDAH1	0.12576	0.52914	0.6908	0.35642	0.64526	0.26644	0.37408	0.048726	0.39926	0.71262	0.16465	0.40238	0.65901	0.10311	0.3912611
DD3051	0.24998	0.68345	0.23952	0.223	0.030022	0.0059381	0.45998	0.59391	0.49648	0.63862	0.38265	0.43848	0.42438	0.053436	0.3514176
PHACTR3	0.091279	0.053313	0.42913	0.28246	0.90329	0.468	0.77692	0.25714	0.50367	0.11124	0.18369	0.33922	0.13735	0.54426	0.3629259
TMEM266	0.13325	0.10444	0.62295	0.20201	0.11437	0.065576	0.66541	0.42808	0.066577	0.41377	0.38572	0.042888	0.42908	0.78249	0.3183294
PLEKHA4	0.51449	0.93547	0.64351	0.36246	0.50026	0.43053	0.22618	0.22956	0.46494	0.15868	0.78047	0.44977	0.12369	0.7092	0.4663721

S8 Table. P-values for the most important 20 pathways and their average across all resistance traits in calves.

Pathways	Salmonella	Escherichia coli	Bovine respiratory syn.	Bovine herpes virus 1	Trichophyton	Cryptosporidium	Coccidia	Myiasis	Bovicola bovis	Average
Reactome pre-notch transcription and translation	2.4308E-06	9.7893E-06	0.002273	0.00079731	0.015709	0.024639	0.063456	0.0088379	0.0083893	0.013790414
Biocarta B-lymphocyte pathway	0.000020051	0.069108	0.15574	0.061525	0.038928	0.036375	0.018815	0.016372	0.020379	0.04636245
Biocarta Cell cycle pathway	0.00011675	0.22572	0.042097	0.064333	0.10518	0.32234	0.08371	0.045446	0.059624	0.105396306
KEGG Graft versus host disease	0.019906	0.27239	0.20962	0.067915	0.15965	0.63295	0.030149	0.04629	0.09647	0.170593333
Biocarta Extrinsic pathway	0.041474	0.30283	0.41536	0.16447	0.25832	0.20204	0.14077	0.24828	0.17822	0.216862667
Reactome nuclear signaling by ERBB4	0.022446	0.30554	0.34867	0.1239	0.15591	0.63132	0.23528	0.24841	0.12016	0.243515111
Biocarta IL22bp pathway	0.074125	0.32177	0.42462	0.17887	0.27838	0.5709	0.21718	0.40292	0.1694	0.293129444
Biocarta IL7 pathway	0.1218	0.32289	0.47148	0.29866	0.29933	0.32584	0.14891	0.52469	0.1591	0.296966667
KEGG Primary immunodeficiency	0.22679	0.46534	0.5563	0.51944	0.40084	0.48651	0.48667	0.56537	0.25226	0.439946667
Reactome common pathway	0.1437	0.39857	0.47978	0.43185	0.39011	0.61268	0.29772	0.53044	0.22528	0.390014444
Reactome Antiviral mechanism by IFN stimulated genes	0.38335	0.53401	0.86233	0.64116	0.43149	0.26072	0.46573	0.60655	0.457	0.515815556
KEGG Vibrio Cholerae infection	0.35626	0.45595	0.70175	0.56798	0.41132	0.37631	0.48071	0.58305	0.31796	0.472365556
KEGG Leishmania infection	0.43003	0.62765	0.63547	0.76829	0.43825	0.694	0.55938	0.69368	0.78863	0.626153333
KEGG Intestinal immune network for IGA production	0.38945	0.56464	0.78119	0.67218	0.44379	0.66933	0.45108	0.64348	0.65207	0.585245556
KEGG T-cell receptor signaling pathway	0.48756	0.67504	0.8378	0.76494	0.4089	0.72299	0.75452	0.71329	0.9333	0.699815556
Reactome APOBEC3G mediated resistance to HIV1 infection	0.54544	0.67122	0.63081	0.80551	0.537	0.75856	0.95213	0.76495	0.98363	0.738805556
Reactome latent infection of homo sapiens with mycobacterium tuberculosis	0.56941	0.77526	0.83566	0.89114	0.57129	0.72969	0.63795	0.81371	0.77386	0.733107778
Reactome translocation of zap 70 to immunological synapse	0.64213	0.7856	0.78966	0.90554	0.8423	0.88517	0.77763	0.82791	0.99837	0.828256667
KEGG Pathogenic Escherichia-coli infection	0.82522	0.97401	0.97517	0.89902	0.93528	0.88414	0.78002	0.95755	0.90677	0.904131111
Reactome immunoregulatory interactions between a lymphoid and a non-lymphoid cell	0.69915	0.97816	0.99893	0.91559	0.98234	0.92213	0.7892	0.92993	0.67315	0.876508889

S9 Table. P-values for the most important 20 pathways and their average across all resistance traits in cows.

