Breeding against infectious diseases in animals

Hamed Rashidi

Thesis committee

Promotor
Prof. Dr J.A.M. van Arendonk
Professor of Animal Breeding and Genomics Centre
Wageningen University, The Netherlands

Co-promotors
Dr H.A. Mulder
Assistant Professor, Animal Breeding and Genomics Centre
Wageningen University, The Netherlands

Dr P.K. Mathur
Senior Geneticist, Topigs Norsvin Research Center
The Netherlands

Other members (assessment committee)
Prof. Dr M.C.M. de Jong, Wageningen University, The Netherlands
Prof. Dr J.A. Stegeman, Utrecht University, The Netherlands
Prof. Dr J.K. Lunney, United States Department of Agriculture, USA
Dr B. Nielsen, Pig Research Center, Denmark

This research was conducted under the auspices of the Graduate School of Wageningen Institute of Animal Sciences (WIAS).

Breeding against infectious diseases in animals

Hamed Rashidi

Thesis

submitted in fulfillment of the requirements for the degree of doctor at
Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Acadamic Board
to be defended in public
on Friday March 18, 2016
at 16.00 p.m. in the Aula

Rashidi, H.
Breeding against infectious diseases in animals,
182 Pages.

PhD thesis, Wageningen University, Wageningen, NL (2016)
With references, with summary in English

ISBN 978-94-6257-645-2

Abstract

Infectious diseases in farm animals are of major concern because of welfare, production costs, and public health. Control strategies, however, are not always successful. Selective breeding for the animals that can defend against infections, therefore, could be an option. Defensive ability of animals against infections consists of resistance (ability to control pathogen burden) and tolerance (ability to maintain performance when pathogen burden increases). When it is difficult to distinguish between resistance and tolerance, defensive ability is measured as resilience that is the ability to maintain performance during an outbreak regardless of pathogen burden. The aims of this thesis were to: 1) estimate the genetic variation in resistance, tolerance, and resilience to infection in order to assess the amenability of these traits for selective breeding in farm animals, 2) estimate the genetic correlation between resistance, tolerance and resilience and 3) identify genomic regions associated with resistance, tolerance, and resilience. To assess the amenability of resistance and tolerance for selective breeding, we studied the genetic variances of resistance and tolerance to nematode infection in sheep. For resistance we used three indicators: faecal nematode egg count (FEC), pepsinogen, and IgA. Tolerance was measured as the reaction norm of body weight on FEC and pepsinogen. The heritabilities for resistance traits ranged from 0.19 to 0.59. There was a significant ($p<0.05$) genetic variation among sheep in tolerance. We also observed a trade-off between resistance and tolerance. To assess the amenability of resilience to selection, we studied variation of sows in the reproduction traits number of piglets born alive and number of piglets born dead before and after porcine reproductive and respiratory syndrome (PRRS) outbreaks. Trait correlations between healthy and disease phases deviated from unity and ranged from 0.57 to 0.87. The repeatabilities of the traits during healthy and disease phases ranged from 0.08 to 0.16. To study the response to selection in resistance and tolerance when using estimated breeding values for resilience we used Monte Carlo simulations along with selection index theory. Selection for resilience in absence of records for pathogen burden resulted in favourable responses in resistance and tolerance, with more emphasis on tolerance than on resistance. To identify genomic regions associated with resistance, tolerance and resilience we studied pigs that were experimentally diseased with PRRS. We identified common genomic regions associated with resistance and resilience to PRRS and other genomic regions (chromosome-wise significant) associated with tolerance to PRRS. From all the chapters in this thesis we conclude that there is genetic variation among animals in response to infection which can be utilized in breeding programs.

For Nikan and Liana

Contents

1

General Introduction

1.1 Costs of infectious diseases in farm animals

Infectious diseases in farm animals are of major concern because of animal welfare, production costs, and public health. Every year, farms undergo huge economic losses due to infectious disease. The costs of infections in farm animals are mainly due to animals' production losses, treatment of infected animals, and disease control strategies. Infectious diseases reduce the average farm production drastically. For example in the case of porcine reproductive and respiratory syndrome (PRRS) in pigs, the disease causes huge reduction in the reproductive performance of sows. Infected sows have reproduction failures such as piglet loss due to stillbirth, mummification and abortion. Because of the reproduction failures, the average production of the farm in terms of number of piglets born alive will decline dramatically. Figure 1.1 is the weekly average production of a commercial pig farm for number of piglets born alive (NBA) from 2004 to 2012. Almost every year the farm was infected with the PRRS virus and during the PRRS outbreak the weekly averages of NBA dropped 10 to 65%. The reduction in the weekly average of NBA means a huge loss in the profit from selling piglets by the farm. In the Netherlands, the mean cost of PRRS including production loss, medication, diagnosis, and labour was estimated to be €126 per sow per PPRS outbreak (18 weeks) (Nieuwenhuis *et al*., 2012), The annual cost of PRRS for US farmers was estimated to be $663.91 with an average of $2.36 reduction in profit per pig weaned per year and $2.24 reduction in profit per pig marketed per year (Holtkamp *et al*., 2013). In cattle, Shaw *et al*. (1998) reported that gastrointestinal nematode infection causes a 155 g/day decrease in weight gain of infected calves. Other studies have also reported 4-5 kg loss in milk production of dairy cattle infected with nematodes (Charlier *et al*., 2014). In sheep, gastrointestinal nematode infection may cause 10-47% reduction in bodyweight gain and up to 21% reduction in wool growth (Charlier *et al*., 2014).

The control strategies of the infectious diseases involve biosecurity, vaccination, sanitation, diagnosis, antibiotics, antiviral medicines, anthelminthic drugs, and culling. For PRRS in US Holtkamp *et al*. (2013) estimated the immunization, pharmaceutical, and diagnosis costs to be $1.71 per pig marketed per year and biosecurity and outbreak related costs to be $3.08 per pig marketed per year. The control strategies, however, are not always effective. For example for PRRS the current vaccines do not provide sustainable disease control due to the antigenic heterogeneity and various immune escaping strategies of the PRRS-virus (Renukaradhya *et al*., 2015; Thanawongnuwech and Suradhat, 2010). Therefore,

selective breeding of the animals that have the ability to mount a response against infection could potentially be a more sustainable approach.

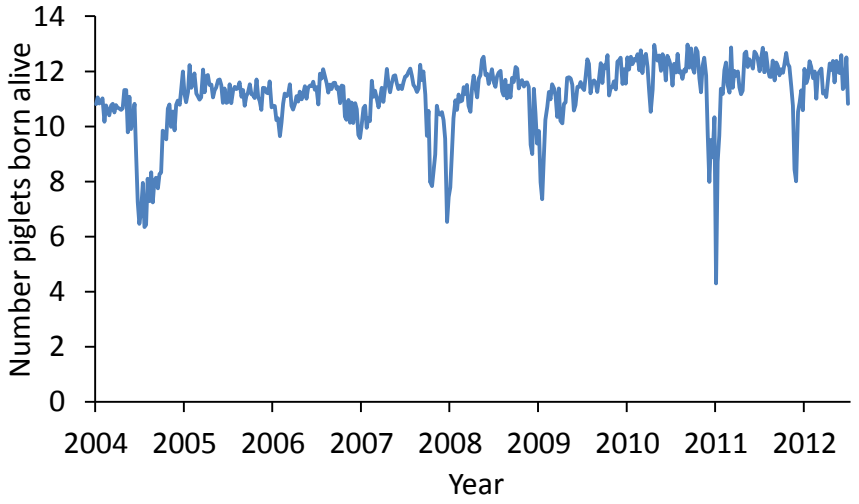

Figure 1.1. Weekly average of number of piglets born alive in a commercial pig farm from 2004 to 2012. The drops in the weekly averages are known to be due to PRRS outbreaks.

1.2 Response to infection: resistance, tolerance, resilience

Response to infections in animals involves two main mechanisms: resistance and tolerance. Resistance is defined as the ability of an animal to limit pathogen burden or resist against the pathogen and harbour less amount of pathogen by e.g. controlling the life cycle of the pathogen. Based on this, resistance could be defined as the inverse of the host's pathogen burden (Råberg *et al.*, 2007): i.e. the more resistant animals will have lower pathogen burden. The advantage of resistance to infection is that a resistant animal will not spread the infection. Resistance, therefore, could be especially helpful in control and eradication of highly infectious disease and zoonotic diseases (Bishop, 2012). Resistance, however, limits the survival and reproduction of the pathogen and as a consequence imposes selection advantages on the pathogens that can overcome resistance (Detilleux, 2011; Kause,

2011; Råberg *et al*., 2007). The selection advantage is more severe on small pathogens with large population size and short generation interval.
The genetic basis of resistance can be modelled as the genetic effect on the phenotype for pathogen burden (y_{PB}):

$$y_{PB} = \mu_{PB} + A_{PB} + e_{PB},$$

where μ_{PB} is the average pathogen burden, A_{PB} is the breeding value for pathogen burden, and e_{PB} is the environmental effect for pathogen burden.

Tolerance is defined as the ability of an individual to maintain its performance in spite of an increase of pathogen burden. In another words, a tolerant individual shows minimum symptoms of the disease despite the infection. The advantage of tolerance to infection is that it does not impose a selection pressure on the pathogen (Rausher, 2001; Read *et al*., 2008). This could be especially an advantage for the pathogens that can overcome the resistance mechanisms. Tolerance, however, does not stop the spread of the infection and therefore is not a suitable control approach for highly infectious pathogens and zoonotic diseases (Bishop, 2012). The genetic basis of tolerance can be modelled as the genetic effect on the reaction norm of animal's performance on its pathogen burden using random regression models (Kause, 2011):

$$y_{ij} = \mu + A_{0j} + A_{1j} \times PB_{ij} + e_{ij,}$$

where y_{ij} is performance of individual *i* from family *j*, μ is mean of population, A_{0j} is the breeding value for intercept for family *j*, A_{1j} is the breeding value for slope for family j, PB_{ij} is the phenotype of pathogen burden for individual *i* from family *j*, e_{ij} is the random error. Figure 1.2 is a schematic illustration of the reaction norm of two individuals performance on their pathogen burden. Individual 1 has a lower pathogen burden compared to individual two and therefore is more resistance. Individual 2, however, has a less steep slope compared to individual 1 and therefore is more tolerant.

If pathogen burden of an individual is unknown, response to infection could be measured as resilience that is the ability to maintain performance regardless of pathogen burden. A resilient animal, therefore, could be resistant or tolerant or both (Doeschl-Wilson *et al*., 2012). The genetic basis of resilience could be

modelled as the genetic effect on performance of an individual during a disease outbreak (Albers *et al*., 1987; Bisset and Morris, 1996).

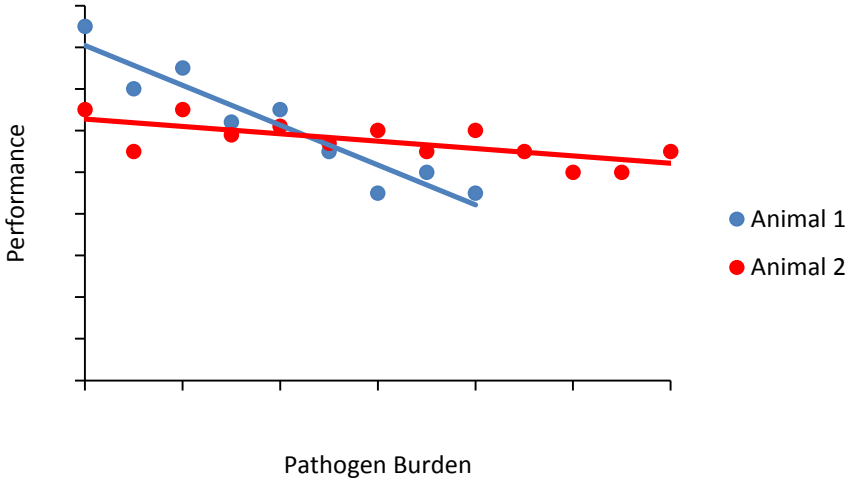

Figure 1.2. Schematic figure showing reaction norms of two animals (red or blue line) on their pathogen burden. Pathogen burden (x-axis) is a measure of resistance and the slope of the reaction norm line is a measure of tolerance.

1.3 Selective breeding for resistance, tolerance and resilience

Selective breeding for the animals that are simultaneously resistant and tolerant could be a pragmatic approach for controlling diseases and prevent production losses due to diseases in farm animals. The first step in developing selection strategies to increase the defensive response traits is to assess the presence of genetic variation for such traits and to quantify the proportion of phenotypic variance explained by genetics (heritability). If the heritability is large, selective breeding can improve the trait rapidly. It is also important to know if there is any trade-off between the traits. Because if for example resistance and tolerance are negatively correlated on the genetic level, improving one by selection will decrease the other one, unless both traits are included in the breeding goal and the selection index.

Existence of genetic variation in resistance to infection has been reported in farm animals such as dairy cattle (Berry *et al*., 2011; Detilleux, 2009; Morris, 2007; Sorensen *et al*., 2009), sheep (Albers *et al*., 1987; Bishop and Morris, 2007; Davies

et al., 2005), pigs (Ait-Ali *et al*., 2007; Boddicker *et al*., 2012; Vincent *et al*., 2005, 2006), poultry (Banat *et al*., 2013; Janss and Bolder, 2000; Jie and Liu, 2011) , and fish (Gjerde *et al*., 2011; Verrier *et al*., 2012). Heritabilities ranged from 0.04 to 0.33.

The existence of genetic variation in resilience, without distinguishing between resistance and tolerance, has also been reported in sheep (Albers *et al*., 1987; Morris *et al*., 2010), pigs (Boddicker *et al*., 2012; Boddicker *et al*., 2014; Lewis *et al*., 2009) and fish (Kuukka-Anttila *et al*., 2010). Heritabilities of resilience being the heritabilities of performance traits during disease periods ranged from 0.09 to 0.46.

The existence of genetic variation in tolerance is greatly overlooked. To date, there are very few studies on the genetic variation of tolerance to infection in farm animals. Hayward *et al*. (2014b) found no genetic variance in tolerance of Soay sheep to strongyle nematode infection. In another study Hayward *et al*. (2014a) found no genetic correlation between resistance and tolerance to strongyle nematode infection in Soay sheep. Råberg *et al*. (2007) showed genetic variation in tolerance and a negative genetic correlation ($-\mathbf{1.0}$) between resistance and to rodent malaria (*Plasmodium chabaudi*) among five different inbred mouse strains. Lough *et al*. (2015) showed genetic variation in tolerance to Listeria among four genetically diverse inbred mouse strains. Corby-Harris *et al*. (2007) reported differences in post-infection mortality, as an indicator of tolerance, for 11 lines (6 inbred lines and 5 wild lines) of *Drosophila melanogaster* infected by a strain of *P. aeruginosa*. In human, there is evidence for variability in tolerance to human malaria. For instance, a monogenic disorder called α^{+} -thalassemia, causing formation of abnormal haemoglobin molecules, tends to reduce the incidence of severe disease causing variability among individuals for disease tolerance (Williams *et al*., 2005).

1.4 This thesis

To date, there are very few studies on genetic aspects of tolerance to infection in farm animals. Furthermore, the genetic relationship between resistance, tolerance and resilience is unknown. The objectives of this thesis, therefore, were to: 1) estimate the genetic variation in resistance, tolerance, and resilience to infection in order to assess the amenability of these traits for selective breeding in farm animals, 2) estimate the genetic correlation between resistance,

tolerance and resilience and 3) detect genomic regions associated with resistance, tolerance, and resilience.

In chapter 2, the objectives were to 1) develop statistical models to detect PRRS outbreaks based on reproduction records of sows, 2) estimate variation among sows in response (resilience) to PRRS using different statistical models, 3) compare predictive ability of the statistical models for estimating variation in response to PRRS. We developed a linear regression method to distinguish healthy and disease phases based on reproduction records of sows. After detecting the outbreaks, we studied variation among sows for reproduction traits during healthy and diseased period of the farm. For that we used two statistical models. We compared the models for their predictive ability for the sow reproduction during healthy and diseased period using cross-validation.

In chapter 3, the objectives were to 1) study the genetic variation in resistance and tolerance of sheep to gastrointestinal nematode infection and 2) to estimate the trade-off between resistance and tolerance to nematode infection. We used a sire model on faecal nematode egg count and pepsinogen to study the genetic variation in resistance. We used a random regression model to study the reaction norm of body weight on faecal nematode egg count to study the genetic variation in tolerance. We finally applied a bivariate model to estimate the genetic correlation between resistance and tolerance to nematode infection.

In chapter 4, the objective was to study the response to selection in resistance and tolerance when using estimated breeding values for resilience. For that we simulated a population of half-sibs with known breeding values with resistance and tolerance using Monte Carlo simulation. We used selection index theory to study genetic gain in resistance and tolerance when pathogen burden is unknown and selection is based on resilience and compared it to the situation when genetic gain in resistance and tolerance is estimated based on known pathogen burden.

In chapter 5 we detected the pig's genomic regions associated with resistance, resilience, and tolerance to PRRS in the data used in chapter 5. We compared the genomic regions associated with the traits.

In chapter 6 I discussed the statistical models to measure response to diseases and the implication of breeding for resistance, tolerance and resilience in farm animals.

1.5 References

Ait-Ali, T., A. D. Wilson, D. G. Westcott, M. Clapperton, M. Waterfall, M. A. Mellencamp, T. W. Drew, S. C. Bishop, and A. L. Archibald. 2007. Innate immune responses to replication of porcine reproductive and respiratory syndrome virus in isolated swine alveolar macrophages. Viral. Immunology 20: 105-118.

Albers, G. A. A., G. D. Gray, L. R. Piper, J. S. F. Barker, L. F. Lejambre, and I. A. Barger. 1987. The genetics of resistance and resilience to haemonchus-contortus infection in young merino sheep. Int. J. Parasitol. 17: 1355-1363.

Banat, G. R., S. Tkalcic, J. A. Dzielawa, M. W. Jackwood, M. D. Saggese, L. Yates, R. Kopulos, W. E. Briles, and E. W. Collisson. 2013. Association of the chicken mhc b haplotypes with resistance to avian coronavirus. Dev. Comp. Immunol. 39: 430-437.

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.

Bishop, S. C. 2012. A consideration of resistance and tolerance for ruminant nematode infections. Front. Genet. 3: 168.

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

Bisset, S. A., and C. A. Morris. 1996. Feasibility and implications of breeding sheep for resilience to nematode challenge. Int. J. Parasitol. 26: 857-868.

Boddicker, N., E. H. Waide, R. R. R. Rowland, J. K. Lunney, D. J. Garrick, J. M. Reecy, and J. C. M. Dekkers. 2012. Evidence for a major qtl associated with host response to porcine reproductive and respiratory syndrome virus challenge. J. Anim. Sci. 90: 1733-1746.

Boddicker, N. J., A. Bjorkquist, R. R. R. Rowland, J. K. Lunney, J. M. Reecy, and J. C. M. Dekkers. 2014. Genome-wide association and genomic prediction for host response to porcine reproductive and respiratory syndrome virus infection. Genet. Sel. Evol. 46.

Charlier, J., M. van der Voort, F. Kenyon, P. Skuce, and J. Vercruysse. 2014. Chasing helminths and their economic impact on farmed ruminants. Trends Parasitol. 30: 361-367.

Corby-Harris, V., K. E. Habel, F. G. Ali, and D. E. Promislow. 2007. Alternative measures of response to pseudomonas aeruginosa infection in drosophila melanogaster. J. Evol. Biol. 20: 526-533.

Davies, G., M. J. Stear, and S. C. Bishop. 2005. Genetic relationships between indicator traits and nematode parasite infection levels in 6-month-old lambs. Anim. Sci. 80: 143-150.

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

Detilleux, J. C. 2011. Effectiveness analysis of resistance and tolerance to infection. Genet. Sel. Evol. 43.

Doeschl-Wilson, A. B., B. Villanueva, and I. Kyriazakis. 2012. The first step towards genetic selection for host tolerance to infectious pathogens: Obtaining the tolerance phenotype through group estimates. Front. Genet. 3.

Gjerde, B., J. Odegard, and I. Thorland. 2011. Estimates of genetic variation in the susceptibility of atlantic salmon (salmo salar) to the salmon louse lepeophtheirus salmonis. Aquaculture 314: 66-72.

Hayward, A. D., R. Garnier, K. A. Watt, J. G. Pilkington, B. T. Grenfell, J. B. Matthews, J. M. Pemberton, D. H. Nussey, and A. L. Graham. 2014a. Heritable, heterogeneous, and costly resistance of sheep against nematodes and potential feedbacks to epidemiological dynamics. Am. Nat. 184: S58-S76.

Hayward, A. D., D. H. Nussey, A. J. Wilson, C. Berenos, J. G. Pilkington, K. A. Watt, J. M. Pemberton, and A. L. Graham. 2014b. Natural selection on individual variation in tolerance of gastrointestinal nematode infection. PloS Biol. 12.

Holtkamp, D. J., J. B. Kliebenstein, E. J. Neumann, J. J. Zimmerman, H. F. Rotto, T. K. Yoder, C. Wang, P. E. Yeske, C. L. Mowrer, and C. A. Haley. 2013. Assessment of the economic impact of porcine reproductive and respiratory syndrome virus on united states pork producers. Journal of Swine Health and Production 21: 72-84.

Janss, L. L. G., and N. M. Bolder. 2000. Heritabilities of and genetic relationships between salmonella resistance traits in broilers. J. Anim. Sci. 78: 2287-2291.

Jie, H., and Y. P. Liu. 2011. Breeding for disease resistance in poultry: opportunities with challenges. Worlds Poult. Sci. J. 67: 687-695.

Kause, A. 2011. Genetic analysis of tolerance to infections using random regressions: a simulation study. Genet. Res. (Camb) 93: 291-302.

Kuukka-Anttila, H., N. Peuhkuri, I. Kolari, T. Paananen, and A. Kause. 2010. Quantitative genetic architecture of parasite-induced cataract in rainbow trout, oncorhynchus mykiss. Heredity 104: 20-27.

Lewis, C. R. G., M. Torremorell, L. Galina-Pantoja, and S. C. Bishop. 2009. Genetic parameters for performance traits in commercial sows estimated before and after an outbreak of porcine reproductive and respiratory syndrome. J. Anim. Sci. 87: 876-884.

Lough, G., I. Kyriazakis, S. Bergmann, A. Lengeling, and A. B. Doeschl-Wilson. 2015. Health trajectories reveal the dynamic contributions of host genetic resistance and tolerance to infection outcome. Proc. Biol. Sci. 282.

Morris, C. A. 2007. A review of genetic resistance to disease in bos taurus cattle. Vet. J. 174: 481-491.

Morris, C. A., S. A. Bisset, A. Vlassoff, M. Wheeler, C. J. West, B. P. Devantier, and A. D. Mackay. 2010. Selecting for resilience in romney sheep under

nematode parasite challenge, 1994–2007. New Zealand Journal of Agricultural Research 53: 245-261.

Nieuwenhuis, N., T. F. Duinhof, and A. van Nes. 2012. Economic analysis of outbreaks of porcine reproductive and respiratory syndrome virus in nine sow herds. Vet. Rec. 170: 225-+.

Råberg, L., D. Sim, and A. F. Read. 2007. Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. Science 318: 812-814.

Rausher, M. D. 2001. Co-evolution and plant resistance to natural enemies. Nature 411: 857-864.

2

Variation among Sows in Response to Porcine Reproductive and Respiratory Syndrome

Hamed Rashidi[1], Han A Mulder[1], Pramod K Mathur[2], Johan AM van Arendonk[1], Egbert F Knol[2]

[1] Topigs Norsvin Research Center, 6640 AA, Beuningen, the Netherlands;
[2] Wageningen University, Animal Breeding and Genomics Centre, 6700 AH, Wageningen, the Netherlands

Journal of Animal Science (2014) 92(1):95-105

Abstract

Porcine reproductive and respiratory syndrome (PRRS) is a viral disease with negative impacts on reproduction of sows. Genetic selection to improve the response of sows to PRRS could be an approach to control the disease. Determining sow response to PRRS requires knowing pathogen burden and sow performance. In practice, though, records of pathogen burden are unavailable. We develop a statistical method to distinguish healthy and disease phases and to develop a method to quantify sows' responses to PRRS without having individual pathogen burden. We analyzed 10,910 sows with 57,135 repeated records of reproduction performance. Disease phases were recognized as strong deviation of herd-year-week estimates for reproduction traits using two methods: method 1 used raw weekly averages of the herd, method 2 used a linear model with fixed effects for seasonality, parity, and year and random effects for herd-year-week and sow. The variation of sows in response to PRRS was quantified using 2 models on the traits "number of piglets born alive" (NBA) and "number of piglets born dead" (LOSS): 1. bivariate model considering the trait in healthy and disease phases as different traits and 2. reaction norm model modeling the response of sows as a linear regression of the trait on herd-year-week estimates of NBA. The linear model for NBA had the highest sensitivity (78%) for disease phases. Residual variances of both were more than doubled in the disease phase compared to the healthy phase. Trait correlations between healthy and disease phases deviated from unity (0.57 ± 0.13 - 0.87 ± 0.18). In the bivariate model repeatabilities were lower in disease phase compared to healthy phase (0.07 ± 0.027 and 0.16 ± 0.005 for NBA; 0.07 ± 0.027 and 0.09 ± 0.004 for LOSS). The reaction norm model fitted the data better than the bivariate model based on Akaike's information criterion and had also higher predictive ability in disease phase based on cross validation. Our results show that the linear model is a practical method to distinguish between healthy and disease phases in farm data. We showed that there is variation among sows in response to PRRS implying possibilities for selection, and the reaction norm model is a good model to study the response of animals towards diseases.

Key words: disease resistance, tolerance to infection, outbreak detection, pig, porcine reproductive and respiratory syndrome, reproduction.

2.1 Introduction

A major problem in the pig industry is the viral disease porcine reproductive and respiratory syndrome (PRRS). The biggest economic impact of PRRS is reproductive failure in sows such as abortion, mummified and stillborn piglets, and pre-weaning mortality in piglets (Murtaugh and Rowland, 2004). Because vaccination against PRRS is not fully successful (Huang and Meng, 2010; Murtaugh and Genzow, 2011), genetic selection on the population level for sows that can mount a defense against PRRS could be an option. The defensive ability of a sow against PRRS has two mechanisms: resistance and tolerance. Resistance occurs when sows prevent the PRRS virus from entering the body or manipulate the virus life cycle to remove it from the body (Read *et al*., 2008; Rowland *et al*., 2012). Tolerance occurs when sows decrease the effects of the PRRS infection on performance despite the infection (Kause, 2011). Most studies on host–pathogen interactions have focused on the genetics of resistance (Vincent *et al*., 2006; Lewis *et al*., 2010) with little known about the genetic aspects of tolerance.

To study resistance and tolerance, the pathogen burden of PRRS along with the record of performance for each sow needs to be known (Kause, 2011; Doeschl-Wilson *et al*., 2012). Recording of pathogen burden in farm animals, however, is laborious and costly. Thus, a new method is needed to quantify sow response to PRRS without knowing the pathogen burden of individuals. Moreover, an effective method of distinguishing between healthy and diseased periods of a farm is required. The objectives of this study were to 1) identify PRRS outbreaks based on reproduction records, 2) develop statistical models to estimate variation among sows in response to PRRS, and 3) assess the goodness of fit and predictive ability of the statistical models.

2.2 Material and methods

2.2.1 Data

Data were collected from a commercial pig farm located in an area in Canada, where PRRS is endemic. There were 68,292 records of 12,441 sows, which had repeated records of reproduction traits from 2003 to 2012. We used the daily reproductive records of the farm including parity, number of piglets born alive (NBA), number of piglets weaned (NWD), number of mummified piglets (MUM), number of stillborn piglets (STB), and abortions (binary: yes/no) (AB). We also created an extra trait for number of piglets born dead by summing the number of

stillborn and mummified piglets (LOSS). Combining mummification and stillbirth into a single trait was done to preclude the problem of misdiagnosis between mummification and stillbirth (Mckay, 1993). Mummification and stillbirth cause a reduction in NBA, and abortion leads to zero NBA. Therefore, reduction in NBA arising from those reproduction failures was expressed at the day of farrowing. After editing the dataset to exclude animals with no observations and also to remove weeks with fewer than 5 observations, a total of 57,135 records for 10,910 sows remained from 2004 through the initial 26 weeks of 2012. The number of parities ranged from 1 to 14 with an average of 5. No pedigree was available. During suspected PRRS outbreaks based on changes in average sow performance, e.g. increased abortion and mummified piglets, blood samples were taken to test for the presence of PRRS virus and to identify the strain of PRRS viruses for veterinary purposes. There was no information available about any negative blood samples for PRRS virus.

2.2.2 Partitioning production periods into healthy and disease phases

Two methods were used to partition periods of production into healthy and disease phases. One method simply used raw weekly averages of the herd in each year. The other method used herd-year-week estimates of a linear model, in which herd-year-week effects were estimated simultaneously with fixed and random effects. In both methods, disease phases were distinguished from healthy phases as strong deviations (1% truncation of a normal distribution, described below) of herd-year-week estimates from the mean of reproduction traits.

The traits NBA, NWD, MUM, STB, AB, and LOSS were used to test which traits are best for outbreak detection compared to virus isolation data. Porcine reproductive and respiratory syndrome decreases NWD and NBA, therefore, herd-year-weeks with an estimate more than 2.326 standard errors below the overall mean of herd-year-week estimates were considered as diseased. The threshold 2.326 corresponds to the 1% truncation point of a normal distribution. Other truncation points like 5% and 10% were tested as well. The 1% truncation point, however, resulted in more precise detection of outbreaks. Porcine reproductive and respiratory syndrome, on the other hand, increases AB, MUM, STB, and LOSS, and therefore, the herd-year-weeks with an estimate more than 2.326 standard errors above the overall mean of herd-year-week estimates were considered as diseased. The linear model was as follows:

$$Y_{ijkl} = \mu + \beta SIN_{ijkl} + PAR_i + YR_j + yw_k + sow_l + e_{ijkl}, \quad [1]$$

where Y_{ijkl} is the phenotypic value for the reproductive traits of l^{th} sow, μ is the overall mean, β is the regression coefficient of the covariate SIN, and SIN is a sinusoidal covariate to account for seasonality in production of the sow, which was calculated as: $sin\{\left[\frac{Farrowing\ Date-"1\ May"}{365.25}\right] \times (2\pi)\}$ (Bergsma and Hermesch, 2012). Bergsma and Hermesch (2012) used March 21 as the point where the sinus-function is zero (day-length is 12 hours). In our study May 1 had the best match with the data. PAR is the fixed effect of the i^{th} parity for sow; YR is the fixed effect of the j^{th} year; and yw is the random effect of the k^{th} herd-year-week with $\mathrm{N}(0, \sigma^2_{yw})$, where σ^2_{yw} is the variance of herd-year-week; sow is the random effect of the l^{th} sow with $\mathrm{N}(0, \mathbf{I}\sigma^2_{sow})$, where $\mathbf{I}$ is the identity matrix, as no pedigree was available, and σ^2_{sow} is the variance of the sow effect; and e_{ijkl} is the random residual term with $e \sim \mathrm{N}(0, \mathbf{I}\sigma^2_e)$, where $\mathbf{I}$ is the identity matrix and σ^2_e is the residual variance. Herd-year-week was included as a random effect because preliminary results showed confounding between sow effects and herd-year-week effects leading to large standard errors of herd-year-week estimates (Visscher and Goddard, 1993).

In both methods, a disease phase was defined in which there were at least two consecutive herd-year-weeks specified as diseased. If there was a one-week gap between two herd-year-weeks specified as diseased, the week in between was also considered to be diseased when the herd-year-week estimate was at least 1.645 standard errors below/above the average (5% truncation point of normal distribution).

The sensitivity of detecting truly positive phases was calculated based on the date of virus isolation. For that, the dates of virus isolation were converted to weeks of virus isolation. A disease phase was considered to be truly positive if there were at most 7 weeks time lag between the disease phase and the week(s) of virus isolation. The 7-week lag between a disease phase and the week(s) of virus isolation was considered because it led to the highest match between the two (see Table 2.1). Transmission of PRRS virus occurs either directly by the pigs through bites, cuts, scrapes, tail and ear-biting or indirectly by instruments in the farm, clothing, water, food and aerosols. Physical obstacles like e.g. sows in different compartments separated by walls, therefore, can delay transmission of PRRS virus within the farm. It is also known that pigs are not equally susceptible to PRRS virus by all routes of exposure (Zimmerman *et al*., 2006). On the other hand, vertical transmission from dams to fetuses mostly happens during the last trimester of pregnancy and results either in dead fetuses or weak new-born piglets that might die before weaning (Zimmerman *et al*., 2006). Depending on the stage of

pregnancy at which the sow is infected, the reproduction symptoms of PRRS vary, which might delay expression of the disease on the herd level. The 7-week, therefore, seemed to be a reasonable time span between disease phase and the week(s) of virus isolation.

Sensitivities of the approaches were calculated as:

$sensitivity = \frac{TP}{TP+FN}$,

where TP is the number of truly positive disease phases and FN is the number of false-negative phases.

Table 2.1 Comparison of linear model and weekly average as two methods for partitioning production periods into healthy and disease phases

		Disease period		Overlap[2]		
Trait[1]	Method	Week[3]	Phase[4]	Week[3]	Phase[4]	Sensitivit[5] (%)
NBA	Linear Model	55	11	10	3	78
	Weekly Average	10	3			21
AB	Linear Model	55	12	4	2	69
	Weekly Average	4	2			7
NWD	Linear Model	68	8	16	5	50
	Weekly Average	16	5			29
STB	Linear Model	24	5	4	2	21
	Weekly Average	4	2			7
LOSS	Linear Model	23	4	2	1	15
	Weekly Average	2	1			0
MUM	Linear Model	8	2	3	1	14
	Weekly Average	3	1			7

[1]Traits: NBA = number piglets born alive; AB = Abortion; NWD = number of piglets weaned; STB = number of stillborn piglets; LOSS = number of piglets born dead due to mummification and/or stillbirth; MUM = number of mummified piglets

[2]Weeks and phases that were detected as disease and outbreak with both linear model and weekly average method

[3]Number of weeks partitioned as diseased, which were in an immediate adjacent of at least one diseased week.

[4]Number of disease phases, which consist at least two consecutive diseased weeks

[5]Sensitivity of the approach in detecting truly positive phases

2.2.3 Estimation of variation among sows in response to PRRS

Two models were used to quantify variation among sows in response to PRRS: a bivariate model and a reaction norm model. In both models, we included traits NBA and LOSS as response variables to assess the ability of each trait in capturing the variation among sows in response to PRRS. Number of piglets born

alive was chosen because most of the reproduction failures arising from PRRS would be expressed as a reduction in NBA at farrowing (Figure 2.1). We hypothesized that NBA is a good option that shows drops in production caused by PRRS. One may think that NWD could be a better trait as compared to NBA because in addition to AB and LOSS, pre-weaning mortality in piglets would be expressed in total number of piglets weaned. In this farm, though, cross-fostering had taken place among sows, which would complicate data analysis and for this reason NWD was not used for estimation of variance among sows in response to PRRS. We also used LOSS because previous studies reported significant increase in number of mummified and stillborn piglets as well as large variance of these two traits during PRRS outbreaks (Lewis *et al*., 2009). Abortion was not used to quantify variation among sows because it was a binary trait and analyses did not converge.

The linear model with NBA as response variable was considered as the best method of partitioning periods into healthy and diseased because it showed the highest sensitivity of detecting truly disease phases (Table 2.1). Herd-year-week estimates of NBA from the linear model are the herd characteristics that are best associated with gradual changes in environment because of PRRS and were used as the environmental parameter in the reaction norm model. The standardized herd-year-week estimates of NBA from the linear model ranged from 2 to −4.4 (healthy to diseased). Using herd-year-week estimates of NBA in the linear model allowed us to quantify variation in sow responses to PRRS with respect to disease severity in each herd-year-week without partitioning records into healthy and diseased phases.

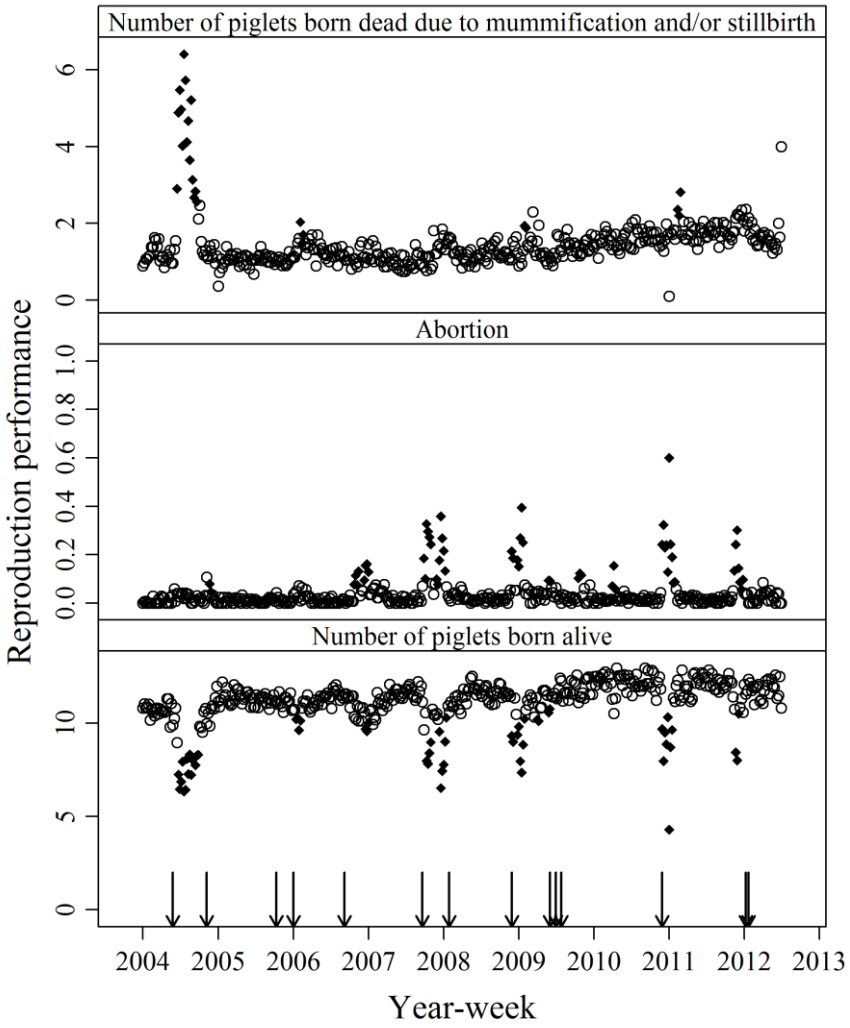

Figure 2.1 Predicted PRRS outbreaks using the linear model based on number of piglets born dead due to mummification and/or stillbirth, abortion, and number of piglets born alive. The solid diamonds show disease herd-year-weeks during an outbreak. The empty circles show healthy herd-year-weeks. The arrows show the weeks in which PRRS viruses were isolated from blood of sows.

2.2.3.1 Bivariate model

With this model [2], sow performances in terms of NBA and LOSS in healthy and disease phases were studied as different traits as proposed by Falconer (1952) to study genotype by environment interaction. The model was as follows:

$$\begin{bmatrix}\mathbf{y}_{\text{healthy}}\\ \mathbf{y}_{\text{diseased}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}_{\text{healthy}} & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_{\text{diseased}}\end{bmatrix}\begin{bmatrix}\mathbf{b}_{\text{healthy}}\\ \mathbf{b}_{\text{diseased}}\end{bmatrix} + \mathbf{Wyw} + \begin{bmatrix}\mathbf{Z}_{\text{healthy}} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_{\text{diseased}}\end{bmatrix}\begin{bmatrix}\mathbf{a}_{\text{healthy}}\\ \mathbf{a}_{\text{diseased}}\end{bmatrix} + \begin{bmatrix}\mathbf{e}_{\text{healthy}}\\ \mathbf{e}_{\text{diseased}}\end{bmatrix} \quad [2]$$

where $\mathbf{y}_{\text{healthy}}$ ($\mathbf{y}_{\text{diseased}}$) is a vector of sow performance in the healthy phase (disease phase); and $\mathbf{b}_{\text{healthy}}$ ($\mathbf{b}_{\text{diseased}}$) is a vector of the fixed effects in healthy phase (disease phase), which were μ, SIN, PAR, YR (see model 1 for description), $\widehat{YW}$ as a covariate for estimated herd-year-weeks from the basic model [1] that corrects for the severity of the infection, and $STATUS$ defined by the herd-year-week solutions of NBA as performance means in healthy and disease phase. The $\widehat{YW}$ was included as a covariate in the bivariate model to make it equivalent to the reaction norm model (see below) in terms of the fixed effects, which enabled us to compare the models using Akaike information criterion (see below). Random effects were $\boldsymbol{yw}$, which is the effect of herd-year-week with $\boldsymbol{yw}\sim\mathrm{N}(\mathbf{0},\sigma_w^2)$, and $\mathbf{a}_{\text{healthy}}$ ($\mathbf{a}_{\text{diseased}}$), which is a vector of sow effects in the healthy phase (disease phase). $\mathbf{X}_{\text{healthy}}$ ($\mathbf{X}_{\text{diseased}}$) and $\mathbf{Z}_{\text{healthy}}$ ($\mathbf{Z}_{\text{diseased}}$) are the design matrices assigning the observations to the levels of fixed and random effects in the healthy phase (disease phase). $\mathbf{W}$ is the design matrix assigning the observations to the levels of $\mathbf{yw}$ random effect. The variance of the residuals and sow effects is:

$$\mathrm{Var}\begin{bmatrix}\mathbf{e}_{\text{healthy}}\\ \mathbf{e}_{\text{diseased}}\end{bmatrix} = \mathbf{R}\otimes\mathbf{I} \text{ where } \mathbf{R} = \begin{bmatrix}\sigma^2_{e_{healthy}} & 0\\ 0 & \sigma^2_{e_{diseased}}\end{bmatrix} \text{ and}$$

$$\mathrm{Var}\begin{bmatrix}\mathbf{a}_{\text{healthy}}\\ \mathbf{a}_{\text{diseased}}\end{bmatrix} = \mathbf{G}\otimes\mathbf{I} \text{ where } \mathbf{G} = \begin{bmatrix}\sigma^2_{a_{healthy}} & \sigma_{a_{healthy}\,a_{diseased}}\\ \sigma_{a_{healthy}\,a_{diseased}} & \sigma^2_{a_{diseased}}\end{bmatrix}, \text{ where}$$

$\otimes$ is the direct matrix product operator.

2.2.3.2 Reaction-norm model

The reaction norms of the reproduction traits NBA and LOSS on standardized herd-year-week estimates of NBA (ranged from 2 to −4.4 with $\mu = 0, \sigma = 1$) were studied with the following model:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Wyw} + \mathbf{Za_i} + \mathbf{Z_{yw}a_s} + \mathbf{e} \quad [3]$$

where $\mathbf{y}$ is a vector of sow performance; $\mathbf{b}$ is a vector of fixed effects (see model 2 for fixed effects). $\widehat{YW}$ was included in the model as a covariate for estimated herd-year-weeks to account for the averages of herd-year-week estimates. $STATUS$ was used to account for the difference in average performance in healthy and disease phases and to make the reaction norm model equivalent to the bivariate model in terms of the fixed effects, which enabled us to compare the models using Akaike information criterion (see below); $\boldsymbol{yw}$ is the random effects of herd-year-week with $\boldsymbol{yw} \sim \mathrm{N}(\mathbf{0}, \sigma^2_{yw})$; $\mathbf{a_i}$ is a vector of sow random effects for intercept with $\mathbf{a_i} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}\sigma^2_{a_i})$, where $\sigma^2_{a_i}$ is the sow variance for intercept and $\mathbf{I}$ is the identity matrix; and $\mathbf{a_s}$ is a vector of sow random effects for slope of the reaction norms of performances on standardized herd-year-week estimates, obtained from the basic model [1], with $\mathbf{a_s} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}\sigma^2_{a_s})$, where $\sigma^2_{a_s}$ is the sow variance for slope and $\mathbf{e}$ is a vector of residuals. $\mathbf{X}$, $\mathbf{W}$, and $\mathbf{Z}$ are the design matrices assigning the observations to the levels of fixed and random effects. $\mathbf{Z_{yw}}$ is the design matrix with standardized herd-year-week estimates of NBA as covariates for the slopes of reaction norms. The variance covariance matrix for intercept and slope is $\mathrm{Var}\begin{bmatrix}\mathbf{a_i}\\ \mathbf{a_s}\end{bmatrix} = \mathbf{G_{RN}}$ where $\mathbf{G_{RN}} = \begin{bmatrix}\sigma^2_{a_i} & \sigma_{a_i,a_s}\\ \sigma_{a_i,a_s} & \sigma^2_{a_s}\end{bmatrix}$, where σ_{a_i,a_s} is the covariance between sow effects for intercept and slope. Preliminary results showed substantial inflated residual variances in disease phase as compared to healthy phases. Heterogeneity of residual variance, therefore, was considered for healthy and disease phases. Based on standardized herd-year-week estimates, records were divided in 10 classes: 9 classes for healthy phases and class 10 for diseased phases. A similar approach was used by Calus *et al*., (2002) and Lillehammer *et al*., (2009) to estimate genotype by environment interaction using reaction norm models. The residual variances of the 10 classes were estimated, so that: $\mathrm{Var}\begin{bmatrix}\mathbf{e_1}\\ \mathbf{e_2}\\ \vdots\\ \mathbf{e_{10}}\end{bmatrix} = \mathbf{R} \otimes \mathbf{I}$ where $\mathbf{R} = \begin{bmatrix}\sigma^2_{e_1} & 0 & \dots & 0\\ 0 & \sigma^2_{e_2} & 0 & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \sigma^2_{e_{10}}\end{bmatrix}$

Variance components were estimated using ASReml (Gilmour *et al*., 2009).

2.2.4 Correlation between performances in healthy and disease phases

To quantify re-ranking of sows between healthy and disease phases, we estimated the correlation of sow performances between healthy and disease

phases. With the bivariate model, the correlation was expressed in the output of the analysis. With the reaction norm model, the correlation between performances of sows in healthy and disease phases was estimated by calculating variances of sows in healthy and disease phases and the covariance between the two phases using the $\mathbf{G_{RN}}$.

2.2.5 Model comparison

2.2.5.1 Akaike information criterion (AIC)

The goodness of fit of the models was assessed with AIC using the following formula (Akaike, 1973):

$AIC = -2logL + 2K,$

where $logL$ is the logarithm of likelihood of the model and K is the number of variables in the model.

2.2.5.2 Cross-validation

The predictive ability of each model was studied using cross validation. Because of lack of pedigree data, sows needed to have some records to predict future performance of them. Prediction of future performance is important for breeding programs as selection is directed towards the future. To assess the effect of including more information on predictive ability, we increased the number of parities included in the reference data. We hypothesized that having more parity of sows will lead to higher accuracy of prediction of their future performance. Four different parity groups were considered: parities 1 through 4 (1–4), 1 through 5 (1–5), 1 through 6 (1–6), and 1 through 7 (1–7). In each group, the last parity was set to missing (validation set), and all other parities before the last one (training set) were used to predict the sow effects in the last parities. First, the base model (1) was run for each trait for each parity group. Then adjusted phenotypes were calculated as the sum of estimated sow effect and estimated residual for each record. The adjusted phenotypes were used as the response variables of each model. Finally, the correlations between the estimated sow effects in the training sets with the adjusted phenotypes in the validation sets were calculated for each model.

2.3 Results

2.3.1 Outbreak detection

Table 2.1 shows the results of partitioning production periods into healthy and disease phases. The number of disease phases detected with the linear model method was much larger than that with the weekly average method, although the threshold based on the normal distribution was the same. During PRRS outbreaks the performance of the farm decreased substantially, which reduced the annual average of production and as a consequence the deviation of a certain year-weak from the annual average was reduced. In the linear model method this problem was solved by correcting the performance for fixed and random effects. With the linear model, the highest numbers of disease phases were detected with NBA, AB, and NWD, respectively. With the weekly average approach, the highest numbers of disease phases were detected with NWD and NBA, respectively. Sensitivity of detecting truly positive disease phases was largest using the linear model with NBA (78%) and AB (69%) and much lower for MUM (14%), LOSS (15%), and STB (21%). Sensitivities were much lower using the weekly average method. Figure 2.1 illustrates partitioned healthy and disease phases using NBA, AB, and LOSS in the linear model as well as the weeks in which virus had been isolated. It shows that there was a good accordance between specified disease phases and weeks of virus isolation. Using NBA in the linear model 10 out of 11 detected disease phases could be assigned to a week of virus isolation, considering at most a 7-week lag between them. It can be concluded that using the linear model with NBA as a response variable is the best approach to detect disease phases related to PRRS.

2.3.2 Comparison of sow performance in healthy and disease phases

The differences in the performance of sows for NBA during healthy and disease phases are summarized in Table 2.2. All the sows had performance in healthy phase and about 50% had at least one performance in disease phase. All the sows in disease phase had at least one record in healthy phase. The mean of NBA in disease phases decreased by 24% whereas the standard deviation increased by 44% as compared to healthy phase. There was a clear decrease in largest (17%) and smallest (50%) values of herd-year-week averages in disease phases as compared to healthy phases. This reduction in NBA during disease phase was expected because the disease phase was basically defined as the reduction in NBA.

Table 2.2 Comparison of number of piglets born alive in healthy and disease phases

Phase	Number of sow	Number of record	Mean	SD	Max[1]	Min[2]
Healthy	10910	50467	11.48	3.63	12.97	8.97
Disease[3]	5374	6668	8.73	5.21	10.74	4.30

[1]Maximum herd-year-week average
[2]Minimum herd-year-week average
[3]Sows in diseased phases had always at least one record in healthy phases.

2.3.3 Variance components, repeatability, and ranking.

Both bivariate and reaction norm models showed that there is variation among sows for NBA and LOSS during healthy and disease phases (Table 2.3). The sow variances in healthy phases were similar between the bivariate and reaction norm models for both NBA and LOSS. The sow variances during disease phases were almost doubled as compared to healthy phases, except for NBA in the bivariate model which decreased by 17%. Residual variances of both models were very similar. For both traits, residual variances during disease phase were more than doubled as compared to healthy phase, using both bivariate and reaction norm models. For LOSS in disease phase the increase in residual variance was larger than that of NBA. In addition, using the reaction norm model showed that for both traits, NBA and LOSS, there was variation in intercepts and slopes of the reaction norms.

The variance of intercept shows the variation in performance of sows at zero estimate of herd-year-week for NBA, which is the average herd-year-week for NBA. The standardized herd-year-week estimates for NBA in healthy phase ranged from 2 to –2.326 in which the average production of the herd is not affected by PRRS. Therefore, the intercept of the reaction norm model is approximately in the average herd-year-week in the healthy phase. The variance of slopes shows variation in responses of sows to gradual changes in herd-year-week estimates associated with PRRS or something else. There were moderate negative correlations between intercept and slope at zero estimate of herd-year-weeks for both traits, NBA (–0.26 ± 0.05) and LOSS (–0.41 ± 0.09). The negative correlations between intercept and slope mean that sows with high intercepts have less steep slopes whereas sows with low intercepts have steep slopes. In other words, sows with high NBA during healthy phases may show smaller reduction in NBA during disease phases, whereas sows with low NBA during healthy phase may show higher reduction in NBA during disease phase. In the case of LOSS, it means that sows with high LOSS during healthy phases may have a small increase in LOSS during diseased phases, whereas sows with low LOSS during healthy phases may have large increase in LOSS during diseased phases.

Table 2.3 Variance components ± SE of the bivariate and reaction norm models for number of piglets born alive (NBA) and number of piglets born dead due to mummification and/or stillbirth (LOSS)

	NBA		LOSS	
Variance	Bivariate	Reaction norm	Bivariate	Reaction norm
Sow in healthy phase	2.05 ± 0.07	1.96 ± 0.06	0.30 ± 0.01	0.28 ± 0.01
Sow in disease phase	1.70 ± 0.68	3.83 ± 0.31	0.57 ± 0.23	0.67 ± 0.09
Residual healthy phase	10.55 ± 0.07	10.56 ± 0.08	2.87 ± 0.02	2.88 ± 0.01
Residual disease phase	22.92 ± 0.77	21.67 ± 0.49	7.43 ± 0.26	7.41 ± 0.16
Corr[1] (healthy, disease)	0.87 ± 0.18	0.81 ± 0.03	0.57 ± 0.13	0.83 ± 0.05
Intercept		2.05 ± 0.07		0.31 ± 0.01
Slope		0.23 ± 0.04		0.04 ± 0.01
Corr[2] (intercept, slope)		-0.26 ± 0.05		-0.41 ± 0.09

[1]Correlation between sows' performance in healthy and disease phase
[2]Correlation between intercept and slope

For NBA, with increasing the class of residual variances from 1 to 10, residual variance increased gradually (Table 2.4), specifically in the first eight classes that were in healthy phase. The increase of residual variance in disease phase ranged from 177% to 47% as compared to the classes 1 to 9 in healthy phase. The last class in healthy phase (class 9) had the highest residual variance (14.72 ± 0.31) among other classes in healthy phase, which might be because some records of diseased sows in a herd-year-week were not classified as disease phase. For LOSS, there were large changes in residual variances in different classes of residual variances. There was not a consistent pattern of change specifically in the first eight classes in healthy phase. In line with NBA, the last class in healthy phase (class 9) had the highest residual variance (4.07 ± 0.08) among other classes in healthy phase, which again suggests that there might have been some records of diseased sows in this class that were not classified in disease phase. For LOSS, increase of residual variance in disease phase ranged from 225% to 82% as compared to classes 1 to 9 of the residual variances in healthy phase. It can be concluded that PRRS severely increased variation in NBA and LOSS.

Table 2.4 Residual variances ± SE of reaction norm model grouped into 10 classes (Class) based on estimates of number of piglets born alive (NBA) and number of piglets born dead due to mummification and/or stillbirth (LOSS). There were 9 classes in healthy phase and 1 classes in disease phase

Phase	Class	NBA	LOSS
	1	7.84 ± 0.18	2.28 ± 0.05
	2	7.76 ± 0.17	2.43 ± 0.05
	3	9.07 ± 0.19	2.41 ± 0.05
	4	9.65 ± 0.20	3.08 ± 0.06
Healthy	5	10.68 ± 0.23	2.83 ± 0.06
	6	10.94 ± 0.23	2.78 ± 0.06
	7	11.67 ± 0.24	2.64 ± 0.05
	8	12.74 ± 0.27	3.39 ± 0.07
	9	14.72 ± 0.31	4.07 ± 0.08
Disease	10	21.67 ± 0.49	7.41 ± 0.16

2.3.4 Correlation between performances of sows in healthy and disease phases

Correlations between performances of sows in healthy and disease phases are shown in Table 2.3. For NBA, correlations between healthy and diseased periods were high using the bivariate (0.87 ± 0.18) and the reaction norm models (0.81 ± 0.03). For LOSS, correlations between healthy and diseased periods were moderate (0.57 ± 0.13) using the bivariate model and high (0.83 ± 0.05) using the reaction norm model. There was a large difference in standard errors of the correlation estimates between the two models. Using the bivariate model, standard errors of the correlation estimates were higher as compared to the reaction norm model. In general, correlations between performances of sows in healthy and disease phases significantly deviated from one (2×SE, Lynch and Walsh, 1998) except for NBA using the bivariate model. These findings imply re-ranking of sows in healthy and disease phases.

Figure 2.2 shows variation among 100 random sows for reaction norm of NBA on herd-year-week estimates across the 10 classes of herd-year-weeks. The x-scale in Figure 2.2 was mirrored to reflect that the diseased phase was on the right side of the figure to have a similar figure as tolerance/resistance as a function of pathogen burden, e.g., Raberg *et al*., (2007). Note that we used the figure only for illustrative purpose to show change in the ranking of sows but not to quantify it. Although most of the sows showed a flat reaction norm, the ranking of sows was different across herd-year-week levels. The difference in ranking of sows was more

pronounced in the lower levels of herd-year-week, which suggests more variation among sows in lower levels of herd-year-weeks.

Repeatabilities of sow performance in healthy and disease phases are shown in Table 2.5 Repeatabilities for LOSS were generally lower than repeatabilities for NBA. In general, repeatabilities of the traits were similar in healthy and disease phases using bivariate and reaction norm models with a slight decrease in disease phase. Using the bivariate model for NBA repeatability in disease phase was almost halved as compared to healthy phase.

Table 2.5 Repeatabilities ± SE of the sows' performances in healthy and disease phases for number of piglets born alive (NBA) and number of piglets born dead due to mummification and/or stillbirth (LOSS)

Traits	Phase	Bivariate	Reaction norm
NBA	Healthy	0.16 ± 0.005	0.16 ± 0.005
	Disease	0.07 ± 0.027	0.15 ± 0.012
LOSS	Healthy	0.09 ± 0.004	0.09 ± 0.040
	Disease	0.07 ± 0.029	0.08 ± 0.011

2.3.5 Model comparison

For both traits, AIC was lower for the reaction norm model, which suggests a better fit of this model for this dataset as compared to the bivariate model (Table 2.6). The predictive abilities of the models were higher for NBA as compared to LOSS (Table 2.7). For both traits in healthy phase, the predictive abilities of the models were similar, and they increased with increasing number of parities. The reaction norm model had higher predictive ability in general. In disease phase, predictive abilities of the basic and reaction norm models were similar, and they also increased with increasing number of parities. In line with healthy phases, the reaction norm had generally higher predictive ability in disease phase as compared to the bivariate and basic model. The bivariate model had a very poor predictive ability in disease phase, which improved with increasing number of parities, but at a lower rate as compared to the healthy phase. These results show that the reaction norm model has a better fit and a better predictive ability as compared to the bivariate model, especially in disease phase.

Table 2.6. Akaike information criterion of the bivariate and reaction norm models for number of piglets born alive (NBA) and number of piglets born dead due to mummification and/or stillbirth (LOSS)

Model	NBA	LOSS
Bivariate	203908	128913
Reaction norm	203181	128614

Table 2.7. Correlations between adjusted phenotypes of sows in validation sets and predicted sow effects with training sets for number of piglets born alive (NBA) and number of piglets born dead due to mummification and/or stillbirth (LOSS) using three statistical models. Four parity groups were considered (1-4, 1-5, 1-6, and 1-7), where training sets include the records of all parities before the last one, and validation sets include the records of the last parity in each group.

Trait	Parity group	Healthy phase			Disease phase		
		[1]Basic	[1]Reaction norm	[1]Bivariate	[1]Basic	[1]Reaction norm	[1]Bivariate
NBA	1-4	0.193	0.202	0.194	0.081	0.077	-0.003
	1-5	0.227	0.240	0.232	0.112	0.113	0.051
	1-6	0.231	0.246	0.242	0.170	0.174	0.058
	1-7	0.251	0.284	0.255	0.177	0.223	0.057
LOSS	1-4	0.118	0.124	0.130	0.020	0.018	-0.012
	1-5	0.162	0.172	0.159	0.078	0.097	-0.046
	1-6	0.174	0.183	0.181	0.137	0.141	0.060
	1-7	0.176	0.186	0.190	0.161	0.148	0.085

[1]Statistical model

2.4 Discussion

2.4.1 Outbreak detection

We introduced the linear model for reproduction performance of pig farms to detect accurately PRRS outbreaks. We obtained a high sensitivity of detecting truly positive outbreaks when considering a maximum time lag of 7 weeks between the weeks of virus isolation and the detected disease phases using the linear model on NBA. Although PRRS infects a herd rapidly (transmission rate R0 > 3) (Nodelijk *et al*., 2001), a major PRRS outbreak may take place several weeks after the introduction of the virus into a population. Studies have reported different time spans between introduction of PRRS virus and occurrence of outbreaks in pig populations. In a study on 4-month PRRS-free gilts (Batista *et al*., 2004), it was observed that 15 out of 15 gilts were positive for the virus around 10 days post-infection. Another study (Houben *et al*., 1995) reported that littermates may seroconvert from 4 to 12 weeks of age when one piglet became infected during the fattening period. Similar to our findings, Lewis *et al*. (2009) observed that the effect of a PRRS outbreak on herd production, in terms of mummified piglets, starts a few weeks after veterinary diagnosis of PRRS in the herd. Based on these evidences, we conclude that the 7-week lag between weeks of virus isolation and the first or last week of detected disease phases is a reasonable approximation for the delay between emergence of the virus and drops in production of the herd because of PRRS. It must be noted that the linear model method is not capable of finding an outbreak at the early and end stages of an epidemic, when not all sows are infected or most of them are already recovered, because the weekly average of the herd is not dramatically influenced by PRRS. It means that in healthy phase, there might be some sows that were diseased but the herd-year-weeks were still partitioned as healthy phases. In addition, during the disease phases of the farm, there might have been healthy sows that were resistant to PRRS infection and sows that already recovered or were recovering from the disease. The possible mixture of diseased and healthy sows during healthy phases is supported by the high residual variance in the 9th class of herd-year-week estimates.

Another method that we used to detect PRRS outbreaks was the weekly average method, which is similar to the Threshold/threshold method applied in Lewis *et al*. (2009). According to them, if the 30-day rolling average of the trait 'mummified piglets' was larger than a 99% confidence threshold, the subsequent litters were considered as being in disease phase. In our study the weekly average

method detected few phases of PRRS outbreaks and also showed low sensitivity of detecting truly positive phases. The linear model, therefore, is a better method for outbreak detection than the weekly average method. Note that the way blood samples were taken may have affected sensitivity, but not the ranking of methods. Furthermore, the presence of diseased animals in specified healthy phases may lead to more number of false negative periods and as a consequence underestimation of sensitivity.

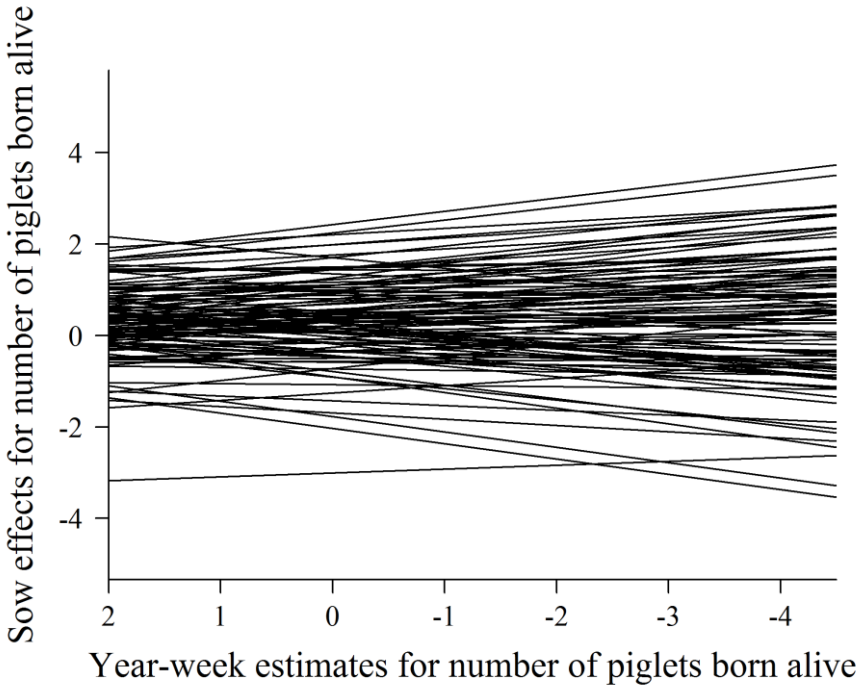

Figure 2.2. Reaction norms of 100 randomly sampled estimated sow effects for number of piglets born alive (NBA) on herd-year-week estimates of NBA. The sow effects were sampled from the sows that had records both in healthy and disease phases. The x-scale was mirrored to reflect that the disease phase was on the right side of the figure.

2.4.2 Modeling variation among sows in response to PRRS

The second objective of this study was to develop a method to quantify variation among sows in response to PRRS. Both bivariate and reaction norm models showed that there is variation among sows during healthy and disease phases and that residual variance during disease phase was more than doubled as compared to healthy phase. This increase in the residual variance shows that the

disease creates a wider range of sow phenotypes because of higher rates of abortion, mummification, and stillbirth. The existence of sow variance in both phases may indicate the presence of additive genetic variance and the possibility for genetic improvement, because the sow variance consists of additive genetic variance, non-additive genetic variance, and permanent environmental variance of sow response to PRRS. In both models, the non-unity correlation between sow effects in healthy and disease phases indicates the re-ranking of sows between phases.

The advantage of the bivariate model is that it directly models heterogeneity of genetic and residual variance as well as re-ranking between healthy and disease phases. Furthermore, the model is conceptually easy. The bivariate model, however, performed worse in terms of predictive ability and model fit. The bad performance is likely because sow effects are estimated in healthy and disease phases separately and not all sows had records in both phases. For some sows, therefore, there was no direct information available in one phase and information came solely from the correlation between sow performances. With the reaction norm model standard errors were smaller because all records contributed to the estimation of variances and covariances of the reaction norm and also to the correlation between the healthy and disease phases. For this reason, the standard errors of the correlation estimates were larger using the bivariate model as compared to the reaction norm model. When pedigree information is available, this problem would be alleviated because relatives would contribute information to the phase in which the sow has no observations available.

Reaction norm models are powerful methods to study host tolerance in response to diseases and have been used in plants (Simms and Triplett, 1994; Simms, 2000) and animals (Raberg *et al*., 2007; Kause, 2011; Kause *et al*., 2012). We showed the merit of reaction norm models to estimate variation among sows in responses to PRRS and re-ranking of sows between healthy and disease phases. We used the estimates of contemporary groups, from a linear model, as the continuous environmental parameter in the reaction norm model. Estimates of contemporary groups have been used as environmental parameter for reaction norm models in other studies such as Pollott and Greeff (2004), Lillehammer *et al*. (2009), and Li and Hermesch (2012). This method provides a practical approach for pork producers to select animals that are robust in production during disease phases. It must be noted that performance of a sow during PRRS outbreaks is a function of 1) performance in PRRS free environment, 2) the degree to which a sow is infected

with PRRS virus (resistance), and 3) the degree to which the sow is performing well despite the infection (tolerance). In the current study these three components underlying performance in disease phases couldn't be disentangled because the individual pathogen burden was not known. Superior performance of a sow during PRRS outbreak, therefore, could be due to higher initial performance in a PRRS free situation (intercept), resistance to PRRS virus, tolerance to PRRS infection or a combination of the three (Kause *et al*., 2012). As a consequence, selection of sows based on performance in disease phase is likely to improve all three components, but it is unknown to which extent. The reaction norm model had the best fit according to AIC and the highest predictive ability in healthy and disease phases. This was because sow effects were estimated on a continuous environmental scale and in the absence of pedigree the model takes advantage of repeated measures in healthy and disease phases along the continuity of the environmental scale. Therefore, the reaction norm model using contemporary group means seems a more powerful selection method to increase performance during disease phases than the bivariate model.

Interpretation of the correlation between intercept and slope is not straightforward in reaction norm models because changing the position of intercept would lead to different correlations between intercept and slope, as shown in Van Tienderen and Koelewijn (1994). They showed that changing the position of the intercept in a reaction norm model could change the correlation between intercept and slope from −1 to 1 with a sigmoid shape. In the current study, the reaction norm model set the intercept at the zero estimate of herd-year-week for NBA. The healthy phase of the farm ranged from 2 to −2.326 herd-year-week estimates of NBA indicating that the intercept was placed almost in the middle of the healthy herd-year-weeks. Therefore, the correlation between intercept and slope can be interpreted as the correlation between performance in healthy phases and the change in performance, e.g. due to PRRS outbreaks. The negative genetic correlation indicates that sows with high performance in average environments have small reduction in performance due to PRRS, whereas sows with a lower performance seem to have a larger reduction in performance.

Interpretation of slopes of individual sows might be affected by the data structure, i.e. not all sows had records in disease phases. For instance about 50% of the sows did not have records in diseased phases and it could well be that these sows have flatter reaction norms. A threshold reaction norm model as used for heat stress (Ravagnolo and Misztal, 2000) might be useful, as it will model the response to PRRS in a more direct manner. Sows without records in diseased

phases, therefore, would have flat reaction norms. On the other hand, the reaction norm model used here does not distinguish between healthy and diseased phases, but uses the whole continuum of fluctuations, e.g. due to mild outbreaks of PRRS, other diseases or other disturbances. The current approach, therefore, makes better use of all data to estimate reaction norms, e.g. to increase general robustness. In the case of diseases, reaction norm models are mainly used in studies on tolerance to infection. As discussed earlier, using average of contemporary groups as environmental parameter in the model could lead to biased estimates of variation in tolerance to infection (Doeschl-Wilson *et al*., 2012; Kause and Odegard, 2012) because using general herd characteristics instead of pathogen burden leads to confounding effects of resistance and tolerance. To obtain an accurate estimate of tolerance to infectious diseases, measuring pathogen burden on an individual basis is needed. Measuring pathogen burden in different stages of pregnancy and its effect on performance, would require monitoring virus load of sows in short sampling intervals (Boddicker *et al*., 2012; Rowland *et al*., 2012). In field studies, collecting blood samples with reasonable intervals would be laborious and costly. With the approach presented in this study, producers can select sows that maintain performance at high levels regardless of resistance and tolerance abilities, which may improve the general robustness of pigs against PRRS and reduce related economic losses.

2.5 Acknowledgement

This research is a part of the NematodeSystemHealth project, financed by Marie Curie Initial Training Networks (ITN), and co-financed by TOPIGS Research Center IPG, The Netherlands. The authors acknowledge contributions from Julie Ménard and Benoît Laplante of F. Ménard, Inc., Canada, for providing the data and background information, Roel Veerkamp and Mario Calus of Wageningen University, The Netherlands, and Maaike Gonggrijp of De Gezondheidsdienst voor Dieren, The Netherlands, for their useful comments.

2.6 References

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Proc. 2nd Int. Symp. Information Theory. Akademiai Kiado, Budapest, Hungary. 267-281.

Batista, L., C. Pijoan, S. Dee, M. Olin, T. Molitor, H. S. Joo, Z. G. Xiao, and M. Murtaugh. 2004. Virological and immunological responses to porcine reproductive and respiratory syndrome virus in a large population of gilts. Can. J. Vet. Res. 68: 267-273.

Bergsma, R., and S. Hermesch. 2012. Exploring breeding opportunities for reduced thermal sensitivity of feed intake in the lactating sow. J. Anim. Sci. 90: 85-98.

Boddicker, N., E. H. Waide, R. R. R. Rowland, J. K. Lunney, D. J. Garrick, J. M. Reecy, and J. C. M. Dekkers. 2012. Evidence for a major QTL associated with host response to porcine reproductive and respiratory syndrome virus challenge. J. Anim. Sci. 90: 1733-1746.

Calus, M. P. L., A. F. Groen, and G. de Jong. 2002. Genotype × environment interaction for protein yield in Dutch dairy cattle as quantified by different models. J. Dairy Sci. 85:3115-3123.

Doeschl-Wilson, A. B., B. Villanueva, and I. Kyriazakis. 2012. The first step towards genetic selection for host tolerance to infectious pathogens: obtaining the tolerance phenotype through group estimates. Front. Genet. 3: 265-275.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Longman, Harlow.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., and Thompson, R. 2009 ASReml user guide release 3.0 VSN International Ltd, Hemel Hempstead, HP1 1ES, UK. www.vsni.co.uk

Houben, S., K. Vanreeth, and M. B. Pensaert. 1995. Pattern of infection with the porcine reproductive and respiratory syndrome virus on swine farms in Belgium. J. Vet. Med. B 42: 209-215.

Huang, Y. W., and X. J. Meng. 2010. Novel strategies and approaches to develop the next generation of vaccines against porcine reproductive and respiratory syndrome virus (PRRSV). Virus Res. 154: 141-149.

Kause, A. 2011. Genetic analysis of tolerance to infections using random regressions: a simulation study. Genet. Res. 93: 291-302.

Kause, A., and J. Odegard. 2012. The genetic analysis of tolerance to infections: a review. Front. Genet. 3: 262-270.

Kause, A., S. van Dalen, and H. Bovenhuis. 2012. Genetics of ascites resistance and tolerance in chicken: a random regression approach. G3 (Bethesda) 2: 527-535.

Kemp, B., and N. M. Soede. 2012. Reproductive issues in welfare-friendly housing systems in pig husbandry: a review. Reprod. Domest. Anim. 47: 51-57.

Lewis, C. R. G., T. Ait-Ali, A. Wilson, D. G. Westcott, J. P. Frossard, B. Naidu, M. A. Mellencamp, M. Torremorell, T. Drew, S. C. Bishop, and A. L. Archibald. 2010. Effects of porcine reproductive and respiratory syndrome virus infection on the performance of pregnant gilts and growing pigs. Anim. Prod. Sci. 50: 890-896.

Lewis, C. R. G., T. Ait-Ali, M. Clapperton, A. L. Archibald, and S. Bishop. 2007. Genetic perspectives on host responses to porcine reproductive and respiratory syndrome (PRRS). Viral Immunol. 20: 343-357.

Lewis, C. R. G., M. Torremorell, L. Galina-Pantoja, and S. C. Bishop. 2009. Genetic parameters for performance traits in commercial sows estimated before and after an outbreak of porcine reproductive and respiratory syndrome. J. Anim. Sci. 87: 876-884.

Li, L. and S. Hermesch. 2012. Genotypes differ in their response to variation in environments experienced by pigs on farm. AGBU Pig Genetics Workshop, University of New England, Armidale: 53-60.

Lillehammer, M., J. Odegard, and T. H. E. Meuwissen. 2009. Reducing the bias of estimates of genotype by environment interactions in random regression sire models. Genet. Sel. Evol. 41: 30-36.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. 1st ed. Sinauer Associates, Inc., Sunderland, MA.

Mckay, R. M. 1993. Preweaning losses of piglets as a result of index selection for reduced backfat thickness and increased growth-rate. Can. J. Anim. Sci. 73: 437-442.

Murtaugh, M. P., and M. Genzow. 2011. Immunological solutions for treatment and prevention of porcine reproductive and respiratory syndrome (PRRS). Vaccine 29: 8192-8204.

Murtaugh, M. P., and R. R. R. Rowland. 2004. Special issue - Immunology and immunopathology of porcine reproductive and respiratory syndrome (PRRS) - Preface. Vet. Immunol. Immunopathol. 102: 105-105.

Nodelijk, G., M. C. M. de Jong, L. A. M. G. van Leengoed, G. Wensvoort, J. M. A. Pol, P. J. G. M. Steverink, and J. H. M. Verheijden. 2001. A quantitative assessment of the effectiveness of PRRSV vaccination in pigs under experimental conditions. Vaccine 19: 3636-3644.

Pollott, G. E., and J. C. Greef. 2004. Genetic relationships between faecal egg count and production traits in commercial Merino sheep flocks. J. Anim. Sci. 79:21–32.

Raberg, L., D. Sim, and A. F. Read. 2007. Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. Science 318: 812-814.

Ravagnolo, O., and I. Misztal. 2000. Genetic component of heat stress in dairy cattle, parameter estimation. J. Dairy Sci. 83:2126–2130.

Read, A. F., A. L. Graham, and L. Raberg. 2008. Animal defenses against infectious agents: is damage control more important than pathogen control? PLoS Biol. 6: 2638-2641.

Rowland, R., J. Lunney, and J. Dekkers. 2012. Control of porcine reproductive and respiratory syndrome (PRRS) through genetic improvements in disease resistance and tolerance. Front. Genet. 3: 260-265.

Simms, E. L. 2000. Defining tolerance as a norm of reaction. Evol. Ecol. 14: 563-570.

Simms, E. L., and J. Triplett. 1994. Costs and benefits of plant-responses to disease - resistance and tolerance. Evolution 48: 1973-1985.

van der Peet-Schwering, C. M. C., B. Kemp, G. P. Binnendijk, L. A. den Hartog, P. F. G. Vereijken, and M. W. A. Verstegen. 2004. Effects of additional starch or fat in late-gestating high nonstarch polysaccharide diets on litter performance anal glucose tolerance in sows. J. Anim. Sci. 82: 2964-2971.

van Tienderen, P. H., and H. P. Koelewijn. 1994. Selection on reaction norms, genetic correlations and constraints. Genet. Res. 64: 115-125.

Vincent, A. L., B. J. Thacker, P. G. Halbur, M. F. Rothschild, and E. L. Thacker. 2006. An investigation of susceptibility to porcine reproductive and respiratory syndrome virus between two genetically diverse commercial lines of pigs. J. Anim. Sci. 84: 49-57.

Visscher, P. M., and M. E. Goddard. 1993. Fixed and random contemporary groups. J. Dairy Sci. 76: 1444-1454.

Zimmerman, J., D. A. Benfield, M. P. Murtaugh, F. Osorio, G. W. Stevenson, and M. Torremorell. 2006. Porcine reproductive and respiratory syndrome virus (Porcine Arterivirus). Pages 387-417 in Diseases of Swine. 9th ed. B. E. Straw, J. J. Zimmerman, S. D'Allaire, and D. J. Taylor, ed. Blackwell Publishing, Ames, IA.

3

A Trade-off between Resistance and Tolerance to Nematode Infection in Domestic Sheep (*Ovis aries*)

H. Rashidi[1], H. A. Mulder[1], L. Matthews[2,3], J. A. M. van Arendonk[1], and M. J. Stear[2,3]

[1] Animal Breeding and Genomics Centre, Wageningen University, 6700 AH, Wageningen, The Netherlands, [2] Institute of Biodiversity, Animal Health & Comparative Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, G61 1QH, Glasgow, United Kingdom, and [3] The Boyd Orr Centre for Population and Ecosystem Health, G61 1QH, Glasgow, United Kingdom

To be submitted

Abstract

Resistance and tolerance are the two main mechanisms of host defence against infection. Resistance is the ability to prevent pathogen entry or control replication of pathogens in body. Tolerance is the ability to reduce pathogen-caused damage. Breeding for resistance is widespread in farm animals. However, there is uncertainty about whether it is better to breed for resistance or tolerance. Though the genetics of resistance to infection has been widely investigated, the genetics of tolerance to infection and, critically, its relationship with resistance remains poorly understood. We applied a random regression model to quantify the relationship between changes in body weight of lambs naturally infected with the nematode *Teladorsagia circumcincta* and changes in their faecal nematode egg count (an indication of nematode burden) and pepsinogenaemia (an indication of damage caused by nematodes in abomasum). We observed a significant additive genetic variation for tolerance, measured as the slopes of the reaction norms. This indicates the possibility of improving tolerance by selective breeding. We also applied a bivariate model to study the genetic correlation between resistance and tolerance, where, resistance was measured as increased Immunoglobulin A (a mucosal antibody that regulates nematode growth and fecundity) and decreased faecal nematode egg count. A negative genetic correlation was observed between tolerance and resistance, indicating that genetically more resistant animals are less tolerant. This is the first study reporting the trade-off between resistance and tolerance to an economically important pathogen in livestock. These findings indicate that unless both traits are included in breeding programs, breeding for increased resistance may decrease tolerance.

Key words: Nematode infection, Resistance, Tolerance, Trade-off, Sheep Text

3.1 Introduction

Resistance and tolerance are the two main mechanisms of host defence against infection. Resistance is the ability of a host to control pathogen burden by, for example, preventing the pathogen from entering the body or stopping the replication of the pathogen within the host. Tolerance is the ability of a host to minimize the impact of infection on performance without influencing the invading pathogen (Bishop, 2012; Painter, 1958; Simms and Triplett, 1994). Improving host resistance will diminish the transmission of the infection in the population (Bishop and MacKenzie, 2003). Resistance, however, limits the survival and reproduction of the pathogen. Resistance imposes selection advantages on pathogens that can overcome resistance. This may lead to an antagonistic co-evolution between the host and pathogen (Parker *et al.*, 2014; Stear *et al.*, 2001; Woolhouse *et al.*, 2002). Tolerance, on the other hand, does not necessarily reduce infection prevalence because a tolerant host can still spread the pathogen in the environment. Therefore, tolerance is an attractive target trait for animal breeders because it does not enforce pathogen to evolve (Råberg *et al.*, 2009; Rausher, 2001; Read *et al.*, 2008). Currently, there is little evidence to support this thought but more evidences to support the opposite. Tolerance can improve the host condition either by limiting the damages in host without influencing the pathogen or by limiting the pathogen's virulence. Vale *et al.* (2014) showed that when there is a trade-off between virulence and transmission, reducing virulence without reducing pathogen burden could lead to pathogen evolution. Tolerance might increase transmission rate as well as virulence (Vale *et al.* 2014). Increased transmission rate of the pathogen is a serious threat for nearby populations or newcomer hosts which are not tolerant. Nevertheless, tolerance reduces the symptoms of the infection in the host. This may provide time for the immune system to clear the infection resulting in decreased pathogen burden. Tolerance could be considered as a complementary to other treatments for eliminating diseases and is worthy to be further studied.

Tolerance could be measured as the slope for the reaction norm of individual's performance on environmental stressors (Simms, 2000). In animals, this approach was firstly used to study tolerance to infections as the norm of reaction for body weight and red blood cells density over density of Plasmodium falciparum in five inbred strains of laboratory mice (Råberg *et al.*, 2007). Random regression models are statistical tools to study reaction norms. When combined with the additive genetic relationship matrix of the individuals, random regression models split the genetic effect on the phenotype into the genetic effect on intercept and

the genetic effect on slopes of the reaction norm curves and estimates the covariance between intercept and slope. In a simulation study Kause (2011) used random regression model to study tolerance to infection as the slope of the individual reaction norms over pathogen burden. He showed that random regression is a powerful approach to study tolerance especially in farm animals with large family size. Random regression models were also used to study tolerance of animals to production diseases like e.g. ascites in domesticated chicken (Kause *et al.*, 2012). Hayward *et al.*, (2014a) and (2014b) studied tolerance as the reaction norm of body weight on gastrointestinal nematode burden in an unmanaged population of Soay sheep. Parker *et al.* (2014) studied variation among pea aphid (Acyrthosiphon pisum) genotypes in tolerance to a fungi infection (Pandora neoaphidis) as the slope of the reaction norm of individuals over the infection dose. In plants, the genetics of tolerance have been more extensively studied than in animals. Several studies have shown variation among morning glory inbred lines in tolerance to damages caused by folivories and herbivores (Fineblum and Rausher, 1995; Simms and Triplett, 1994; Tiffin and Rausher, 1999).

Instead of distinguishing between resistance and tolerance, other studies have reported the genetic basis of resilience. Resilience is measured as the ability to maintain performance during an infected period without measuring individual pathogen burdens (Albers *et al.*, 1987; Bisset *et al.*, 1994; Rashidi *et al.*, 2014). Resilience, therefore, cannot distinguish between resistance and tolerance due to the absence of individual pathogen burdens (Doeschl-Wilson *et al.*, 2012; Kause and Ødegård, 2012), because a healthy looking animal during an outbreak may be resistant, tolerant or even unexposed to infection. To study tolerance, therefore, records of individual pathogen burdens are necessary.

The genetic basis of resistance has been explored in many studies of host-pathogen interactions in livestock (Bishop and Morris, 2007; Chang *et al.*, 2014; Detilleux, 2009; Kuukka-Anttila *et al.*, 2010; Lewis *et al.*, 2010; Stear *et al.*, 2009). Despite the availability of the statistical tools, however, tolerance to infection and its correlation with resistance to infection have been largely overlooked in livestock. To breed for tolerance or resistance, it is important to know if there is any genetic correlation between these traits, because if for example there is an unfavourable genetic correlation between tolerance and resistance, improving one would decrease the other one.

We studied the trade-off between resistance and tolerance to nematode infection in sheep as an example of a globally important livestock system for which selective breeding is a key control measure. Natural infection with gastrointestinal nematodes reduces growth in grazing sheep (Coop *et al.*, 1977; Coop *et al.*, 1982).

Infected sheep excrete nematode eggs in faeces, which hatch and moult to become infective larvae and are ingested by grazing sheep. The larvae migrate to the abomasum and mature in the gastric glands. The mature female nematodes mate and produce eggs and the cycle continues. In response to gastrointestinal infection, lambs produce Immunoglobulin A which regulates parasite growth and fecundity (Stear *et al.*, 1995). In addition, in response to the damage caused by the nematodes in the abomasum, which is one of the reasons for reduced growth of animals, pepsinogen levels rise in the bloodstream of the infected sheep (Stear *et al.*, 1999). To maintain animal health and prevent economic losses, anthelmintic drugs are widely used to control nematode infection but this method of control is threatened by resistance of nematodes against anthelmintics (Jackson *et al.*, 2009). Selective breeding is an attractive option for disease control (Stear *et al.*, 2001) but there is debate about whether to select for resistance or tolerance (Bishop, 2012; Morris *et al.*, 2010). In this study, therefore, our objectives were: 1) to quantify genetic variation in tolerance of sheep to nematode infection in terms of reaction norm of body weight on faecal egg count and pepsinogen, and 2) to examine the genetic correlation between resistance and tolerance to nematode infection.

3.2 Material and methods

3.2.1 Data

From a commercial flock of Scottish Blackface sheep 962 lambs from 38 rams and 492 ewes were studied. The relationships between parents were not known. Lambs were born outside from 1992 to 1996 during the last two weeks of April and the first week of May. Lambs grazed on pasture and were continuously exposed to natural mixed nematode infection. The most obvious sign of nematode infection in sheep is excretion of nematode eggs in faeces. Faecal egg count (FEC), therefore, is known as an adequate estimate of worm burden (Davies *et al.*, 2005; Hayward *et al.*, 2014a; Stear *et al.*, 1995). To measure FEC of lambs, faecal samples were collected from the rectum of lambs at four weeks of age and thereafter at four-week intervals until 20 weeks of age. A full description and analysis of FEC has been provided by Bishop *et al.* (1996). Body weight (lb) was measured at each of the first five faecal sampling dates each year. Plasma activity of Immunoglobulin A (IgA) against mature larvae from *T. circumcincta* was measured in the blood at four, five and six months of age in August, September and October of each year, except for October 1992 and 1993 and August 1995, using an indirect enzyme-linked immunosorbant assay (Strain *et al.*, 2002). Studies have shown that IgA activity is

associated with reduced nematode growth and fecundity (Stear *et al.*, 1995; Strain *et al.*, 2002). High IgA level, therefore, is an indication of reduced nematode replication and consequently reduced FEC. Plasma pepsinogen concentrations were measured in the blood at five months of age in September of each year except 1996, following the method of Paynter (1992) adapted for small quantities. Pepsinogen activity is associated with increased damage to the epithelial barrier (Stear *et al.*, 1999) which is triggered by mast cell degranulation and as a consequence reduced growth (Stear *et al.*, 2003). Higher pepsinogen level, therefore, is an indication of damage in abomasum caused by nematodes and can be used as an indication of nematode burden (Davies *et al.*, 2005). The traits FEC, IgA, and pepsinogen were used as indicator traits for resistance to nematode infection. After collection of each faecal sample, to prevent the lambs dying from an overwhelming infection, all lambs were treated with the broad spectrum anthelmintic "Albendazole sulfoxide". Lambs from the same year were given anthelmintic at the same time. The anthelmintic was given at the recommended dose rate of 5 mg/kg body weight (Bishop *et al.*, 1996). Albendazole sulfoxide disrupts formation of microtubules in nematodes and kills them immediately. The formulation of Albendazole sulfoxide, however, was short-lived and sheep were producing parasite eggs 3-4 weeks after treatment. Lambs were slaughtered six weeks after the final anthelmintic treatment when they were six to seven months old.

3.2.2 Data used for the analysis

From the records of body weight, records at five months of age in September were analysed because growth reduction caused by previous nematode infection are captured in the final record of body weight. For FEC, records at five months of age were used in the random regression model because preliminary results showed that the likelihood of the random regression model using FEC records of five months compared with its corresponding model without slope was largest amongst the models including FEC recorded at younger ages. For IgA, the genetic correlation between records at five month of age and the two other recorded in August and October were 0.70±0.24 and 0.84±0.16, respectively. The high genetic correlations between IgA recorded in three consecutive months indicate that IgA activity in each of the months is genetically highly associated with IgA activity in other months. Therefore, the IgA records in September were also used because they were most complete and closest to the FEC and BW records used for analysis. In general, records at 5 month of age were the most complete ones as compared to the records of the other ages. We used only the records at

five month of age and not the records at other ages to avoid the problem of heterogeneity of genetic variance across ages. The genetic effect on body weight might be different in different ages due to the change in genetic architecture of the trait (Hayward *et al.*, 2014a). After removing the missing and incomplete records at 5 month of age, about 700 lambs remained from 29 rams and 381 ewes. The number of offspring per sire ranged from 2 to 72 with mean 23 and SD of 17 (for more details see table A1 in Appendix 3.1). In general, sires had offspring in high and low levels of FEC and pepsinogen. Summary statistics of the traits and number of records available per analysis are in Table 3.1. Prior to analyses, traits were log-transformed as ln(trait+1) for FEC (lnFEC), IgA (lnIgA), and pepsinogen (lnPeps) to make them normally distributed.

Table 3.1. Summary statistics of the traits body weight (BW), faecal egg count (FEC), log-transformed FEC (lnFEC), plasma IgA activity against mature larvae from *T. circumcincta* (IgA), log-transformed IgA (lnIgA), plasma pepsinogen concentrations, and log-transformed pepsinogen (lnPeps). Traits were recorded at 5 month of age and log-transformed as ln(trait+1).

Trait[1]	Mean	Standard deviation	Minimum	Maximum	Number of records
BW (lbs)	29.13	4.44	17	43	687
FEC	213.47	299.21	0	2700	673
lnFEC	3.98	2.33	0	7.9	673
IgA	21.11	17.12	0	109.69	699
lnIgA	2.75	0.94	0	4.71	699
Pepsinogen	29.75	29.49	0	250.32	685
lnPeps	2.80	1.39	0	5.53	685

[1]Number of records available in analyses with more than one trait (multivariate analysis) ranged from 662 to 688.

3.2.3 Statistical analysis

Four statistical mixed models were used. The first model was a univariate mixed model [1] for heritability estimation of the traits. The second model was a random regression model [2] to estimate the change of body weight concomitant with the change in FEC and Pepsinogen. The third model was a trivariate mixed model [3] to estimate the genetic variance of body weight with low, medium, and high level of FEC and Pepsinogen. The fourth model was a bivariate mixed model [4] to estimate the genetic correlation between tolerance and resistance to nematode infection.

3.2.4 Heritability estimation of body weight and resistance traits

To study the genetic variation and heritability of body weight and resistance traits, we applied a univariate mixed model in which the sire effect on each trait was estimated. The model was as follows:

$$Y_{ijkl} = \mu + SEX_i + AGE_j + YEAR_k + sire_l + e_{ijkl}, \quad [1]$$

where Y_{ijkl} is the phenotype of the l^{th} lamb (body weight, lnFEC, lnIgA, and lnPeps), μ is the overall mean, SEX is the fixed effect of the i^{th} sex of lamb, AGE is the fixed effect of the j^{th} age of lamb, $YEAR$ is the fixed effect of the k^{th} year. For body weight, a fixed covariate of lnFEC or lnPeps was included in the univariate mixed model to make it comparable to the random regression model (see below), which enabled us to compare the models using Akaike information criterion and likelihood ratio test (see below). Random effects were: sire, which is the effect of the l^{th} sire with $\mathbf{N}(\mathbf{0}, \mathbf{A}\sigma^2_{sire})$, where $\mathbf{A}$ is the additive genetic relationship matrix, and σ^2_{sire} is the variance of the sire effect; and e_{ijkl} is the random residual term with $e \sim \mathbf{N}(\mathbf{0}, \mathbf{I}\sigma^2_e)$, where $\mathbf{I}$ is the identity matrix and σ^2_e is the residual variance. We did not include a maternal effect in the model because preliminary results showed that in a univariate analysis of body weight the maternal effect absorbs all the genetic variation so that the direct genetic effect becomes insignificant. The reason was that there were not enough records per ewe, which makes the model unable to disentangle direct genetic effects from maternal effects. Furthermore, Hayward *et al.* (2014a) found insignificant maternal effect on body weight of sheep. Heritability (h^2) was calculated as $h^2 = \frac{\sigma^2_A}{\sigma^2_P} = \frac{4\sigma^2_{sire}}{\sigma^2_P}$, where, σ^2_A is the additive genetic variance for each trait and σ^2_P is the phenotypic variance for each trait (calculated as $\sigma^2_{sire} + \sigma^2_e$).

3.2.5 Genetic analysis of tolerance

We studied tolerance of nematode infection as the sire effect on the slope of the reaction norm of body weight on lnFEC and on lnPeps. We used a random regression model as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za_{int}} + \mathbf{Z_w a_{sl}} + \mathbf{e}, \quad [2]$$

where $\mathbf{y}$ is a vector of body weight; $\mathbf{X}$ is the incidence matrix for fixed effects; $\mathbf{Z}$ is the incidence matrix for the intercept random effect; $\mathbf{Z_w}$ is the incidence matrix for lnFEC or lnPeps as a covariate for the slopes of the reaction norms; $\mathbf{b}$ is a vector of fixed effects (see description of model [1] for the fixed effects), lnFEC or lnPeps

were fixed covariates to account for the average of lnFEC or lnPeps, respectively; $\mathbf{a_{int}}$ is a vector of random sire effect on intercept and $\mathbf{a_{sl}}$ is a vector of random sire effects on slope, with $\begin{bmatrix} \mathbf{a_{int}} \\ \mathbf{a_{sl}} \end{bmatrix} \sim \mathbf{N}\left(\mathbf{0}, \mathbf{G_{RN}} \otimes \mathbf{A}\right)$, where $\mathbf{G_{RN}} = \begin{bmatrix} \sigma^2_{a_{int}} & \sigma_{a_{int},a_{sl}} \\ \sigma_{a_{int},a_{sl}} & \sigma^2_{a_{sl}} \end{bmatrix}$, $\sigma^2_{\boldsymbol{a_{int}}}$ is the sire variance for $\mathbf{a_{int}}$, $\sigma^2_{\boldsymbol{a_{sl}}}$ is the sire variance for $\mathbf{a_{sl}}$, and $\sigma_{a_{int},a_{sl}}$ is the covariance between $\mathbf{a_{int}}$ and $\mathbf{a_{sl}}$; $\mathbf{e}$ is the vector of residuals. Heterogeneous residual variances were considered in the model to account for the possibility that residual variance may change with pathogen burden. Data were sorted from small to large based on lnFEC (lnPeps) and grouped into three classes of equal size with low, medium, and high levels of lnFEC (lnPeps). A similar approach was used by (Calus *et al.*, 2002) for estimation of genotype by environment interaction. Table 3.3 shows the number of records in each class for lnFEC and for lnPeps. The additive genetic variance for body weight at each level of infection ($\sigma^2_{A_{BW_i}}$) was calculated as:

$$\sigma^2_{A_{BW_i}} = \sigma^2_{\boldsymbol{a_{int}}} + X_i^2 \times \sigma^2_{\boldsymbol{a_{sl}}} + 2X_i \times (\sigma_{a_{int},a_{sl}}),$$

where, X_i is the level of lnFEC or lnPeps.

The covariance between body weights at two levels of lnFEC or lnPeps ($\sigma_{A_{BW1,BW2}}$) was calculated as:

$$\sigma_{A_{BW1,BW2}} = \sigma^2_{\boldsymbol{a_{int}}} + (X_i X_j) \times \sigma^2_{\boldsymbol{a_{sl}}} + (X_i + X_j) \times (\sigma_{\boldsymbol{a_{int}},\boldsymbol{a_{sl}}}),$$

where, X_i and X_j are the levels i or j of lnFEC or lnPeps.

The genetic correlation between body weights at different levels of lnFEC or lnPeps ($r_{g_{BW1,BW2}}$) was calculated as:

$$r_{g_{BW1,BW2}} = \frac{\sigma_{A_{BW1,BW2}}}{\sqrt{\sigma^2_{A_{BW1}} \sigma^2_{A_{BW2}}}},$$

We used sire random regression models because preliminary analysis of body weight with a univariate animal model did not converge because of singularities in the average information matrix. This happened because the animal model did not have enough information to disentangle the residual from the genetic effects for intercept and slope. Therefore, we used therefore a sire model, which did not have this problem because of multiple offspring per sire.

To check the results of the random regression model and make sure that the obtained heritabilities and correlations were not artefacts of the model, a trivariate

model was applied. In the trivariate model the genetic effect on body weight was estimated at low, medium, and high levels of lnFEC and of lnPeps. The genetic correlations between body weights at low, medium, and high levels of lnFEC and of lnPeps were also estimated. The model was:

$$\begin{bmatrix} \mathbf{BW}_l \\ \mathbf{BW}_m \\ \mathbf{BW}_h \end{bmatrix} = \begin{bmatrix} \mathbf{X}_l & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{X}_h \end{bmatrix} \begin{bmatrix} \mathbf{b}_l \\ \mathbf{b}_m \\ \mathbf{b}_h \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_l & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Z}_h \end{bmatrix} \begin{bmatrix} \mathbf{a}_l \\ \mathbf{a}_m \\ \mathbf{a}_h \end{bmatrix} + \begin{bmatrix} \mathbf{e}_l \\ \mathbf{e}_m \\ \mathbf{e}_h \end{bmatrix}, \quad [3]$$

where $\mathbf{BW}_l$, $\mathbf{BW}_m$, $\mathbf{BW}_h$ are the vectors of body weight in environments with low, medium, and high levels of lnFEC or lnPeps; $\mathbf{X}_l$, $\mathbf{X}_m$, and $\mathbf{X}_h$ are the incidence matrices for fixed effects in environments with low, medium, and high levels of lnFEC or lnPeps; $\mathbf{Z}_l$, $\mathbf{Z}_m$, and $\mathbf{Z}_h$ are the incidence matrix for random effects in environments with low, medium, and high levels of lnFEC or lnPeps; $\mathbf{b}_l$, $\mathbf{b}_m$, and $\mathbf{b}_h$ are the vectors of the fixed effects (see model [1] for fixed effects) in environments with low, medium, and high levels of lnFEC or lnPeps; $\mathbf{a}_l$, $\mathbf{a}_m$, and $\mathbf{a}_h$ are the vectors of the sire random effect with $\begin{bmatrix} \mathbf{a}_l \\ \mathbf{a}_m \\ \mathbf{a}_h \end{bmatrix} \sim N(\mathbf{0}, \begin{bmatrix} \sigma^2_{a_l} & \sigma_{a_l,a_m} & \sigma_{a_l,a_h} \\ \sigma_{a_l,a_m} & \sigma^2_{a_m} & \sigma_{a_m,a_h} \\ \sigma_{a_l,a_h} & \sigma_{a_m,a_h} & \sigma^2_{a_h} \end{bmatrix} \otimes \mathbf{A})$; $\sigma^2_{a_l}$, $\sigma^2_{a_m}$, and $\sigma^2_{a_h}$ are the sire variances for $\mathbf{a}_l$, $\mathbf{a}_m$, and $\mathbf{a}_h$; σ_{a_l,a_m}, σ_{a_m,a_h} and σ_{a_l,a_h} are the sire covariances between $\mathbf{a}_l$ and $\mathbf{a}_m$, $\mathbf{a}_m$ and $\mathbf{a}_h$, and $\mathbf{a}_l$ and $\mathbf{a}_h$; $\mathbf{e}_l$, $\mathbf{e}_m$, and $\mathbf{e}_h$ are the vectors of the residuals for low, medium, and high levels of lnFEC or lnPeps.

3.2.6 Tolerance coheritability

We calculated the tolerance coheritability as the coheritability for the slope of the reaction norm of body weight on FEC, applying the following formula described in Sae-Lim *et al.* (2015):

$$h^2_{slope} = \frac{r_g \sigma_{a,E_1} \sigma_{a,E_2} - \sigma^2_{a,\text{int}}}{\sigma^2_{P,\text{int}}},$$

where, h^2_{slope} is the slope coheritability, r_g is the genetic correlation between body weight in two levels of FEC, σ_{a,E_1} and σ_{a,E_2} are the genetic standard deviations of body weight in each of the two levels of FEC, $\sigma^2_{a,\text{int}}$ is the additive genetic variance of body weight at the intercept point, and $\sigma^2_{P,\text{int}}$ is the phenotypic variance of body weight at the intercept point. The coheritability describes the heritable association between BW and slope in one environment. The coheritability sign elucidates the

change in correlated response of tolerance when selecting for higher phenotypic value in one environment (Sae-Lim *et al.*, 2015). We calculated the coheritability between 3 environments: zero and medium level of FEC (lnFEC=4.7), zero and high level of FEC (lnFEC=6.1), and medium and high level of FEC. For all the scenarios we calculated the coheritability when the selection environment (intercept) was considered at both mildest and harshest environments in terms of FEC level.

3.2.7 Genetic relationship between tolerance and resistance

The genetic correlation between tolerance and resistance of lambs to nematode infection was estimated using a bivariate model. The first response variable in the bivariate model was body weight, with a random slope applied to it, and the second response variable was FEC or IgA, without a random slope applied to it. The model was as follows:

$$\begin{bmatrix} \mathbf{y}_{\mathrm{BW}} \\ \mathbf{y}_{\mathrm{R}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{\mathrm{BW}} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{\mathrm{R}} \end{bmatrix} \begin{bmatrix} \mathbf{b}_{\mathrm{BW}} \\ \mathbf{b}_{\mathrm{R}} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{\mathrm{int}} & \mathbf{Z}_{\mathrm{sl}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Z}_{\mathbf{R}} \end{bmatrix} \times \begin{bmatrix} \mathbf{a}_{\mathrm{int}} \\ \mathbf{a}_{\mathrm{sl}} \\ \mathbf{a}_{\mathrm{R}} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_{\mathrm{BW}} \\ \mathbf{e}_{\mathrm{R}} \end{bmatrix}, \qquad [4]$$

where $\mathbf{y}_{\mathrm{BW}}$ is the vector of body weight and $\mathbf{y}_{\mathrm{R}}$ is the vector of resistance traits (lnFEC or lnIgA); $\mathbf{X}_{\mathrm{BW}}$ and $\mathbf{X}_{\mathrm{R}}$ are the incidence matrices for fixed effects for body weight and resistance traits; $\mathbf{b}_{\mathrm{BW}}$ and $\mathbf{b}_{\mathrm{R}}$ are the vectors of the fixed effects for body weight and resistance traits (see description of model [1] for the fixed effects); $\mathbf{Z}_{\mathrm{int}}$ is the incidence matrix for intercept, $\mathbf{Z}_{\mathbf{sl}}$ is the matrix with the environmental parameter lnFEC as a covariate for the sire effects on slope, and $\mathbf{Z}_{\mathbf{R}}$ is the incidence matrices for sire effect on resistance traits; and $\mathbf{a}_{\mathrm{int}}$, $\mathbf{a}_{\mathrm{sl}}$ and $\mathbf{a}_{\mathrm{R}}$ are the vectors for sire effects on intercept, slope, and resistance traits, respectively, with $\begin{bmatrix} \mathbf{a}_{\mathrm{int}} \\ \mathbf{a}_{\mathrm{sl}} \\ \mathbf{a}_{\mathrm{R}} \end{bmatrix} \sim \mathrm{N}\left(\mathbf{0}, \mathbf{G} \otimes \mathbf{A}\right)$, where $\mathbf{G} = \begin{bmatrix} \sigma^2_{a_{int}} & \sigma_{a_{int}\, a_{sl}} & \sigma_{a_{int}\, a_{\mathrm{R}}} \\ \sigma_{a_{int}\, a_{sl}} & \sigma^2_{a_{sl}} & \sigma_{a_{sl}\, a_{\mathrm{R}}} \\ \sigma_{a_{int}\, a_{\mathrm{R}}} & \sigma_{a_{sl}\, a_{\mathrm{R}}} & \sigma^2_{a_{\mathrm{R}}} \end{bmatrix}$,

where $\sigma^2_{a_{int}}$, $\sigma^2_{a_{sl}}$, and $\sigma^2_{a_{\mathrm{R}}}$ are the sire variances for $\mathbf{a}_{\mathrm{int}}$, $\mathbf{a}_{\mathrm{sl}}$ and $\mathbf{a}_{\mathrm{R}}$, respectively; $\sigma_{a_{int}\, a_{sl}}$, $\sigma_{a_{int}\, a_{\mathrm{R}}}$, and $\sigma_{a_{sl}\, a_{\mathrm{R}}}$ are the sire covariances between $\mathbf{a}_{\mathrm{int}}$ and $\mathbf{a}_{\mathrm{sl}}$, $\mathbf{a}_{\mathrm{int}}$ and $\mathbf{a}_{\mathrm{R}}$, and $\mathbf{a}_{\mathrm{sl}}$ and $\mathbf{a}_{\mathrm{R}}$, respectively; $\mathbf{e}_{\mathrm{BW}}$ and $\mathbf{e}_{\mathrm{R}}$ are the vector of residuals for body weight and resistance traits, respectively, with

$$\begin{bmatrix} \mathbf{e}_{\mathrm{BW}} \\ \mathbf{e}_{\mathrm{R}} \end{bmatrix} \sim \mathrm{N}\left(\mathbf{0}, \begin{bmatrix} \sigma^2_{e_{BW}} & \sigma_{e_{BW},e_{\mathrm{R}}} \\ \sigma_{e_{BW},e_{\mathrm{R}}} & \sigma^2_{e_{\mathrm{R}}} \end{bmatrix} \otimes \mathbf{I} \right)$$, where $\sigma^2_{e_{BW}}$ and $\sigma^2_{e_{\mathrm{R}}}$ are the variances of $\mathbf{e}_{\mathrm{BW}}$ and $\mathbf{e}_{\mathrm{R}}$; $\sigma_{e_{BW},e_{\mathrm{R}}}$ is the covariance between $\mathbf{e}_{\mathrm{BW}}$ and $\mathbf{e}_{\mathrm{R}}$.

All the variance components were estimated using ASReml (Gilmour *et al.*, 2009).

3.2.8 Model comparison

Akaike information criterion. The goodness of fit for the univariate mixed model of body weight and the random regression models was studied with the Akaike information criterion **(AIC**) using the following formula (Akaike, 1973):

$AIC = -2logL + 2K$,

where $logL$ is the logarithm of likelihood for the model, and K is the number of variables in the model.

Likelihood ratio test. Likelihood ratio tests **(LRT)** were used to compare the fit of the univariate mixed model of body weight and the random regression models. It was also used to compare the bivariate models with and without the genetic correlation between slope and the resistance traits (lnFEC and lnIgA). The following formula was used (Lynch and Walsh, 1998):

$$LRT = -2logL_1 + 2logL_2,$$

where, $logL_1$ is the logarithm of the likelihood for the univariate mixed models of body weight (or bivariate model with zero correlation between the slope and the resistance indicator traits), and $logL_2$ is the logarithm of likelihood for the random regression models (or bivariate model with correlation between the slope and the resistance traits). To compare the univariate mixed model with the random regression model, the LRT was assumed to follow a 50%-50% mixture of χ^2_1 and χ^2_2 distribution, leading to a 5% threshold of 5.14 (Stram and Lee, 1994). To compare the bivariate models with and without genetic correlations between the slope and resistance traits, LRT was assumed to follow χ^2_1 distribution, leading to a 5% threshold of 3.84.

3.3 Results

3.3.1 Heritabilities for body weight and resistance traits

Heritability for body weight was 0.21, for lnFEC was 0.19, for lnIgA was 0.59, and for lnPeps was 0.28 (Table 3.2). The high heritability of the parasite-specific IgA

response means that most of the observed variation is due to genetic variation in the host and a relatively small part is due to variation in exposure, i.e. the environmental component of the phenotypic variation.

3.3.2 Genetic analysis of tolerance

The association between body weight and nematode infection was generally negative (Fig. 3.1 and Fig. 3.2). At population level, the decrease in body weight with increased lnFEC or pepsinogen was negligible showing hardly any sign of growth retardation due to nematode infections. In general, rams showed decreases in body weight as lnFEC or lnPeps increased. Interestingly, there were some rams showing improvements in body weight when nematode burden increases.

The variance components for the random regression model are shown in Table 3.3. There were genetic variances of intercept for both lnFEC and lnPeps environmental factors. The genetic variance of intercept indicates the genetic variation among sheep in body weight when FEC and Peps are zero. The distributions of sire estimates for slope are in Fig. A1 in Appendix 3.1. There was genetic variance for slope when using lnFEC and lnPeps as environmental covariates in the random regression. The genetic variance in slope indicated the genetic variation among sheep in tolerance to nematode infection. There were high negative genetic correlations between intercept and slope. The negative genetic correlation indicates that animals with high body weight at zero FEC or zero pepsinogen were most likely to show greater reductions in body weight at higher FEC and pepsinogen. In contrast animals with lower body weights at zero FEC and zero pepsinogen were least likely to show reduced body weights at higher FEC and pepsinogen. The residual variance in body weight increased as the level of FEC increased whereas, when the level of pepsinogen increased, the residual variance for body weight decreased.

Table 3.2. Heritabilities from the univariate mixed models for 5 traits: body weight (BW), log-transformed faecal egg count (lnFEC), log-transformed Plasma IgA activity against mature larvae from *T. circumcincta* (lnIgA), and log-transformed plasma pepsinogen concentration (lnPeps). Traits were recorded at 5 month of age and log-transformed as ln(trait+1). (Standard errors in parentheses)

Trait	Heritability
Body weight	0.21 (0.11)
lnFEC	0.19 (0.10)
lnIgA	0.59 (0.20)
lnPeps	0.28 (0.13)

Table 3.3. Variance components from the random regression model for reaction norm of body weight on faecal egg count and plasma pepsinogen concentrations. Traits were recorded at 5 month of age. Before the analysis, the log-transformation ln(trait+1) was used on faecal egg count (lnFEC) and pepsinogen (lnPeps). Heterogeneous residual variance was considered at 3 levels (low, medium, and high) of faecal egg count and pepsinogen. (Standard errors in parenthesis)

Source	lnFEC		lnPeps	
	Variance	Number of records[1]	Variance	Number of records[1]
Genetic variance in intercept	1.77 (1.16)		1.37 (1.34)	
Genetic variance in slope	0.09 (0.06)		0.16 (0.15)	
Genetic correlation	-0.85 (0.13)		-0.76 (0.24)	
Residual 1	11.60 (1.15)	224	15.69 (1.58)	221
Residual 2	13.00 (1.27)	224	12.13 (1.19)	221
Residual 3	15.36 (1.51)	225	12.10 (1.20)	220

[1] Number of records per classes of lnFEC and lnPeps

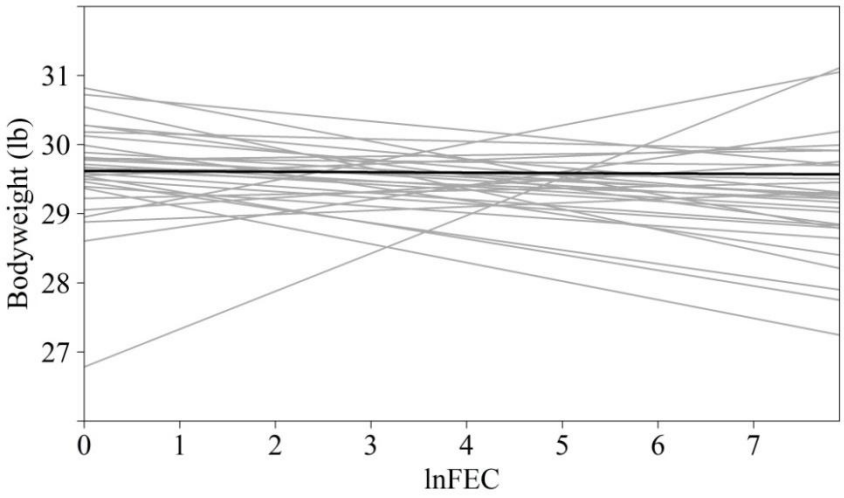

Figure 3.1. Sire effects on the reaction norm of body weight on log-transformed faecal egg count (lnFEC) obtained from the random regression model. The reaction norms (grey lines) are shown as the deviation from the mean reaction norm line (bold line).

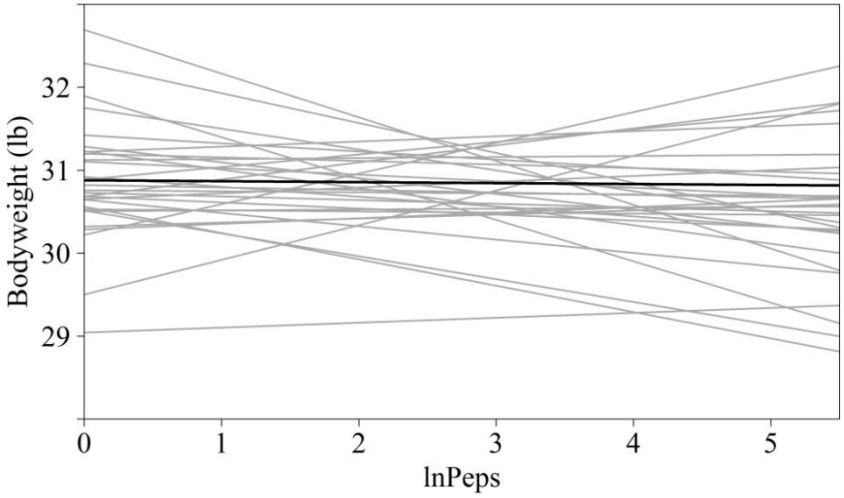

Figure 3.2. Sire effects on the reaction norm of body weight on log-transformed pepsinogen (lnPeps) obtained from the random regression model. The reaction norms (grey lines) are shown as the deviation from the mean reaction norm line (bold line).

The AIC was smaller for the two random regression models with either lnFEC (AIC = 2469.02) or lnPeps (AIC = 2423.04) as a covariate, compared to the univariate mixed models (for lnFEC AIC = 2474.92 and for lnPeps AIC = 2425.30), indicating a better fit of the random regression models than the univariate mixed models. The LRT indicated a significant variance of slope for both random regression models with lnFEC ($P = 0.0037$) or lnPeps ($P = 0.025$) as covariates.

Using the genetic variance-covariance matrix ($\mathbf{G_{RN}}$) from the random regression models, the heritabilities of body weight across different levels of FEC and pepsinogen were calculated. The heritabilities formed a parabolic curve (Fig. 3.3 and Fig. 3.4). For the random regression model of body weight on lnFEC: at zero FEC, the heritability of body weight was 0.53±0.31, dropping to 0.15±0.10 at moderate levels of FEC (lnFEC ≈ 4). At high levels of FEC (lnFEC ≈ 8) the heritability of body weight increased to 0.45±0.27. A similar curve was seen in the random regression model of body weight on lnPeps: at a zero pepsinogen value, the heritability of body weight was 0.32±0.30, dropping to 0.14±0.10 at moderate levels of pepsinogen (lnPeps ≈ 2.5). At high levels of pepsinogen (lnFEC ≈ 5.5), the heritability of body weight increased to 0.62±0.40.

Using the random regression model, the breeding value for body weight was shown to be dependent on the infection level in the animal. Consequently, the changes in the heritability of body weight across different level of lnFEC and lnPeps, was due to changes in genetic variance of body weight while environmental variances were stable. To further explore the changes in the genetic variance of body weight as FEC and pepsinogen vary, a trivariate analysis was performed, where genetic and residual variances were estimated at low, medium and high levels of FEC and pepsinogen (Table 3.4). Heritabilities from random regression models were lower than the heritabilities from the trivariate models. The heritabilities of body weight from the trivariate model for three levels of FEC showed the same trend as random regression model, i.e. higher heritability at low and high levels of FEC. Therefore, the trivariate analysis supported the parabolic shape of the heritability for body weight in different levels of FEC obtained from the random regression models. The heritabilities of body weight from the trivariate model for three levels of pepsinogen were lower at the low to moderate pepsinogen compared to the high level of pepsinogen, which does not agree with the heritabilities from the random regression model.

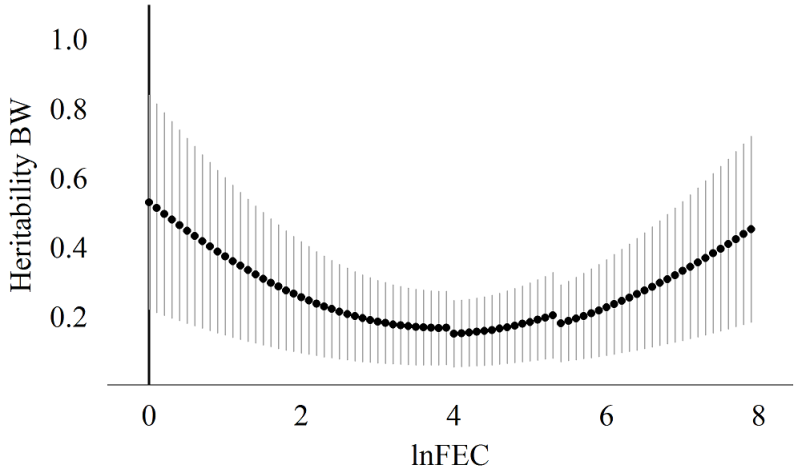

Figure 3.3. Heritability of body weight (BW) recorded at 5 months of age with heterogeneous residual variances for 3 levels of faecal egg count obtained from the random regression model. Faecal egg count was log-transformed (lnFEC) as ln(trait+1). Vertical lines show the ±standard error of each estimate. The breaks in the heritability graph are because of heterogeneous residual variance for three environments with low, medium, and high level of faecal egg count.

Table 3.4. Heritability of body weights for three levels (Low, medium, and high) of log-transformed faecal egg count (lnFEC) and log-transformed plasma pepsinogen concentrations (lnPeps) generated by the random regression and trivariate models. Traits were recorded at 5 month of age and log-transformed as ln(trait+1). (Standard errors in parentheses)

Trait	Level[1]	Model	
		Reaction norm	Trivariate
lnFEC	1	0.35 (0.22)	0.39 (0.25)
	2	0.17 (0.10)	0.25 (0.21)
	3	0.23 (0.14)	0.32 (0.22)
lnPeps	1	0.19 (0.16)	0.25 (0.22)
	2	0.22 (0.12)	0.28 (0.23)
	3	0.33 (0.19)	0.41 (0.26)

[1]low (1), medium (2), and high (3) level of lnFEC and lnPeps

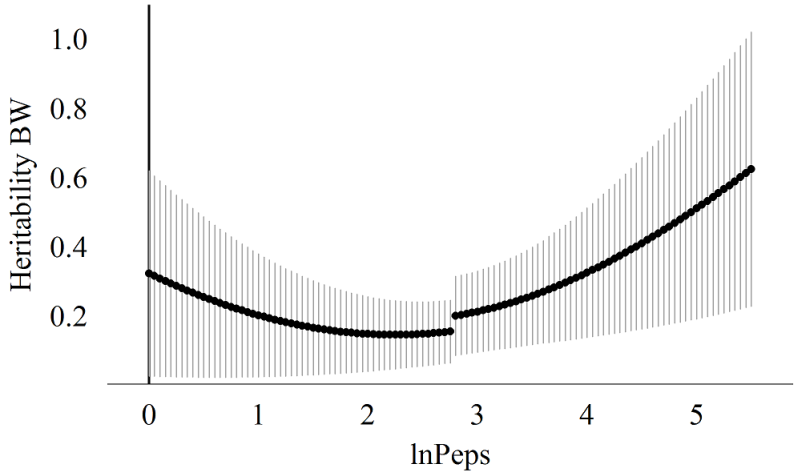

Figure 3.4. Heritability of body weight (BW) recorded at 5 months of age with heterogeneous residual variances for 3 levels of pepsinogen obtained from the random regression model. Pepsinogen was log-transformed (lnPeps) as ln(trait+1). Vertical lines show the ±standard error of each estimate. The breaks in the heritability graph are because of heterogeneous residual variance for three environments with low, medium, and high level of pepsinogen. Residual variances for environments with medium and high level of Pepsinogen were very similar, that is why only one break is visible in this heritability graph.

Genetic correlations between body weight at different levels of FEC or pepsinogen are shown in Fig. 3.5 and Fig. 3.6. The genetic correlations were close to unity for low FEC and pepsinogen and decreased as FEC and pepsinogen increased. At very high levels of FEC and pepsinogen the genetic correlation became negative. The changes in genetic correlation between body weights at different levels of FEC and pepsinogen were checked using the trivariate analysis (Table 3.5). The genetic correlations were calculated from the average of lnFEC and lnPeps at each level. The trends in estimated genetic correlations between body weight in different levels of FEC and pepsinogen were similar between random regression and trivariate models. However, the genetic correlations were generally higher in the random regression model as compared to the trivariate model. For pepsinogen, the genetic correlation between body weight at low and moderate levels was moderately high and positive with the random regression model, whereas for the trivariate model it was moderately low and negative.

Table 3.5. Correlations between body weights in three levels (Low, medium, and high) of log-transformed faecal egg count (lnFEC) and log-transformed plasma pepsinogen concentrations (lnPeps) generated by the random regression and trivariate models. Traits were recorded at 5 month of age and log-transformed as ln(trait+1). (Standard errors in parentheses)

Trait	Level[1]	Model	
		Reaction norm	Trivariate
lnFEC	(1, 2)	0.37 (0.42)	0.10 (0.30)
	(2, 3)	0.92 (0.07)	0.51 (0.61)
	(1, 3)	-0.02 (0.49)	-0.17 (0.57)
lnPeps	(1, 2)	0.57 (0.40)	-0.19 (0.62)
	(2, 3)	0.96 (0.05)	0.70 (0.44)
	(1, 3)	0.30 (0.53)	0.41 (0.60)

[1]low (1), medium (2), and high (3) level of lnFEC and lnPeps

3.3.3 Tolerance coheritability

The slope coheritabilities ranged from -0.61 to 0.03 (Table 3.6). The strongest slope coheritability was at high level of FEC, when the selection environment was at zero FEC. The weakest slope coheritability was at high level of FEC, when the selection environment was at medium level of FEC. The magnitude of the slope coheritability in one environment depended on the selection environment. For example the slope coheritability at medium level of FEC was moderately strong (-0.47±0.31) when the selection environment was at zero FEC, and was weak (-0.07±0.06) when the selection environment was at high level of FEC. The existence of coheritability for slope indicates the possibility for selective breeding for tolerance. The magnitude of the slope coheritability indicates the magnitude of the correlated response and accuracy of selection in one environment when selection is performed in another environment. For example the strong slope coheritability at high level of FEC, when the selection environment is at zero FEC, indicates a strong negative response at high level of FEC with high accuracy of selection. In contrast, the weak slope coheritability at high level of FEC, when the selection environment is at medium level of FEC, indicates almost zero response to selection in slope. The slope coheritabilities were in general negative indicating that the correlated responses for tolerance are generally negative in the response environment.

Table 3.6. Coheritabilities of slope in zero, medium (lnFEC=4.7), and high (lnFEC=6.1) level of FEC when selection environments are different.

Environment (FEC level)		Slope coheritability
Selection	Response	
Zero	Medium	-0.47 (0.31)
Medium	Zero	-0.11 (0.15)
Zero	High	-0.61 (0.41)
High	Zero	-0.30 (0.24)
Medium	High	0.03 (0.05)
High	Medium	-0.07 (0.06)

3.3.4 Genetic correlation between tolerance and resistance

The variance components from the bivariate analysis with lnFEC and lnIgA as resistance traits are shown in Table 3.7 and 3.8, respectively. The variances for intercept and slope and the correlation between intercept and slope were similar to those of the random regression model. The genetic correlation between intercept and lnFEC was strongly negative (-0.76±0.32). There was a moderate positive genetic correlation between the intercept and lnIgA (0.48±0.32). The genetic correlations between intercept and lnFEC and lnIgA indicated that animals with high body weight at zero FEC were genetically more resistant. There was a moderately high positive genetic correlation between the slope and lnFEC (0.60±0.33), indicating that genetically the slopes become less steep when FEC increases. There was a moderately strong negative genetic correlation between the slope and lnIgA (-0.63±0.25), indicating that genetically the slopes become steeper when IgA increases. As FEC is unfavourable and IgA is favourable indicators of resistance, together all these results show that resistance and tolerance are unfavourably genetically correlated. A likelihood ratio test (LRT) showed that the genetic correlation between slope and lnIgA was significantly different from zero ($\boldsymbol{P = 0.046}$) but the genetic correlation between slope and lnFEC was not ($\boldsymbol{P = 0.126}$).

Table 3.7. Genetic variances from random regression model for intercept and slope of the reaction norm of body weight on log-transformed faecal egg count (lnFEC) and lnFEC (on the diagonal) generated by the bivariate random regression model; Genetic correlations on upper off-diagonal. The residual variances for body weight and FEC are at the bottom. Traits were recorded at 5 month of age. (Standard errors in parentheses)

	Intercept	Slope	FEC
Intercept	1.69 (1.16)	-0.86 (0.13)	-0.76 (0.32)
Slope		0.10 (0.07)	0.60 (0.33)
FEC	Sym.		0.21 (0.11)
Residual		13.30[1] (0.75)	4.12 (0.23)

[1]the residual variance for intercept and slope end up in the overall residual variance.

Table 3.8. Genetic variances from random regression model of intercept and slope for the reaction norm of body weight on log-transformed faecal egg count (lnFEC), and log-transformed IgA (lnIgA) (on the diagonal); Genetic correlations on upper off-diagonal. The residual variances for body weight and FEC are at the bottom. Traits were recorded at 5 month of age. (standard errors in parentheses)

	Intercept	Slope	IgA
Intercept	1.93 (1.28)	-0.88 (0.11)	0.48 (0.32)
Slope		0.13 (0.07)	-0.63 (0.24)
IgA	Sym.		0.13 (0.05)
Residual		13.28[1] (0.75)	0.77 (0.04)

[1]the residual variance for intercept and slope end up in the overall residual variance.

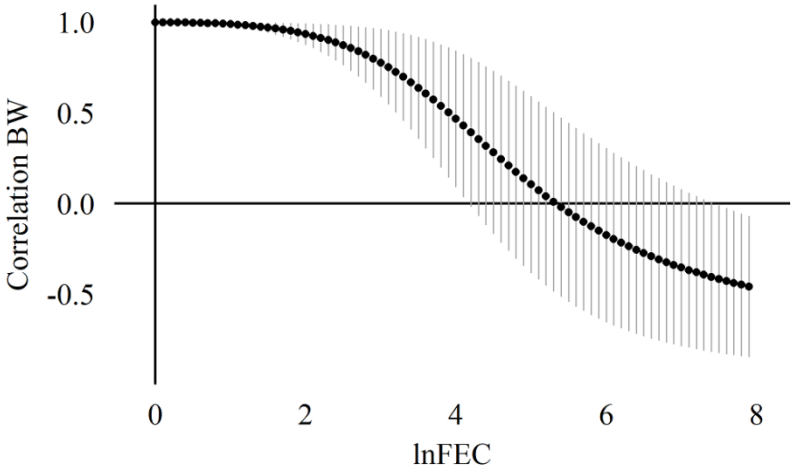

Figure 3.5. Genetic correlation between body weight (BW) at zero faecal egg count and body weight at other levels of faecal egg count obtained from the random regression model. Traits were recorded at 5 month of age. Faecal egg count was log transformed (lnFEC) as ln(trait+1). Vertical lines show the ±standard error of each estimate.

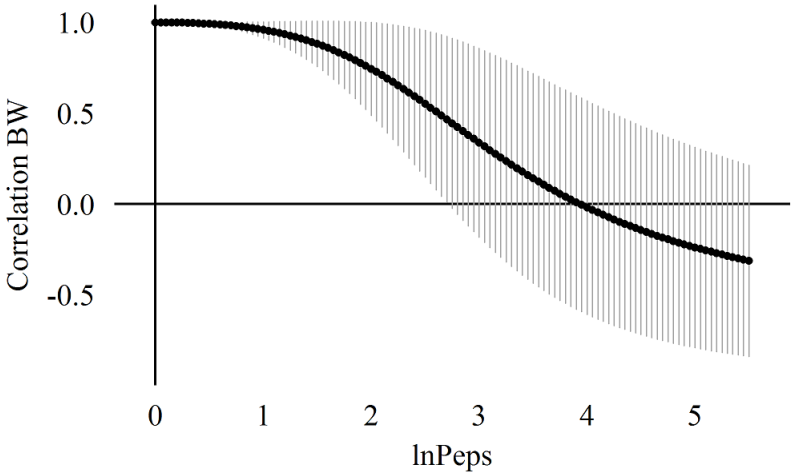

Figure 3.6. Genetic correlation between body weight (BW) at zero pepsinogen and body weight at other levels of pepsinogen obtained from the random regression model. Traits were recorded at 5 month of age. Pepsinogen was log transformed (lnPeps) as ln(trait+1). Vertical lines show the ±standard error of each estimate.

3.4 Discussion

This study has found significant genetic variation in tolerance to nematode infections. In addition, there was a strong negative genetic correlation between resistance and tolerance to nematode infection, which demonstrates a trade-off between tolerance and resistance.

3.4.1 Heritability of body weight and resistance traits

The heritability estimate for body weight was in the range (0.18-0.33) reported by other studies (Borg *et al.*, 2009; Mortimer *et al.*, 2014; Riggio *et al.*, 2008; Rose *et al.*, 2013). The heritability estimates for faecal nematode egg counts were similar to other studies in different breeds (0.11-0.48) in France (Gruner *et al.*, 2004) Australia (Pollott and Greeff, 2004), and New Zealand (Shaw *et al.*, 1999). Our flock, therefore, is representative of genetic variation in response to nematode infection. The heritability of IgA activity against *Teladorsagia circumcincta* was remarkably high (0.59±0.20, Table 3.2). Other studies have also shown high levels of genetic variation in immune responses to nematode infection (Strain *et al.*, 2002). For pepsinogen, the heritability was between the values reported by Davies *et al.* (2005) (0.56±0.16) and Gutierrez-Gil *et al.* (2010) (0.21±0.04). Davies *et al.* (2005) used the same data set as ours but used animal models with different fixed effects.

3.4.2 Genetic analysis of resistance and tolerance

Our results indicated significant genetic variation in both resistance and tolerance to nematode infection. This indicates that it would be possible to improve resistance or tolerance to nematode infections by selective breeding. Previous studies reported genetic variance in tolerance: Råberg *et al.* (2007) found variation in tolerance to *Plasmodium falciparum* in five inbred strains of mice. Kause *et al.* (2012) found genetic variance in tolerance to ascites in domesticated chicken. Ascites, however, is a metabolic disorder and not an infection. Simms and Triplett (1994) and Tiffin and Rausher (1999) found genetic variance in tolerance to parasites in the plant Common Morning Glory. In contrast, Hayward *et al.* (2014b) found no genetic variance in tolerance (measured as the slope of body weight on FEC) to strongyle nematode infection in feral Soay sheep. They, however, found that there is positive phenotypic correlation between tolerance and lifetime breeding success (defined as the lifetime number of lambs born to females or sired by males) suggesting that tolerance was under positive selection. Given that the previous studies of tolerance are on laboratory animals, natural populations, or

metabolic diseases our study is the first to report genetic variation for tolerance to infection in livestock.

We found a negative genetic correlation between intercept (body weight at zero level of FEC or pepsinogen) and slope of the reaction norm (tolerance). The negative genetic correlation between intercept and slope indicates that sheep with high body weight at zero nematode burden may show a severe reduction in body weight when nematode burden increases. This means that for example, when FEC is the environmental factor, improving the intercept by one genetic standard deviation by genetic selection, without considering the slope (tolerance), would result in a decrease of the slope mean by 0.85 genetic standard deviations in the next generation. Selection of sheep based on a nematode free situation, therefore, has negative consequences on the performance when there is infection. We found a positive genetic association between intercept and resistance, meaning that the zero FEC at the intercept point is because the lambs were resistant and not that the environment was nematode free. The genetic variation in body weight at zero FEC (intercept), therefore, cannot be interpreted as variation in vigour of the animals. The genetic correlation between intercept and slope might be different when the intercept is placed in a nematode free environment.

3.4.3 Tolerance coheritability

The novelty of this study is that using this data allowed us to estimate the heritability for tolerance as the coheritability for slope. The coheritability of slope explains to what extent tolerance would change due to mass selection in a certain environment, similar to the classical definition of heritability. The coheritabilities were in general negative, indicating negative correlated responses. We observed that the coheritability depends on the selection environment. For example the tolerance coheritability at medium level of FEC, when the selection environment was at zero FEC, was different from the coheritability at zero FEC, when the selection environment was at medium level of FEC. This is because the genetic variance of body weight differs at different levels of nematode infection. In some cases, the coheritability was low. The magnitude of the coheritability also reflects the accuracy of selection because in mass selection the square root of heritability determines the accuracy. A low coheritability for tolerance, therefore, would indicate a low accuracy of selection, meaning many records of sibs or offspring are needed to obtain accurate breeding values for slope. We observed the highest absolute coheritability at high level of FEC, when the selection environment is at zero FEC. The high negative coheritability at high level of FEC when selection is performed at zero FEC, is in accordance with the high negative genetic correlation

between body weights at the two extreme levels of nematode infection. The existence of coheritabilities for slope indicates possibility of selective breeding for tolerance. The general negative signs for coheritabilities suggest that selecting animals in one environment leads to lower tolerance to another environment. These findings are in accordance with Sae-Lim *et al.* (2015) where they also found negative. Our slope coheritabilities are generally stronger than the slope coheritabilities found by Sae-Lim *et al.* (2015).

3.4.4 Genetic correlation between resistance and tolerance

We found a strong negative genetic correlation between resistance and tolerance to nematode infection in sheep. The negative genetic correlation between resistance and tolerance was observed using a bivariate model, where the slope for the reaction norm of body weight on FEC (tolerance) had a strong positive correlation (0.60±0.33) with FEC (Table 3.7) and a strong negative correlation (-0.63±0.25) with IgA (Table 3.8). Additionally, there was a negative genetic correlation (-0.76±0.32) between intercept and FEC (Table 3.7) and a positive genetic correlation (0.48 ±0.32) between intercept and IgA (Table 3.8) indicating that animals with high body weight at zero level of FEC (intercept) are genetically more resistant. The negative genetic correlation between intercept and slope, therefore, suggests that animals with higher body weights (more resistant) show more severe reduction in growth when the level of infection is high (less tolerant). In our population, there were several rams showing an upward slope indicating improvement in body weight despite the nematode infection (Fig 3.1. and Fig 3.2.). The upward slope of the rams may suggest that offspring of those rams do not invest energy on clearance of infection but invest energy in tolerating the infection. Previous studies also reported a negative genetic correlation between resistance and tolerance: Råberg *et al.* (2007) showed a negative genetic correlation between resistance and tolerance to *Plasmodium falciparum* in 5 strains of inbred mice. Fineblum and Rausher (1995) showed a negative genetic correlation between resistance and tolerance in the plant Common Morning Glory. In contrast, Hayward *et al.* (2014a) did not find a significant genetic correlation between resistance and tolerance to mixed nematode infections in feral Soay sheep using similar methodology as us. As compared to the study on Soay sheep (Hayward *et al.*, 2014a, 2014b), our study may have been statistically more powerful because our data was more homogeneous in terms of the age of the animals at the time of sampling. One explanation for the negative genetic correlation between resistance and tolerance is that resistant animals are less likely to become heavily infected

during an outbreak hence tolerance is a less valuable trait for them. On the other hand, susceptible animals are more likely to become heavily infected and therefore tolerance is a more important trait for these less resistant animals (Bishop, 2012). The negative genetic correlation between resistance and tolerance indicates that a resistant animal is not tolerant and vice versa. Therefore, selective breeding for resistance would result in reduced mean tolerance of the population, unless both resistance and tolerance are included in the selection index. In a breeding context, it means that for example when IgA is the resistance trait, selection for increased IgA by one genetic standard deviation, without considering tolerance, would result in reduced tolerance mean by 0.63 of its genetic standard deviations in the next generation.

Body weights at low level of nematode infection were shown to be genetically different traits from body weights at high level of nematode infection. One interpretation is that at low levels of infection, the immune response controls the infection and animals grow quickly. At high levels of infection, the ability of the host to repair or minimize the damages caused by infection or the immune system is the major determinant of growth. Therefore, resistance is likely to be more related to immune response to resist the infection, whereas tolerance is more related to the ability to repair or minimize the damages caused by the infection (Vale *et al.* 2014).

One advantage of sheep and their nematodes as a model system is that the mechanisms of protection and pathology are well understood (Stear *et al.*, 2003; Stear *et al.*, 2009). Protection is mediated by IgA and IgE dependent mechanisms while pathology is due to IgE mediated hypersensitivity in addition to damage caused directly by nematodes. The relative contributions of IgE and IgA may vary (Stear *et al.*, 2009). In lambs, IgA is usually the major resistance mechanism and resistance is associated with increased weight gain. In older animals or at higher intensities of infection, IgE is the dominant mechanism and IgE activity is associated with decreased weight gain. Therefore, it is possible that the genetic relationships between resistance and tolerance may differ in older animals and more research is needed to explore this possibility.

3.4.5 The effect of anthelmintic treatment on resistance and tolerance

Lambs in this study were treated with anthelmintic drugs every 4 weeks to prevent death of animals or severe growth retardation due to high nematode infection intensity. The function of anthelmintic drug was to eliminate the nematodes en masse. The effect of anthelmintic, however, was temporary and right after elimination of the nematode population a new population started

growing in the intestines of lambs. The obvious effect of the regular anthelmintic treatment is increased resistance of sheep because it does not allow the nematode burden to go to very extreme levels. The anthelmintic treatment, therefore, may change the range of nematode burden in the population as compared to the situation where animals are not treated. Consequently the genetic variance in resistance might be different in untreated populations. The change in nematode burden may change the genetic variance for body weight. The change in genetic variation for slope as a consequence of the change in pathogen burden remains unclear. On one hand slopes are estimated as linear reaction norms, which keep the slope variance constant across different levels of nematode infection. On the other hand change in the environment might influence the amount of information available to estimate the slope. We observed a small decrease in the average body weight as FEC and pepsinogen increased. The small decrease of average body weight might be due to the anthelmintic treatment, which prevented extreme increase in nematode burden. The limited increase of nematode burden in our population might have caused increased average tolerance. Nevertheless, as anthelmintic treatment is common practice in sheep husbandry our results are applicable for the sheep industry. However, care must be taken to generalize these results to wild populations where animals are not treated with anthelmintic.

3.4.6 Standard errors of variance components

The standard errors for the variance components in our study are generally large. The reason for the large standard errors is the relatively small dataset. We used only the records at five month of age and not the records at other ages to avoid the problem of heterogeneity of genetic variance across ages. To deal with the problem of heterogeneous genetic variance in repeatability models, one could include the interaction between animal identity and age (Hayward *et al.*, 2014a). Adding the random interaction of animal and age, however, would create more complexity to the analysis and gives no additional power to the analysis because more parameters need to be estimated.

In a simulation study of random regression with 100 sires, Kause (2011) showed that small family size leads to bias in estimation of variance components. As a similar scenario to our data, he showed that with the simulated slope heritability of 0.30 and 30 offspring per sire the slope variance was overestimated by 28%, the intercept variance was overestimated by 22%, and the genetic correlation between intercept and slope was overestimated by 25%. In the study of Kause (2011), the accuracy, calculated as a Pearson correlation coefficient between

true breeding values and estimated breeding values of the families, for the slope estimates was 65% and for the intercept estimates was 73%. The estimated slope variance was overestimated by more than threefold in the scenario with the same family size but heritability of 0.05. Therefore, our estimates may be biased and genetic variance in tolerance may be smaller than estimated here. Nevertheless, comparison of the random regression model with the univariate model show strong evidence for existence of genetic variation in tolerance. Despite the small size, an important strength of our data is that it contains different immunological measures on nematode infection. This allowed us to study tolerance and resistance based on different immunological aspects. Different measures for infection, i.e. FEC, IgA and pepsinogen, as well as the homogeneity in terms of the age of the animals at the time of sampling added to the statistical power of our analysis. Therefore, our study still provides useful insights in the genetic mechanisms of resistance and tolerance to nematode infections.

3.5 Conclusion

This study provides insight into the genetics of resistance and tolerance to nematode infection. This is the first study reporting the genetic correlation between resistance and tolerance to infection in pedigreed animals. Using the random regression models, we showed that different markers of infection can be used to study tolerance. This approach could easily be implemented for studying tolerance to infection in humans and wild animals by replacing the pedigree with the genomic relation matrix (Yang *et al.*, 2010). We showed that there is genetic variation among lambs in both resistance and tolerance to nematode infection indicating the possibility for selective breeding for both traits. We also showed that there is a negative genetic correlation between resistance and tolerance meaning a trade-off between these two traits. These findings indicate that breeding schemes need to include both resistance and tolerance in index selection (Hazel, 1943) to avoid inadvertently decreasing tolerance while improving resistance.

3.6 Acknowledgements

This research was a part of the NematodeSystemHealth project, financed by Marie Curie Initial Training Networks (FP7-People-2010-ITN), and co-financed by TOPIGS Research Center IPG, The Netherlands, and Dutch Ministry of Economic Affairs, Agriculture, and Innovation (Public-private partnership “Breed4Food” code

KB-12-006.03-004-ASG-LR and KB-12-006.03-005-ASG-LR). The work was also supported by the BBSRC (grant number L004070/1).

3.7 References

Akaike, H., 1973 Information theory and an extension of the maximum likelihood principle. In: B. N. Petrov and F. Csaki, editors, Proc. 2nd Int. Symp. Information Theory. Akademiai Kiado, Budapest, Hungary.**:** p. 267**–**281.

Albers, G. A. A., G. D. Gray, L. R. Piper, J. S. F. Barker, L. F. Lejambre *et al.*, 1987 The genetics of resistance and resilience to haemonchus-contortus infection in young Merino sheep. Int. J. Parasitol. 17**:** 1355-1363.

Bishop, S. C., 2012 A consideration of resistance and tolerance for ruminant nematode infections. Front. Genet. 3**:** 168.

Bishop, S. C., K. Bairden, Q. A. McKellar, M. Park and M. J. Stear, 1996 Genetic parameters for faecal egg count following mixed, natural, predominantly *Ostertagia circumcincta* infection and relationships with live weight in young lambs. Anim. Sci. 63**:** 423-428.

Bishop, S. C., and K. A. MacKenzie, 2003 Genetic management strategies for controlling infectious diseases in livestock populations. Genet. Sel. Evol. 35**:** S3-S17.

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

Bisset, S. A., C. A. Morris, D. R. Squire, S. M. Hickey and M. Wheeler, 1994 Genetics of resilience to nematode parasites in Romney sheep. New Zealand Journal of Agricultural Research 37**:** 521-534.

Borg, R. C., D. R. Notter and R. W. Kott, 2009 Phenotypic and genetic associations between lamb growth traits and adult ewe body weights in western range sheep. J. Anim. Sci. 87**:** 3506-3514.

Calus, M. P., A. F. Groen and G. de Jong, 2002 Genotype × environment interaction for protein yield in Dutch dairy cattle as quantified by different models. J. Dairy Sci. 85**:** 3115-3123.

Chang, S., Q. Xie, J. R. Dunn, C. W. Ernst, J. Song *et al.*, 2014 Host genetic resistance to Marek's disease sustains protective efficacy of herpesvirus of turkey in both experimental and commercial lines of chickens. Vaccine 32**:** 1820-1827.

Coop, R. L., A. R. Sykes and K. W. Angus, 1977 The effect of a daily intake of Ostertagia *circumcincta* larvae on body weight, food intake and concentration of serum constituents in sheep. Res. Vet. Sci. 23**:** 76-83.

Coop, R. L., A. R. Sykes and K. W. Angus, 1982 The effect of three levels of intake of *Ostertagia Circumcincta* Larvae on growth rate, food intake and body composition of growing lambs. J. Agric. Sci. 98**:** 247--255.

Davies, G., M. J. Stear and S. C. Bishop, 2005 Genetic relationships between indicator traits and nematode parasite infection levels in 6-month-old lambs. Anim. Sci. 80**:** 143-150.

Detilleux, J. C., 2009 Genetic factors affecting susceptibility to udder pathogens. Vet. Microbiol. 134**:** 157-164.

Doeschl-Wilson, A. B., B. Villanueva and I. Kyriazakis, 2012 The first step towards genetic selection for host tolerance to infectious pathogens: Obtaining the tolerance phenotype through group estimates. Front. Genet. 3.

Fineblum, W. L., and M. D. Rausher, 1995 Tradeoff between resistance and tolerance to herbivore damage in a morning glory. Nature 377: 517-520.

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. 2009. Asreml user guide release 3.0. VSN International, Hemel Hempstead, UK.

Gruner, L., J. Bouix, and J. C. Brunel. 2004. High genetic correlation between resistance to haemonchus contortus and to trichostrongylus colubriformis in inra 401 sheep. Veterinary Parasitology 119: 51-58.

Gutierrez-Gil, B., J. Perez, L. F. de la Fuente, A. Meana, M. Martinez-Valladares, F. San Primitivo, F. A. Rojo-Vazquez, and J. J. Arranz. 2010. Genetic parameters for resistance to trichostrongylid infection in dairy sheep. Animal 4: 505-512.

Hayward, A. D., R. Garnier, K. A. Watt, J. G. Pilkington, B. T. Grenfell, J. B. Matthews, J. M. Pemberton, D. H. Nussey, and A. L. Graham. 2014a. Heritable, heterogeneous, and costly resistance of sheep against nematodes and potential feedbacks to epidemiological dynamics. Am. Nat. 184: S58-S76.

Hayward, A. D., D. H. Nussey, A. J. Wilson, C. Berenos, J. G. Pilkington, K. A. Watt, J. M. Pemberton, and A. L. Graham. 2014b. Natural selection on individual variation in tolerance of gastrointestinal nematode infection. PloS Biol. 12.

Hazel, L. N. 1943. The genetic basis for constructing selection indexes. Genetics 28: 476-490.

Jackson, F., D. Bartley, Y. Bartley, and F. Kenyon. 2009. Worm control in sheep in the future. Small Rum. Res. 86: 40-45.

Kause, A. 2011. Genetic analysis of tolerance to infections using random regressions: a simulation study. Genet. Res. (Camb) 93: 291-302.

Kause, A., and J. Ødegård. 2012. The genetic analysis of tolerance to infections: a review. Front. Genet. 3: 262.

Kause, A., S. van Dalen, and H. Bovenhuis. 2012. Genetics of ascites resistance and tolerance in chicken: a random regression approach. G3 (Bethesda) 2: 527-535.

Kuukka-Anttila, H., N. Peuhkuri, I. Kolari, T. Paananen, and A. Kause. 2010. Quantitative genetic architecture of parasite-induced cataract in rainbow trout, oncorhynchus mykiss. Heredity (Edinb) 104: 20-27.

Lewis, C. R. G., T. Ait-Ali, A. Wilson, D. G. Westcott, J. P. Frossard, B. Naidu, M. A. Mellencamp, M. Torremorell, T. Drew, S. C. Bishop, and A. L. Archibald. 2010. Effects of porcine reproductive and respiratory syndrome virus

infection on the performance of pregnant gilts and growing pigs. Anim. Prod. Sci. 50: 890-896.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, Mass.

Morris, C. A., S. A. Bisset, A. Vlassoff, M. Wheeler, C. J. West, B. P. Devantier, and A. D. Mackay. 2010. Selecting for resilience in romney sheep under nematode parasite challenge, 1994–2007. New Zealand Journal of Agricultural Research 53: 245-261.

Mortimer, S. I., J. H. J. van der Werf, R. H. Jacob, D. L. Hopkins, L. Pannier, K. L. Pearce, G. E. Gardner, R. D. Warner, G. H. Geesink, J. E. H. Edwards, E. N. Ponnampalam, A. J. Ball, A. R. Gilmour, and D. W. Pethick. 2014. Genetic parameters for meat quality traits of australian lamb meat. Meat Science 96: 1016-1024.

Murphy, L., P. D. Eckersall, S. C. Bishop, J. J. Pettit, J. F. Huntley, R. Burchmore, and M. J. Stear. 2010. Genetic variation among lambs in peripheral ige activity against the larval stages of *Teladorsagia circumcincta*. Parasitol. 137: 1249-1260.

Painter, R. H. 1958. Resistance of plants to insects. Annu. Rev. Entomol. 3: 267-290.

Parker, B. J., J. R. Garcia, and N. M. Gerardo. 2014. Genetic variation in resistance and fecundity tolerance in a natural host-pathogen interaction. Evolution 68: 2421-2429.

Pollott, G. E., and J. C. Greeff. 2004. Genotype x environment interactions and genetic parameters for fecal egg count and production traits of merino sheep. J. Anim. Sci. 82: 2840-2851.

Råberg, L., A. L. Graham, and A. F. Read. 2009. Decomposing health: tolerance and resistance to parasites in animals. Proc. R. Soc. Lond. B. Biol. Sci. 364: 37-49.

Råberg, L., D. Sim, and A. F. Read. 2007. Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. Science 318: 812-814.

Rashidi, H., H. A. Mulder, P. Mathur, J. A. M. van Arendonk, and E. F. Knol. 2014. Variation among sows in response to porcine reproductive and respiratory syndrome. J. Anim. Sci. 92: 95-105.

Rausher, M. D. 2001. Co-evolution and plant resistance to natural enemies. Nature 411: 857-864.

Read, A. F., A. L. Graham, and L. Råberg. 2008. Animal defenses against infectious agents: Is damage control more important than pathogen control? PloS Biol. 6: 2638-2641.

Riggio, V., R. Finocchiaro, and S. C. Bishop. 2008. Genetic parameters for early lamb survival and growth in scottish blackface sheep. J. Anim. Sci. 86: 1758-1764.

Rose, G., A. Kause, H. A. Mulder, J. H. J. van der Werf, A. N. Thompson, M. B. Ferguson, and J. A. M. van Arendonk. 2013. Merino ewes can be bred for

body weight change to be more tolerant to uncertain feed supply. J. Anim. Sci. 91: 2555-2565.

Sae-Lim, P., H. Mulder, B. Gjerde, H. Koskinen, M. Lillehammer, and A. Kause. 2015. Genetics of growth reaction norms in farmed rainbow trout. PLoS ONE 10: e0135133.

Shaw, R. J., C. A. Morris, R. S. Green, M. Wheeler, S. A. Bisset, A. Vlassoff, and P. G. C. Douch. 1999. Genetic and phenotypic relationships among trichostrongylus colubriformis-specific immunoglobulin E, anti-trichostrongylus colubriformis antibody, immunoglobulin G(1), faecal egg count and body weight traits in grazing romney lambs. Livest. Prod. Sci. 58: 25-32.

Simms, E. L. 2000. Defining tolerance as a norm of reaction. Evol. Ecol. 14: 563-570.

Simms, E. L., and J. Triplett. 1994. Costs and benefits of plant-responses to disease resistance and tolerance. Evolution 48: 1973-1985.

Stear, M. J., K. Bairden, Q. A. McKeller, I. Scott, S. Strain, and S. C. Bishop. 1999. The relationship between the number and size of nematodes in the abomasum and the concentration of pepsinogen in ovine plasma. Res. Vet. Sci. 67: 89-92.

Stear, M. J., S. C. Bishop, M. Doligalska, J. L. Duncan, P. H. Holmes, J. Irvine, L. McCririe, Q. A. McKellar, E. Sinski, and M. Murray. 1995. Regulation of egg production, worm burden, worm length and worm fecundity by host responses in sheep infected with *Ostertagia circumcincta*. Parasite Immunol. 17: 643-652.

Stear, M. J., S. C. Bishop, N. G. Henderson, and I. Scott. 2003. A key mechanism of pathogenesis in sheep infected with the nematode *Teladorsagia circumcincta*. Anim Health Res. Rev. 4: 45-52.

Stear, M. J., S. C. Bishop, B. A. Mallard, and H. Raadsma. 2001. The sustainability, feasibility and desirability of breeding livestock for disease resistance. Res. Vet. Sci. 71: 1-7.

Stear, M. J., B. Boag, I. Cattadori, and L. Murphy. 2009. Genetic variation in resistance to mixed, predominantly *Teladorsagia circumcincta* nematode infections of sheep: From heritabilities to gene identification. Parasite Immunol. 31: 274-282.

Strain, S. A. J., S. C. Bishop, N. G. Henderson, A. Kerr, Q. A. McKellar, S. Mitchell, and M. J. Stear. 2002. The genetic control of iga activity against *Teladorsagia circumcincta* and its association with parasite resistance in naturally infected sheep. Parasitology 124: 545-552.

Stram, D. O., and J. W. Lee. 1994. Variance-components testing in the longitudinal mixed effects model. Biometrics 50: 1171-1177.

Tiffin, P., and M. D. Rausher. 1999. Genetic constraints and selection acting on tolerance to herbivory in the common morning glory ipomoea purpurea. Am. Nat. 154: 700-716.

Vale, P. F., A. Fenton, and S. P. Brown. 2014. Limiting damage during infection: Lessons from infection tolerance for novel therapeutics. PloS Biol. 12.

Woolhouse, M. E. J., J. P. Webster, E. Domingo, B. Charlesworth, and B. R. Levin. 2002. Biological and biomedical implications of the co-evolution of pathogens and their hosts. Nat. Genet. 32: 569-577.

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: 565-569.

Appendix 3.1

Table A1. Number of offspring and range of faecal egg count and pepsinogen per sire.

Sire	Number offspring	Faecal egg count				Pepsinogen			
		Min	Max	Mean	SD	Min	Max	Mean	SD
88	2	75.0	600.0	337.5	371.2	6.4	93.8	50.1	61.8
G178	2	50.0	150.0	100.0	70.7	0.0	11.4	5.7	8.1
G196	3	112.5	237.5	175.0	62.5	0.0	29.5	10.6	16.4
G197	5	87.5	350.0	167.5	112.0	7.5	48.0	26.3	16.9
G177	8	100.0	600.0	301.6	205.0	0.0	82.7	25.5	28.1
G198	9	175.0	562.5	363.9	133.7	0.0	30.5	7.5	13.1
G200	9	62.5	275.0	133.3	68.2	0.0	50.6	21.4	21.4
R94	9	0.0	325.0	111.7	93.4	0.0	33.9	9.0	10.5
62	10	0.0	2700.0	360.0	828.9	17.4	148.7	54.3	39.3
R93	10	25.0	800.0	436.3	285.9	0.0	25.7	5.3	8.8
G174	13	0.0	400.0	165.4	137.5	0.0	45.0	15.9	14.2
G179	16	0.0	737.5	303.1	243.9	0.0	52.5	13.4	16.7
R95	16	0.0	450.0	151.6	133.4	1.3	69.4	25.2	20.2
R99	16	0.0	812.5	252.3	248.8	18.2	88.9	43.5	20.9
G199	18	25.0	562.5	257.6	159.6	0.0	73.7	18.7	22.0
R97	23	0.0	2350.0	244.6	522.4	0.0	103.0	34.4	24.5

P55	24	0.0	675.0	190.6	188.6	0.0	68.0	22.9	22.4
66	26	0.0	2400.0	401.9	687.8	0.0	165.0	37.3	47.4
96	27	0.0	1875.0	364.8	417.1	0.0	56.2	27.8	17.0
65	30	0.0	900.0	228.3	232.0	0.0	140.9	41.1	39.1
77	33	0.0	1250.0	192.4	298.7	0.0	98.8	42.5	24.6
G176	36	0.0	300.0	44.4	78.4	0.0	57.2	24.2	15.3
R98	38	0.0	1200.0	382.2	327.7	0.0	250.3	45.9	46.4
P90	39	0.0	900.0	210.3	227.8	0.0	68.5	15.3	17.5
R83	40	0.0	550.0	121.9	142.6	0.0	84.9	31.4	21.7
R74	42	0.0	825.0	120.2	172.8	0.0	246.7	37.9	44.3
G175	44	0.0	887.5	196.9	219.4	0.0	75.9	18.2	21.3
R92	53	0.0	912.5	189.9	204.2	0.0	73.3	27.7	19.1
R78	72	0.0	1550.0	147.9	269.2	0.0	123.8	34.8	27.3

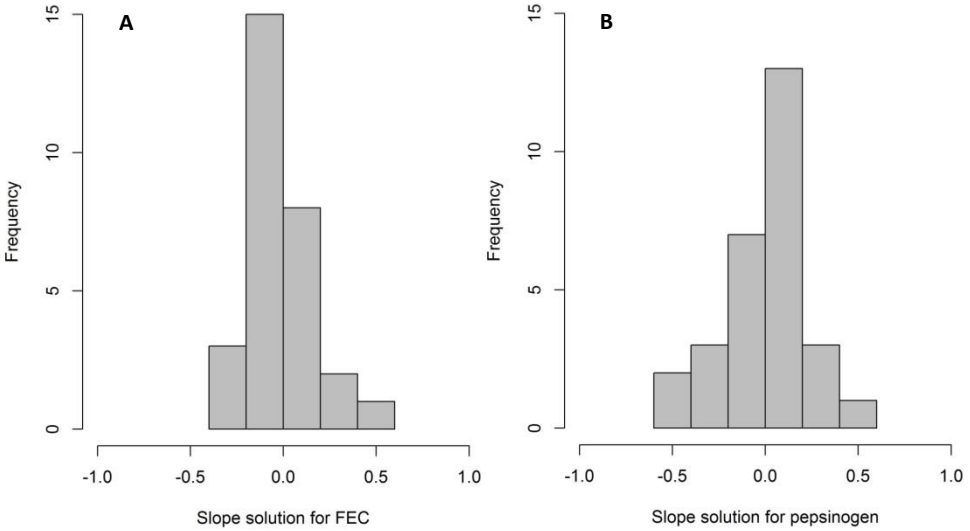

Figure A1. Distribution of sire effects on slope of body weight reaction norm on faecal egg count (FEC) (**A**) and pepsinogen (**B**)

4

Selection on Resilience Improves Disease Resistance and Tolerance to Infections

H. Rashidi[1], H. A. Mulder[1]

[1] Wageningen University, Animal Breeding and Genomics Centre, 6700 AH, Wageningen, the Netherlands

To be submitted

Abstract

Response to infection in animals has two main mechanisms: resistance (ability to control pathogen burden) and tolerance (ability to maintain performance given the pathogen burden). Selection on disease resistance and tolerance to infections seems a nice avenue to increase productivity of animals in the presence of disease infections, but it is hampered by lack of records of pathogen burden of infected animals. Selection on resilience (ability to maintain performance regardless of pathogen burden) may therefore be alternative pragmatic approach because it does not need records of pathogen burden. Therefore, the aim of this study was to assess response to selection in resistance and tolerance when selecting on resilience compared to direct selection on resistance and tolerance. Monte Carlo simulation was used combined with selection index theory to predict responses to selection. Using EBV for resilience in absence of records for pathogen burden resulted in favourable responses in resistance and tolerance to infections, with more emphasis on tolerance than on resistance. If resistance and tolerance were unfavourably correlated, lower selection responses were obtained, especially in resistance. Although using EBV for resilience resulted mostly in favourable responses in resistance and tolerance, more genetic gain could be achieved when pathogen burden was recorded.

Key words: Resilience, resistance, tolerance, pathogen burden, genetic gain

4.1 Introduction

Infectious diseases in farm animals impose costs on the farmer, reduce the welfare of animals, and create public concerns about the animal products. The conventional control strategies for disease such as biosecurity, vaccination, antibiotic treatment, and culling might not be fully successful and selective breeding might be a more promising approach. Selective breeding for improved response to infection has been a part of breeding programs for decades targeting (Albers *et al*., 1987; Bishop and Morris, 2007; Morris *et al*., 2010; Bishop, 2012). Response to infection in animals has two main mechanisms: resistance and tolerance. Resistance is the ability of animals to restrict the invading pathogen's life cycle. A resistant animal will have minimum pathogen burden during an infection period. Tolerance is the animal's ability to minimize the symptoms of infection at a given pathogen burden. A tolerant animal maintains the performance despite the pathogen burden. Tolerance, therefore, is measured as the regression of the animal performance on pathogen burden, i.e. reaction norm (Kause, 2011). If pathogen burden is known, breeding values for resistance and tolerance can be estimated. In practice, however, pathogen burden is not recorded at the individual level because it is laborious and costly. Breeders, therefore, measure resilience instead of tolerance. Resilience is the animal's ability to maintain performance during a disease outbreak and does not need the records of pathogen burden (Albers *et al*., 1987; Bisset *et al*., 1996). A resilient animal shows minimum reduction in performance during a disease outbreak. Resilience is measured as the performance variation during an outbreak irrespective of the pathogen burden (Doeschl-Wilson *et al*., 2012). A resilient animal might, therefore, be resistance, tolerant or both. It is unknown, however, to which extent resistance and tolerance can be improved in breeding programs when selecting on breeding values for resilience. The aim of this study was to assess response to selection in resistance and tolerance when using estimated breeding values for resilience when pathogen burden is not recorded compared to using estimated breeding values for resistance and tolerance when pathogen burden is recorded. We used Monte Carlo simulation and investigated the effects of genetic parameters and the proportion of animals diseased on selection responses in resistance and tolerance.

4.2 Material and methods

4.2.1 Outline of breeding scheme

A pig-breeding scheme in a dam line was considered focusing to increase dam performance traits such as litter size, piglet birth weight and survival. Due to disease outbreaks such as porcine reproductive and respiratory syndrome (PRRS), the breeding goal is extended with resistance and tolerance to infections to reduce the loss in performance due to infections. However, breeding values for resistance and tolerance can only be measured when a measure of pathogen burden is recorded. In absence of a measure of pathogen burden, a breeding value for resilience may be used. Because sow traits can only be measured in females, boars were selected based on half-sib information. Sows were selected based on own performance and half-sib information. To predict responses to selection in resistance and tolerance, we used Monte Carlo simulation to simulate true breeding values and estimate breeding values using ASReml (Gilmour *et al*., 2009). Monte Carlo simulation was used because deterministic prediction equations are not available for this complex case of resistance and tolerance to infections.

4.2.2 Monte Carlo simulation

We simulated here a half-sib family structure, although in practice small full-sib may exist as well. In the base generation, we simulated 100 sires and 10000 dams that are all unrelated. These 100 sires were mated each with 100 dams and each produced one female offspring, which resulted in 100 half-sib families. We sampled breeding values for the base generation and the generation of offspring, whereas only the phenotypes of the offspring were used in breeding value estimation. We simulated a performance trait, i.e. litter size that is affected by pathogen burden when the animal is diseased and the trait pathogen burden, which is the inverse of disease resistance, i.e. if pathogen burden is lower (higher) the animal is more (less) resistant. When animals were not infected, the performance P_{int} was a function of the breeding value when the animal is not infected (A_{int}), i.e. performance when pathogen burden is zero, and an environmental effect (E_{int}) following the classical genetic model (Falconer and Mackay, 1996):

$$P_{int} = \mu_{int} + A_{int} + E_{int} \quad (1)$$

When animals were infected, the performance of the animal does not only depend on the components in equation 1, but also on the response of the animal (P_{sl}) to

pathogen burden P_{PB}, i.e. the reaction norm on infection, which indicates the tolerance to infection (Kause, 2011):

$$P_{inf} = P_{int} + P_{sl} \times P_{PB} \quad (2)$$

In this equation the phenotypes for tolerance (P_{sl}) and pathogen burden (P_{PB}) are:

$$P_{sl} = \mu_{sl} + A_{sl} + E_{sl} \quad (3)$$

$$P_{PB} = \mu_{PB} + A_{PB} + E_{PB} \quad (4)$$

where μ_{sl} is the average decrease in performance of infected animals compared to not infected animals, μ_{PB} is the average pathogen burden, A_{sl} is the breeding value for slope, i.e. tolerance, E_{sl} is the environmental effect for slope, A_{PB} is the breeding value for pathogen burden, and E_{PB} is the environmental effect for pathogen burden. The three breeding values were assumed to follow a multivariate normal distribution $\begin{bmatrix} \mathbf{a_{int}} \\ \mathbf{a_{sl}} \\ \mathbf{a_{PB}} \end{bmatrix} \sim \text{MVN}\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \mathbf{A} \otimes \begin{bmatrix} \sigma^2_{A_{int}} & \sigma_{A_{int},A_{sl}} & \sigma_{A_{int},A_{PB}} \\ & \sigma^2_{A_{sl}} & \sigma_{A_{sl},A_{PB}} \\ symmetric & & \sigma^2_{A_{PB}} \end{bmatrix}\right)$, where $\mathbf{A}$ is the numerator relationship matrix, $\sigma^2_{A_{int}}$, $\sigma^2_{A_{sl}}$, $\sigma^2_{A_{PB}}$, $\sigma_{A_{int},A_{sl}}$, $\sigma_{A_{int},A_{PB}}$ and $\sigma_{A_{sl},A_{PB}}$ are the additive genetic variances and covariances among the three breeding values. The three environmental effects were assumed to follow a multivariate normal distribution $\begin{bmatrix} \mathbf{e_{int}} \\ \mathbf{e_{sl}} \\ \mathbf{e_{PB}} \end{bmatrix} \sim \text{MVN}\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \mathbf{I} \otimes \begin{bmatrix} \sigma^2_{E_{int}} & \sigma_{E_{int},E_{sl}} & \sigma_{E_{int},E_{PB}} \\ & \sigma^2_{E_{sl}} & \sigma_{E_{sl},E_{PB}} \\ symmetric & & \sigma^2_{E_{PB}} \end{bmatrix}\right)$, where $\mathbf{I}$ is the identity matrix, $\sigma^2_{E_{int}}$, $\sigma^2_{E_{sl}}$, $\sigma^2_{E_{PB}}$, $\sigma_{E_{int},E_{sl}}$, $\sigma_{E_{int},E_{PB}}$ and $\sigma_{E_{sl},E_{PB}}$ are the environmental variances and covariances among the three environmental effects. The phenotypic variances for P_{int}, P_{sl} and P_{PB} were always one. The offspring were randomly allocated to 100 contemporary groups of equal size. In the default situation, half of the contemporary groups were not infected and their phenotype for performance was simulated according to equation 1, whereas for infected animals the phenotype for performance was simulated according to equation 2. We simulated 100 replicates. Table 4.1 shows all the parameters used with their default and varied values.

Table 4.1. Parameter values used in the basic situation and alternative situations

	Parameter values	
Parameter	Basic	Alternative
$\sigma^2_{A_{int}}$	0.3	-
$\sigma^2_{A_{slope}}$	0.05	0, 0.1, 0.2, 0.3, 0.4, 0.5
$\sigma^2_{A_{PB}}$	0.3	0, 0.1, 0.2, 0.3, 0.4, 0.5
$\sigma^2_{p_{int}}$, $\sigma^2_{A_{slope}}$, $\sigma^2_{A_{PB}}$	1	-
$r_g(int, slope)$	0	-0.75, -0.50, -0.25, 0.25, 0.50, 0.75
$v_{A_{int}}$	1	0
$v_{A_{slope}}$	1	0
$v_{A_{PB}}$	-1	0
Number of sires	100	-
Number of dams	10000	-
Number of half-sib progeny	100	-
Selected proportions sires	0.05	-
Selected proportions dams	0.2	-
Number of contemporary groups	100	
Number of infected contemporary groups	50	40, 30, 20, 10

4.2.3 Breeding value estimation scenarios

We considered two scenarios: (1) both performance and pathogen burden were recorded on each animal and (2) only performance was recorded. In the first scenario, pathogen burden was known for each diseased animal. Therefore, we used a bivariate model for performance and pathogen burden. The model for performance (y_{Perf}) was an animal random regression model in ASReml (Gilmour *et al.*, 2009) to estimate breeding values for A_{int}, and A_{sl}; the model for pathogen burden (y_{PB}) was a simple animal model:

$$y_{Perf} = \mu + PB + A_{int} + A_{sl} \times P_{PB} + e_{Perf} \quad (5)$$

$$y_{PB} = \mu_{PB} + A_{PB} + e_{PB} \quad (6)$$

where μ is the overall mean, PB is a fixed covariate of pathogen burden for the average slope, and e_{Perf} is the residual. Breeding values were assumed trivariate

normally distributed as shown before. The residual variance was assumed to be heterogeneous for records belonging to infected contemporary groups and not infected contemporary groups. The residuals e_{Perf} and e_{PB} for infected animals were assumed bivariate normally distributed $\begin{bmatrix}\mathbf{e_{Perf}}\\ \mathbf{e_{PB}}\end{bmatrix} \sim \text{MVN}\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}, \mathbf{I}\otimes\begin{bmatrix}\sigma^2_{e_{Perf}} & \sigma_{e_{Perf},e_{PB}}\\ symmetric & \sigma^2_{e_{PB}}\end{bmatrix}\right)$. In the second scenario, pathogen burden was considered not recorded and therefore in equation 5 replaced by the average performance of the contemporary groups ($\overline{CG}$) as an indirect measure of infection:

$$y_{Perf} = \mu + \overline{CG} + A_{int,2} + A_{res} \times \overline{CG} + e_{Perf} \quad (7)$$

where $\overline{CG}$ is a fixed covariate of contemporary groups average for the average slope, $A_{int,2}$ is the breeding value for intercept, which is different than in equation 5, A_{res} is the breeding value for resilience to infection that is the slope of the reaction norm on the average of the contemporary group. The breeding values $A_{int,2}$ and A_{res} were assumed bivariate normally distributed $\begin{bmatrix}\mathbf{a_{int}}\\ \mathbf{a_{res}}\end{bmatrix} \sim \text{MVN}\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}, \mathbf{A}\otimes\begin{bmatrix}\sigma^2_{A_{int,2}} & \sigma_{A_{int},A_{res}}\\ symmetric & \sigma^2_{A_{res}}\end{bmatrix}\right)$.

4.2.4 Evaluation of scenarios

Per replicate, we estimated the average correlations between estimated and true breeding values as well as the average correlations among estimated breeding values for boars based on its offspring and for sows based on own performance and her half-sibs per replicate. The average correlation and its standard deviation were calculated across the 100 replicates. These average correlations were the inputs for response to selection calculations.

4.2.5 Response to selection

The breeding goal was to increase performance in contemporary groups with and without infection and therefore the aim was to increase performance in absence of disease outbreak (A_{int}), to increase tolerance to infections (A_{sl}) and to decrease pathogen burden (A_{PB}), i.e. to increase resistance:

$$H = v_{A_{int}}A_{int} + v_{A_{sl}}A_{sl} + v_{A_{PB}}A_{PB} = \mathbf{v}'\mathbf{a} \quad (8)$$

where $v_{A_{int}}$ is the economic value to increase performance in the absence of disease outbreak, $v_{A_{sl}}$ is the economic value for tolerance to infections and $v_{A_{PB}}$ is the economic value for pathogen burden or resistance. We used three breeding goals: (1) only increase tolerance, (2) only increase resistance, i.e. reduce pathogen burden and (3) to increase performance in absence of disease outbreak, increase tolerance and increase resistance. In the last breeding goal, we assumed for simplicity that all traits had an equal absolute economic value (1.0, 1.0 and -1.0 for A_{int}, A_{sl} and A_{PB}), because a formal economic analysis was beyond the scope of this study.

Selection was based on an index using the estimated breeding values (EBV). In scenario 1, we used the EBV $\widehat{A_{int}}$, $\widehat{A_{sl}}$ and $\widehat{A_{PB}}$ in index I_1. In scenario 2, we used the EBV $\widehat{A_{int,2}}$ and $\widehat{A_{res}}$ in index I_2:

$$I_1 = b_{11}\widehat{A_{int}} + b_{12}\widehat{A_{sl}} + b_{13}\widehat{A_{PB}} = \mathbf{b_1}'\mathbf{a_1} \quad (9)$$

$$I_2 = b_{21}\widehat{A_{int,2}} + b_{22}\widehat{A_{res}} = \mathbf{b_2}'\mathbf{a_2} \quad (10)$$

The optimal selection index weights $\mathbf{b_1}$ and $\mathbf{b_2}$ were calculated using selection index theory (Hazel, 1943):

$$\mathbf{b} = \mathbf{P^{-1}Gv} \quad (11)$$

The **P**-matrices $\mathbf{P_1}$ and $\mathbf{P_2}$ contain the variances and covariances between EBV in the selection indices I_1 and I_2. In this case, we assumed that these estimated breeding values were scaled towards a variance of 1. The covariances are then equal to the correlations between EBV:

$$\mathbf{P_1} = \begin{bmatrix} 1 & r_{\widehat{A_{int}},\widehat{A_{sl}}} & r_{\widehat{A_{int}},\widehat{A_{PB}}} \\ & 1 & r_{\widehat{A_{sl}},\widehat{A_{PB}}} \\ symmetric & & 1 \end{bmatrix} \quad (12)$$

$$\mathbf{P_2} = \begin{bmatrix} 1 & r_{\widehat{A_{int,2}},\widehat{A_{res}}} \\ symmetric & 1 \end{bmatrix} \quad (13)$$

where $r_{\widehat{A_{int}},\widehat{A_{sl}}}$, $r_{\widehat{A_{int}},\widehat{A_{PB}}}$ and $r_{\widehat{A_{sl}},\widehat{A_{PB}}}$ are the correlations between $\widehat{A_{int}}$, $\widehat{A_{sl}}$ and $\widehat{A_{PB}}$ and $r_{\widehat{A_{int,2}},\widehat{A_{res}}}$ is the correlation between $\widehat{A_{int,2}}$ and $\widehat{A_{res}}$ (resilience). The G-matrices $\mathbf{G_1}$ and $\mathbf{G_2}$ contain the covariances between the EBV $\widehat{A_{int}}$, $\widehat{A_{sl}}$ and $\widehat{A_{PB}}$ or $\widehat{A_{int,2}}$ and $\widehat{A_{res}}$ with the true breeding values (TBV) A_{int}, A_{sl} and A_{PB} in the

breeding goal. Because the EBV were standardized with variance 1, matrices $\mathbf{G_1}$ and $\mathbf{G_2}$ were calculated as:

$$\mathbf{G_1} = \begin{bmatrix} r_{\widehat{A_{int}}A_{int}}\sigma_{A_{int}} & r_{\widehat{A_{int}}A_{sl}}\sigma_{A_{sl}} & r_{\widehat{A_{int}}A_{PB}}\sigma_{A_{PB}} \\ r_{\widehat{A_{sl}}A_{int}}\sigma_{A_{int}} & r_{\widehat{A_{sl}}A_{sl}}\sigma_{A_{sl}} & r_{\widehat{A_{sl}}A_{PB}}\sigma_{A_{PB}} \\ r_{\widehat{A_{PB}}A_{int}}\sigma_{A_{int}} & r_{\widehat{A_{PB}}A_{sl}}\sigma_{A_{sl}} & r_{\widehat{A_{PB}}A_{PB}}\sigma_{A_{PB}} \end{bmatrix} \quad (14)$$

$$\mathbf{G_2} = \begin{bmatrix} r_{\widehat{A_{int,2}}A_{int}}\sigma_{A_{int}} & r_{\widehat{A_{int,2}}A_{sl}}\sigma_{A_{sl}} & r_{\widehat{A_{int,2}}A_{PB}}\sigma_{A_{PB}} \\ r_{\widehat{A_{res}}A_{int}}\sigma_{A_{int}} & r_{\widehat{A_{res}}A_{sl}}\sigma_{A_{sl}} & r_{\widehat{A_{res}}A_{PB}}\sigma_{A_{PB}} \end{bmatrix} \quad (15)$$

We calculated the selection responses for trait j, i.e. $A_{int,}$, A_{sl} and A_{sl} as:

$$\Delta G_j = \frac{R_{s,j}+R_{d,j}}{L_s+L_d} \quad (16)$$

Where $R_{s,j}$ and $R_{d,j}$ are the genetic selection differentials for sires and dams and L_s and L_d are the relative generation intervals of sires and dams. We aimed to simulate a pig breeding program for a dam line based on sib testing scheme, although for boars the EBV in the Monte Carlo simulation were based on offspring, i.e. a progeny testing scheme. However, in this simplified case in absence of Bulmer effect (Bulmer, 1976) a sib testing scheme and a progeny scheme would yield equal selection responses when L_s is set to 2 and L_d is set to 1, because the accuracy based on half-sibs is exactly half of the accuracy with half-sib offspring. The genetic selection differentials were calculated as:

$$R_j = \frac{i\mathbf{b'g_j}}{\sigma_I} \quad (17)$$

where i is the selection intensity and $\sigma_I = \sqrt{\mathbf{b'Pb}}$ is the standard deviation of the index. Selection intensities were calculated assuming an infinite population of selection candidates without correction for correlated index values among relatives (Hill, 1976; Meuwissen, 1991). The selected proportions in boars and sows were assumed 5% and 20%, respectively. Selection responses are presented in genetic standard deviations to facilitate comparison across traits.

4.3 Results

4.3.1 Correlations between estimated and true breeding values

Correlations between estimated (EBV) and true breeding values (TBV) of boars are shown in Table 4.2. Correlation between EBV and TBV for pathogen

burden was zero when heritability of pathogen burden was zero and increased by increasing the heritability for pathogen burden. When pathogen burden was recorded, correlations between EBV and TBV for the intercept ($\widehat{A_{int}}$ and A_{int}) and for slope ($\widehat{A_{sl}}$ and A_{sl}) were about 0.9 and not changing when the heritability of pathogen burden was increased. However, when pathogen burden was not known and EBV for resilience were estimated, correlations between EBV and TBV for intercept ($\widehat{A_{int,2}}$ and A_{int}) and for slope and resilience ($\widehat{A_{res}}$ and A_{sl}) were decreasing with increasing heritability of pathogen burden. In addition, the correlation between TBV for pathogen burden and EBV for resilience ($\widehat{A_{res}}$ and A_{PB}) were increasing with increasing heritability of pathogen burden. The key message is that the EBV for resilience is correlated both with the TBV for pathogen burden (i.e. resistance) and the TBV for slope (i.e. tolerance).

4.3.2 The effect of genetic variance in resistance and tolerance on selection responses

The Figures 4.1, 4.2 and 4.3 show selection responses in slope (tolerance) and in pathogen burden (resistance) as a function of the heritability of pathogen burden and slope for different breeding goals. If the breeding goal was to select only on tolerance (Figure 4.1) or only on pathogen burden (Figure 4.2), there was no response in pathogen burden or slope if pathogen burden was recorded and used in breeding value estimation. This is expected, because the genetic correlation between pathogen burden and slope was zero. However, when selecting on the EBV for resilience because pathogen burden was not recorded, both slope and pathogen burden responded in the favourable directions (Figure 4.1 and 4.2), because in both cases the same animals were selected based on the EBV for resilience. When selecting on resilience and the heritability of pathogen burden was increasing, the response in pathogen burden was increasing (in absolute terms) at the cost of a lower response in slope. When selecting on resilience and the heritability of slope was increasing, the response in slope was increasing at the cost of a lower response in pathogen burden. In general when selecting on resilience, the response in slope was higher, i.e. between 0.7 and 1.0 genetic standard deviation, than for pathogen burden, i.e. between 0.2 and 0.7 genetic standard deviation. In other words, selection on resilience places a greater selection pressure on slope than on pathogen burden. If the breeding goal contained all three traits (Figure 4.3), the selection response in tolerance (slope) was higher and in resistance (pathogen burden) was lower when using the EBV for resilience than when EBV for slope and pathogen burden were estimated. The response in slope increased not only when the heritability for slope was higher, but

also when the heritability of pathogen burden increased. In conclusion, when the EBV for resilience is used in index selection when pathogen burden is not recorded, high selection response in tolerance and moderate selection responses in resistance can be achieved.

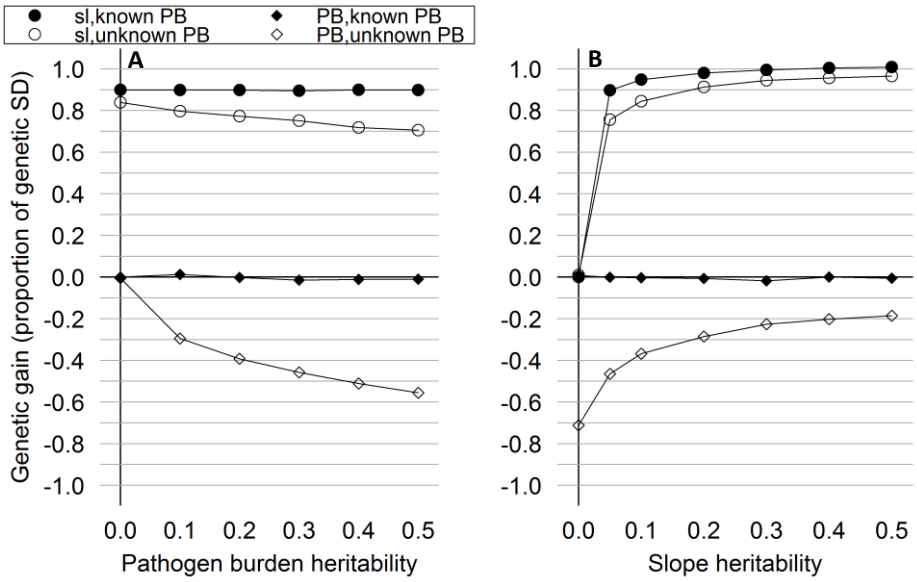

Figure 4.1. Genetic gain in slope (sl) and pathogen burden (PB) shown as the proportion of the genetic standard deviation of the traits after one generation of index selection in sib testing schemes when PB phenotype is either known or unknown. Economic values for intercept (int), slope, and pathogen burden were: $v_{A_{int}} = 0,\ v_{A_{sl}} = 1, v_{A_{PB}} = 0$. Different values of $\sigma^2_{A_{PB}}$ (0, 0.1, 0.2, 0.3, 0.4, 0.5) (Panel **A**) and $\sigma^2_{A_{sl}}$ (0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5) (Panel **B**) were simulated. Parameters value: $\sigma^2_{P_{PB}} = 1$, $\sigma^2_{P_{sl}} = 1$, $r_A = 0$. Number of progeny per sire = 100. Selected proportion sires = 0.05, selected proportion dams = 0.20.

4.3.3 The effect of genetic correlations on selection responses

Figure 4.4 shows the effect of the genetic correlation between slope and pathogen burden. A negative genetic correlation is favourable and positive correlation is unfavourable. When the genetic correlation was negative, i.e. -0.75, selection responses were very similar when pathogen burden was known or unknown. When the genetic correlation increased, selection responses decreased or became unfavourable, especially in pathogen burden when selecting on resilience. When using the EBV for resilience, selection responses in slope were higher than in pathogen burden, as seen before. With increasing the genetic correlation, the difference in selection responses with or without pathogen burden increased for all breeding goals. A peculiar result occurred when the breeding goal

was solely to decrease pathogen burden (Figure 4.4C) and selecting on the EBV for resilience, the selection response in slope became suddenly negative when the genetic correlation was 0.75. The direction of selection on the EBV for resilience suddenly changed from selecting the animals with the highest EBV to animals with the lowest EBV. In conclusion, the genetic correlation between slope and pathogen burden has a high impact on the selection responses in resistance and tolerance and selection on resilience may lead to an unfavourable response in resistance.

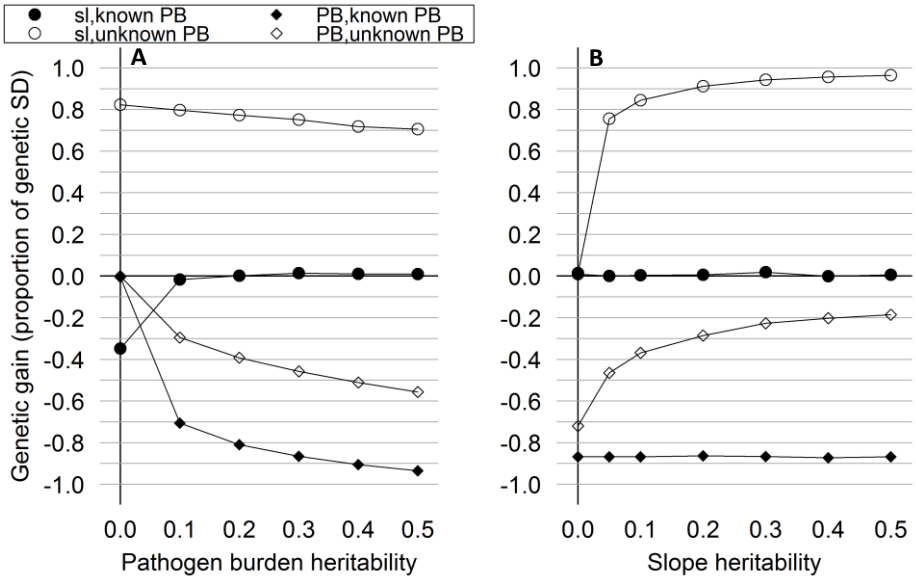

Figure 4.2. Genetic gain in slope (sl) and pathogen burden (PB) shown as the proportion of the genetic standard deviation of the traits after one generation of index selection in sib testing schemes when PB phenotype is either known or unknown. Economic values for intercept (int), slope, and pathogen burden were: $v_{A_{int}} = 0,\ v_{A_{sl}} = 0, v_{A_{PB}} = -1$. Different values of $\sigma^2_{A_{PB}}$ (0, 0.1, 0.2, 0.3, 0.4, 0.5) (Panel **A**) and $\sigma^2_{A_{sl}}$ (0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5) (Panel **B**) were simulated. Parameters value: $\sigma^2_{P_{PB}} = 1$, $\sigma^2_{P_{sl}} = 1$, $r_A = 0$. Number of progeny per sire = 100. Selected proportion sires = 0.05, selected proportion dams = 0.20.

4.3.4 The effect of the proportion of animals infected on selection responses

In the previous part, it was assumed that half of the contemporary groups were infected and the other half were not, but fortunately disease outbreaks are less frequent. Therefore, we investigated the effect of the percentage of contemporary groups infected and varied from 10% to 50% (Figure 4.5). As expected, selection responses in slope and pathogen burden were higher when the proportion of infected contemporary groups was higher. If 10% of the contemporary groups was infected the responses were about 57-73% of the

responses when 50% of the contemporary groups were infected. The figures show that the proportion of infected contemporary groups did not affect the efficiency of selection on EBV for resilience compared to selection on EBV for pathogen burden and slope.

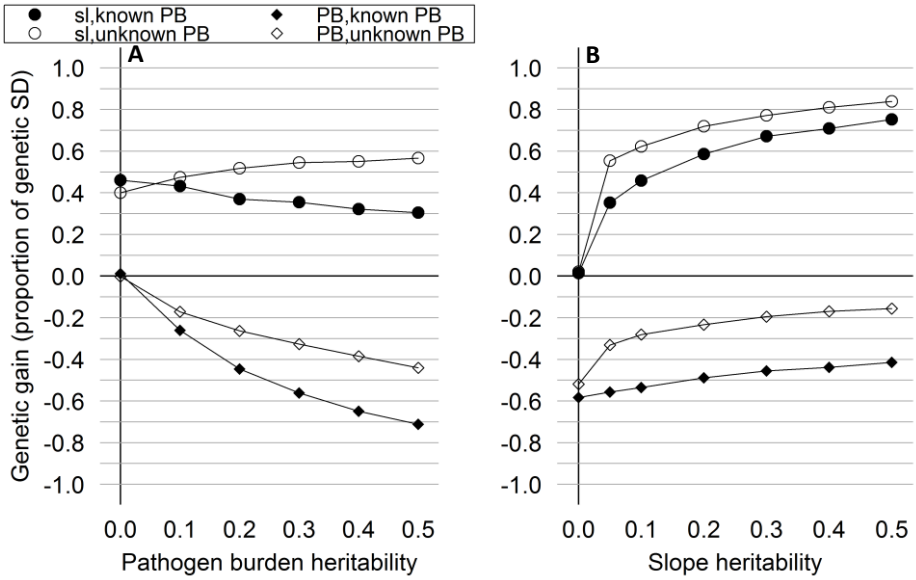

Figure 4.3. Genetic gain in slope (sl) and pathogen burden (PB) shown as the proportion of the genetic standard deviation of the traits after one generation of index selection in sib testing schemes when PB phenotype is either known or unknown. Economic values for intercept (int), slope, and pathogen burden were: $v_{A_{int}} = 1,\ v_{A_{sl}} = 1, v_{A_{PB}} = -1$. Different values of $\sigma^2_{A_{PB}}$ (0, 0.1, 0.2, 0.3, 0.4, 0.5) (Panel **A**) and $\sigma^2_{A_{sl}}$ (0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5) (Panel **B**) were simulated. Parameters value: $\sigma^2_{P_{PB}} = 1$, $\sigma^2_{P_{sl}} = 1$, $r_A = 0$. Number of progeny per sire = 100. Selected proportion sires = 0.05, selected proportion dams = 0.20.

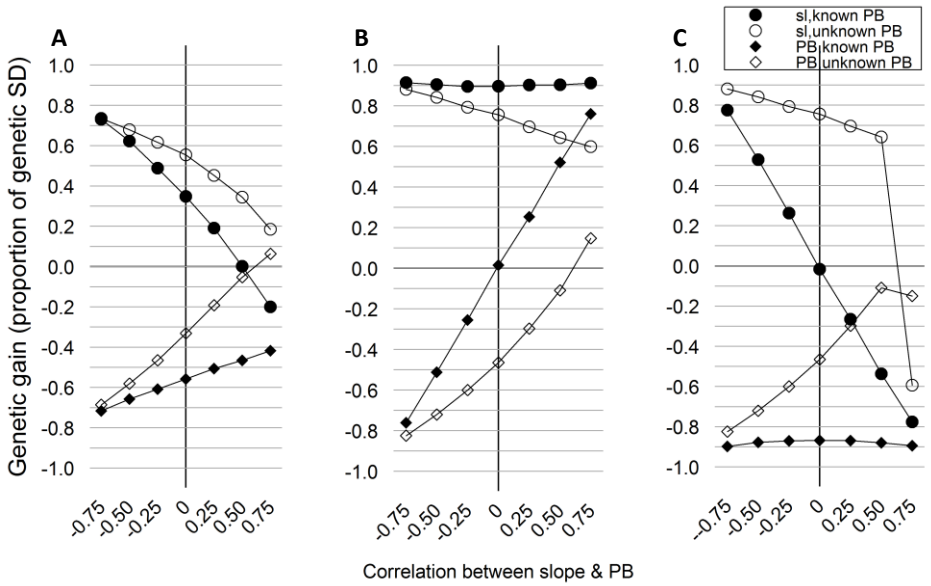

Figure 4.4. Genetic gain in slope (sl) and pathogen burden (PB) shown as the proportion of the genetic standard deviation of the traits after one generation of index selection in sib testing schemes when PB phenotype is either known or unknown as a function of the genetic correlation between slope and pathogen burden. Economic values for intercept (int), slope (sl), and pathogen burden (PB): $v_{A_{int}} = 1,\ v_{A_{sl}} = 1, v_{A_{PB}} = -1$ (Panel **A**), $v_{A_{int}} = 0,$ $v_{A_{sl}} = 1, v_{A_{PB}} = 0$ (Panel **B**), and $v_{A_{int}} = 0,\ v_{A_{sl}} = 0, v_{A_{PB}} = -1$ (Panel **C**). Parameters value: $\sigma^2_{A_{PB}} = 0.3,\ \sigma^2_{A_{sl}} = 0.05,\ \sigma^2_{P_{PB}} = 1,\ \sigma^2_{P_{sl}} = 1,\ r_A = 0$. Number of progeny per sire = 100. Selected proportion sires = 0.05, selected proportion dams = 0.20.

4.4 Discussion

4.4.1 Model and results

The aim of the current study was to assess response to selection in resistance and tolerance when selecting on resilience (unknown pathogen burden) compared to when selecting on resistance and tolerance (known pathogen burden). We found that using the EBV for resilience resulted in favourable selection responses in resistance and tolerance, but it was not as effective as selection on EBV for resistance and tolerance. Selection responses in resistance were smaller than in tolerance. The selection responses in resistance and tolerance depended on the genetic variances in these traits as well as the genetic correlation between the two. However, the comparison of index selection using the EBV for resilience compared to index selection on EBV for resistance and tolerance was not much affected by the genetic variances in pathogen burden or slope. This indicates that

selection on the EBV for resilience was quite robust and yielded favourable responses at least in tolerance and mostly also in resistance.

To our knowledge, this is the first study investigating the efficiency of selection on resilience on the underlying genetics of resistance and tolerance. The results indicate that selection on an EBV for resilience can be considered being index selection for resistance and tolerance. In this study, we used a combination of Monte Carlo simulation and selection index theory to predict responses to selection. This was a fast and accurate way of predicting selection responses, because deterministic predictions for elements in the P and G matrix were not needed, which are expected to be complex because of involvement of a product of tolerance and pathogen burden. We simulated one generation of selection and did not account for gametic phase disequilibrium (Bulmer, 1976). Ignoring gametic phase disequilibrium or the so-called Bulmer-effect, however, is expected to do not affect the comparison of using EBV for resilience compared to EBV for tolerance and resistance, because the breeding scheme was considered constant in this study. Accounting for the Bulmer-effect is especially important when comparing different breeding schemes such as sib testing, progeny testing or genomic selection (Mulder and Bijma, 2005; Van Grevenhof *et al.*, 2012).

Table 4.2. Correlation between estimated and true breeding values when pathogen burden (PB) is known or unknown. When PB was known, correlation between estimated (EBV) and true breeding values (TBV) for PB ($r_{\widehat{A_{PB}}A_{PB}}$), intercept ($r_{\widehat{A_{int}}A_{int}}$), slope ($r_{\widehat{A_{sl}}A_{sl}}$), TBV for PB and EBV for intercept ($r_{\widehat{A_{int}}A_{PB}}$), and TBV for PB and EBV for slope ($r_{\widehat{A_{sl}}A_{PB}}$) was calculated. When PB was unknown correlation between EBV and TBV for intercept ($r_{\widehat{A_{int2}}A_{int}}$), TBV for slope and EBV for resilience ($r_{\widehat{A_{res}}A_{sl}}$), TBV for PB and EBV for intercept ($r_{\widehat{A_{int2}}A_{PB}}$), and TBV for PB and EBV for resilience ($r_{\widehat{A_{res}}A_{PB}}$) was calculated. Correlations were calculated for sires based on 100 half-sib offspring when $\sigma^2_{A_{PB}}$ varies. Parameters values: $\sigma^2_{A_{int}} = 0.3, \sigma^2_{A_{sl}} = 0.05, \sigma^2_{P_{int}} = 1, \sigma^2_{P_{sl}} = 1, \sigma^2_{P_{PB}} = 1.$

$\sigma^2_{A_{PB}}$	$r_{\widehat{A_{PB}}A_{PB}}$	$r_{\widehat{A_{int}}A_{int}}$	$r_{\widehat{A_{int2}}A_{int}}$	$r_{\widehat{A_{sl}}A_{sl}}$	$r_{\widehat{A_{res}}A_{sl}}$	$r_{\widehat{A_{int}}A_{PB}}$	$r_{\widehat{A_{int2}}A_{PB}}$	$r_{\widehat{A_{sl}}A_{PB}}$	$r_{\widehat{A_{res}}A_{PB}}$
0.0	-0.04 $_{0.12}$	0.90 $_{0.02}$	0.73 $_{0.05}$	0.90 $_{0.02}$	0.86 $_{0.03}$	-0.02 $_{0.10}$	0.01 $_{0.10}$	0.00 $_{0.10}$	0.00 $_{0.10}$
0.1	0.75 $_{0.04}$	0.90 $_{0.02}$	0.72 $_{0.05}$	0.90 $_{0.02}$	0.82 $_{0.03}$	-0.02 $_{0.09}$	-0.19 $_{0.10}$	-0.01 $_{0.10}$	0.28 $_{0.09}$
0.2	0.84 $_{0.03}$	0.90 $_{0.02}$	0.70 $_{0.05}$	0.90 $_{0.02}$	0.79 $_{0.04}$	0.02 $_{0.09}$	-0.28 $_{0.09}$	0.00 $_{0.10}$	0.39 $_{0.10}$
0.3	0.89 $_{0.02}$	0.90 $_{0.02}$	0.69 $_{0.06}$	0.90 $_{0.02}$	0.76 $_{0.04}$	0.00 $_{0.10}$	-0.34 $_{0.08}$	0.01 $_{0.10}$	0.47 $_{0.07}$
0.4	0.92 $_{0.02}$	0.90 $_{0.02}$	0.67 $_{0.06}$	0.90 $_{0.02}$	0.73 $_{0.05}$	0.00 $_{0.10}$	-0.38 $_{0.08}$	-0.01 $_{0.10}$	0.51 $_{0.08}$
0.5	0.94 $_{0.01}$	0.90 $_{0.02}$	0.64 $_{0.06}$	0.90 $_{0.02}$	0.71 $_{0.05}$	0.00 $_{0.10}$	-0.42 $_{0.09}$	0.00 $_{0.10}$	0.56 $_{0.08}$

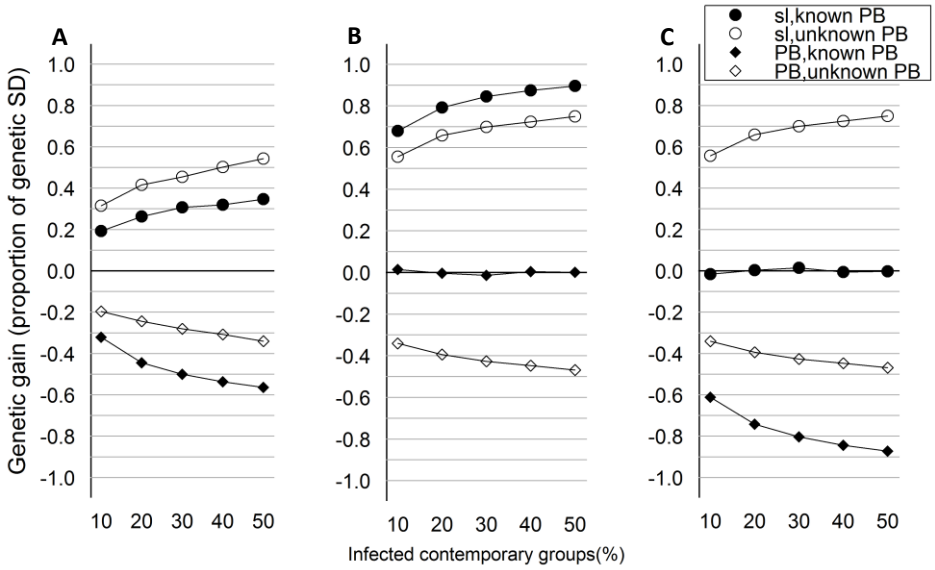

Figure 4.5. Genetic gain in slope (sl) and pathogen burden (PB) shown as the proportion of the genetic standard deviation of the traits after one generation of index selection in sib testing schemes when PB phenotype is either known or unknown as a function of proportion of infected contemporary groups. Economic values for intercept (int), slope (sl), and pathogen burden (PB): $v_{A_{int}} = 1,\ v_{A_{sl}} = 1, v_{A_{PB}} = -1$ (Panel **A**), $v_{A_{int}} = 0,\ v_{A_{sl}} = 1, v_{A_{PB}} = 0$ (Panel **B**), and $v_{A_{int}} = 0,\ v_{A_{sl}} = 0, v_{A_{PB}} = -1$ (Panel **C**). Parameters value: $\sigma^2_{A_{PB}} = 0.3$, $\sigma^2_{A_{sl}} = 0.05$, $\sigma^2_{P_{PB}} = 1$, $\sigma^2_{P_{sl}} = 1$, $r_A = 0$. Number of progeny per sire = 100. Selected proportion sires = 0.05, selected proportion dams = 0.20.

4.4.2 The value of recording pathogen burden

Until recently, genetics of tolerance to infections has not attracted much attention in animal breeding, whereas for a long time breeding for resistance has been on the research agenda in animal breeding for several years (Doeschl-Wilson *et al.*, 2012). The main hurdle with breeding for tolerance is that we need a measure of pathogen burden. Our study shows that recording of pathogen burden would yield a 22% increase in selection response in tolerance when the breeding goal contains only tolerance compared to using the EBV for resilience. Furthermore, for resistance it would yield an 87% increase in selection response when the breeding goal contains only resistance compared to using the EBV for resilience. When the breeding goal is to increase performance in periods without infection, tolerance and resistance, recording pathogen burden and using it in breeding value estimation would increase the genetic gain in the breeding goal by 8% compared to when pathogen burden is not recorded. These results show clearly

that measuring pathogen burden can increase genetic gain in breeding programs compared to selection on EBV for resilience.

Although measuring pathogen burden has clear advantages for genetic improvement, it is often difficult to obtain measures of pathogen burden in commercial animals, because animals get infected at different time points, registration of diseases is limited and measuring pathogen burden is costly and laborious. In some cases, it is feasible to obtain some indication for pathogen burden such as faecal egg count for nematode infections in sheep (Albers *et al*., 1987; Stear *et al*., 1995; Bishop *et al*., 1996) or somatic cell count in milk as an indication for the severity of mastitis infection (Detilleux *et al*., 2012). Even though pathogen burden may be difficult to obtain, recording which animals are infected and which are not infected would be already of great value, especially for endemic diseases, such as mastitis. Generally, using field data leads to underestimation of the incidence of diseases due to imperfect sensitivity and incomplete data recording (Bishop and Woolliams, 2010).

With challenge experiments, for instance the PRRS host consortium trials at Kansas State University (Lunney *et al*., 2011; Rowland *et al*., 2012), it is feasible to obtain measures of viremia at different points after infection (Boddicker *et al*., 2012; Islam *et al*., 2013). In other words, these measures of viremia can be used as pathogen burden to estimate genetic variation in tolerance. Even though many data have been recorded on these infected pigs, it has proven be difficult to find genetic variation in tolerance (Lough *et al*. manuscript in preparation). One important issue is unbiased estimation of the intercept of the reaction norm when animals are not infected (Kause, 2011; Doeschl-Wilson *et al*., 2012). Therefore, the experiment should ideally contain partly relatives that are infected and another part that are not infected. Although challenge experiments are very useful for research on genetics of disease resistance and tolerance to infections, the value for commercial breeding programs may be limited because the challenge environment may still be very different from the commercial environments.

4.4.3 Selection on resilience to infections

Although selection for increased tolerance seems to be still challenging because of lack of pathogen burden, our study shows that using EBV for resilience is an effective way to increase tolerance by selection and at the same time also improving resistance. This is in contrast to Albers *et al*. (1987), who concluded that the heritability of resilience is too small to obtain direct selection responses to mass selection. Indeed mass selection will yield small selection responses in resilience (Kolmodin and Bijma, 2004; Sae-Lim *et al*., 2015), but using information

of sibs that are infected and not infected can greatly increase the selection responses in resilience, as observed in this study.

An important drawback of using the EBV for resilience is that resilience is a 'black-box': the emphasis on tolerance and resistance depends on the parameters. Furthermore, obtaining correct measures of the contemporary group mean may be statistically challenging and may lead to biased estimates of the genetic variance in resilience, especially to disentangle genetic trend from the contemporary group means (Knap and Su, 2008). If contemporary groups are large, which is generally the case in pig breeding, bias is expected to be small or absent. Mixed model estimates of contemporary group means could be used (Rashidi *et al*., 2014; Silva *et al*., 2014). In previous studies, we showed that such contemporary group means could be used well to detect disease outbreaks (Mathur *et al*., 2014; Rashidi *et al*., 2014). Another drawback is that the EBV for resilience will mainly pick up resistance and tolerance to epidemic diseases. For endemic diseases, the approach is less useful because there are continuously animals infected and therefore the contemporary group mean is not a good indicator for presence of infections. In those cases, presence or absence of infection at animal level could be used as a covariate in the random regression model.

An important advantage of using the EBV for resilience is that it is aiming to select for general resilience (Guy *et al*., 2012). Multiple diseases may decrease the contemporary group means. Therefore, selection on the EBV for resilience will target general disease tolerance and resistance rather than specific disease resistance or tolerance. In addition to diseases, there may be other environmental factors that decrease performance, such as heat stress (Bloemhof *et al*., 2008) or seasonality (Sevillano *et al*.). It is likely that general mechanisms related to dealing with stress situations are involved. In laboratory species, heat-shock proteins are found to be controlling effects of stress (Queitsch *et al*., 2002; Sangster *et al*., 2008). Genome-wide associations can help unravelling the genetic background of resilience (Sell-Kubiak *et al*.; Silva *et al*., 2014).

4.5 Conclusion

In this study, we showed that using EBV for resilience in absence of pathogen burden recorded led to favourable responses in resistance and tolerance to infections. The selection responses in resistance and tolerance depended on the heritabilities of resistance and tolerance and the genetic correlation between resistance and tolerance. If resistance and tolerance were unfavourably correlated, responses decreased, especially in resistance. Although using EBV for resilience

resulted mostly in favourable responses in resistance and tolerance, more genetic gain could be achieved when pathogen burden is recorded. Selecting on resilience is targeting, however, general resilience rather than specific tolerance to a certain disease.

4.6 Acknowledgement

This research was a part of the NematodeSystemHealth project, financed by Marie Curie Initial Training Networks (FP7-People-2010-ITN), and co-financed by Topigs Norsvin, the Netherlands, and Dutch Ministry of Economic Affairs, Agriculture, and Innovation (Public-private partnership "Breed4Food" code KB-12-006.03-004-ASG-LR and KB-12-006.03-005-ASG-LR).

4.7 References

Albers, G. A. A. *et al*. 1987. The Genetics of resistance and resilience to haemonchus-contortus infection in young merino sheep. Int. J. Parasitol. 17: 1355-1363.

Bishop, S. C. 2012. Possibilities to breed for resistance to nematode parasite infections in small ruminants in tropical production systems. Animal 6: 741-747.

Bishop, S. C., K. Bairden, Q. A. McKellar, M. Park, and M. J. Stear. 1996. Genetic parameters for faecal egg count following mixed, natural, predominantly Ostertagia circumcincta infection and relationships with live weight in young lambs. Anim. Sci. 63: 423-428.

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. 2010. On the Genetic interpretation of disease data. Plos One 5.

Bisset, S. A., C. A. Morris, D. R. Squire, and S. M. Hickey. 1996. Genetics of resilience to nematode parasites in young Romney sheep-use of weight gain under challenge to assess individual anthelmintic treatment requirements. New Zeal. J. Agr. Res. 39: 313-323.

Bloemhof, S., E. H. van der Waaij, J. W. M. Merks, and E. F. Knol. 2008. Sow line differences in heat stress tolerance expressed in reproductive performance traits. J. Anim. Sci. 86: 3330-3337.

Boddicker, N. *et al*. 2012. Evidence for a major QTL associated with host response to Porcine Reproductive and Respiratory Syndrome Virus challenge. J. Anim. Sci. 90: 1733-1746.

Bulmer, M. G. 1976. The effect of selection on genetic variability: a simulation study. Genet. Res. (Camb) 28: 101-117.

Detilleux, J., L. Theron, J. M. Beduin, and C. Hanzen. 2012. A structural equation model to evaluate direct and indirect factors associated with a latent measure of mastitis in Belgian dairy herds. Prev. Vet. Med. 107: 170-179.

Doeschl-Wilson, A. B., B. Villanueva, and I. Kyriazakis. 2012. The first step towards genetic selection for host tolerance to infectious pathogens: Obtaining the tolerance phenotype through group estimates. Front. Genet. 3.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Longman, Harlow.

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. 2009. ASReml user guide release 3.0. VSN International, Hemel Hempstead, UK.

Guy, S. Z., P. C. Thomson, and S. Hermesch. 2012. Selection of pigs for improved coping with health and environmental challenges: breeding for resistance or tolerance? Front. Genet. 3: 281.

Hazel, L. N. 1943. The genetic basis for constructing selection indexes. Genetics 28: 476-490.

Hill, W. G. 1976. Order statistics of correlated variables and implications in genetic selection programmes. Biometrics 32: 889-902.

Islam, Z. U. *et al*. 2013. Quantitative analysis of porcine reproductive and respiratory syndrome (PRRS) viremia profiles from experimental infection: a statistical modelling approach. Plos One 8.

Kause, A. 2011. Genetic analysis of tolerance to infections using random regressions: a simulation study. Genet. Res. (Camb) 93: 291-302.

Knap, P. W., and G. Su. 2008. Genotype by environment interaction for litter size in pigs as quantified by reaction norms analysis. Animal 2: 1742-1747.

Kolmodin, R., and P. Bijma. 2004. Response to mass selection when the genotype by environment interaction is modelled as a linear reaction norm. Genet. Sel. Evol. 36: 435-454.

Lunney, J. K. *et al*. 2011. Probing genetic control of swine responses to PRRSV infection: current progress of the PRRS host genetics consortium. BMC Proc. 5 Suppl 4: S30.

Mathur, P. K. *et al*. 2014. Estimating challenge load due to disease outbreaks and other challenges using reproduction records of sows. J. Anim. Sci. 92: 5374-5381.

Meuwissen, T. H. 1991. Reduction of selection differentials in finite populations with a nested full-half sib family structure. Biometrics 47: 195-203.

Morris, C. A. *et al*. 2010. Selecting for resilience in Romney sheep under nematode parasite challenge, 1994–2007. New Zeal. J. Agr. Res. 53: 245-261.

Mulder, H. A., and P. Bijma. 2005. Effects of genotype x environment interaction on genetic gain in breeding programs. J. Anim. Sci. 83: 49-61.

Queitsch, C., T. A. Sangster, and S. Lindquist. 2002. Hsp90 as a capacitor of phenotypic variation. Nature 417: 618-624.

Rashidi, H., H. A. Mulder, P. Mathur, J. A. M. van Arendonk, and E. F. Knol. 2014. Variation among sows in response to porcine reproductive and respiratory syndrome. J. Anim. Sci. 92: 95-105.

Rowland, R., J. Lunney, and J. Dekkers. 2012. Control of porcine reproductive and respiratory syndrome (PRRS) through genetic improvements in disease resistance and tolerance. Front. Genet. 3.

Sae-Lim, P. *et al*. 2015. Genetics of growth reaction norms in farmed rainbow trout. Plos One 10.

Sangster, T. A. *et al*. 2008. HSP90 affects the expression of genetic variation and developmental stability in quantitative traits. Proc. Natl. Acad. Sci. U. S. A. 105: 2963-2968.

Sell-Kubiak, E. *et al*. Genome-wide association study reveals novel loci for litter size and its variability in a Large White pig population. submitted.

Sevillano, C. A., H. A. Mulder , H. Rashidi, and E. F. Knol. Estimation of variance components for farrowing rate in pigs in response to photoperiod and ambient temperature. submitted.

Silva, F. F. *et al*. 2014. Sire evaluation for total number born in pigs using a genomic reaction norms approach. J. Anim. Sci. 92: 3825-3834.

Stear, M. J. *et al*. 1995. Regulation of egg production, worm burden, worm length and worm fecundity by host responses in sheep infected with Ostertagia circumcincta. Parasite Immunol. 17: 643-652.

Van Grevenhof, E. M., J. A. M. Van Arendonk, and P. Bijma. 2012. Response to genomic selection: The Bulmer effect and the potential of genomic selection when the number of phenotypic records is limiting. Genet. Sel. Evol. 44.

5

Identification of Genomic Regions Associated with Resistance, Resilience, and Tolerance to Porcine Reproductive and Respiratory Syndrome

H. Rashidi[1], M. S. Lopes[2], G. Lough[3], N. Duijvesteijn[2], J. C. M. Dekkers[4], J.A.M van Arendonk[1], P. K. Mathur[2], H. A. Mulder[1]

[1] Animal Breeding and Genomics Centre, Wageningen University, PO Box 338, 6700 AH Wageningen, the Netherlands, [2] Topigs Norsvin Research Center, PO Box 43, 6640 AA Beuningen, the Netherlands, [3] The Roslin Institute & R(D)SVS, University of Edinburgh, Edinburgh, Midlothian, United Kingdom, [4] Department of Animal Science, Iowa State University, Ames, IA, USA

To be submitted

Abstract

Porcine reproductive and respiratory syndrome (PRRS) is a major viral disease in pigs causing reproduction loss in sows and decreased growth in growing pigs. Conventional strategies for controlling PRRS have not been successful so far. Therefore, selective breeding for pigs that mount a defence against PRRS could be an option. Defensive ability of a pig against PPRS might have two mechanisms: resistance (ability to limit the PRRS viral load in the body) and tolerance (ability to minimize performance loss at a certain level of PRRS viral load). When it is not possible to distinguish between resistance and tolerance, defensive ability is measured as resilience, which is the ability to maintain performance during a PRRS outbreak regardless of viral load. In this study, area under the viremia curve up to 14 dpi (AUC14) was used as a measure of resistance, average daily gain up to 28 dpi (ADG28) as a measure of resilience and regression of ADG28 on AUC14 as a measure of tolerance. Our aim was to identify genomic regions associated with resistance, tolerance and resilience to PRRS. Data on 1,320 crossbred pigs that were experimentally infected with PRRS virus were analysed. Animals were genotyped using the Illumina 60K SNP chip. After quality control, 44,787 SNP were used in a genome-wide association study (GWAS). Genome-wide associations for resistance and resilience were detected as significant SNP effect on AUC14 and ADG28, respectively, while genome-wide associations for tolerance were detected as significant SNP effects on the regression of ADG28 on AUC14.

The heritabilities were 0.20 for AUC14, 0.26 for ADG28, and 0.21 for tolerance estimated as the heritability of ADG28 at average AUC14. For AUC14 and ADG28, a significant region (FDR < 0.20) was identified on chromosome 4 in which the most significant SNP for AUC14 and ADG28 explained 4.54 and 4.64% of the phenotypic variances, respectively. We also identified a region on chromosome 11 for AUC14 and a region on chromosome 16 for ADG28. For tolerance, significant regions were identified on chromosomes 1, 9, and 18. The most significant SNP was on chromosome 1 and explained 0.88% of the phenotypic variance. These associations indicate that tolerance is under genetic control and may play an important role in host response to PRRS, alongside resistance. This is the first study to detect genomic regions associated with tolerance to PRRS that can be used in marker-assisted selection for improving resistance, resilience, and tolerance to PRRS.

Key words: Genetics, genomic regions, pigs, porcine reproductive and respiratory syndrome, resistance, tolerance

5.1 Introduction

Porcine reproductive and respiratory syndrome (PRRS), caused by the PRRS virus (PRRSv), is a major problem in pig industry. In addition to welfare problems, PRRS causes huge economic losses. Different control strategies, such as vaccination and biosecurity, have been practiced in pig farms but not all of these have been fully successful so far (Thanawongnuwech and Suradhat, 2010; Renukaradhya et al., 2015). Therefore, selective breeding for pigs that can mount a defence against PRRS could be a promising approach to control PRRS. The defensive ability of animals against infection may have two mechanisms: resistance and tolerance. Resistance is the ability to prevent entry of pathogen or inhibiting replication of pathogen in the body and it is measured as the genetic effect on pathogen burden following exposure (Albers et al., 1987; Raberg et al., 2007). Tolerance is the ability to show minimal decrease in performance given a certain pathogen burden (Raberg et al., 2007). The genetic basis of tolerance can be modelled as the genetic effect on the reaction norm of performance on its pathogen burden (Kause, 2011). Resilience, on the other hand, is defined as the ability of animals to maintain performance during a disease outbreak regardless of pathogen burden. Resilience is modelled as the genetic effect on performance during an outbreak (Albers et al., 1987; Bisset et al., 1996; Rashidi et al., 2014). Boddicker et al. (2012) found heritabilities of 0.31 for PRRS viral load (resistance) and 0.30 for bodyweght gain during PRRS infection period (resilience), indicating the existence of genetic basis for resistance and resilience to PRRS. Boddicker et al. (2012 and 2014a) also identified a genomic region on chromosome 4 associated with PRRS viremia (resistance) and average daily gain of the infected pigs (resilience), indicating the possibilities for marker-assisted selection to reduce the impact of PRRS. However, the genetic basis of tolerance to PRRS is not known and genome-wide association studies (GWAS) on tolerance to PRRS have not been performed yet. The aims of this study were to 1) estimate the genetic parameters for resistance, resilience, and tolerance to PRRS, and to 2) identify genomic regions associated with resistance, resilience, and 3) tolerance to PRRS and investigate potential overlap in associated regions.

5.2 Material and methods

The Kansas State University Institutional Animal Care and Use Committee approved all experimental protocols for this study.

5.2.1 Animals

From the first 8 trials and trial 15 of the PRRS host genetics consortium (PHGC), data on 1,320 crossbred pigs were analysed. A general overview of the PHGC trials is described in Lunney et al. (2011). In short, from high health multiplier farms that were free of PRRS, swine influenza, and Mycoplasma hyopneumoniae, weaned pigs from different crossbred lines were sent to the experimental centre of Kansas State University. Pigs were between 11 and 21 days old upon arrival at the experimental facility. In each trial, pigs were from the same cross and from the same farm, except for trial 5 and 8 in which pigs were from two farms. The pedigree for the first three trials included three generations. The pedigree for trials 4-8 and 15 included sire and dam only. Upon arrival, pigs received broad spectrum antibiotics. After one week of acclimatization, pigs (between 17 to 32 weeks old) were challenged with 105 (TCID50) of NVSL-97-7985, a highly virulent PRRSv strain. In order to measure the virus load, blood samples were collected at -6, 0, 7, 10, 14, 21, 28, 35, and 42 days post infection (DPI). Body weights of pigs in kg were recorded at 0, 7, 1, 21, 28, 35, and 42 DPI. Pigs were euthanized at 42 DPI.

5.2.2 Genotypes

Pigs were genotyped for single nucleotide polymorphism (**SNP**) using the Illumina Porcine 60k Beadchip (San Diego, California). After genotyping, unmapped SNPs and SNPs located on sex chromosomes, according to the Sscrofa10.2 assembly of the reference genome (Groenen et al., 2012), were excluded from the dataset. In addition, SNPs with call rate <0.95, minor allele frequency <0.01, strong deviation from Hardy Weinberg Equilibrium ($\chi 2$ values>600), and with one of the genotypes having a frequency <0.02 were excluded. After all quality control procedures, missing genotypes were imputed across all trials using Beagle (Browning and Browning, 2007). Finally, out of the 64,232 initial SNPs, genotypes on 44,787 SNPs were available for the GWAS.

5.2.3 Phenotypic traits

The cumulative PRRS viral load of the pigs was calculated as the area under the PRRS viremia curve (Islam et al., 2013) from 0 to 14 DPI (**AUC14**). The average daily gain (kg/day) of the pigs from 0 to 28 DPI (**ADG28**) was also calculated. Average daily gain records after 28 DPI were not used because about 20% of the pigs showed a rebound in viremia levels mostly after 28 DPI (Boddicker et al., 2012). In addition, we preliminarily studied the reaction norms of average daily gain (**ADG**) on the areas under the PRRS viremia curve (**AUC**) up to 7, 14, 21, 28, 35,

and 42 DPI using sire random regression models. Comparison of the random regression models with their corresponding sire models showed that the difference between the log-likelihoods for the random regression of ADG28 on AUC14 (2801.48) and the univariate sire model of ADG28 (2800.90) was highest among other comparisons of random regression model with the univariate sire model. The records for AUC14 were normally distributed and ranged from 51.27 to 106.70 with a mean of 82.38 and a SD of 5.95. The records for ADG28 were also normally distributed and ranged from -0.09 to 0.77 kg/day with a mean of 0.33 kg/day and a SD of 0.12 kg/day. The genetic correlation between AUC14 and ADG28 was -0.85±0.20.

5.2.4 Statistical models for GWAS

To identify genomic regions associated with PRRS resistance and tolerance, we used a single-SNP analysis in which each SNP was modelled individually as a fixed class variable.

Three statistical models applied for performing the GWAS were: 1) a sire model for AUC14 to identify genomic regions associated with PRRS resistance; 2) a sire model for ADG28 to identify genomic regions associated with PRRS resilience; and 3) a sire model where the genetic effect on ADG28 was estimated as the random regression of sire on viremia (AUC14) and two fixed SNP effects were included to identify genomic regions associated with intercept (vigour) and slope (tolerance). All three models were implemented in ASReml 3 (Gilmour et al., 2009).

5.2.5 SNP effects on resistance (AUC14)

AUC14 was used as a measure of resistance to PRRS and therefore a sire model for AUC14 was used to identify genomic regions associated with resistance to PRRS. The model was as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Z_t p} + \mathbf{Z_l u} + \mathbf{e}, \qquad [1]$$

where, $\mathbf{y}$ is a vector of phenotypes for AUC14; $\mathbf{X}$ is the incidence matrix for fixed effects; $\mathbf{b}$ is the vector of fixed effects including age of the pigs upon arrival as a covariate, sex of the pigs as a class variable, virus rebound (yes/no) as a class variable for the pigs showing a second viremia peak after 21 DPI, trial as a class variable, parity of the dam as a class variable, and SNP as a class variable coded as 0 (AA), 1 (AB) ,or 2 (BB); $\mathbf{Z}$ is the incidence matrix for random sire effects; $\mathbf{a}$ is a vector of random effects for sire, with $\mathrm{N}(\mathbf{0}, \mathbf{A}\sigma_a^2)$, where $\mathbf{A}$ is the pedigree-based average numerator relationship matrix, and σ_a^2 is the variance of the sire effect; $\mathbf{Z_t}$

is the incidence matrix for the trial by pen interaction; $\mathbf{p}$ is a vector of the trial by pen random interaction with $N(\mathbf{0}, \mathbf{I}\sigma_p^2)$, where $\mathbf{I}$ is the identity matrix and σ_p^2 is the variance of trial by pen interaction; $\mathbf{Z_l}$ is the incidence matrix for litter; $\mathbf{u}$ is a vector of the litter random effects with $N(\mathbf{0}, \mathbf{I}\sigma_l^2)$, where σ_l^2 is the litter variance; and $\mathbf{e}$ the random residual term with $N(0, \mathbf{I}\sigma_e^2)$, and σ_e^2 is the residual variance.

5.2.6 SNP effects on resilience (ADG28)

ADG28 was used as a measure of resilience to PRRS and therefore a sire model for ADG28 was used to identify genomic regions associated with PRRS resilience. The model was the same as model [1], but with $\mathbf{y}$ being a vector of phenotypes for ADG28.

5.2.7 SNP effects on vigour (intercept) and tolerance (slope)

Tolerance is conventionally measured as the genetic effect on the reaction norm of host's performance on pathogen burden in a random regression model (Kause, 2011). If the intercept is put in the infection free environment, the genetic effect on the intercept would be the genetic effect on hosts' performance level when healthy, which is also known as vigour. The genetic effect on the slope of the reaction norm is the genetic effect on tolerance. For the current data, however, we observed that the full random regression model was not significantly better (P>0.05) than a model with only the intercept or only the slope and yielded almost equal likelihoods (Lough et al. manuscript in preparation). The random regression model showed a perfect negative genetic correlation between intercept and slope. Consequently, the analysis was not able to disentangle the genetic effects of intercept and slope. Because of interest in regions associated with tolerance, we used a sire model for ADG28 in which two fixed SNP effects were included for intercept and slope and slope was modelled as the random regression of sire on AUC14. Despite the inability to disentangling the genetic effects on intercept and slope with the full random regression model, we hypothesised the presence of genetic variance in tolerance and that including separate SNP effects for intercept and slope might help to distinguish genomic regions affecting ADG without disease (intercept) and genomic regions related to tolerance to PRRS. The model was as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_{AUC}a} + \mathbf{Z_t p} + \mathbf{Z_l u} + \mathbf{e}, \qquad [2]$$

where $\mathbf{y}$ is the vector of phenotypes for ADG28; the fixed effects were mostly the same as the fixed effects in model 1, except that AUC14 was added as a fixed

covariate and two fixed SNP effects were added as class variables for the intercept and slope; $\mathbf{Z_{AUC}}$ is the incidence matrix for AUC28; $\mathbf{a}$ is a vector of sire effects for slope on AUC14, assumed distributed $N(\mathbf{0}, \mathbf{A}\sigma_a^2)$, where σ_a^2 is the variance of $\mathbf{a}$. The additive genetic variance for ADG28 was calculated as $4 \times \sigma_a^2 \times (\overline{AUC14})^2$, where $\overline{AUC14}$ was the average AUC14 (82.38).

From the two SNP effects in the model, one was the SNP effect for the intercept (ADG28 at zero AUC14) that estimated the SNP effects on vigour of the pigs, and one was the SNP effect for the regression of SNP on AUC14 that measured the SNP effect on tolerance. Biologically, the SNP effect for intercept (vigour) is the SNP effect for ADG28 in absence of disease. Note that the data does not contain zero AUC14 and is therefore a mathematical extrapolation to the situation without disease.

5.2.8 Genome- and chromosome-wide associations

The inflation factor (λ) for the distribution of P-values from the GWAS was estimated using the estlambda() function of the R package GenABEL (Aulchenko et al., 2007). P-values were adjusted for inflation following the genomic control approach described by Devlin and Roeder (1999). For $\lambda > 1.1$ (**WTCCC, 2007**), the F-values were divided by λ and P-values were recalculated. Further, to control the number of false positives due to the large number of tests (44,787), the false discovery rate (**FDR**) was calculated using the R package qvalue (Storey and Tibshirani, 2003). An FDR ≤0.20 was used to indicate significant genome-wide and chromosome-wide associations.

5.2.9 Genetic variance of the SNP

The variance explained by a QTL region for resistance and resilience was defined as the genetic variance of the significant SNPs in this region across all animals. The variance explained by a QTL region for vigour and tolerance was defined as the total genetic variance of the SNP effect on ADG28 ($var(b_{i_{ADG28}})$), where $b_{i_{ADG28}}$ is the total effect of the i[th] SNP on ADG28 calculated as:

$$b_{i_{ADG28}} = b_{i_{int}} + b_{i_{slope}} \times \overline{AUC14},$$

where $b_{i_{int}}$ is the effect of the i[th] SNP on intercept, $b_{i_{slope}}$ is the effect of the i[th] SNP on slope and $\overline{\mathrm{AUC14}}$ is the average AUC14, which was equal to 82.38. $b_{i_{int}}$ and $b_{i_{slope}}$ were obtained from the ASReml output. The total genetic variance explained by the SNP, therefore, was calculated as:

$var(b_{i_{ADG28}}) = var(b_{i_{int}}) + (\overline{AUC14})^2 * var(b_{i_{slope}}) + (2 * \overline{AUC14}) * cov(b_{i_{int}}, b_{i_{slope}})$,

where $var(b_{i_{ADG28}})$ and $var(b_{i_{int}})$ are, respectively, the variances of $b_{i_{ADG28}}$ and $b_{i_{int}}$ across all animals and $cov(b_{i_{int}}, b_{i_{slope}})$ is the covariance between $b_{i_{ADG28}}$ and $b_{i_{int}}$ across all animals. This approach allowed us to take the covariance between the SNP effects on intercept and slope into account for calculating the total genetic variance of the SNP effect on ADG28. Note that the genetic variance due to each SNP includes the additive genetic variance as well as the dominance variance due to that SNP.

5.2.10 Candidate genes

The 0.2 Mb left and right flanking region of the most significant SNPs were searched for associated genes using BIOMART in Ensembl Sus scrofa 10.2 (http://www.ensembl.org/biomart). The genes functions were searched for using NCBI genes (http://www.ncbi.nlm.nih.gov/gene) and GeneCards (http://www.genecards.org).

5.3 Results

5.3.1 Variance components

Variance components from the models are in Table 5.1. Heritabilities for resistance (0.20) and tolerance (0.21) at average AUC14 were lower than the heritability for resilience (0.26). Although the response variable for resilience and tolerance is ADG28 in both cases, the heritability of resilience was higher than heritability for tolerance because AUC14 was included as a fixed covariate in the model for tolerance, whereas for resilience it was not included in the model. Including AUC14 in the model explained more of the genetic variance for ADG28 rather than the residual variance and consequently the heritability of tolerance was lower than that of the resilience (Table 5.1).

5.3.2 Genomic control

The λ for SNP effects on the traits was generally larger than one except for tolerance. For SNP effects on resistance, the λ was 1.21, which became 0.95 after adjusting for the inflation (Fig. 5.1A). For SNP effects on resilience, the inflation factor λ was 1.28 and 0.93 after adjusting for the inflation (Fig. 5.1B). For the SNP effects on vigour and tolerance, the inflation factor λ for the SNP effects on vigour

was 1.22 and 0.94 after adjusting (Fig. 5.1C). The inflation factor λ for SNP effects on tolerance was 0.97, which did not require any adjustment (Fig. 5.1D).

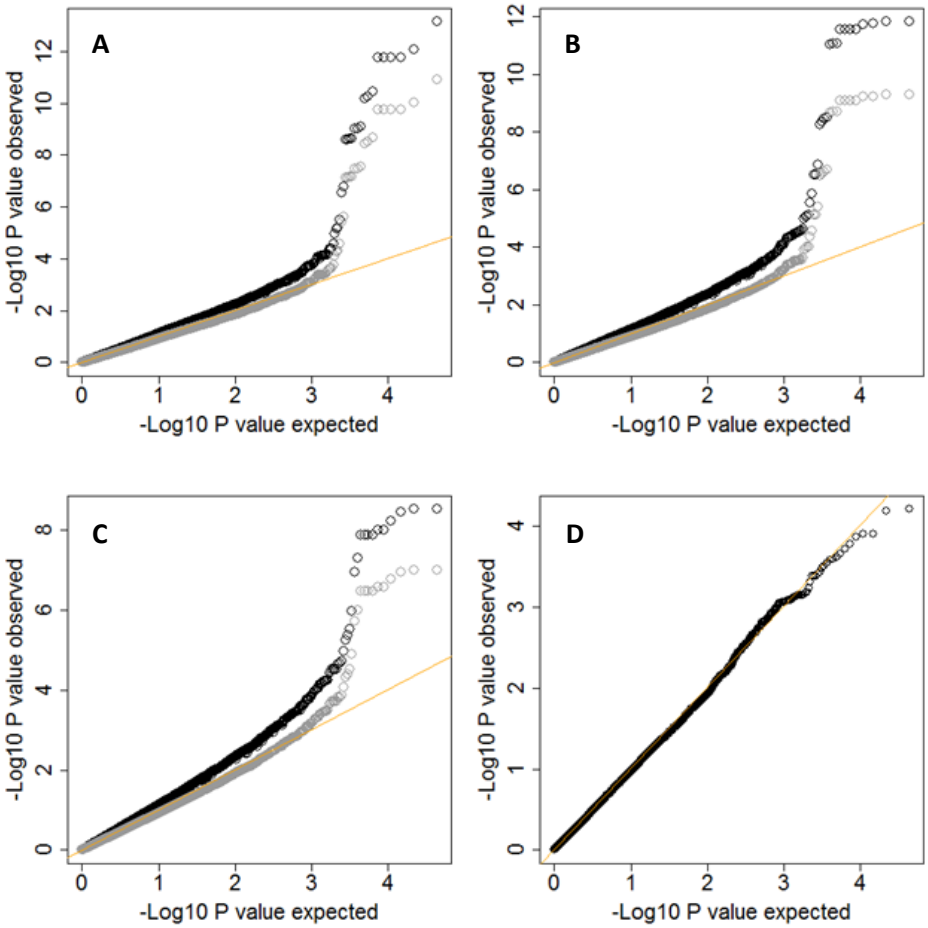

Figure 5.1. QQ-plots of the SNP P-value from sire model on area under the viremia curve up to 14 days post infection (A), average daily gain up 28 days post infection (B), vigour for average daily gain before infection (C), and tolerance (D). The black circles are the observed P-values, the grey circles are the adjusted P-values after genomic control, and the straight line shows the expected P-values under the null hypothesis. For tolerance P-values were not adjusted.

Table 5.1. Variance components (standard errors as subscripts) from the sire model on resistance defined as the genetic effect on area under the viremia curve at 14 days post infection (AUC14), resilience defined as the genetic effect on the average daily gain up to 28 days post infection (ADG28 (g/day)), and the tolerance defined as the genetic effect on ADG28 (g/day) at average AUC14 (82.34) modelled as the interaction of sire and AUC14 (Slope).

Variance	Resistance	Resilience	Tolerance
Genetic	$4.78_{2.80}$	$28.24_{11.50}$	$21.14_{9.37}$
Residual	$13.32_{0.64}$	$73.66_{3.57}$	$71.90_{3.47}$
Trial.Pen	$1.10_{0.36}$	$11.65_{2.81}$	$11.51_{2.74}$
Litter	$8.58_{1.07}$	$14.23_{3.29}$	$11.89_{3.00}$
Phenotypic	$24.19_{1.24}$	$106.60_{5.16}$	$100.59_{4.75}$
Heritability	$0.20_{0.11}$	$0.26_{0.10}$	$0.21_{0.09}$[1]

[1]Genetic and phenotypic variances and heritability for tolerance were calculated at the average AUC14 (82.34)

5.3.3 Associated SNPs and candidate genes

5.3.3.1 Significant SNPs for resistance, resilience and vigour

For resistance, resilience, and vigour, one region was identified on chromosome 4 from 139.26 to 140.42 Mb (Fig. 5.2-5.4, Table 5.2) in which the SNPs were in strong linkage disequilibrium as the r^2 ranged from 0.5 to 1. For resistance and resilience, 17 significant SNPs and for vigour, 12 significant SNPs were in the region on chromosome 4. For resistance, the most significant SNP was INRA0017729 (139.50 Mb), which explained 4.54% of the phenotypic variance. For resilience and vigour the most significant SNPs were ASGA0023349 (139.88 Mb) and ALGA0029538 (139.94 Mb), which explained 4.64% of the phenotypic variance for resilience and 3.76% of the phenotypic variance for vigour. In addition, one significant SNP (ASGA0050951) was identified on chromosome 11 (56.97 Mb) for resistance and one significant SNP (ASGA0073613) was identified on chromosome 16 (60.18 Mb) for resilience and vigour.

Twelve genes were identified near the region on chromosome 4 for resistance and resilience and 10 genes were identified near the region on chromosome 4 for vigour (Table 5.4). In addition, one gene was identified near the significant SNP on chromosome 11 for resistance and one gene near the significant SNP on chromosome 16 for resilience and vigour.

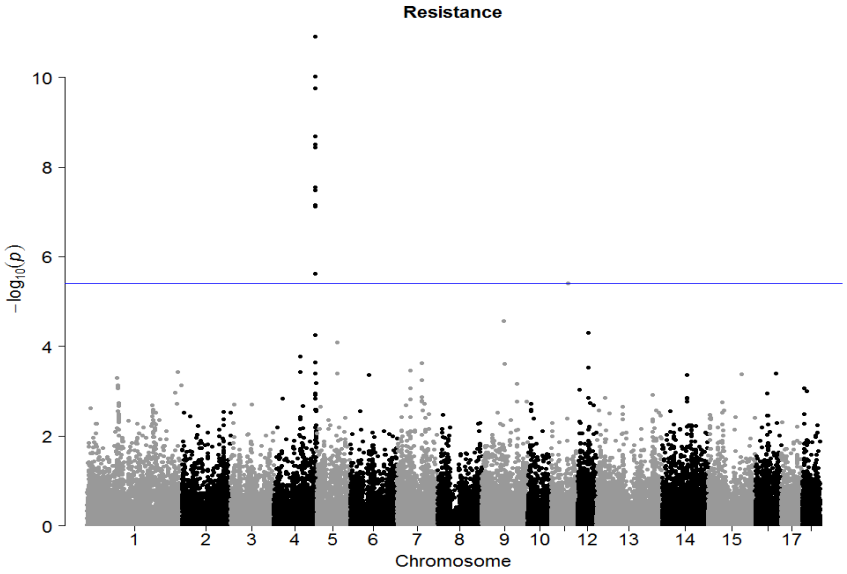

Figure 5.2. Genome-wide association between the area under the viremia curve up to 14 days post infection (AUC14) and 44,787 mapped SNP across 18 autosomes using a univariate sire model. The straight line is the cut-off value of 5.41 which equals a FDR q-value ≤ 0.20.

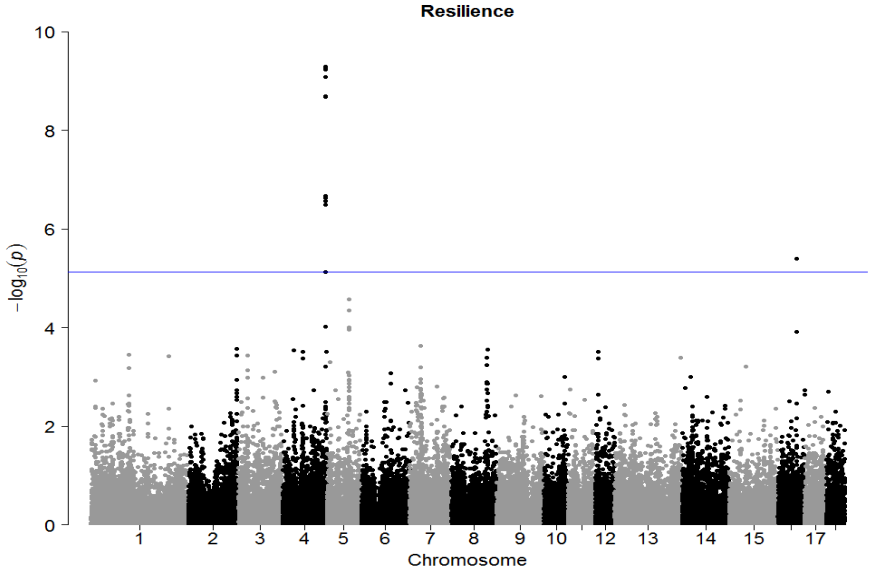

Figure 5.3. Genome-wide association between resilience measured as the average daily gain up to 28 days post infection (ADG28) and 44787 mapped SNP across 18 autosomes using a univariate sire model. The straight line is the cut-off value of 5.12 which equals a FDR q-value ≤ 0.20.

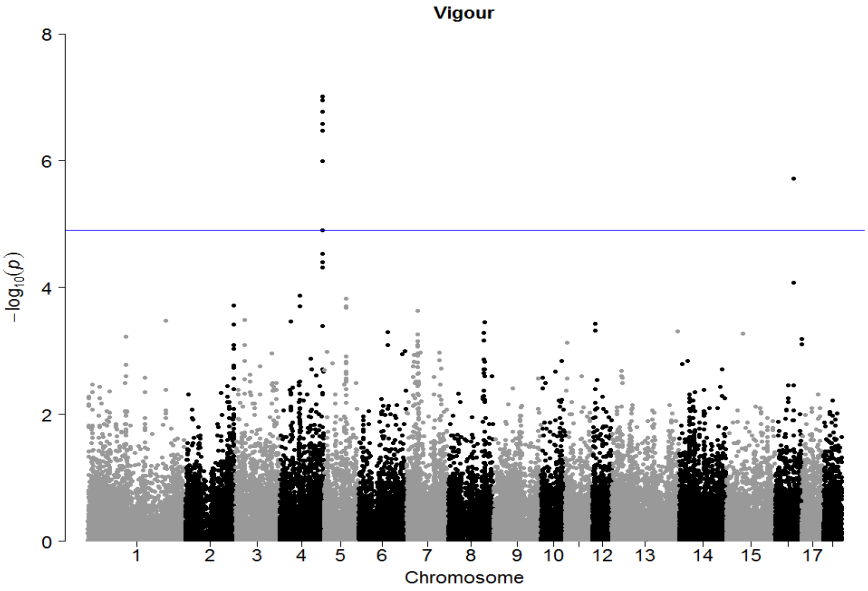

Figure 5.4. Genome-wide associations for vigour measured as the intercept of average daily gain at zero viremia and 44787 mapped SNP across 18 autosomes. The straight line is the cut-off value of 4.90 which equals a FDR q-value ≤ 0.20.

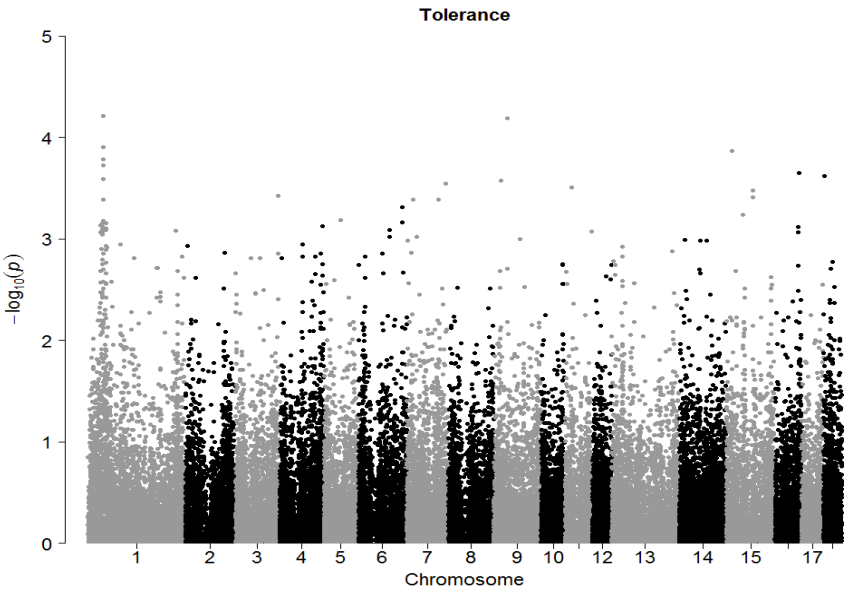

Figure 5.5. Genome-wide association for tolerance measured as the slope of average daily gain up to 28 days post PRRS infection (ADG28) on area under the viremia curve up to 14 days post infection (AUC14).

Table 5.2. The position and proportion of phenotypic variance ($\boldsymbol{\sigma_p^2}$) explained by significant SNPs for resistance, resilience, and vigour. Results for the most significant SNP for each trait are bold.

SNP	SSC	Position (Mb)	SNP variance (% of σ_p^2)		
			Resistance	Resilience	Vigour
ASGA0023314		139.26	2.48	3.45	2.68
INRA0017729		139.50	**4.54**	4.28	3.67
ASGA0023322		139.60	4.32	3.40	-
MARC0056249		139.64	4.09	4.44	3.63
WUR10000125		139.67	4.09	4.44	3.63
ALGA0029524		139.69	3.62	4.57	3.74
ASGA0023335		139.74	3.21	4.59	3.76
ASGA0023344		139.77	4.09	4.44	3.63
MARC0014819	4	139.80	4.09	4.44	3.63
ASGA0023349		139.88	3.21	**4.64**	**3.76**
ALGA0029538		139.94	3.21	**4.64**	**3.76**
ASGA0023354		139.97	3.11	4.38	3.61
DRGA0005385		140.01	3.11	4.38	3.61
M1GA0006784		140.08	4.05	3.60	-
MARC0000425		140.20	4.16	3.52	-
ASGA0023397		140.38	3.80	3.04	-
MARC0040196		140.42	3.80	3.04	-
ASGA0050951	11	56.97	2.62	-	-
ASGA0073613	16	60.18	-	2.86	3.46

5.3.3.2 Significant SNPs for tolerance

For tolerance, we did not identify any significant SNP at the genome-wide significance level. However, a suggestive association for tolerance on chromosome 1 was observed (Fig. 5.5) and the chromosome-wise significance level (FDR ≤0.20) revealed significant regions on chromosomes 1, 9, and 18 (Table 5.3). The SNPs in the suggestive QTL region on chromosome 1 formed three LD blocks. The first block was at 39.69 Mb to 40.58 Mb and the r^2 ranged from 0.3 to 1. The second block was located at 44.23 Mb to 49.47 Mb and the r^2 ranged from 0.1 to 1. The third block was located at 57.08 Mb to 58.06 Mb and ther^2 ranged from 0.6 to 1. The most (chromosome-wise) significant SNP for tolerance was ALGA0003292 (57.40 Mb), which explained 0.88% of the phenotypic variance of ADG28 at average AUC14. In addition, there were 8 (chromosome-wise) significant SNPs on chromosome 1 located at 105.38 Mb, 149.05 Mb, 223.21 Mb, 225 Mb, 285 Mb, 290.49-290.66 Mb, and 305 Mb. The (chromosome-wise) significant SNP on chromosome 9 (MARC0010165) was located at 42.23 Mb and the (chromosome-

wise) significant SNP on chromosome 18 (ASGA0078711) was located at 52.35 Mb (Table 5.3).

There were 14 genes near the three regions genes on chromosome 1. One gene was near the significant SNP on chromosome 9, and two genes near the (chromosome-wise) significant SNP on 18 (Table 5.4).

In summary, the same genomic region on chromosome 4 was associated with resistance, resilience, and vigour. For vigour and resilience, the same SNP was detected on chromosome 16. For tolerance, a chromosome-wide genomic region was found on chromosome 1 and 2 regions on chromosome 9 and 18, which were not found for resistance, resilience and vigour.

Table 5.3. Chromosome-wide significant SNP of pigs for tolerance and the proportion of phenotypic variance ($\boldsymbol{\sigma_p^2}$) of average daily gain up to 28 post infection explained by the SNP.

Chromosome	First SNP	Last SNP	Number of SNP	Position (Mb)		Variance Explained (% of σ_p^2)	
				Start	End	Min	Max
1	H3GA0001452	ASGA0002324	7	39.69	40.58	0.26	0.41
	MARC0021005	ASGA0002556	23	44.23	49.47	0.27	0.76
	ALGA0003278	MARC0047693	12	57.10	62.13	0.31	0.88
	ALGA0115211	ALGA0115211	1	105.38	105.38	0.32	0.32
	DRGA0001536	DRGA0001536	1	149.00	149.00	0.24	0.24
	ALGA0007558	ALGA0007558	1	223.21	223.21	0.48	0.48
	INRA0005754	INRA0005754	1	225.00	225.00	0.48	0.48
	ALGA0009447	ALGA0009447	1	285.62	285.62	0.57	0.57
	ALGA0009785	DRGA0002408	2	290.49	290.66	0.75	0.75
	ASGA0008088	ASGA0008088	1	305.52	305.52	3.29	3.29
9	MARC0010165	MARC0010165	1	42.23	42.23	0.42	0.42
18	ASGA0078711	ASGA0078711	1	5.24	5.24	0.47	0.47

Table 5.4. Candidate genes associated with resistance, resilience, vigour, and tolerance of pigs to porcine reproductive and respiratory syndrome.

Gene	Chromosome	Resistance	Resilience	Vigour	Tolerance
TBC1D32	1				✓
FAM184A					✓
MCM9					✓
ASF1A					✓
CEP85L					✓
PLN					✓
SLC35F1					✓
SMAP1					✓
B3GAT2					✓
PRRC2B					✓
SNORD62					✓
POMT1					✓
UCK1					✓
RAPGEF1					✓
PKN2	4	✓	✓	✓	
BARHL2		✓	✓		
ZNF326		✓	✓		
LRRC8C		✓	✓	✓	
LRRC8B		✓	✓	✓	
GBP4		✓	✓	✓	
GBP6		✓	✓	✓	
GBP5		✓	✓	✓	
GBP2		✓	✓	✓	
GBP1		✓	✓	✓	
CCBL2		✓	✓	✓	
GTF2B		✓	✓	✓	
C11orf87	9				✓
POU4F1	11	✓			
RNF219		✓			
RBM26		✓			
NDFIP2		✓			
SPRY2		✓			
TENM2	16		✓	✓	
XRCC2	18				✓
CCT8L2					✓

5.3.4 The effects of the most significant SNP for intercept and slope on ADG28

The overall effects of the genotypes for the most significant SNPs on for intercept (ASGA0023349 on chromosome 4) and slope (ALGA0003292 on chromosome 1) on ADG28 are in Fig. 5.6 and 5.7. The effects of the homozygous genotypes at low AUC14 (51.27) were positive and the difference between their genotypic values (2a, following the notation of Falconer and Mackay (1996)) was 0.09 kg/day (a=0.045kg/day). Therefore, selecting pigs that have the AA genotype will improve ADG28 at low AUC14. At high AUC14 (106.7) the effects of the both homozygote genotypes became negative and the difference between them decreased to almost zero ($2a = 0.007$ kg/day), meaning that both genotypes had equal ADG28 at high AUC14. The deviation of the AB genotypic value from the average of the AA and BB genotypes (d, following the notation of Falconer and Mackay 1996) at low AUC14 was -0.08 kg/day, indicating a negative dominance effect. The deviation of the AB genotypic value from the average of the AA and BB genotypes (d) at high AUC14 was 0.1 kg/day and a positive dominance effect, indicating that the heterozygote had higher ADG28 than the homozygotes.

The overall effects of the genotypes for the most significant SNP for slope (ALGA0003292 on chromosome 1) on ADG28 are in Fig. 5.7. The effect of the BB genotype at low AUC14 was positive and the effect of the AA genotype was negative. The difference between the average phenotypic values of BB and AA genotypes (2a) was 0.26 kg/day (a=0.13 kg/day). Therefore, selecting pigs that have the BB genotype will improve ADG28 at low AUC14. At high AUC14 the effects of both homozygous genotypes became negative and the difference between them slightly decreased (a=0.10 kg/day). The deviation of the AB genotypic value (d) from the average of the AA and BB genotypes at low AUC14 was zero and 0.18 kg/day at high AUC14, indicating no dominance effect at low AUC14 and a positive dominance effect at high AUC14. Therefore, the heterozygotes had an advantage at high viremia levels compared to the homozygotes.

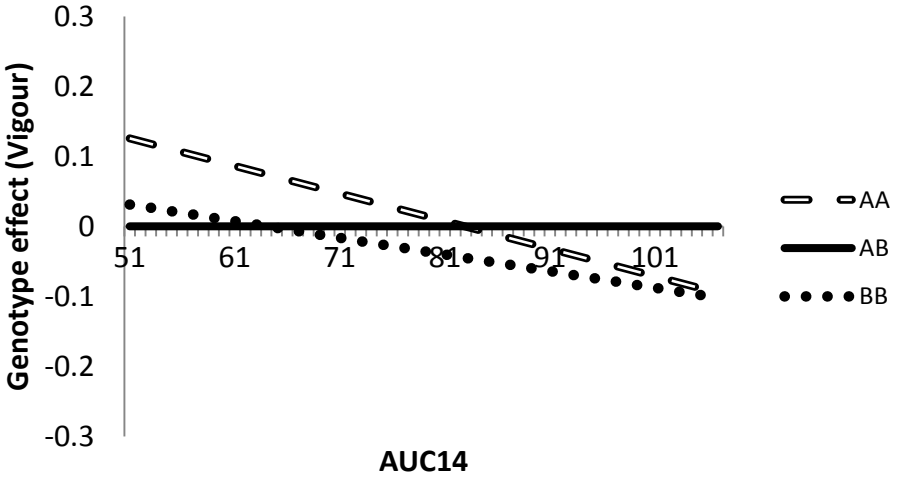

Figure 5.6. Genotype effects of the most significant SNP for intercept (ASGA0023349) on average daily gain of pigs up to 28 days post infection (ADG28) at each viremia level (area under the viremia curve) up to 14 days post infection (AUC14). The frequency for AA was 0.03, for BB was 0.72 and for AB was 0.25.

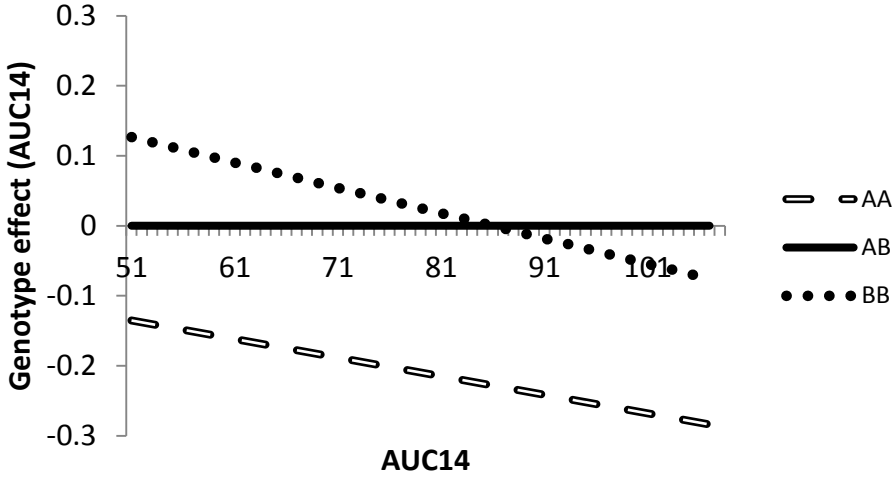

Figure 5.7. Genotype effects of the most significant SNP for slope (ALGA0003292) on average daily gain of pigs up to 28 days post infection (ADG28) at each viremia level (area under the viremia curve) up to 14 days post infection (AUC14). The frequency for AA was 0.29, for BB was 0.52 and for AB was 0.19.

5.4 Discussion

The aim of this study was to identify genomic regions associated with resistance (AUC14), tolerance (slope of the change in ADG28 over AUC14) and resilience (ADG28) to PRRS. We identified significant genomic regions associated with resistance, tolerance and resilience of pigs to PRRS.

Resistance, tolerance and resilience had moderate heritabilities indicating that selective breeding can improve response of pigs to PRRSv. The heritability for resistance (AUC14) in our study was similar to the heritability for viral load at 14 DPI in a subset of the current data in Boddicker et al. (2014b). The heritability for resilience (ADG28) in our study, however, was smaller than the heritabilities for bodyweight gains at 21 and 42 DPI in a subset of the current data in Boddicker et al. (2014b). One reason for the different heritabilities for resilience is that, Boddicker et al. (2014b) used an animal model for variance component estimation, whereas in the current study we used a sire model for that. Another reason is that Boddicker et al. (2014b) used a subset of the data that we used.

5.4.1 Common genomic region for resistance, resilience, and vigour on chromosome 4

We identified one genomic region on chromosome 4 associated with resistance, resilience, and vigour. The overlap between the SNPs on chromosome 4 for resistance and resilience was expected as the preliminary results showed a strong favourable genetic correlation between AUC14 and ADG28. For vigour, however, fewer SNP were found that overlapped with the SNP for resistance and resilience. The overlap between the associated genomic region on chromosome 4 for resistance, resilience, and vigour are in agreement with the high genetic correlations found and are in agreement with Boddicker et al. (2012) and (2014a).

The significant SNP on chromosome 4 explained a considerable proportion of the phenotypic variance for resistance and ADG28. The region on chromosome 4 was previously reported by Boddicker et al. (2012); Boddicker et al. (2014a). Boddicker et al. (2012) used pigs of the same crossbred lines from the first three trials of the PHGC and identified the genomic regions associated with area under the viremia curve 0 to 21 DPI (VL in their paper) and average daily gain from 0 to 21 DPI (WG21 in their paper) or 42 DPI (WG42 in their paper). For VL two genomic regions on chromosome 4 and chromosome X and for WG42 four genomic regions on chromosome 1, 4, 7, and 17 were found (Boddicker et al., 2012). In a follow up study, Boddicker et al. (2014b) validated the effect of the genomic region on

chromosome 4 on WG21, WG42 and VL21of the pigs from batch 4 and 5 of the PHGC. Then in another study Boddicker et al. (2014a) analysed pigs from trials 1 to 8 of the PHGC and redetected the genomic region on chromosome 4 associated with PRRS viremia and bodyweight gain. The re-detection of the region on chromosome 4 in our study, using even more data, validates the major impact of this region on resistance, resilience, and vigour of different pig breeds.

We detected 12 candidate genes located within the 0.2 Mb distance of the region on chromosome 4 based on ENSEMBL (http://www.ensembl.org/biomart) and the databases NCBI (http://www.ncbi.nlm.nih.gov/gene) and GeneCards (http://www.genecards.org) to search the gene functions. Among the genes, there was the guanylate-binding protein family genes (*GBP1*, *GBP2*, *GBP4*, *GBP5*, *GBP6*), which play an important role in anti-viral activities of the immune system by inducing cytokines. The importance of cytokines for PRRS response is reported by Lunney et al. (2010) and Miller et al. (2004). The association of guanylate-binding protein family genes with PRRS response has also been reported by Boddicker et al. (2012). Another relevant gene on chromosome 4 that we found was general transcription factor IIB (*GTF2B*), which in human is involved in disease pathway and has antiviral effects (Lund et al., 2007). The *GTF2B*-gene is a novel candidate gene for PRRS response and was present for resistance and resilience to PRRS as well as vigour. Another candidate gene for resilience and vigour on chromosome 4 is Cysteine Conjugate-Beta Lyase 2 (*CCBL2*) that is involved in metabolic pathways. Other genes near the significant SNP on chromosome 4 were involved in cell cycle (*PKN2*), sequence-specific DNA and RNA binding polymerase (*BARHL2* and *ZNF326*), and component of the volume-regulated anion channel (*LRRC8B*).

The genes near the significant SNP for resistance on chromosome 11 were involved in neural transcription factors that help developing sensory nervous system (*POU4F1*), RNA and nucleotide binding (*RBM26*), signal transducer activity (*NDFIP2*), and protein kinase binding and protein serine (*SPRY2*).

The gene (*TENM2*) near the significant SNP for resilience and vigour on chromosome 16 is involved in protein homodimerization activity and receptor binding.

5.4.2 Genomic regions for tolerance

We identified three regions on chromosome 1 that were significantly associated with tolerance at a chromosome-wide level (FDR<0.20). Chromosome-wide associations could be considered as suggestive as the significance threshold is set per chromosome and therefore we were able to find SNP that are significantly

associated per chromosome (Duijvesteijn et al., 2014). The additive variance for most of the SNP on chromosome 1 explained a small proportion of the phenotypic variance indicating small effects of those SNP on ADG28. Boddicker et al. (2012 and 2014a) also reported two significant regions on chromosome 1 associated with WG42 (located on 123.33-124.67 Mb) and VL (located on 292 Mb).

We detected 14 candidate genes located within the 0.2 Mb distance of the SNP on chromosome 1 based on (http://www.ensembl.org/biomart). Among the genes near the chromosome-wide significant SNP for tolerance on chromosome 1, the most interesting one was the Beta-1,3-Glucuronyltransferase 2 gene (*B3GAT2*) that encodes a protein that is involved in the synthesis of the human natural killer-1 (*HNK-1*) carbohydrate epitope (Kahler et al., 2011). This gene is involved in disease and metabolic pathways (www.genecards.org). Other genes on chromosome 1 have a broad range of functions such as cell cycle (*MCMC9*), cellular senescence (*ASF1A*), embryonic cell signal carrier (*TBC1D32*), signalling by epidermal growth factor receptor (*EGFR*), protein kinase in cardiac muscle (*PLN*), breast cancer antigen (*CEP85L*), producing red blood cells (*SMAP1*), brain development (*PRRC2B*), and muscle structure (*POMT1*).

The gene near the significant SNP for tolerance on chromosome 9 was a protein coding gene, the function of which is not clear. The genes near the significant SNP for tolerance on chromosome 18 were involved in stability and repair DNA damage (*XRCC2*) and unfolded protein binding and anion channel activity (*CCT8L2*).

5.4.3 Model for detecting genomic regions associated with tolerance

Tolerance is conventionally measured as the slope for the reaction norm of performance on pathogen burden. Random regression models are powerful approaches to estimate the genetic effects on intercept and slope of the reaction norms (Kause, 2011). Applying random regression model in our study, however, was not able to distinguish between intercept and slope. This was mainly due to a perfect negative correlation between intercept and slope (Lough et al. manuscript in preparation). Nevertheless, including two SNP effects in the model for intercept and slope, distinguished between genomic regions associated with vigour and tolerance. Finding different regions associated with vigour and tolerance approved our hypothesis that the model could differentiate SNP with effects on intercept and slope. We observed that the correlation between SNP effects on intercept and slope were almost -1 for all the SNPs. The completely negative correlation between

SNP effects for intercept and slope proved the inability to model sire effects for intercept and slope simultaneously.

Another way of modelling SNP effects on tolerance would be a 2-step approach where the significant SNP effects on intercept and slope are modelled separately (Streit et al., 2013). In the first step, breeding values for intercept and slope are estimated in a random regression model. In the second step, genome-wide association studies are performed using de-regressed breeding values for intercept and slope to detect significant SNP effects. This approach is particularly useful in dairy cattle with reliable breeding values for sires based on large offspring groups, but not suitable in our study because of the relative small number of animals with genotypes and phenotypes. We, therefore, estimated the SNP effects on intercept and slope in one model.

5.4.4 Genetic variance of SNPs for vigour and tolerance

The genetic variance of each significant SNP for vigour and tolerance was calculated as the total genetic variance of SNP effects on ADG28. This was done to account in a simple but accurate way for complete negative correlation between the SNP effect on intercept and the SNP effect on slope. As a consequence, calculating the genetic variance of the SNP effect on intercept only or slope only without taking into account the genetic variance of the other SNP effect and the covariance between the two SNP effects, would result in a genetic variance of more than 100% of the phenotypic variance.

The complete negative correlation makes marker-assisted selection for vigour and tolerance complex, because of the complete re-ranking of genotypes across the range of viremia. Index selection will give proper index weights for these SNP and determine the direction of selection. Selection will target the additive allele substitution effects and not the dominance deviations. However, the evidence for dominance on tolerance can be utilized in the crossbreeding part of the pig-breeding pyramid.

This is the first study that reports genomic regions associated with tolerance to PRRS. We also detected genomic regions associated with resistance and resilience to PRRS, confirming earlier studies. The most significant SNP on chromosome 4 detected for resistance to PRRS explained a considerable proportion of the phenotypic variance for PRRS viremia (AUC14). Also the significant SNP detected for resilience and tolerance explained a considerable proportion of the phenotypic variance of growth (ADG28). These results provide biological knowledge on resistance, tolerance and resilience to PRRSv. These

markers identified in this study can potentially be used for marker-assisted selection to improve pigs response to PRRS.

5.5 Acknowledgement

This research was a part of the NematodeSystemHealth project, financed by Marie Curie Initial Training Networks (FP7-People-2010-ITN), and co-financed by Topigs Norsvin, the Netherlands, and Dutch Ministry of Economic Affairs, Agriculture, and Innovation (Public-private partnership "Breed4Food" code KB-12-006.03-004-ASG-LR and KB-12-006.03-005-ASG-LR). The PRRS Host Genomics Consortium that generated all data analyzed herein is led by R.R.R. Rowland, J. Lunney and J.C.M. Dekkers and was supported by the USDA NIFA PRRS CAP Award 2008-55620-19132, the National Pork Board, the NRSP-8 Swine Genome and Bioinformatics Coordination projects, USDA ARS, Kansas State University, Iowa State University, PIC/Genus, Choice Genetics, Fast Genetics, Genetiporc, Inc., Genesus, Inc., PigGen Canada, Inc., and Topigs.

5.6 References

Albers, G. A. A. et al. 1987. The Genetics of resistance and resilience to Haemonchus rontortus rnfection in roung Merino rheep. Int. J. Parasitol. 17: 1355-1363.

Aulchenko, Y. S., S. Ripke, A. Isaacs, and C. M. Van Duijn. 2007. GenABEL: an R library for genome-wide association analysis. Bioinformatics 23: 1294-1296.

Bisset, S. A., C. A. Morris, D. R. Squire, and S. M. Hickey. 1996. Genetics of resilience to nematode parasites in young Romney sheep - Use of weight gain under challenge to assess individual anthelmintic treatment requirements. New Zeal. J. Agr. Res. 39: 313-323.

Boddicker, N. et al. 2012. Evidence for a major QTL associated with host response to porcine reproductive and respiratory syndrome virus challenge. J. Anim. Sci. 90: 1733-1746.

Boddicker, N. J. et al. 2014a. Genome-wide association and genomic prediction for host response to porcine reproductive and respiratory syndrome virus infection. Genet. Sel. Evol. 46.

Boddicker, N. J. et al. 2014b. Validation and further characterization of a major quantitative trait locus associated with host response to experimental infection with porcine reproductive and respiratory syndrome virus. Anim. Genet. 45: 48-58.

Browning, S. R., and B. L. Browning. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am. J. Hum. Genet. 81: 1084-1097.

Devlin, B., and K. Roeder. 1999. Genomic control for association studies. Biometrics 55: 997-1004.

Duijvesteijn, N., E. F. Knol, and P. Bijma. 2014. Boar taint in entire male pigs: a genomewide association study for direct and indirect genetic effects on androstenone. J. Anim. Sci. 92: 4319-4328.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Longman, Harlow.

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. 2009. ASReml user guide release 3.0 VSN International Ltd, Hemel Hempstead, HP1 1ES, UK. www.vsni.co.uk.

Groenen, M. A. M. et al. 2012. Analyses of pig genomes provide insight into porcine demography and evolution. Nature 491: 393-398.

Islam, Z. U. et al. 2013. Quantitative analysis of porcine reproductive and respiratory syndrome (PRRS) viremia profiles from experimental infection: a statistical modelling approach. Plos One 8.

Kahler, A. K. et al. 2011. Candidate gene analysis of the human natural killer-1 carbohydrate pathway and perineuronal nets in schizophrenia: B3GAT2 is associated with disease risk and cortical surface area. Biol. Psychiat. 69: 90-96.

Kause, A. 2011. Genetic analysis of tolerance to infections using random regressions: a simulation study. Genet. Res. (Comb) 93: 291-302.

Lough, G., H. Rashidi, I. Kyriazakis, N. Deeb, J. C. M. Dekkers, A. Kause, P. K. Mathur, H. A. Mulder, A. B. Doeschl-Wilson. Identifying genetic variances in resistance and tolerance to porcine reproductive & respiratory syndrome. (Manuscript in preparation)

Lund, R. J. et al. 2007. Genome-wide identification of novel genes involved in early Th1 and Th2 cell differentiation. J. Immunol. 178: 3648-3660.

Lunney, J. K., D. A. Benfield, and R. R. R. Rowland. 2010. Porcine reproductive and respiratory syndrome virus: an update on an emerging and re-emerging viral disease of swine. Virus. Res. 154: 1-6.

Lunney, J. K. et al. 2011. Probing genetic control of swine responses to PRRSV infection: current progress of the PRRS host genetics consortium. BMC Proc 5 Suppl. 4: S30.

Miller, L. C., W. W. Laegreid, J. L. Bono, C. G. Chitko-McKown, and J. M. Fox. 2004. Interferon type I response in porcine reproductive and respiratory syndrome virus-infected MARC-145 cells. Arch. Virol. 149: 2453-2463.

Raberg, L., D. Sim, and A. F. Read. 2007. Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. Science 318: 812-814.

Rashidi, H., H. A. Mulder, P. Mathur, J. A. M. van Arendonk, and E. F. Knol. 2014. Variation among sows in response to porcine reproductive and respiratory syndrome. J. Anim. Sci. 92: 95-105.

Renukaradhya, G. J., X. J. Meng, J. G. Calvert, M. Roof, and K. M. Lager. 2015. Live porcine reproductive and respiratory syndrome virus vaccines: current status and future direction. Vaccine 33: 4069-4080.

Storey, J. D., and R. Tibshirani. 2003. Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. U.S.A. 100: 9440-9445.

Streit, M., F. Reinhardt, G. Thaller, and J. Bennewitz. 2013. Genome-wide association analysis to identify genotype x environment interaction for milk protein yield and level of somatic cell score as environmental descriptors in German Holsteins. J. Dairy Sci. 96: 7318-7324.

Thanawongnuwech, R., and S. Suradhat. 2010. Taming PRRSV: revisiting the control strategies and vaccine design. Virus. Res. 154: 133-140.

6

General Discussion

In this thesis we aimed to: 1) estimate the genetic variation in resistance, tolerance, and resilience to infection in order to assess the amenability of these traits for selective breeding in farm animals, 2) estimate the genetic correlation between resistance, tolerance and resilience and 3) detect genomic regions associated with resistance, tolerance, and resilience. We used simple linear random regression models to study the reaction norm of animal performance on indicators of infection severity. We presented the possibility for selective breeding for resistance, tolerance, and resilience to two important infectious diseases in farm animals: porcine reproductive and respiratory syndrome in pigs and gastrointestinal nematode infection in sheep. We showed that selective breeding for resilience is expected to improve both resistance and tolerance. We also showed that there might be a genetic correlation between resistance and tolerance. The presence and sign of the genetic correlation between resistance and tolerance, however, might be different for different species and diseases. We detected genomic regions associates with resistance and tolerance to porcine reproductive and respiratory syndrome in pigs, indicating the possibility for marker assisted selection for the disease

In the current chapter, I will discuss the following topics: 1) response to infection as a special case of genotype by environment interaction, 2) random regression model as a statistical tool for studying response to disease, 3) advantages and requirements of random regression models, and 4) selective breeding of farm animals for resistance, tolerance, and resilience to infections.

6.1 Response to infection as a special case of genotype by environment interaction

Genotype by environment (G×E) interaction is the phenomenon of a genotype responding differently in different environments (Falconer and Mackay, 1996). In the case of G×E, the best performing organism in one environment might not be the best performing one in another environment, which may cause inefficiency in animal breeding. When G×E is present, breeders should not only consider the traits but also the environment in which the animal will perform. This way, breeders can breed animals for their appropriate environments. Response to infections in animals is a special case of G×E in two ways. The first is when ranking of animals on performance within a farm during different health status (healthy/diseased) changes. This could happen because the presence of a disease in a farm might change the genetic response of an animal to the environment. In

this case, one could apply a bivariate analysis in which the performance is considered as different traits in the presence and absence of a disease outbreak (Chapter 2), as proposed by Falconer (1952) to study genotype by environment interaction. The genetic correlation between the trait in the presence and absence of a disease outbreak would be an estimate of G×E. The non-unity genetic correlation reveals the re-ranking of the animals due to the presence of G×E in different health status of the farm. In Chapter 2 we showed that the correlation of sows for the reproduction traits, number of piglets born alive and piglet loss between healthy and diseased periods of the pig farm were 0.87 and 0.57, respectively. In line with our results, Herrero-Medrano *et al*. (2015) found the genetic correlations of 0.75 for number of piglets born alive, 0.74 for piglet loss, and 0.51 for number of weaned pigs between high and low production phases of pig farms.

The second is when during a disease outbreak animals respond differently to infection. The genetic response of the animals might be different per level of infection. Some animals might show variation in their ability to control the infection. Animals that are able to prevent the pathogen from entry or limit the pathogen burden within the body are called resistant. Some other animals might show variation in the change in the level of performance during the infection outbreak. Animals that are able to keep up the performance during an infection outbreak could be tolerant, resistant or both (discussed in details in the followings). The variation in ability of animals in responding to disease may lead to re-ranking in different levels of infection. For that, random regression models (RRM) could be applied whereby the reaction norm of performance is measured along the continuous change in the environment due to infection (de Jong and Bijma, 2002; Kause, 2011). The variation in the slope of the reaction norms indicates the presence of G×E.

Throughout this thesis, we extensively used random regression models to study response to infections in animals as the reaction norms of performance on infection severity or indictors of that (Chapter 2, 3, and 4). In the current chapter, I discuss the use of random regression models to study variation among animals in response to infections for direct and indirect measures of infection severity.

6.2 Random regression model for studying variation among animals in response to infections

Random regression models measure G×E as the change in phenotype over the change in the environment. The main characteristic of RRM is its ability for simultaneous estimation of breeding values of animals in different environments and the ability to estimate the persistency in production (as the slope of reaction norm) over different environments. To study response to infections in animals, RRM are used to measure the reaction norm of animal performance on severity of disease, such as pathogen burden and year-week estimates of the performance. When the records of pathogen burden are used in the RRM, the reaction norm of animal performance on pathogen burden is a measure of tolerance to infection. In Chapter 3, we studied tolerance of sheep to gastrointestinal nematode infection as the reaction norm of sheep bodyweight on two measures of nematode burden: faecal egg count (FEC) and pepsinogen. We showed the presence of genetic variation among sheep in tolerance to nematode infection and feasibility of selective breeding for tolerance in sheep. We also showed that the additive genetic variance and consequently the heritability of sheep bodyweight depend on the level of nematode infection. We showed that RRM is significantly better as compared to the sire model using a likelihood ratio test.

In farm animals, however, the records for individual pathogen burden or indications of that might not be available because recording pathogen burden or its related traits is laborious and costly. In that case, the mean phenotypic performance of the farm, like e.g. herd-year-week estimates for performance, that indirectly represents the infection severity can be used as a random environmental covariate in the RRM. When an indirect indication of disease severity is used in RRM, the slopes of the performance reaction norms measure the ability of the animals in response to infection (resilience) without distinguishing between resistance and tolerance. In Chapter 2, we studied resilience of sows to porcine reproductive and respiratory syndrome (PRRS) as the slope of the reaction norm of reproduction (number of piglets born alive or piglet loss) over the herd-year-week estimates for number of piglets born alive. We showed that there is variation among sows in resilience to PRRS. We compared the predictive ability of RRM during healthy and disease phase of the farm with a univariate animal model and a bivariate animal model. We showed that the predictive ability of RRM in both healthy and diseased environment is better than the other two models. In line with our results, other studies reported superiority of RRM over conventional models

for breeding value estimation. Huisman and van Arendonk (2004) studied the genetic basis of daily feed intake of pigs in different age. They showed that a RRM has a better fit as compared to a multivariate model based on the Akaike's information criterion and the Bayesian-Schwarz information criterion. Kranis *et al*. (2007) studied the genetic basis of egg production for turkeys in different production periods and compared RRM with a conventional multivariate model. They found a higher predictive ability for RRM compared with multivariate models. These findings show that when the environmental range is continuous, such as infection level in our case, RRM are better models compared to the conventional models for estimating breeding values. In the following paragraphs, I will discuss the advantages and requirements for RRM for unbiased and accurate estimation of breeding values and genetic variances.

6.3 Strength of RRM when G×E is present

Accurate estimation of breeding values

In Chapter 2, we showed that both in healthy and diseased periods of the pig farm, RRM has the highest accuracy of estimating the sow effects on number of piglets born alive (NBA) compared with univariate and bivariate models. In a validation study on 612,186 sow records from TopigsNorsvin, Mathur (personal communication, 2015) studied the predictive ability of RRM. He studied the reaction norm of total number of piglets born in one parity (TNB) on different environmental factors including seasonality, ambient temperature (heat), and the year-week estimates for TNB (YW). He compared the predictive ability of the RRM with a univariate animal model. For that, records in the training set (511,325 records) were used to predict the breeding values in the validation set (100,861 records). The same fixed effects were used as in the RRM and animal model. The correlation between the phenotype and predicted breeding values showed the predicting ability of each model. He observed that the predictive ability of RRM using different environmental factor is higher compared with that of the animal model (Table 6.1). These findings indicate that breeding value estimation using RRM is more accurate. Silva *et al*. (2014) studied G×E using RRM whereby the genetic effects on the number of piglet born alive in different herd-year-weeks (HYW in their paper) were estimated. They also showed that RRM models provide more accurate estimates of breeding values compared to a conventional univariate sire model.

Table 6.1. The correlation of sow phenotypes for total number of piglets born (TNB) with the breeding values from univariate animal model and random regression model (RRM). In the RRM, three different environmental factors were used: seasonality, ambient temperature (heat), and the year-week estimates for TNB (YW). The improvement is shown relative to the animal model.

Model	Environmental factor	Correlation	Improvement
Animal		0.201	
	Seasonality	0.207	3.00%
Random regression	Heat	0.218	8.50%
	YW	0.208	3.50%

RRM requires few parameters to describe data

When there are more than two environments, RRM requires fewer parameters to describe the genetic variation in each environment as compared to a multivariate analysis. For instance, when there are three environments, a multivariate model would require 3 genetic variances for each environment and 3 genetic covariances between the environments to describe the data. By increasing the number of environments (n), the number of genetic variances and covariances increases with n^2. This issue is clearly manifested with disease data where each level of infection can be viewed as an environment. A linear RRM requires two genetic variances for intercept and slope and one genetic covariance between them to estimate the genetic variation in each level of infection. The additive genetic variance of the trait in each environment and the covariance between the environments could be easily calculated using the (co)variance matrix of RRM. In chapter 3, we showed that using genetic (co)variance matrix from the RRM, the additive genetic of sheep bodyweight at each level of nematode burden (faecal egg count or Pepsinogen) is different. Furthermore, in comparison to a multivariate model, using an adequate amount of data, a RRM estimates variances and covariances smoother and with less bias (Kirkpatrick *et al*., 1990).

In this thesis, we only used linear RRM, which assumed a linear relationship between the change in phenotype and environments. Linear relationship between the change in phenotype and environments has been found for milk production traits and fertility traits in dairy cattle (Calus *et al*., 2002; Kolmodin *et al*., 2002; Lillehammer *et al*., 2009). Linear RRM are straightforward for calculating the breeding values in each environment, as they are simply functions of the environment, breeding value for intercept, breeding value for slope, and the genetic covariance between intercept and slope. It is, however, possible that the

relationship between change in phenotype and environment is non-linear (Streit *et al*., 2012; Herrero-Medrano *et al*., 2015). The non-linear RRM require more parameters and therefore are more complex for estimating breeding values as compared to linear RRM.

6.4 Requirements for RRM

Despite the clear advantages of RRM over other conventional models, there are some requirements to be considered when applying the model.

Large data for accurate estimation of breeding values in extreme environments

For unbiased and accurate estimation of breeding values and variances and covariances using RRM, count of animals as well as the number of records per animal per environment is important. One way of measuring the bias in estimation of breeding values is to look at the regression coefficient of phenotypes on estimated breeding values. Regression coefficient of one would indicate unbiased estimates of breeding values. When the regression coefficient is smaller than one or larger than one, breeding values are biased and the variance of estimated breeding values is either overestimated (b<1) or underestimated (b>1). To estimate the bias in breeding value estimation using RRM, I calculated the regression coefficient of sow phenotypes for NBA on estimated sow effects on NBA during the healthy and diseased phase of the farm. For that, I used a univariate animal model and a RRM on 57,135 records of 10,910 sows (data from Chapter 2). In the RRM, the reaction norm of sow phenotype for NBA on the year-week estimates for NBA was assessed. The sow effects on NBA from RRM were calculated at the average year-week estimates in healthy or diseased periods using the random regression (co)variance matrix (Chapter 3). To assess the effect of including more information on bias, four parity groups were considered: parities 1 through 4 (1–4), 1 through 5 (1–5), 1 through 6 (1–6), and 1 through 7 (1–7). In each group, the last parity was set to missing (validation set), and all other parities before the last one (training set) were used to predict the sow effects in the last parities. The regression coefficients were calculated for the regression of the sow phenotypes in the validation sets on the sow effects estimated from the training set. For the healthy period, the regression coefficients were generally closer to one compared with the diseased period, indicating that estimated sow effects are less biased during healthy periods (Table 6.2). This is probably because the incidence of the healthy periods was high and the majority of animals had records during the healthy periods. The regression coefficient in the parity groups 1–4 and 1–5 for healthy

period, however, were slightly larger than one indicating an over-estimation of breeding values. By including more parity, the regression coefficients tended to be closer to one especially for RRM, indicating an unbiased estimation of breeding values. For the disease period, however, the regression coefficients from the RRM were generally below one and smaller than those of the animal model, indicating an under-estimation of breeding values using RRM. This is probably because the incidence of diseased periods was lower than the healthy periods and fewer animals had records during the diseased periods. By increasing the number of records (including more parity), especially during the diseased phase, the regression coefficient tended to be closer to one, meaning that breeding values were less biased. This indicated that using RRM, the breeding values in the environments with fewer records were more biased unless adequate amount of data was provided. In line with our finding, in a simulation study Kause (2011) showed that small sample size can lead to biased estimation of variance components as the variances of intercept and slope and the genetic covariance between them were over-estimated compared with their simulated values. The over-estimation of the variance components is because with small sample size, it is most likely that the data is not a representative of the true distribution and genetic parameters are strongly influenced by single observations.

The small sample size at the environment also leads to more inaccuracy in estimating breeding values as well as the genetic variance. In Chapter 3, we observed that at the two extreme (very low/very high) levels of faecal egg count and pepsinogen, the SEs on the estimated breeding values for bodyweight and consequently heritabilities are much larger than the intermediate level of faecal egg count and pepsinogen. This is because fewer sheep had records at the very low and very high levels of faecal egg count and pepsinogen. Knap and Su (2008) showed that for RRM of litter-size in pigs on hear-year-season effects, increasing the family size resulted in smaller SE for slope estimates, indicating more accurate estimation of slope.

The linear RRM assumes constant environmental sensitivity over environments. Therefore, even for the environments with very few observations, the slope could be estimated. This might enforce the breeding values and correlation between them towards meaningless limits leading to imprecision in variance components and breeding value estimation. In other words, linear RRM may extrapolate the breeding values for slope outside the trajectory in which parameters are estimated. Adequate numbers of records in each environment, therefore, would improve the accuracy of estimation of slope breeding values, because it provides more data points along the environmental covariate. Accurate

estimates of slope breeding values would result in accurate heritability estimates per environment.

To obtain unbiased and accurate estimates of breeding values in each environment using RRM, the amount of records per environment and the connectedness between the records are crucial. The availability of adequate amounts of data per environment is challenging, especially for diseases in farm animals that may have low incidence.

Table 6.2. Regression coefficients of sow phenotypes for number of piglet born alive (NBA) on the sow effects on NBA in healthy and diseased periods of farm, from a univariate animal model (Animal) and a random regression model (RRM). Four parity groups were considered to assess the effect of sample size on bias.

Parity group	Healthy		Diseased	
	Animal	RRM	Animal	RRM
Parity 1-4	1.14	1.11	0.76	0.59
Parity 1-5	1.17	1.16	0.86	0.75
Parity 1-6	1.08	1.07	1.04	0.98
Parity 1-7	1.06	1.02	1.09	0.91

Dependency of breeding values on the environmental factor

Using RRM, the genetic variances are dependent on the range of environmental factor, which provides the opportunity to calculate breeding values per environment. The risk is, however, that estimated heritabilities and correlations are enforced towards meaningless limits. In addition, the changes in the heritability across different environments obtained from the RRM are not certain, because heritabilities with linear RRM always follow a parabolic shape across sequential environments. Heritability estimates from RRM, therefore, need to be validated with other models such as multivariate models (Chapter 3). This can be done by making ordinal subsets of data with sufficient number of animals and estimating the heritability in each subset, which could be of course at the expense of losing G×E information (Calus *et al*., 2004). For low incidence diseases, however, removing environments with few observations from the data is not an option because it can drop the important information about the disease. For low incidence diseases, therefore, random regression models might not be useful as it may cause over-estimation of breeding values (Calus *et al*., 2004; Kause, 2011). In that case, data should be analysed with other models that provide accurate estimates of breeding values. Silva *et al*. (2014) showed that employing genomic relationship

matrix in RRM might alleviate the uncertainty in estimated breeding values in the environments with few observations. They applied a genomic random regression models (RRM with genomic relationship matrix) on NBA of pigs to study the genetic effects in different herd-year-weeks. Comparing the genomic RRM with the RRM that used pedigree relationship matrix, they observed a greater accuracy of breeding value estimation across herd-year-week, which was more pronounced for the herd-year-weeks with fewer numbers of observations.

6.5 Environmental factors in RRM

To study the genetic components of animal responses to diseases using RRM, an important element is the environmental factor. The genetic effect on the slope is a function of the environmental factor and the total additive genetic effect per environment is a function of the genetic effects on intercept and slope. To study genetics of tolerance to infection, the reaction norm of performance over the records of individual pathogen burdens is measured in RRM (Chapter 3). In practice, though, records of pathogen burden might not be available. Using the environmental factor that represents the infection severity, therefore, is important for accurate breeding value estimation of response to infections using RRM. There are two types of environmental factor: direct and indirect environmental factors. The direct environmental factors are the covariates that are direct representative of the environment in which animal is performing. The most obvious example of a direct environmental factor in case of infections is pathogen burden. Pathogen burden is a direct indicator of the infection severity in an animal. The inverse of pathogen burden in an animal is a measure of resistance to infection) (Raberg *et al*., 2007. In most cases of infectious diseases, however, the direct measure of pathogen burden is difficult. For example, in the case of gastrointestinal nematode infection in sheep, in order to measure the exact number of nematodes the animal needs to be autopsied. In that case, an indication of nematode burden such as faecal egg count and IgA will be used (Chapter 3). Another example of direct environmental factors is the ambient temperature used in RRM for studying genetics of heat stress in pigs (Bloemhof *et al*., 2012). To estimate the heat tolerance in pigs, it is desired to use the outside temperature recorded at the location of the farm as the environmental factor in the RRM. In practice, though, the outside temperature near the farm might not be available and the temperature recorded at the nearest weather station could be used as representative of the temperature on farm (Freitas *et al*., 2006).

Other types of environmental covariates are indirect approximations for the environmental challenges. During a disease outbreak, the individual pathogen burden or its indicator traits might not be known. Recording pathogen burden or its indicator traits require regular sampling of animals that is costly and laborious. For example, in the case of PRRS in pigs in order to have an accurate measure of viral load during an outbreak weekly sampling of the animals is needed, which requires a lot of labour. In field studies on response to infection, therefore, an environmental parameter that represents the infection severity could be used in RRM. The most common environmental parameters that are used in RRM are raw mean phenotypic performance in each environment (Calus and Veerkamp, 2003) and herd-year-week estimates for phenotypic performance in each environment (Li and Hermesch, 2012) and (Chapter 2). In the studies on response to infection, herd-year-weeks of the performance during outbreaks are first estimated with simple linear models. Then the estimated herd-year-week is used as the environmental covariate in a RRM (Chapter 2). The slope for the reaction norm of performance on the herd-year-week estimates is a combined measure of resistance and tolerance to infection and is known as resilience (Chapter 4).

There are two arguments against using approximations of environmental factor in RRM. One argument is that using an environmental estimate that is based on the data may result in incorrect estimation of breeding values due to the presence of genetic trend in the environmental estimates (Su *et al*., 2006). They proposed a RRM with Bayesian Markov Chain Monte Carlo that estimates environmental values simultaneously with other parameters of the model. They compared that method with two RRM in which the true environmental values was used as a covariate and in another one the phenotypic mean of the herd-year was used as a covariate. They observed that the correlation between true values of herd-year effect and herd-year averages was smaller (0.901) than the correlation between true values of herd-year effect and estimated means of herd-year effects in their proposed method (0.97). They also observed that the estimated variance components from the RRM with phenotypic mean of the herd-year as covariate were biased as the genetic variance of intercept was over-estimated and the genetic variance in slope was under-estimated compared to the true (simulated) variances. They conclude that applying the method that estimates environmental values simultaneously with other parameters of the RRM is a more appropriate approach as it provides unbiased estimates of variance components. For estimation of the phenotypic mean of the herd-year, however, Su *et al*. (2006) did not correct for genetic effect. Indeed, without correcting the phenotypic mean of the herd-year, the estimates could bear genetic trends. Calus *et al*. (2004) suggested that

estimating environmental parameters using large number of animals per herd-year might result in unbiased estimation of breeding values using RRM. To test the effect of sample size on the accuracy of estimating the environmental parameter, I estimated the accuracy of the year-week estimates for NBA, using the data in chapter 2. The year-weeks were from the first week of 2004 through the week 28 of 2012. The number of animals per year-week ranged from 2 to 190. The year-week estimates were obtained from a univariate animal model on NBA corrected for fixed effects and random effects of year-week and sow (Chapter 2). I calculated the accuracy of the year-week estimates using the following formula:

$$r = \sqrt{1 - (\frac{se^2}{\sigma^2_{yw}})},$$

where, r is the accuracy of each year-week estimate, se is the standard error for each year-week estimate, and σ^2_{yw} is the variance of the year-week estimates. The accuracy increased by increasing the number of observation per year-week (Fig. 6.1). The lowest accuracy was 0.12 for 2 sows per year-week and the highest one was 0.94 for 190 sows per year-week. The number of 25 sows per year-week resulted in the accuracy of 0.75 and 80 sows per year-week resulted in the accuracy of 0.90. This finding shows that high accuracy of the year-week estimates are easily achievable in practice with adequate number of observations per year-week. We showed that RRM using year-week estimates of NBA as approximations for disease severity predict the future performance of sows better than univariate and bivariate models (Chapter 2). These findings show that using an appropriate statistical model and sufficient amount of data can result in accurate estimation of environmental factor. Accurate estimation of environmental factor in RRM provides more accurate estimation of breeding values in different environment compared with the univariate and bivariate models.

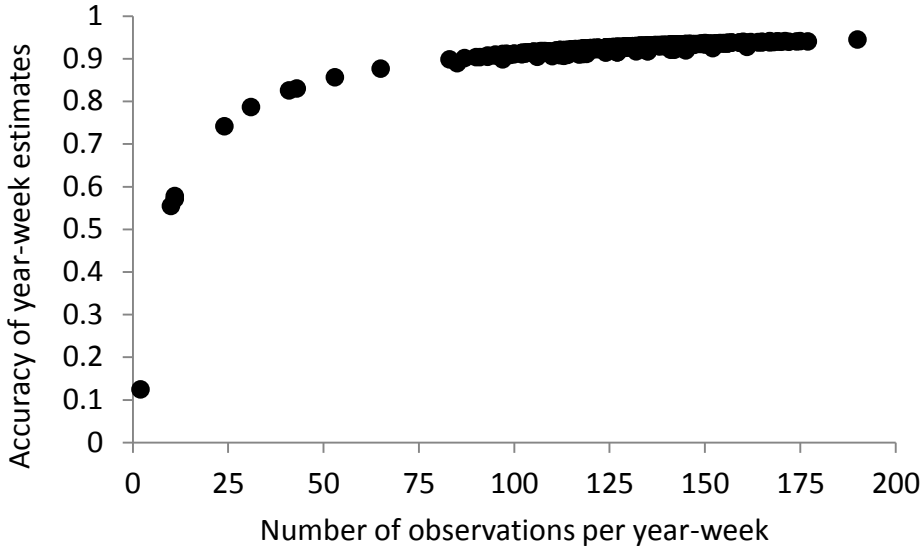

Figure 6.1. The accuracy of year-week estimates for number of piglets born alive.

Another argument against using environmental estimates in RRM is that the selection environment of animals may be different from their response environment. Environmental estimates that are based on the selection environment may therefore be bad predictors for environmental factors of the response environment. In other words, if high performing animals are selected based on the severity of infection during a disease outbreak but the severity of infection is different in the response environment, then it is uncertain what the performance will be. Therefore, there might be re-ranking of animals between the selection environment and the response environment. To study the re-ranking of animals between the selection environment and the response environment, I estimated the correlation of sow effects on NBA between different year-week estimates of NBA (Fig. 6.2). Based on the PRRS infection severity, three environments (no infection, medium infection, severe infection) were considered as selection environments. Then, the correlation between sow effects in the selection environments and the sow effects in other year-week estimates was calculated. The results showed that the correlations between the sow effects in similar environments (nearby year-week estimates) are high (Fig. 6.2). By increasing the differences between the environments, the correlation tends to decline. This result indicates that as long as the selection environments and the response

environments are similar, e.g. similar severity of PRRS infection in the farm, breeders should not worry about re-ranking of animals.

Using RRM, breeders can estimate the breeding values of the animals for different environments. Breeders can also decide which animals are suitable for each environment.

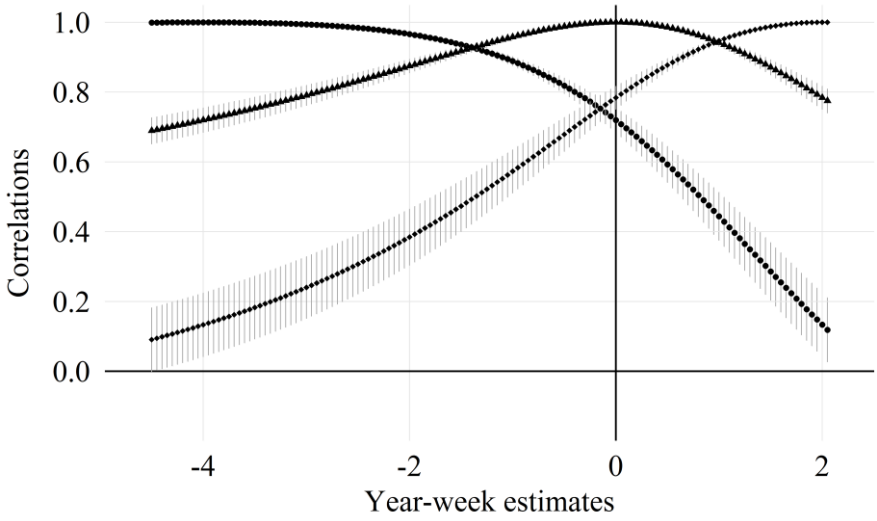

Figure 6.2. Correlation between the sow effects on number of piglets born alive (NBA) in the selection environments and the response environments. Three selection environments are considered based on the year-week estimates for NBA: No infection where the year-week estimate is 2 (diamonds), medium infection where the year-week estimate is 0 (triangles), and severe infection when the year-week estimate is -4 (circles). The vertical lines are the standard errors.

6.6 Breeding for response to infection

Controlling the infectious diseases in farm animals is an important part of the farm management. Farms undergo huge economic costs due to infectious diseases. Farm animals severely suffer from infectious diseases. The conventional approaches for disease control including vaccination, antibiotics, antiviral drugs, culling, sanitation, and biosecurity have been practiced for a long time in farm animals. Nowadays, the public concern about the excessive use of medicines and vaccines and contamination of animal products on one hand, and unsustainability of the conventional control strategies on the other hand have persuaded the

farmers to include response to infections in their breeding programs. Response to infections in any host consists mainly of the mechanism of controlling the replication of the invading pathogen in the host called “resistance”, and mechanism of minimizing the symptoms of the infection in the host called “tolerance”. When the distinction between these two mechanisms is difficult, the response to infections is referred to as resilience. In farm animals, studies have mainly reported the feasibility of selective breeding for resistance and resilience to infection for sheep (Albers *et al*., 1987; Bishop and Morris, 2007; Morris *et al*., 2010), cows (Heringstad *et al*., 2000; Bermingham *et al*., 2014), poultry (Cheng *et al*., 2008; Wolc *et al*., 2013), and fish (Kuukka-Anttila *et al*., 2010; Gjerde *et al*., 2011). Furthermore, there are several reviews on the importance and implication of resistance and resilience to infection in farm animals (Bishop and Morris, 2007; Bishop, 2012a, b; Bishop and Woolliams, 2014). For tolerance studies have discussed the methodology to estimate the trait (Kause, 2011; Kause and Odegard, 2012), the difficulties to estimate tolerance (Doeschl-Wilson *et al*., 2012; Bishop and Woolliams, 2014), and the importance of tolerance in animals (Raberg, 2014). To date, apart from our studies (Chapter 3, 4), there are only two studies on the genetics of tolerance to infection (Raberg *et al*., 2007; Hayward *et al*., 2014).

In this thesis, we showed that there is genetic variation in farm animals in response to infections. We estimated the genetic variation in resistance, tolerance and resilience to two economically important infectious diseases in farm animals: PRRS in pigs and gastrointestinal nematode infection in sheep. We showed the possibility of selective breeding for both diseases. We also showed that there might be a trade-off between resistance and tolerance, which requires careful consideration when including these traits in breeding programs. For including resistance, tolerance, and resilience in breeding programs, nevertheless, there are several matters to be considered. In the following paragraphs, I will discuss situations in which breeding for any of these three traits is beneficial.

Resistance

Resistance is the ability of a host to prevent pathogens from entry, restrain the replication of the invading pathogen or control the life cycle of the invading pathogen. Resistance is conventionally measured as the inverse of pathogen burden in the animal. The genetic effect on pathogen burden, therefore, is the genetic effect on resistance. The main advantage of resistance mechanism is that a resistant animal is well able to combat the disease quickly and therefore does not

spread the infective agent in the population. Consequently, resistance might lead to clearance of the infection in the population.

We showed the possibility of selective breeding for resistance to nematode infection in sheep (Chapter 3) and resistance to PRRS in pigs (Chapter 5, Lough *et al.*, manuscript in preparation). Selection for resistance, however, is not a novel strategy as nature has been using it for a long time. There are examples of naturally resistant livestock species to different types of infection (Bishop *et al*., 2002). Breeding for resistance has the major advantage that the resistant animal will not spread the infection. Breeding for resistance, therefore, could be especially a promising approach for controlling diseases with high transmission rates. In case of diseases with high transmission rate, selective breeding for resistance could help stopping the infection to spread to other animals or populations. Measuring resistance, however, requires records of pathogen burden on the individual level, which is not always possible in practice. Genomic selection could be a solution for lack of records on individual pathogen burden. By recording the pathogen burden of the genotyped animals in a reference population, the genomic breeding values for resistance in the target population can be estimated.

Resistance mechanisms are known to be pathogen specific meaning that resistance to one type of pathogen may not work for another type. Resistance mechanism may put a selection pressure on the pathogens to overcome the resistance mechanism. The selection pressure is especially high in small pathogens with short generation interval, such as viruses, with high mutation rates and short generation intervals. Before including resistance into the breeding programs, therefore, animal breeders should be convinced that including resistance in the selection index would add considerably to the overall value of genetic progress as compared to the breeding programs without resistance in it.

Tolerance

Tolerance is the ability of a host to show minimum decrease in performance despite a certain amount of pathogen in body. Tolerance involves the mechanisms that minimize the damages caused by infection. The immunological aspects of tolerance are less known.

We showed the possibility of selective breeding for tolerance to nematode infection in sheep (Chapter 3) and tolerance to PRRS in pigs (Chapter 5, Lough *et al.*, manuscript in preparation). There are also evidences for naturally selected tolerant livestock species (Baker *et al*., 2004). Unlike resistance, tolerance is a non-

specific mechanism against diseases. Selecting animals for tolerance can improve the responses of the farms to a wide range of diseases. Tolerance does not affect pathogen burden *per se* and therefore is presumed to put no selection pressure on the pathogen. If that is true, tolerance could be a promising approach to control diseases in the case when there is a risk that pathogen evolves against resistance by the animal. At the moment there is no scientific evidence for this thought.

A tolerant animal might still spread the pathogen. Selection for tolerance, therefore, should take into account the type of the disease. In case of highly transmissible infections and zoonotic diseases, stopping the infection transmission has the highest priority. In that case breeding for tolerance is not an option because it does not stop the spread of the disease. Furthermore, improving the average tolerance may increase the transmission rate of the pathogen, which is a serious threat for nearby populations or newcomer hosts that are not tolerant. Further studies are needed on epidemiological aspects of tolerance and its effects on the pathogen. Therefore, animal breeders have to make sure that including tolerance in the breeding program has obvious advantages compared to a breeding program without tolerance in it.

It is clear that breeding for resistance alone or tolerance in farm animals have advantages and disadvantages. Obviously, simultaneous improvement of resistance and tolerance in farm animals would be highly beneficial because it will improve the health status of the farms both in terms of controlling the infection and minimizing the symptoms. Including both resistance and tolerance into the breeding programs, though, has to consider the possible trade-off between these two traits (see Chapter 3 and 4).

Resilience

In field studies response to infections is mostly referred to as resilience. Resilience is the ability of a host to keep up performance during an outbreak. The difference between resilience and tolerance is that for resilience the pathogen burden in the animal's body is not known while for tolerance it is known. In chapter 2, we studied the variation among sows in resilience to PRRS and showed that selective breeding for sows which show minimal change in their performance during the PRRS outbreaks is possible.

Resilience is a more practical way of measuring response to infections, compared with resistance and tolerance. The main advantage it that measuring resilience does not require the records of the individual pathogen burden. Because the individual pathogen burdens are not recorded, resilience does not distinguish

between resistance and tolerance. A resilient animal, therefore, could be resistant, tolerant, or both. In chapter 4, we showed that resilience is genetically correlated with both resistance and tolerance and that breeding for resilience will improve both resistance and tolerance to infections. This would allow selection of robust animals irrespective of resistance or tolerance.

Breeding for resilience, is a more pragmatic approach in farm animals where robustness and production life of animals is an important trait in the breeding program. Applying RRM models to study resilience would allow estimation of breeding values of animals for different levels of disease severity. Although breeding for resilience is a pragmatic approach, it does not differentiate between resistance and tolerance and generally the realized selection responses in resistance and tolerance are lower compared to when having pathogen burden measured for each animal (Chapter 3).

To conclude, we showed that selective breeding for resistance, tolerance, and resilience to disease is possible in farm animals. Tolerance to infection has more scientific interests, whereas, resilience to infections have more practical interests and is much easier to apply in breeding programs.

6.7 References

Albers, G. A. A. et al. 1987. The genetics of resistance and resilience to Haemonchus-contortus infection in young Merino sheep. Int. J. Parasitol. 17: 1355-1363.

Baker, R. L., J. M. Mugambi, J. O. Audho, A. B. Carles, and W. Thorpe. 2004. Genotype by environment interactions for productivity and resistance to gastro-intestinal nematode parasites in Red Maasai and Dorper sheep. Anim. Sci. 79: 343-353.

Bermingham, M. L. et al. 2014. Genome-wide association study identifies novel loci associated with resistance to bovine tuberculosis. Heredity 112: 543-551.

Bishop, S. C. 2012a. A consideration of resistance and tolerance for ruminant nematode infections. Front. Genet. 3: 168.

Bishop, S. C. 2012b. Possibilities to breed for resistance to nematode parasite infections in small ruminants in tropical production systems. Animal 6: 741-747.

Bishop, S. C., M. C. M. De Jong, and D. Gray. 2002. Opportunities for incorporating genetics elements into the management of farm animal diseases. FAO Commission on Genetic Resources for Food and Agriculture. Background Study Paper No. 18. Available:

ftp://ftp.fao.org/docrep/fao/meeting/015/aj629e.pdf. Accessed Sep 22, 2003.

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: 190-198.

Bloemhof, S., A. Kause, E. F. Knol, J. A. M. Van Arendonk, and I. Misztal. 2012. Heat stress effects on farrowing rate in sows: Genetic parameter estimation using within-line and crossbred models. J. Anim. Sci. 90: 2109-2119.

Calus, M. P., P. Bijma, and R. F. Veerkamp. 2004. Effects of data structure on the estimation of covariance functions to describe genotype by environment interactions in a reaction norm model. Genet. Sel. Evol. 36: 489-507.

Calus, M. P. L., A. F. Groen, and G. de Jong. 2002. Genotype x environment interaction for protein yield in Dutch dairy cattle as quantified by different models. J. Dairy. Sci. 85: 3115-3123.

Calus, M. P. L., and R. F. Veerkamp. 2003. Estimation of environmental sensitivity of genetic merit for milk production traits using a random regression model. J. Dairy Sci. 86: 3756-3764.

Cheng, H. et al. 2008. Using integrative genomics to elucidate genetic resistance to Marek's disease in chickens. In 'Animal genomics for animal health. Developments in biologicals. Vol. 132'. (Eds M-H Pinard, C Gay, P-P Pastoret, B Dodet) pp. 365–372. (Karger: Basel, Switzerland).

de Jong, G., and P. Bijma. 2002. Selection and phenotypic plasticity in evolutionary biology and animal breeding. Livest. Prod. Sci. 78 (2002) 195–214.

Doeschl-Wilson, A. B., B. Villanueva, and I. Kyriazakis. 2012. The first step towards genetic selection for host tolerance to infectious pathogens: Obtaining the tolerance phenotype through group estimates. Front. Genet. 3.

Falconer, D. S. 1952. The problem of environment and selection. Am. Nat. 86: 293-298.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Longman, Harlow.

Freitas, M. S., I. Misztal, J. Bohmanova, and J. West. 2006. Utility of on- and off-farm weather records for studies in genetics of heat tolerance. Livest. Sci. 105: 223-228.

Gjerde, B., J. Odegard, and I. Thorland. 2011. Estimates of genetic variation in the susceptibility of Atlantic salmon (Salmo salar) to the salmon louse Lepeophtheirus salmonis. Aquaculture 314: 66-72.

Hayward, A. D. et al. 2014. Natural selection on individual variation in tolerance of gastrointestinal nematode infection. Plos Biol. 12.

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. Livest. Prod. Sci. 64: 95-106.

Herrero-Medrano, J. M. et al. 2015. Estimation of genetic parameters and breeding values across challenged environments to select for robust pigs. J. Anim. Sci. 93: 1494-1502.

Huisman, A. E., and J. A. M. van Arendonk. 2004. Genetic parameters for daily feed intake patterns of growing Dutch Landrace gilts. Livest. Prod. Sci. 87: 221-228.

Kause, A. 2011. Genetic analysis of tolerance to infections using random regressions: a simulation study. Genet. Res. (Camb) 93: 291-302.

Kause, A., and J. Ødegård. 2012. The genetic analysis of tolerance to infections: a review. Front. Genet. 3: 262.

Kirkpatrick, M., D. Lofsvold, and M. Bulmer. 1990. Analysis of the inheritance, selection and evolution of growth trajectories. Genetics 124: 979-993.

Knap, P. W., and G. Su. 2008. Genotype by environment interaction for litter size in pigs as quantified by reaction norms analysis. Animal 2: 1742-1747.

Kolmodin, R., E. Strandberg, P. Madsen, J. Jensen, and H. Jorjani. 2002. Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. Acta Agric. Scand. A Anim. Sci. 52: 11-24.

Kranis, A., G. Su, D. Sorensen, and J. A. Woolliams. 2007. The application of random regression models in the genetic analysis of monthly egg production in turkeys and a comparison with alternative longitudinal models. Poult. Sci. 86: 470-475.

Kuukka-Anttila, H., N. Peuhkuri, I. Kolari, T. Paananen, and A. Kause. 2010. Quantitative genetic architecture of parasite-induced cataract in rainbow trout, Oncorhynchus mykiss. Heredity 104: 20-27.

Lillehammer, M., B. J. Hayes, T. H. E. Meuwissen, and M. E. Goddard. 2009. Gene by environment interactions for production traits in Australian dairy cattle. J. Dairy. Sci. 92: 4008-4017.

Morris, C. A. et al. 2010. Selecting for resilience in Romney sheep under nematode parasite challenge, 1994–2007. New Zeal. J. Agr. Res. 53: 245-261.

Raberg, L. 2014. How to live with the enemy: understanding tolerance to parasites. Plos Biol. 12.

Raberg, L., D. Sim, and A. F. Read. 2007. Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. Science 318: 812-814.

Silva, F. F. et al. 2014. Sire evaluation for total number born in pigs using a genomic reaction norms approach. J. Anim. Sci. 92: 3825-3834.

Streit, M., F. Reinhardt, G. Thaller, and J. Bennewitz. 2012. Reaction norms and genotype-by-environment interaction in the German Holstein dairy cattle. J. Anim. Breed. Genet. 129: 380-389.

Su, G. et al. 2006. Bayesian analysis of the linear reaction norm model with unknown covariates. J. Anim. Sci. 84: 1651-1657.

Wolc, A. et al. 2013. Genome-wide association study for Marek's disease mortality in layer chickens. Avian Dis. 57: 395-400.

Summary

Summary

Infectious diseases in farm animals are of major concern because of animal welfare, production costs, and public health. Farms undergo huge economic losses due to infectious disease. The costs of infections in farm animals are mainly due to production losses, treatment of infected animals, and disease control strategies. Control strategies, however, are not always successful. Selective breeding for the animals that can mount a defence against infection could therefore be a promising approach. Defensive ability of an animal has two main mechanisms: resistance (ability to control the pathogen burden) and tolerance (ability to maintain performance when pathogen burden increases). When it is difficult to distinguish between resistance and tolerance, defensive ability is measured as resilience that is the ability to maintain performance during a disease outbreak regardless of pathogen burden. Studies have focused on the genetics of resistance and resilience with little known about the genetics of tolerance and its relationship with resistance and resilience. The objectives of this thesis were to: 1) estimate the genetic variation in resistance, tolerance, and resilience to infection in order to assess the amenability of these traits for selective breeding in farm animals, 2) estimate the genetic correlation between resistance, tolerance and resilience and 3) detect genomic regions associated with resistance, tolerance, and resilience.

In chapter 2, we studied the variation among sows in response to porcine reproductive and respiratory syndrome (PRRS). First a statistical method was developed to detect PRRS outbreaks based on reproduction records of sows. The method showed a high sensitivity (78%) for disease phases. Then the variation of sows in response to PRRS was quantified using 2 models on the traits number of piglets born alive (NBA) and number of piglets born dead (LOSS): 1) bivariate model considering the trait in healthy and disease phases as different traits, and 2) reaction norm model modelling the response of sows as a linear regression of the trait on herd-year-week estimates of NBA. Trait correlations between healthy and disease phases deviated from unity (0.57±0.13 – 0.87±0.18). The repeatabilities ranged from 0.07±0.027 to 0.16±0.005. The reaction norm model had higher predictive ability in disease phase compared to the bivariate model.

In chapter 3 we studied 1) the genetic variation in resistance and tolerance of sheep to gastrointestinal nematode infection and 2) the genetic correlation between resistance and tolerance. Sire models on faecal nematode egg count (FEC), IgA, and pepsinogen were used to study the genetic variation in resistance. Heritability for resistance traits ranged from 0.19±0.10 to 0.59±0.20. A random regression model was used to study the reaction norm of sheep body weight on

FEC as an estimate of tolerance to nematode infection. We observed a significant genetic variance in tolerance ($P<0.05$). Finally a bivariate model was used to study the genetic correlation between resistance and tolerance. We observed a negative genetic correlation (-0.63±0.25) between resistance and tolerance.

In chapter 4, we studied the response to selection in resistance and tolerance when using estimated breeding values for resilience. We used Monte Carlo simulation to generate 100 half-sib families with known breeding values for resistance (pathogen burden) and tolerance. We used selection index theory to predict response to selection for resistance and tolerance: 1) when pathogen burden is known and selection is based on true breeding values for resistance and tolerance and 2) when pathogen burden is unknown and selection is based on estimated breeding values for resilience. Using EBV for resilience in absence of records for pathogen burden resulted in favourable responses in resistance and tolerance to infections, with more emphasis on tolerance than on resistance. However, more genetic gain in resistance and tolerance could be achieved when pathogen burden was known.

In chapter 5 we studied genomics regions associated with resistance, resilience, and tolerance to PRRS. Resistance was modelled as sire effect on area under the PRRS viremia curve up to 14 days post infection (AUC14). Resilience was modelled as sire effects on daily growth of pigs up to 28 days post infection (ADG28). Tolerance was modelled as the sire effect on the regression of ADG28 on AUC14. We identified a major genomics region on chromosome 4 associated with resistance and resilience to PRRS. We also identified genomics regions on chromosome 1 associated with tolerance to PRRS.

In the general discussion (chapter 6) I discussed: 1) response to infection as a special case of genotype by environment interaction, 2) random regression model as a statistical tool for studying response to disease, 3) advantages and requirements of random regression models, and 4) selective breeding of farm animals for resistance, tolerance, and resilience to infections. I concluded that random regression is a powerful approach to estimate response to infection in animals. If the adequate amount of data is available random regression model could estimate breeding values of animals more accurately compared to other models. I also concluded that before including resistance and tolerance into breeding programs, breeders should make sure about the added values of including these traits on genetic progress. Selective breeding for resilience could be a pragmatic approach to simultaneously improve resistance and tolerance.

Publication List

Peered reviewed papers

Herrero-Medrano, J. M., P. K. Mathur, J. t. Napel, H. Rashidi, P. Alexandri, E. F. Knol, and H. A. Mulder. 2015. Estimation of genetic parameters and breeding values across challenged environments to select for robust pigs. J Anim Sci 93(4):1494-1502

Mathur, P. K., J. M. Herrero-Medrano, P. Alexandri, E. F. Knol, J. ten Napel, H. Rashidi, and H. A. Mulder. 2014. Estimating challenge load due to disease outbreaks and other challenges using reproduction records of sows. J Anim Sci 92(12):5374-5381.

Rashidi, H., H. A. Mulder, P. K. Mathur, J. A. M. van Arendonk, and E. F. Knol. Variation among sows in response to porcine reproductive and respiratory syndrome. Journal of Animal Science 92: 95-105.

Manuscripts in preparation

Lough, G., H. Rashidi, I. Kyriazakis, N. Deeb, J. C. M. Dekkers, A. Kause, P. K. Mathur, H. A. Mulder, A. Doeschl-Wilson. Identifying Genetic Variances in Resistance and Tolerance to Porcine Reproductive & Respiratory Syndrome. *In preparation*.

Rashidi, H, M. S. Lopes, G. Lough, A. B. Doeschl-Wilson, J. C. M. Dekkers, P. Mathur, H. A. Mulder. 2015. Identification of genomic regions associated with resistance and tolerance to porcine reproductive and respiratory syndrome. *In preparation.*

Rashidi, H., H. A. Mulder, L. Matthews, J. A. M. van Arendonk, M. J. Stear. A trade-off between resistance and tolerance to nematode infection in sheep. *In preparation.*

Rashidi, H., H. A. Mulder, J. A. M. van Arendonk. An estimation of resistance and tolerance to infection when individual pathogen burdens are not recorded: a simulation study. *In preparation.*

Sevillano, C. A., H. Mulder, H. Rashidi, E.F. Knol. Estimation of variance components for farrowing rate in sows in response to day-length and temperature. *Submitted*.

Conference proceedings

Rashidi, H., H. A. Mulder, L. Matthews, J. A. M. van Arendonk, M. J. Stear. Estimating the genetic correlation between resistance and tolerance in sheep. 66th EAAP, Warsaw, Poland, 2015.

Rashidi, H., H. A. Mulder, L. Matthews, J. A. M. van Arendonk, M. J. Stear. Trade-off between resistance and tolerance in sheep. 10th WCGALP, Vancouver, Canada, 2015.

Rashidi, H., H. A. Mulder, L. Matthews, J. A. M. van Arendonk, M. J. Stear. Estimating the trade-off between resistance and tolerance. NematodeSystemHealth final conference, Glasgow, Scotland, 2014.

Rashidi, H., H. A. Mulder, P. K. Mathur, J. A. M. van Arendonk, and E. F. Knol. Variation among sows in resilience to porcine reproductive and respiratory syndrome. WIAS science day, Wageningen, The Netherlands, 2014.

Rashidi, H., H. A. Mulder, P. K. Mathur, J. A. M. van Arendonk, and E. F. Knol. Between-sow variation in tolerance in response to porcine reproductive and respiratory syndrome. 63rd EAAP, Bratislava, Slovakia, 2012.

Training and Supervision Plan

Training and Supervision Plan

The Basic Package (3 ECTS)	
WIAS Introduction Course	2011
Course on philosophy of science and/or ethics	2011
Scientific Exposure (10.7 ECTS)	
International conferences	
63rd EAAP Annual Meeting, Bratislava, Slovak Republic	2012
NematodeSystemHealth final conference, Glasgow, Scotland	2014
10th WCGALP, Vancouver, Canada	2014
66th EAAP Annual Meeting, Warsaw, Poland	2015
Seminars and workshops	
WIAS science day	2012
WIAS science day	2013
WIAS science day	2014
WIAS science day	2015
Presentations	
63rd EAAP Annual Meeting, Bratislava, Slovak Republic (oral)	2012
WIAS Science day (oral)	2014
NematodeSystemHealth final conference, Glasgow, Scotland (oral)	2014
10th WCGALP, Vancouver, Canada (oral)	2014
66th EAAP Annual Meeting, Warsaw, Poland (oral)	2015
In-Depth Studies (11.6 ECTS)	
Disciplinary and interdisciplinary courses	
Quantitative genetics with integration of genomic selection (NOVA course, Mustiala, Finland)	2011
Mixed models in quantitative genetics (EISG, Edinburgh, Scotland)	2012
Identity by descent approaches to genomic analyses of genetic traits (Wageningen, the Netherlands)	2012

Advanced methods and algorithms in animal breeding with focus on genomic selection (Wageningen, the Netherlands)	2012
Introduction to theory and implementation of genomic selection (Wageningen, the Netherlands)	2014
PhD students' discussion groups	
Weekly meeting of Quantitative Genetics Discussion group (QDG)	2011-2014
Monthly meeting of Animal immunology and health discussion group	2015
Professional Skills Support Courses (3.8 ECTS)	
Course Techniques for Scientific Writing and Presenting a Scientific Paper	2012
Career assessment	2015
Mobilising your scientific network	2015
Effective behaviour in your professional surroundings	2015
Research Skills Training (2.9 ECTS)	
Preparing own PhD research proposal	2011
Getting Started with ASReml	2012
External training period, University of Glasgow (3 months)	2013
Didactic Skills Training (5.6 ECTS)	
Supervising practicals and excursions	
Student assistance in the course Animal breeding and genetics	2012
Student assistance in the course Animal breeding and genetics	2013
Supervising theses	
Daily supervision of an MSc major thesis	2014
Education and Training Total	**43 ECTS**

Acknowledgement

When I started the MSc of Animal Breeding and Genetics in 2008 in the old Zodiac, I thought that I would be looking for a job outside the academia after 2 years. However, the MSc thesis that I did under the supervision of Piter, along with his nice support and complements, encouraged me to start the PhD in 2011. First of all, I have to thank Piter for his trust and providing the opportunity to do a PhD.

Han, I am proud that I was the first PhD candidate whose 4-year project was fully supervised by you. Thanks for all the support and friendly discussions. I was always welcomed to your office. I must admit you are a wonderful supervisor with a complex brain ☺. I am always fascinated by your ability in inventing brilliant ideas and solutions for problems. You are incredible at remembering literature that contains details of whatever small topic in quantitative genetics. How do you do that?! I learned a lot about G×E and the statistical modeling related to that. You always made me bringing my thoughts to a higher level. I learned a lot from you and I will apply all in my future carriers. I am deeply happy and grateful that I had you as my supervisor. Thank you, thank you, thank you!

Johan, I always imagined you as a conductor whose orchestra is always in harmony. I was always amazed, given your busy schedule, how you managed to go into the details of each PhD project and get actively involved in them. I felt your support during my project, which encouraged me to overcome the challenges. It was a good feeling having such a nice promotor. Thanks for your punctual criticisms and complements. Also thanks for your friendly guidance and supports for finding a job. I would like to wish you a lot of success in your new carrier.

Pramod, I must admit at the beginning of my PhD that I had to come to Topigs, I felt a little bit lost. Your openness to discuss issues helped me a lot to find my way. The challenge was to bring the scientific and practical part of my PhD together; your supervision was greatly helpful connecting these two parts. Our “health room” was the best place to discuss modeling and the issues related to the diseases. I enjoyed and learned a lot during our health meetings. I proudly look back at my Topigs visits and working with you.

Egbert, thanks for your support from the beginning onwards. You always helped me to think about the statistical modeling in a practical way. I will always apply your golden mother-in-law-rule for explaining complex stuff. Thanks for the nice encouragements that you gave me from the beginning.

My dear friends at Topigs Norsvin, it was a great pleasure to meet you almost every week. I always enjoyed my time at Topigs Norsvin.

Mike, it was a great pleasure to be part of the NematodeSystemHealth project and have the opportunity to come to Glasgow and work with you. That period was one of the most fruitful periods of my PhD life. I felt very welcome in your group and I was very well received by you. Thanks for all the supports and the nice social events that I had with you and your group. I would also like to express my deepest gratitude to the friendly people at your lab. I miss them all.

Ada, Monique and Lisette, thank you very much for your practical supports. Ada, you have been amazingly kind and supportive to me. Whenever I entered the secretariat office I faced your kind smile and willingness to help. Thank you for taking care of all administrations related to my project. Monique, I came to you several times with an issue and you solved them very quickly. Your speed at getting things done is amazing. I still remember your support at the beginning of my PhD. Lisette, your willingness to help is incredible. I have a lot to be thankful for when it comes to you. Throughout my PhD you helped a lot with all the travel expenses, hotels, and tickets. I am specifically very thankful for your support in preparing my thesis. You are always responsible and caring for the PhD issues as if they are yours.

My dear family at ABGC, I don't mention names because I don't want to miss anyone. Spending more than 6 years at ABGC has made me so attached to the group and leaving the group reminds me the first time that I left my homeland. I am greatly thankful to each and every one of you for every single minute at ABGC. It was a great feeling to chat with you all at the coffee breaks. I've learned something from every single person in the group. Thank you all for the nice scientific meetings and social events. I am very proud to have many friends from many different countries. Thanks for letting me being part of the ABGC family. I hope I stay in touch with all of you.

At the end I would like to thank my family for all of their supports from the start until the end of my PhD.

Hamed Rashidi

Wageningen, March 2016

Colophon

The research described in this thesis was financially supported by Marie Curie Initial Training Networks (FP7-People-2010-ITN), Topigs Norsvin research center (Beuningen, The Netherlands), and Dutch Ministry of Economic Affairs, Agriculture, and Innovation (Public-private partnership "Breed4Food" code KB-12-006.03-004-ASG-LR and KB-12-006.03-005-ASG-LR).

The cover of this thesis was created by Piotr Sell (www.pcell.it)

This thesis was printed by GVO drukkers & vormgevers B.V. | Ponsen & Looijen Ede, The Netherlands