Pathways	Salmonella	Escherichia coli	Staph. Aureus	Staph. Haemolyticus	Strep. Agalactiae	Strep. Dysgalactiae	Strep. Uberis	Clostridium perfringens	Mycobac. Paratuberculosis	Rotavirus	Trichophyton	Dictyocaulus viviparus	Bovicola bovis	Chorioptic scabies	Average
Reactome pre-notch transcription and translation	0.74594	0.014893	0.49738	0.000040635	0.02444	0.029709	0.097002	0.000024432	5.0433E-06	0.02931	0.00086113	0.0105	0.03542	0.029929	0.108246731
Biocarta B-lymphocyte pathway	0.15878	0.02894	0.92604	0.018786	0.050133	0.089777	0.29431	0.027686	0.000071998	0.056063	0.11244	0.046084	0.14984	0.077424	0.145455357
Biocarta Cell cycle pathway	0.22462	0.05119	0.023953	0.48307	0.11252	0.43575	0.36756	0.0060457	0.0092438	0.12729	0.25765	0.058557	0.12671	0.1614	0.174682821
KEGG Graft versus host disease	0.29368	0.11039	0.79752	0.17685	0.14849	0.38065	0.16671	0.079421	0.59903	0.13254	0.38758	0.085076	0.028686	0.19244	0.255647357
Biocarta Extrinsic pathway	0.34998	0.16355	0.41279	0.033994	0.26126	0.99205	0.41629	0.14597	0.39925	0.18825	0.15633	0.25121	0.24849	0.22836	0.303412429
Reactome nuclear signaling by ERBB4	0.095096	0.11582	0.87672	0.41734	0.16066	0.89889	0.23151	0.066422	0.33767	0.16519	0.32097	0.092761	0.15491	0.21968	0.2966885
Biocarta IL22bp pathway	0.51964	0.23579	0.042018	0.057481	0.21679	0.84831	0.057883	0.16775	0.54699	0.19332	0.42147	0.29436	0.34185	0.48337	0.316215857
Biocarta IL7 pathway	0.63583	0.22201	0.60872	0.61142	0.51163	0.76854	0.11635	0.53327	0.28451	0.26775	0.199	0.37087	0.38002	0.55233	0.433017857
KEGG Primary immunodeficiency	0.067279	0.26895	0.72893	0.53976	0.32599	0.94006	0.23947	0.53639	0.25416	0.25214	0.63202	0.42203	0.549	0.61757	0.455267786
Reactome common pathway	0.51033	0.29132	0.46628	0.67299	0.64875	0.83523	0.40815	0.48911	0.6104	0.27858	0.59325	0.24344	0.51541	0.57919	0.510173571
Reactome Antiviral mechanism by IFN stimulated genes	0.090802	0.37802	0.10882	0.33746	0.54515	0.84944	0.59218	0.29143	0.040895	0.36864	0.22957	0.8364	0.62159	0.88243	0.440916214
KEGG Vibrio Cholerae infection	0.70462	0.44029	0.59568	0.44486	0.49254	0.77842	0.52819	0.33635	0.54005	0.35842	0.5947	0.43246	0.55654	0.887	0.549294286
KEGG Leishmania infection	0.28479	0.9995	0.27891	0.52822	0.40238	0.30837	0.75797	0.70964	0.54697	0.42578	0.35257	0.53126	0.6702	0.73509	0.537975
KEGG Intestinal immune network for IGA production	0.64654	0.46303	0.8611	0.35243	0.60876	0.62164	0.65445	0.63606	0.4528	0.39234	0.69198	0.89903	0.62338	0.80661	0.622153571
KEGG T-cell receptor signaling pathway	0.80494	0.99351	0.26316	0.45737	0.28734	0.70379	0.85136	0.6545	0.47069	0.5473	0.15517	0.90949	0.67836	0.99865	0.626830714
Reactome APOBEC3G mediated resistance to HIV1 infection	0.88437	0.66365	0.19787	0.02905	0.53416	0.89435	0.90875	0.61016	0.40099	0.55149	0.65404	0.77542	0.74384	0.87666	0.6232
Reactome latent infection of homo sapiens with mycobacterium tuberculosis	0.8453	0.66146	0.08106	0.50765	0.7086	0.67248	0.91766	0.67509	0.37569	0.56483	0.77676	0.7594	0.78607	0.96484	0.664063571
Reactome translocation of zap 70 to immunological synapse	0.96087	0.93041	0.47126	0.60107	0.77769	0.6085	0.93579	0.76052	0.69729	0.69616	0.85881	0.95228	0.85163	0.90963	0.786565
KEGG Pathogenic Escherichia-coli infection	0.83422	0.61988	0.042437	0.51516	0.92306	0.9491	0.99266	0.79534	0.81669	0.87692	0.83387	0.62753	0.84024	0.99332	0.761459071
Reactome immunoregulatory interactions between a lymphoid and a non-lymphoid cell	0.85319	0.70511	0.78458	0.18774	0.97445	0.88942	0.99721	0.96022	0.73425	0.91282	0.90907	0.52593	0.90577	0.82974	0.797821429

DECLARATION

"Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der "Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis" niedergelegt sind, eingehalten."

I declare: this dissertation submitted is a work of my own, written without any illegitimate help by any third party and only with materials indicated in the dissertation. I have indicated in the text where I have used texts from already published sources, either word for word or in substance, and where I have made statements based on oral information given to me. At any time during the investigations carried out by me and described in the dissertation, I followed the principles of good scientific practice as defined in the "Statutes of the Justus Liebig University Giessen for the Safeguarding of Good Scientific Practice".

Full name:

Medhat Ahmed Abbas Mahmoud

Signature: