THE UNIVERSITY
of EDINBURGH

Should we aim for genetic improvement of host resistance or tolerance to infectious disease?

Graham Lough

MSc (University of Glasgow)

BSc (University of Glasgow)

This thesis is presented for the degree of
Doctor of Philosophy

College of Medicine and Veterinary Medicine
The Roslin Institute and Royal (Dick) School of Veterinary Studies
University of Edinburgh

2017

Declaration

I hereby declare that I am the author of this thesis and that I did all the work described herein, unless otherwise specified. This work has not been submitted for any other degree or professional qualification except as specified. I have read and understood The University of Edinburgh guidelines on Plagiarism and declare that this written dissertation is all my own work except where I indicate otherwise by proper use of quotes and references. This thesis is an account of work conducted by me whilst studying for the degree of Doctor of Philosophy at the University of Edinburgh.

Graham Lough

Date: 25/08/17

This thesis is dedicated to the animals who gave their lives for this research.

Acknowledgements

To my first supervisor, the ever-wonderful Andrea Doeschl-Wilson: thank you for the truly above-and-beyond approach to managing me. You've been a guiding beacon through the last 4 years of torment; a passionate and inspirational scientist and individual; a helping hand to drag me up the steep learning curve; and a deeply understanding, caring and excellent support through difficult times. You're like a boss, sister and friend rolled into one. After every meeting we had together, I would always come out in a good mood, ready to take on every challenge and achieve every goal. That really has been the most wonderful thing about you as a supervisor. Sorry for gushing. I didn't come from a genetic, statistical or mathematical background. The fact that I've learned anything at all is a real testament to your guidance. Thank you.

To my second supervisor, Ilias Kyriazakis: thank you for your time, guidance and insight into the ups and downs of life as an academic. Your continued interest in my progress, work and well-being was always a motivation, even if my presentations sometimes put you to sleep. You have been a wonderful individual to have on board in the project, and a huge thank you for the support!

Thesis examination committee: To Johann Detilleux, Kenny Baillie and Andreas Lengeling, thank you for finding time to come together and watch me condense and justify 4 years of work in 3 hours.

PIC: So long, and thanks for all the funding! I fully appreciate the numerous opportunities I had to visit the office and do some work in sunny Hendersonville, TN, a welcome break from the grey skies of Scotland, even though country music really is

truly awful. I learned a lot from these visits, and enjoyed the unlimited free coffee very much, which I dutifully abused. Thanks to Selma Forni and Joseph Deeb for their input into my project. I hope you are both well, wherever you are!

Iowa State University: Though I will never wear red and gold and shout "go Cyclones" and talk about agriculture like it's humanity's last hope, I enjoyed my stay. Thank you to Jack Dekkers and Chris Tuggle for giving me a scholarship and hosting me at ISU for 2 months. I learned a lot about the animal breeding side of things. And the depth of knowledge of the grad students there pretty much gave me a panic attack. A testament to your teaching. Andrew Hess needs a special thanking in particular. You took so much interest and offered me a lot of wisdom and guidance on my project, helping me decide which route to take, and of course, for cooking me a veggie burger on the grill, much to your disgust. My palette for beer learned more in that 2 months than I have learned over my entire PhD. I also learned that glass bottles don't melt on a fire. Who knew?! Mel Hess, thank you for taking time out for helping me with coding problem solving, and sharing some Long Islands with me on the occasional Thursday. Also, a shout out to my bromance Nick Serão and his wonderful family for hosting me on my first visit to ISU, even though his English no so good. My guitar and your bass are meant to be together. Heart.

Wageningen: Thank you to Han Mulder for inviting me out to work on my tolerance paper in Wageningen and providing excellent and timely feedback and advice. I thoroughly enjoyed my time there. And to Hamed Rashidi: You taught me a lot: how to use AsREML, how to do tables in Excel, and how to ride a bike in the dark after a couple of beers. I thoroughly enjoyed every moment we spent together, no matter how

frustrating the work, and enjoyed recalling with you the “endeavours” I had been up to on my trip.

Family: Thank you to my wonderful family for the support over the last 4 years. Particularly the latter two. To my mother and sister, your emotional support over a very difficult time in life was greatly appreciated. To my ol’ man, your ongoing interest in my work and, dare I say, a wee twinkle of pride in your eye always kept me going. I’m lucky to have someone who understood the whole process, and who took such interest in my progress.

Louise: Thank you to Louise for your support in the early years, and your understanding in the later years. You got a trip to Tennessee and Canada out of it, so it’s not all bad, right? And going back to Tennessee didn’t feel the same without you.

Kayleigh: You have been a wonderful support in the last 18 months and a creepy joy to be around. Seeing the red car pull up in the rain is the happiest feeling one can have whilst trudging through the PhD.

Band: To the boys in my band, Goodbye Blue Monday, (look us up on Facebook, voyeuristic readers), thank you for helping me spill out my anxieties in an artful form. You helped me cope with the shame of mental illness. Particularly Jack, my brother. What you lack in talent you make up for in heart. I can’t think of anything I’d rather do with life than scream my lungs out with you. Better to fall together rather than fall apart, eh? Glad we got a record deal before I finished, so I could boast about it in this thesis.

Office mates: Kate, Andrew and Beanz: Thanks for putting up with my handsome appearance, relentless conversations and wonderful sense of humour. It was a “fun” change from open office and the best time I had in the academic misery. I hope the

ugly xmas elf haunts your dreams. And also, cheers to all the friends I've made over the years.

Animals: Being vegan, it's especially difficult to find ethical grounds to justify your work, and you learn, as a scientist, to see a lot of grey areas. But, I have never thought of any life given as a number or a test statistic. Perhaps one day, the scientific community will have a paradigm shift and put more funding and resources into finding alternatives. Until then, simulation > experimentation.

List of Publications

Research articles (peer-reviewed)

G. Lough, I. Kyriazakis, S. Bergmann, A. Lengeling, A. Doeschl-Wilson. (2015) Health Trajectories reveal the dynamic contributions of host genetic resistance and tolerance to infection outcome. *Proc. R. Soc. B* **282** (based on chapter 4).

G. Lough, H. Rashidi, I. Kyriazakis, J.C.M. Dekkers, A. Hess, M. Hess, N. Deeb, A. Kause, J.K. Lunney, R.R.R. Rowland, H.A. Mulder, A. Doeschl-Wilson. (2017) Use of multi-trait and random regression models to identify genetic variation in tolerance to Porcine Reproductive and Respiratory Syndrome Virus. *Genetics Selection Evolution* **49:37**. (Chapter 1).

G. Lough, A. Hess, I. Kyriazakis, J.C.M. Dekkers, M. Hess, H. Rashidi, A. Kause, J.K. Lunney, R.R.R. Rowland, H.A. Mulder, A. Doeschl-Wilson. (2017) Using repeated measures to analyse the relationship between resistance and tolerance of pigs to Porcine Reproductive and Respiratory Syndrome Virus over different stages of infection. Manuscript in preparation. (Based on chapters 2 and 3).

Conference Proceedings (peer-reviewed)

G. Lough, I Kyriazakis, A. Lengeling, S. Bergmann, A. Doeschl-Wilson. (2013) Selecting for improved host tolerance: an alternative strategy for disease control? *BSAS Annual Conference 2013 Proceedings.*

G. Lough, I. Kyriazakis, S. Forni, A. Doeschl-Wilson. (2014) Dynamic and Genetic Signatures of Resistance and Tolerance of pigs to PRRS. *Proceedings, 10th World Congress of Genetics Applied to Livestock Production.*

G. Lough, I. Kyriazakis, J.C.M. Dekkers, J.K. Lunney, R.R.R. Rowland, A. Doeschl-Wilson. (2014) Novel phenotypes for capturing genetic variation in resistance and tolerance of pigs to PRRS. *Proceedings, North American PRRS Symposium.*

G. Lough, A. Hess, H. Rashidi, I. Kyriazakis, J.C.M. Dekkers, M. Hess, A. Kause, J.K. Lunney, R.R.R. Rowland, H.A. Mulder, A. Doeschl-Wilson. (2016) Can we select for both resistance and tolerance of pigs to PRRS? *EAAP Conference Proceedings.*

Other

A. Doeschl-Wilson and **G. Lough**. (2015) Inferring genetic resilience of animals to infectious
pathogens – opportunities and pitfalls. *Breeding Focus Workshop: Building Resilience book chapter.*

Table of contents

Abstract

A host can adopt two strategies when facing infection: *resistance*, where host immune responses prevent or reduce pathogen replication; or *tolerance*, which refers to all mechanisms that reduce the impact of the infection on host health or performance. Both strategies may be under host genetic control, and could thus be targeted for genetic improvement. Although there is ample evidence of genetic variation in resistance to infection, there is limited evidence to suggest that individuals also differ genetically in tolerance. Furthermore, although resistance and tolerance are typically considered as alternative host defense mechanisms, relatively little is known about the genetic relationship between them and how they change together over time and jointly determine infection outcome. In this thesis, two datasets from experimental challenge infection experiments were considered for investigating tolerance genetics: Porcine Reproductive & Respiratory Syndrome (PRRS), an endemic viral disease which causes loss of growth and mortality in growing pigs; and *Listeria monoctyogenes* (Lm), a bacterium which causes food-borne infections in mammals. The two datasets differed substantially in size and genetic structure; the PRRS dataset consists of thousands of records from outbred commercial pig populations, whereas the Listeria dataset comprises much fewer records from genetically diverse highly inbred strains of a mice as a model species. The aims of this thesis were to: 1) Identify if genetic variation in host tolerance to infection exists, with case studies in PRRS and listeria, using conventional reaction-norm methodology; 2) Identify if host tolerance, along with resistance, changes longitudinally as infection progresses; 3) Identify whether the WUR genotype is associated with tolerance slope; 4) Analyse the dynamic relationship between host performance and pathogen load over the time-course of infection by

examining the relationship at different stages of infection using GWAS; 5) Develop novel trajectory methodology to offer insight into health-infection dynamics, and identify whether there is genetic variation in trajectories; 6) Develop novel trajectory-derived phenotypes that analyse changes in host performance with respect to changes in pathogen load, as an alternative to tolerance, and identify whether genetic variation exists. This study found that conventional reaction-norm methodology is limited to capture genetic variation in tolerance in outbred populations without measures of performance in the absence of infection. However, by utilising repeated longitudinal data on the same dataset, stages of infection (early, mid and late) were defined for each individual, based on host pathogen load. Using these stages of infection, genetic variation in tolerance was identified over all stages of infection and at mid to late stage of infection. Genetic correlation between resistance and tolerance was strong and positive over all stages of infection, and evidence suggested that resistance and tolerance may be under pleiotropic control. Furthermore, this research found that genetic correlations between resistance and growth changed considerably over time, and that individuals who expressed high genetic resistance early in infection tended to grow slower during that time-period, but were more likely to clear the virus by late stage, and thus recover in growth. However, at mid-late stage of infection, those with high virus load also had high growth, indicating potential epidemiological problems with genetic selection of host resilience to infection. Furthermore, genome wide association studies for pathogen load and growth associated with different stages of infection did not identify novel genetic loci associated with these traits than those previously reported for the whole infection period. By adopting conventional methodology, this study found genetic variation in tolerance of genetically diverse

mouse strains to Lm and pigs to PRRS, despite statistical problems. The relationship between resistance and tolerance indicated that both traits should be considered in genetic selection programs. By adopting novel trajectory analysis, this study demonstrated that level of expression of resistance and tolerance changed throughout the experimental infection period and, furthermore, that expression of resistance, followed by tolerance, determined survival of infection. Survivors and non-survivors followed different infection trajectories, which were partially determined by genetics. By deriving novel phenotypes from trajectories that explained changes in growth in relation to change in pathogen load at specific time points, and applying these to the PRRS data, this study did not identify genetic variation in these phenotypes. The genetic signal in these phenotypes may have been masked by the fact that individuals were likely at different stages of infection. In summary, this study has shown that genetic improvement of tolerance, in addition to resistance may be desirable, but could be difficult to achieve in practice due to shortcomings in obtaining accurate and unbiased tolerance estimates based on conventional reaction-norms. Infection trajectories have proven to be a promising tool for achieving an optimally timed balance between resistance and tolerance, but further work is needed to incorporate them in genetic improvement programs.

Lay Summary

Infectious diseases are a major concern for livestock production. Sometimes it is not possible to control spread of a pathogen by conventional methods e.g. vaccination. Genetic selection to improve host response to infection has gained increased attention. When facing infection, a host can adopt two strategies: *resistance*, where the host immune response prevents or reduces pathogen replication; or *tolerance*, where the host reduces the impact of the pathogen on health. To date, it is not clear whether either of these defense strategies, or both together, are important for limiting the impact of infection, and determining survival. Here, we considered two case studies of infections that affect livestock:

1) *Listeria monocytogenes* (Lm), which is a zoonotic, bacterial pathogen that can cause severe food-borne infections in humans and animals using genetically different inbred strains of mice as a model species.
2) Porcine Reproductive & Respiratory Syndrome (PRRS), which is a viral disease that can lead to severe weight loss and mortality in growing pigs.

Individuals are known to differ in resistance in both Lm and PRRS, and that the outcome of infection is partly determined by the host's genetics. However, it is not known whether hosts also vary genetically in tolerance. Our research identified that pigs may exhibit tolerance to PRRS, but that the methodology normally used to define and estimate tolerance were impractical to estimate tolerance of pigs to PRRS in this study, and that large amounts of data are required to estimate tolerance. When using repeated data for each individual, we found that pigs do vary genetically in tolerance to PRRS, and that resistance and tolerance to PRRS were positively related, meaning that the two traits may be controlled by similar genes. Furthermore, we found mice

exhibit different degrees of both resistance and tolerance to Lm based on their genetics. We also examined changes in host health and resistance at different stages of infection. We found that the relationship between health and resistance changed throughout infection, and that pigs that exhibit resistance at early stage of infection could recover health better later in infection. Furthermore, we established a novel method which allowed us to investigate how resistance and tolerance change together over the time-course of infection. For Lm, only mice exhibiting high resistance early in infection survived. We also found that a host can show one of four unique types of response to Lm infection based on both on bacterial load and health outcome over time. These types are partly determined by genetic factors. However, when applying this method to examining how pigs differed in their response to PRRS, we could not determine any genetic basis. Our study provides new insights into the complex relationship between resistance and tolerance and helps to understand what matters to recover from and survive infection challenges, and the need to use new methods to capture this relationship.

Chapter 1

General Introduction

1.1 Infectious disease threat to livestock production

Infectious diseases pose a severe threat to livestock production worldwide. Management challenges and high population density on farms provide an environment in which pathogens can transmit easily between hosts, evolve, and lead to, often severe, epidemics. Pathogens affecting livestock production have been diverse, from viruses and prions to bacteria and parasites which have led to severe economic losses to the livestock industry (with costs up to 20% and 50% of turnover in developed and developing countries respectively), and severe animal welfare implications [1].

In some cases, conventional infection control measures, such as vaccination programmes and culling, have been successful. However, there are some pathogens which cannot be controlled or eliminated, due to e.g. high pathogen mutation rate or lack of vaccine efficacy. The question, then, is how to control the spread of infectious disease on, or across, farms, and, how to mitigate the impact of the infection on animal health and production. One possible solution that has gained attention in recent years is genetic selection for improved host response to infection. Recent advances in genetic and genomic technologies have provided opportunities to implement genetic selection programs to improve host response to infectious disease which, when combined with conventional infection control measures, could provide a strategy to control infectious disease spread in livestock production [2,3].

1.2 Resilience, Resistance & Tolerance

A host can respond to infection in different ways. Generally, *resilience* refers to the ability of a host to cope with stressors, such as infectious challenge [4]. Resilience thus describes how host fitness, health or performance, such as body weight or growth, is affected by being exposed to infection. An individual that performs better than another when exposed to infectious pathogens is therefore more resilient (i.e. performance is less impacted by infection). Improving host resilience is an obvious goal for genetic selection [2]. However, resilience is a composite trait, comprising two underlying strategies that a host can adopt when infected with a pathogen. The first strategy is *resistance*, which is well-defined in the literature, and considered as the ability of the host to prevent or inhibit pathogen replication [5]. Resistance can be viewed as a function of the host immune system, and its involvement in the detection, neutralisation, destruction or removal of pathogens [6]. As such, resistance is often defined as the inverse of within-host pathogen load, i.e. the lower the pathogen load, the more resistant the host (figure 1A) [5].

Another strategy a host can adopt when infected with a pathogen is *tolerance*. Similar to resilience, tolerance refers to the ability of the host to maintain fitness, health or performance in the face of pathogen challenge. However, unlike resilience, tolerance to pathogens is only defined for infected animals and describes this ability in relation to changes in within-host pathogen load [5,7,8]. Similar to resilience, tolerance may be defined as a reaction norm of performance with respect to pathogen load, but where the x-axis for a resilience reaction norm refers to exposure i.e. external pathogen load, the x-axis for a tolerance reaction norm refers to within-host pathogen load [8,9] (see figure 1B & C for illustration, and **1.3 Conventional methodology for estimation of**

genetic variation in tolerance section). As such, tolerance does not impact on the pathogen life-cycle, per se. Rather, tolerance mechanisms work on host repair and damage control. Although the immunological mechanisms underlying tolerance are largely unknown, it has been suggested that tolerance can be related to tissue damage caused by the pathogen (*direct* tolerance) or by the host immune response (*indirect* tolerance) which can be achieved through tissue protection and repair [6,10]. Furthermore, possible mechanisms for a host to express tolerance have been suggested, such as: the ability of the host to minimise damage caused by host inflammatory response, to take into consideration the costs of mounting an immune response to a pathogen, in terms of nutrient resources, and to minimise the cost of damage caused by the pathogen [11]. For example, a host can respond to infection by undergoing anorexia, where infected individuals can increase tolerance to infection by reducing their intake, thus diverting essential nutrients required for digestion to manage infection, assuming the costs of digestion are high [12,13].

The first study to demonstrate genetic variation in tolerance in animals was provided by Raberg et al [5], where mice were shown to differ in both tolerance and resistance to *Plasmodium chabaudi*. However, tolerance terminology has been ambiguous in the literature. For example, ruminants with improved ability to counteract the lifecycle of pathogenic trypanosomes have been referred to as trypano-tolerant ruminants [14]. In this case, tolerance actually referred to resistance mechanisms or encompassed both resistance and tolerance mechanisms. Indeed, many scientists and animal breeders consider resilience and tolerance as interchangeable terms regarding the same trait, as both are concerned with the impact of pathogens on host performance [4]. Assuming individuals are equally exposed to pathogens, measures of performance under

infection are often used as measures of resilience [15]. The crucial difference between resilience and tolerance, that tolerance quantitatively accounts for pathogen burden, has been largely ignored by the literature. This may explain why few studies to date have explicitly investigated the relationship between resilience and tolerance.

Despite tolerance mechanisms being studied extensively in plants [16–18], and being acknowledged as playing a vital role in ecology and immunology [6,7,11,19], to date, genetic analyses of host response to infection have focused primarily on the mechanisms underlying host resistance [3,20–24], the rationale being that improvement of host resistance mechanisms should, in turn, improve host resilience to infectious disease. This is because hosts that harbor fewer pathogens are expected to be healthier and thus perform better. Estimates and genetic parameters of resistance have been explored, with genetic variation in resistance identified to different infectious diseases and across different livestock species [25,22,26,27]. However, despite interest in improving host resilience through resistance mechanisms, few studies focused on improving host resilience through tolerance mechanisms [28]. It is important to distinguish between resistance and tolerance as selection for one may affect the other, and may have different epidemiological and evolutionary outcomes.

There have been conceptual theories and supporting evidence to suggest an antagonistic relationship between resistance and tolerance [29]; that cost and trade-off associated with both tolerance and resistance would create a negative genetic correlation between the traits [18,30].

In addition to the trade-off theory promoting negative genetic correlation between tolerance and resistance, there has been both discussion and evidence to suggest that tolerance and resistance co-occur together, contributing additively to host fitness [31].

It has been argued that expression of tolerance is conditional upon host resistance [26] and that, in the presence of disease, tolerance and resistance are expressed pleiotropically, with tolerance only expressed when the host experiences damage [9]. Furthermore, from an immunological perspective, resistance may require immunological mechanisms to reduce pathogen load, which may be damaging to the host. In this case, resistance would be related to indirect tolerance (to damage caused by immune response), which would be advantageous to the host over direct tolerance (to damage caused directly by the pathogen), thus there would be an expected positive genetic correlation between resistance and indirect tolerance. The genetic correlation between direct tolerance and resistance would thus depend on the genetic relationship between indirect and direct tolerance [10]. The authors of the former study suggested that resistance and tolerance could not evolve independently, using the rationale that resistance is only beneficial when disease reduces performance, which would not occur if the host expressed tolerance, and that tolerance is only expressed when infection causes damage, which would not occur if the host was resistant.

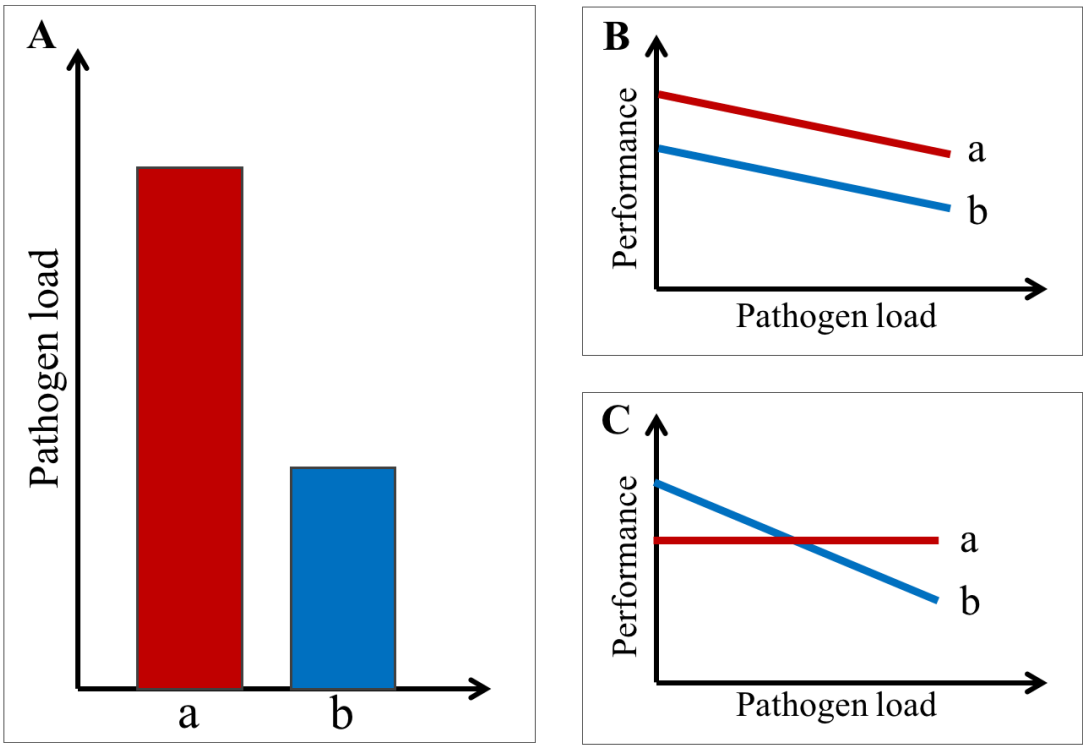
Additionally, there has been support for the argument that resistance and tolerance have alternative evolutionary outcomes and different degree of genetic variation; notably that tolerance will tend to fixation, whereas genetic variation in resistance is more likely to be maintained [32]. It has also been argued that hosts with resistant genotypes to one pathogen may still be susceptible to another, where a tolerant genotype may allow tolerance to multiple pathogens given that tolerance mechanisms would limit the damage caused by pathogens [33]. With that in mind, selecting for host tolerance may be a viable infection control strategy where elimination of a pathogen is unlikely, such as where the pathogen is highly prevalent in a population [26]. Unlike

resistance, tolerance should not, in theory, impose selection pressure on the pathogen, as it does not limit pathogen replication, at least directly [32]. However, tolerance may be driven to fixation in a population, given that it would be a selective advantage to carry the gene [32]. Some studies have identified host genetic variation in resistance, but not tolerance, concluding that lack of genetic variation in tolerance meant that the host had evolved "maximum tolerance" [34].

1.3 Conventional methodology for estimation of genetic variation in tolerance

As stated above, tolerance is typically defined as a reaction norm of performance with respect to pathogen load. Assuming a linear relationship, tolerance can be estimated by regression of host performance against pathogen load (figure 1B/C), where the corresponding regression slope provides a measure of tolerance. Thus, slopes of equal value would indicate equal tolerance between genotypes (figure 1B); a slope value of zero would confer complete tolerance, that is, no performance is lost with increasing pathogen load (figure 1C); and a negative value would then indicate a reduction in performance due to infection. The steeper the slope, then, the less tolerant an individual (figure 1C). A positive value would confer overcompensation, where an increase in host performance would occur, despite being infected with a pathogen. Evidence for genetic variation in tolerance has been provided by showing a significant statistical interaction between family (or genotype, or breed) and infection severity in an ANCOVA [5].

Figure 1. Graphical illustration of A) Resistance, and B/C) Tolerance. Figure 1A) shows 2 genotypes (a and b) infected with a pathogen. Genotype b has a lower pathogen load, and as therefore the more resistant genotype. Figure 1B shows two genotypes (a & b) of equal tolerance, where performance is lost with respect to pathogen load at an equal rate. However, genotype a has higher vigour (performance in absence of infection) than genotype b. Figure 1C shows two genotypes (a & b) that differ in tolerance to infection. Genotype b has a steeper slope then genotype a, and is thus less tolerant than genotype b, suffering greater loss in performance with respect to pathogen load. Despite having lower vigour, genotype a exhibits complete tolerance (slope value of zero), thus suffers no loss in performance related to pathogen load.



For outbred populations with a complex pedigree structure, tolerance slopes of related individuals can be estimated by random regression models, which, when combined with pedigree or genomic information, provide estimates for genetic variance of

tolerance, and for genetic variance in host performance as a function of pathogen load [35]. Such a random regression model can provide additive genetic variance estimates for vigour (intercept – performance in the absence of infection), tolerance (the regression slope), and the covariance between them. However, due to the large amount of data required to obtain unbiased variance estimates for reaction norm slopes [4,35,36], very few studies to date have applied this methodology to gain insight into the genetic basis of disease tolerance in outbred populations [37]. To our knowledge, there is currently no conclusive evidence for genetic variation in tolerance of domestic livestock to infections.

1.4 Novel methodology for estimation of genetic variation in tolerance

Using the reaction norm approach, tolerance can be estimated at the group level [8], but cannot be estimated at the individual level, as estimation of regression slopes requires multiple pairwise measures of the response variables [38]. There are disadvantages of estimating tolerance with the reaction norm methodology. Firstly, the current methodology can only provide unbiased tolerance slope estimates if hosts are assumed to become infected at the same time, or if infection time is known, which will not be the case in farms. Secondly, in animal breeding, there is restricted selection accuracy and response to selection if phenotypes are on the group level [4].

Resistance and tolerance are comprised of many different immunological and pathological processes and are thus likely highly dynamic. As such, expression of host resistance and tolerance, and their relationship are likely to change over the time-course of infection. It has been reported that genetic variation of tolerance and resistance are time-dependent [39], with genetic variation in both tolerance and

resistance identified at acute stage of infection, and genetic variation in only tolerance during recovery [31]. As a potential method of capturing this dynamic relationship and simultaneously improving host resistance and tolerance, 2D infection trajectories that describe how changes in performance change with respect to changes pathogen load throughout the time course of infection, may be able to provide valuable insight. It has been suggested that such infection trajectories could be used to describe changes in health dynamics in disease space [38] and that host resilience can be predicted based on partial trajectory analysis [40]. Doeschl-Wilson et al., introduced the theory of trajectories to the context of animal breeding, suggesting different trajectory classes may exist, depending on the dynamics of host health-pathogen interactions [41].

1.5 Genome Wide Association Studies

Another approach to analysing the relationship between health and pathogen load is to adopt genome wide association studies (GWAS). To identify genomic regions that are associated with phenotypic traits, GWAS have been be adopted in animal breeding [42]. GWAS enables the statistical association of genetic markers with phenotypic traits of interest by linking the marker genotypes of each individual to the phenotype [43–45]. For the pig dataset in this study, Single Nucleotide Polymorphisms (SNPs) were previously used as genetic markers to detect QTL. Currently, high-density SNP chip panels provide genotypes for hundreds of thousands of genetic loci [46], across the whole genome of an individual, in this case, the *Sus scrofa* genome. By identifying genomic regions associated with a phenotypic trait, chromosomal position can be further investigated to identify potential target genes for selection, and the associated biochemical pathways the gene encodes. Numerous studies have used GWAS to

demonstrate genetic basis for traits associated with Porcine Reproductive and Respiratory Syndrome (see ***1.6.1 Case study 1: Porcine Reproductive and Respiratory Syndrome (PRRS)*** section of thesis).

Genomic association studies, such as GWAS, rely on linkage disequilibrium (LD), between the genetic markers and QTL that is, marker and SNP genotypes are not independent of each other. Furthermore, if SNPs are in high LD with each other, it is more appropriate to identify genomic regions comprising several SNPs associated with a phenotypic trait, rather than carrying out the GWAS at the level of individual SNPs. There are different methods that can be adopted to identify genomic regions associated with a trait. GWAS using Bayesian variable selection can be used, which is a method adapted from genomic prediction methodology [47–50].

1.6 Case studies considered in this thesis

To identify genetic variation in tolerance through conventional and novel methodology, two case studies will be considered in this thesis: Porcine Reproductive and Respiratory Syndrome (PRRS) and Listeria. The PRRS data were from an outbred population of pigs, and was used for conventional estimation of tolerance, whereas the Listeria data came from genetically diverse inbred mice, as a model species, which were used to develop and test the novel trajectory methodology.

1.6.1 Case study 1: Porcine Reproductive and Respiratory Syndrome (PRRS)

Porcine Reproductive and Respiratory Syndrome (PRRS) is a viral disease, endemic in pigs worldwide. PRRS causes respiratory problems and considerable reduction in the growth rate of piglets, with estimates ranging between 10% to 20% [51], with

resulting production losses amounting to an annual cost of $493.57 million to the U.S. swine industry [52]**.** PRRS is caused by a single-stranded positive sense RNA virus, which has a high mutation rate causing antigenic and genic shift in both structural and non-structural proteins of the virus, allowing the virus to avoid innate immune response [53]. As such, vaccination has been largely unsuccessful [54], and so host genetic improvement in response to PRRS has gained increased attention [3,55,56]. Large-scale PRRSV challenge studies carried out by the PRRS Host Genetics Consortium (PHGC) have demonstrated considerable genetic variation in resistance of pigs to PRRSV infection, as well as in growth of infected piglets [57,58,55]. Genetic correlations between resistance and growth (resilience) have been shown to be positive and strong [55], indicating that selection for improved resistance is expected to simultaneously improve resilience to infection and vice versa. However, it is currently unknown whether there is genetic variation in tolerance to PRRS. The PRRS Host Genetics Consortium (PHGC) data, which provide simultaneous measures of growth and viral load for over 1,500 pedigreed pigs infected with the same PRRS virus load offer a unique opportunity to estimate genetic parameters for tolerance.

1.6.2 Case study 2: Listeria

Listeria monocytogenes is a zoonotic facultative anaerobe. The natural route of transmission is by epithelial cells in the intestine, from where dissemination is possible and can lead to systemic infection. Animals do not generally show clinical signs when infected with listeria, with 6% carrying listeriae as part of their faecal flora [59]. However, in a small proportion of cases, severe listeriosis can occur, where it may be possible for listeria to cross the blood-brain barrier causing meningitis, or crossing the

placenta, causing abortion in pregnant animals. Theoretically, this points to variation in tolerance, rather than in resistance to this disease. A recent listeria infection challenge experiment collected data from multiple repeated measures for individual host mice from genetically diverse inbred strains. This dataset provided individual longitudinal measures for both performance and infection intensity prior and during infection thus constitute the ideal dataset to test whether the previously proposed infection trajectories provide valid new insight into the genetic and dynamic regulation of host resistance and tolerance.

1.7 Aims of this study

This study aimed to:

1) Identify if genetic variation in host tolerance to infection exists, with case studies in PRRS and listeria, using conventional reaction-norm methodology.
2) Identify if host tolerance, along with resistance, changes longitudinally as infection progresses.
3) Identify whether and to what extent a previously identified genomic region associated with resistance is also associated with tolerance.
4) Analyse the dynamic relationship between host performance and pathogen load over the time-course of infection by examining the relationship at different stages of infection using GWAS.
5) Develop novel trajectory methodology to offer insight into health-infection dynamics, and identify whether there is genetic variation in trajectories.

6) Develop novel trajectory-derived phenotypes that analyse changes in host performance with respect to changes in pathogen load, as an alternative to tolerance, and identify whether genetic variation exists.

1.8 Structure of this thesis

This thesis consists of 5 research chapters: In chapter 2, random regression methodology is applied to the PHGC data to estimate genetic variation in tolerance of pigs to PRRS at both the acute phase of infection (defined as 0 to 21 days post-infection (dpi)) and over a prolonged observation period (defined as 0 to 42 dpi). In chapter 3, the PHGC data are partitioned into stages of infection (early, mid and late stages), based on individual viremia profile characteristics to allow for more biologically relevant models. Random regression models are then applied to each stage of infection to identify if genetic variation in tolerance exists at different stages of infection. Furthermore, by harnessing information on growth and serum virus load for each stage of infection for each individual, a random regression model is applied across all stages of infection to identify if repeated measures could help identification of genetic variation in tolerance. Chapter 3 also assesses the association between the previously identified resistance genotype and tolerance. In chapter 4, the stages of infection defined in chapter 3 are used to analyse the relationship between growth and resistance by examining genetic correlations and heritabilites of the traits throughout different stages of infection. Additionally, a GWAS is adopted to identify genomic regions associated with growth and resistance to PRRSV infection over time. After thorough inspection of the shortcomings of random regression models for estimating genetic parameters for tolerance in chapters 2-3, a novel approach based on previously

described individual infection trajectories is proposed in chapter 5. The methodology is first applied to infection data from a mouse model, as these are less noisy and thus easier to analyse. In line with the previous chapters, a conventional reaction norm approach is first used to identify genetic variation in tolerance of inbred strains of mice to listeria. Subsequently, bacterial infection trajectories of mice are constructed and analysed, in particular with regards to the question whether trajectories are genetically determined and could be used to discriminate the infection paths between animals that eventually succumb or recover from infection. In chapter 6, trajectories are then implemented for the pigs of the PHGC PRRSV infection trials. Finally, in chapter 7, the general discussion section combines the results from each chapter and discusses their implication to animal breeding and future scientific studies.

Chapter 2

Use of multi-trait and random regression models to identify genetic variation in tolerance to Porcine Reproductive and Respiratory Syndrome virus

G. Lough, H. Rashidi, I. Kyriazakis, J. C. M. Dekkers, A. Hess, M. Hess, N. Deeb, A. Kause, J.K. Lunney, R.R.R. Rowland, H.A. Mulder, A. Doeschl-Wilson

Authors' contributions

GL and HR conducted the preliminary statistical analyses and interpretation of results. GL ran extensive simulations and wrote the manuscript. AH and MH aided statistical analyses. RRRR conceived the experimental trials and led the animal infection trials and sample collection. JKL conceived the experimental trials and coordinated the handling, storage, and sample preparation for DNA genotyping. JCMD helped conceive the study and collated the data. IK, JCMD, AK and ND all aided interpretation of results. ADW & HM helped to conceive the study, coordinated and oversaw statistical analysis of the data and contributed to the interpretation of results and writing the manuscript. All co-authors reviewed and contributed to development of the manuscript. All authors read and approved the final manuscript.

2.1 Abstract

A host can adopt two response strategies to infection: resistance (reduce pathogen load) and tolerance (minimize impact of infection on performance). Both strategies may be under genetic control and could thus be targeted for genetic improvement. Although there is evidence that supports a genetic basis for resistance to porcine reproductive and respiratory syndrome (PRRS), it is not known whether pigs also differ genetically in tolerance. We determined to what extent pigs that have been shown to vary genetically in resistance to PRRS also exhibit genetic variation in tolerance. Multi-trait linear mixed models and random regression sire models were fitted to PRRS Host Genetics Consortium data from 1320 weaned pigs (offspring of 54 sires) that were experimentally infected with a virulent strain of PRRS virus to obtain genetic parameter estimates for resistance and tolerance. Resistance was defined as the inverse of within-host viral load (VL) from 0 to 21 (VL_{21}) or 0 to 42 (VL_{42}) days post-infection and tolerance as the slope of the reaction-norm of average daily gain (ADG_{21}, ADG_{42}) on VL_{21} or VL_{42}.

Multi-trait analysis of ADG associated with either low or high VL was not indicative of genetic variation in tolerance. Similarly, random regression models for ADG_{21} and ADG_{42} with a tolerance slope fitted for each sire did not result in a better fit to the data than a model without genetic variation in tolerance. However, the distribution of data around average VL suggested possible confounding between level and slope estimates of the regression lines. Augmenting the data with simulated growth rates of non-infected half-sibs (ADG_0) helped resolve this statistical confounding and indicated that genetic variation in tolerance to PRRS may exist if genetic correlations between ADG_0 and ADG_{21} or ADG_{42} are low to moderate.

Evidence for genetic variation in tolerance of pigs to PRRS was weak when based on data from infected piglets only. However, simulations indicated that genetic variance in tolerance may exist and could be detected if comparable data on uninfected relatives were available. In conclusion, of the two defense strategies, genetics of tolerance is more difficult to elucidate than genetics of resistance.

2.2 Introduction

Infectious challenges in domestic livestock not only raise health and welfare concerns, but also have detrimental effects on livestock production. The impact of infections on an animal's productive performance is controlled by two alternative (albeit not mutually exclusive) host traits that may be amenable to genetic improvement: resistance and tolerance. Resistance is defined as the ability of a host to prevent pathogen entry or inhibit replication of the pathogen, whereas tolerance refers to the ability of a host to limit the impact of infection on health or performance without interfering with the pathogen life-cycle per se [7]. Thus, animals with greater resistance are expected to harbor fewer pathogens that can inflict loss in performance. In contrast, animals with greater tolerance may harbor high within-host pathogen load but are able to prevent or repair the damage of infection on health and performance [5,60]. To date, most efforts to control infectious disease have targeted primarily improvement of host resistance. More recently, the focus has expanded towards boosting host tolerance as an alternative means to counteract the detrimental impact of infection on health and performance [6,28]. However, little is known about the extent to which tolerance is genetically controlled and thus suitable for genetic improvement.

Porcine Reproductive and Respiratory Syndrome (PRRS) is an endemic virus, which causes one of the most devastating swine diseases worldwide. PRRS causes considerable reduction in the growth rate of piglets, with estimates ranging from 10 to 20%, depending on pig breed and virus strain [51], with resulting production losses amounting to an annual cost of $493.57 million to the U.S. swine industry alone [52]. As vaccination has been largely unsuccessful [54], genetic solutions to PRRS have gained increased attention [3,55,56]. Recent large-scale PRRSV challenge studies carried out by the PRRS Host Genetics Consortium (PHGC) have demonstrated considerable genetic variation in resistance of pigs to PRRSV (virus) infection, as well as in weight gain of infected piglets [57,58,55]. Furthermore, genetic correlations between resistance and weight gain have been shown to be positive and strong (ranging from 0.57 to 0.75 for two different PRRSV strains) [55], indicating that selection for improved resistance is expected to simultaneously improve growth under infection and vice versa. However, it is not currently known whether pigs also differ genetically in their tolerance to PRRSV infection, or whether pigs with greater genetic resistance to PRRSV are also genetically more tolerant to PRRS.

Resistance can be measured as the inverse of within-host pathogen load, whereas tolerance is related to the degree to which performance is reduced by infectious pathogens. Tolerance is mathematically defined as a reaction-norm of performance with respect to pathogen load [5,9]. Assuming a linear relationship, reaction-norms can be modelled by linear regression of performance against pathogen load, where the regression slope provides a measure of tolerance (Figure 1). Thus, a slope of zero indicates complete tolerance, while a more negative slope indicates lower tolerance. Statistically significant differences in reaction-norm slopes associated with, e.g.,

different breeds or families are indicative of genetic variation in tolerance. For outbred populations, tolerance slopes for groups of related individuals can be estimated by random regression models, which provide estimates for genetic variance of tolerance and for genetic variance in host performance as a function of pathogen load when combined with pedigree or genomic information [35]. However, due to the large amount of data required to obtain unbiased variance estimates for reaction norm slopes [4,35,36], very few studies have applied this methodology to gain insight into the genetic basis of tolerance in outbred populations [37]. The PRRS Host Genetics Consortium (PHGC) data, which provide simultaneous measures of growth and viral load for over 1,500 pedigreed pigs infected with the same PRRS virus load offer a unique opportunity to estimate genetic parameters for tolerance.

The main aim of this study was to determine whether pigs previously found to differ genetically in resistance to a virulent strain of PRRSV also differ genetically in tolerance. Furthermore, by augmenting the data with simulated data, novel insights into data requirements for accurately estimating genetic variance in tolerance using random regression models were obtained.

2.3 Methods

2.3.1 Infection experiment and data

Data were provided by the PRRS Host Genetics Consortium (PHGC) from 9 different PRRSV challenge trials with an identical infection protocol [55,3], comprising altogether 1569 pigs supplied by various commercial breeding companies, as outlined in Table 1.

Table 1 Animal, pedigree and breed composition of the PHGC trials

Trial	Number of animals	Number of sires	Number of dams	Breed cross
1	174	6	70	LW × LR
2	164	10	72	LW × LR
3	115	7	47	LW × LR
4	191	6	33	Duroc × LW/LR
5	182	10	38	Duroc × LR/LW
6	109	26	53	LR × LR
7	186	6	27	Pietrain × LW/LR
8	158	15	43	Duroc × LW/LR
15	166	11	49	Pietrain × LW

LW: large white breed; LR: landrace breed

All experimental protocols for these trials were approved by the Kansas State University Institutional Animal Care and Use Committee. In each trial, approximately 200 commercial crossbred piglets were transferred from high health farms at weaning age (mean age=26 days, range=17 to 32 days) to a research facility at Kansas State University. The source farms were determined to be free of PRRSV, *Mycoplasma hyopneumoniae*, and swine influenza virus. Pigs were randomly placed in pens of 10-15 individuals. Following a 7-day acclimation period, pigs between 17 and 32 days of age were experimentally infected both intramuscularly and intranasally with 10^5 (TCID50) of NVSL-97-7985, a highly virulent PRRSV isolate [61]. Body weight (BW) and blood samples were collected at 0, 7, 14, 21, 28, 35 and 42 days-post-infection (dpi). Pigs were then euthanized at 42 dpi and ear notches were collected for genotyping. Trials 7 and 8 were terminated at 35 dpi due to facility availability. Estimates for Average Daily Gain (ADG) from 0 dpi until day of measurement were obtained by dividing the difference in body weight between the day of observation and 0 dpi by the corresponding time period. Note that measurements of ADG for these pigs

prior to infection were not available, nor were ADG measurements for non-infected relatives.

Serum viremia, measured using a semi-quantitative TaqMan PCR assay for PRRSV RNA, provided repeated (bi-weekly up to 14 dpi, then weekly) measures for $\log_{10}$-transformed qPCR viremia, as described in Boddicker et al. [22,57,58]. Mathematical functions were previously fitted to these $\log_{10}$ transformed viremia measures to smooth the data and to obtain continuous viremia estimates over the 42-day observation period [62]. As outlined in Islam et al, the uni-modal Woods function and the extended bi-modal Woods function provided a good fit to the individual's data with either uni-modal ($y(t) = a_1 t^{b_1} e^{-c_1 t}$) (~67%) or bi-modal ($y(t) = a_1 t^{b_1} e^{-c_1 t} + \max(0, a_2 (t - t_0)^{b_2} e^{-c_2 (t - t_0)})$) (~33%) viremia profiles, respectively, with strong correlations between model predictions of VL and actual viremia measures (genetic and phenotypic correlation estimates were 0.98±0.03 and 0.90±0.01, respectively) [55].

Across all trials, 198 pigs died before 42 dpi. PRRS was identified as the primary cause of mortality, except for trial 6, for which mortality was higher (46% by 42 dpi) and was potentially caused by secondary bacterial infections [58]. These pigs were included in the analyses until their time of death.

Only offspring from sires with more than 10 progeny with phenotypes were considered in this study to reduce the risk of bias in tolerance estimates [35]. As such, the number of animals included was 1320 from 0 to 21 dpi and 1001 from 0 to 42 dpi, originating from 54 sires.

Pedigree information and genomic information using genotypes from Illumina's Porcine SNP60 Beadchip v.1 [46], was available for all pigs. The pedigree-based

numerator relationship matrix (A) and genomic relationship G-matrix (G_m), were constructed in ASReml 3.0 [63] using the VanRaden method for all animals used in the analysis, where A contains the additive genetic relationship between animals, and G_m contains the genotypes for each animal. For the G-matrix, SNPs that were fixed in a trial were removed. Trials 1, 2, and 3 had the most extensive pedigree information, with pedigree up to two generations back, while the rest of the trials only had sire and dam recorded. As such, there were no relationships between animals in different trials, except for trials 1, 2, and 3, which consisted of animals from consecutive parities of the same breeding company (Table 1). Pedigree was corrected using parental genotypes for all trials, as described by Boddicker et al. [58] and Hess et al. [55]. The G-matrix was constructed using the VanRaden method [64], and included relationships between animals across trials regardless of breed, as outlined by Hess et al. [55]. The A-matrix was used for all the following statistical models, unless otherwise noted.

2.3.2 Resistance, tolerance, and performance without infection

Resistance is often quantified by a measure of within-host pathogen load, whereby lower pathogen load reflects higher host resistance [4–6]. In this study resistance to PRRS was defined as the inverse of serum viral load, whereby VL_{42} represents the cumulative log-transformed viral load from 0 to 42 dpi from the Woods curve. As viremia had reduced to undetectable levels within 21 to 28 dpi for a large proportion of pigs, cumulative viral load (and thus resistance) was not only calculated for the entire observation period from 0 to 42 dpi, but also for the period from 0 to 21 dpi. This represents the acute phase of infection, yielding two indicator traits for resistance (VL_{21} and VL_{42}).

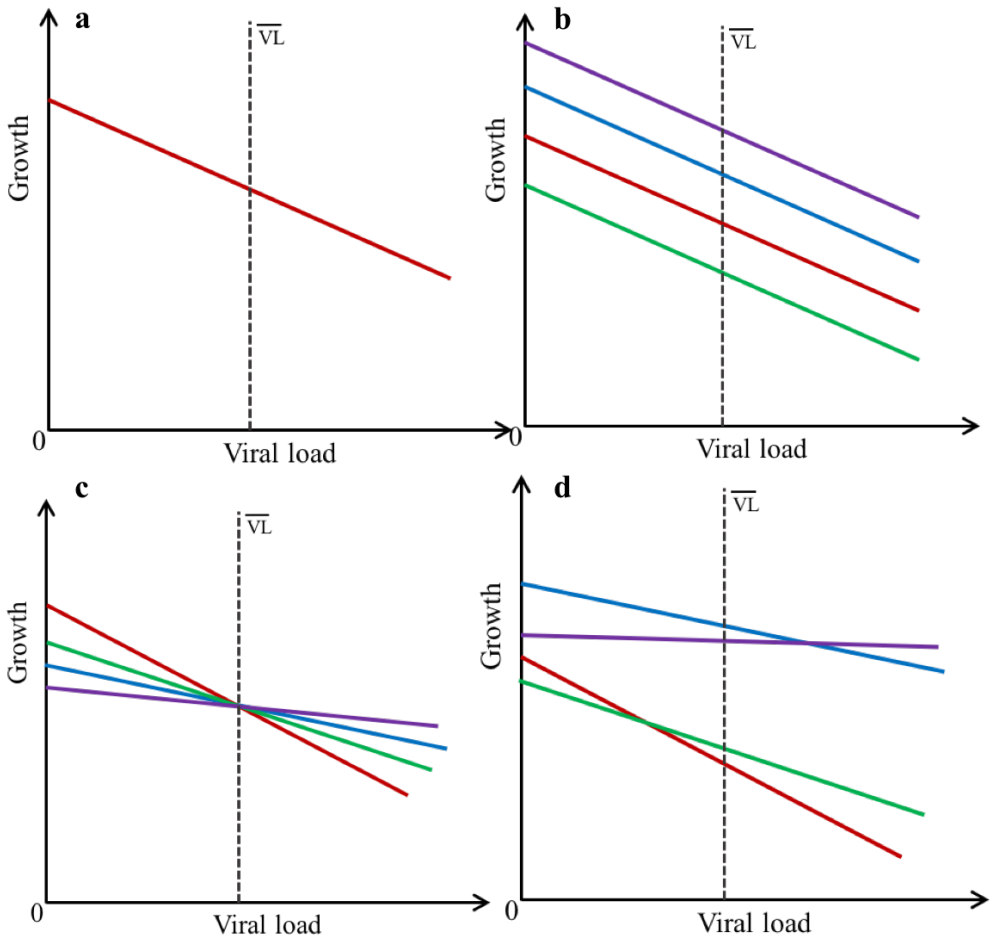
In this study, tolerance was assessed by regressing performance measures (i.e. ADG_{21} or ADG_{42} on y-axis) on pathogen load (i.e. VL_{21} or VL_{42}, respectively on x-axis). The regression of ADG_{42} on VL_{21} was also evaluated to account for the possibility of a time lag in growth response with respect to changes in pathogen load.

Growth performance of an infected individual is likely to depend on both their response to infection and performance in the absence of infection. Performance in absence of infection (i.e. when pathogen load is equal to zero), commonly denoted in the tolerance literature as vigor [65], constitutes the intercept of the linear reaction-norms (Figure 1). Previous simulation studies indicated performance measures in the absence of infection are important for obtaining unbiased tolerance slope estimates [35]. However, information on performance of the PRRSV challenged pigs in absence of infection was not available in this study.

Two approaches were adopted to overcome this lack of performance measures without infection. 1) in line with the standard approach of quantitative genetic studies of reaction-norms, the origin of the explanatory variable VL was shifted to the mean VL; this 'shifted intercept' for ADG is referred to as the 'level', in contrast to vigor [36,66,67] (Figure 1). Note that this approach does not provide accurate information about the genetic relationship between tolerance and vigor, as the genetic correlation between level and slope is not equal to the genetic correlation between performance at VL = 0 and slope [68,69]. Furthermore, individual body weight at the start of the infection (BW_0) was included as a fixed covariate in the corresponding statistical models to partially account for differences in vigor. 2) to gain better insight into data requirements for accurately estimating genetic parameters for tolerance, and about how these estimates depend on the genetic relationship between growth in absence or

presence of infection, growth records of infected pigs were augmented with simulated growth records of non-infected half-siblings, as outlined in step 4 below.

Figure 1. Graphical illustration of reaction norms for analysis of tolerance.



Mean VL is indicated by the stippled line in each graph. Each line corresponds to one of four hypothetical sires. Figure 1A shows the null model, where all sires have equal tolerance and equal overall growth level. As such, there is only one (average) tolerance slope. Figure 1B shows reaction-norms of sires with equal tolerance. Sires differ in intercept (growth where VL=0) and level (growth at mean VL), but have equal tolerance slopes. No re-ranking of sires occurs between growth associated with low and high VL, and genetic correlation between intercept and level is 1.00. Figure 1C

illustrates reaction norms of sires with variation in intercept and tolerance slopes, but no variation in level. Re-ranking of sires occurs depending on whether offspring harbor low or high VL, respectively, as indicated by crossing over of lines before and after mean VL. Figure 1D describes reaction norms of sires where variation occurs at intercept, level and tolerance slope. Sire re-ranking occurs between low and high VL, and genetic correlation between intercept and level is below one.

2.3.3 Statistical analyses

All statistical analyses were carried out using ASReml 3.0 [63]. Random regression reaction-norm models are prone to biased estimates if data requirements to disentangle intercept from slope are not met [35,36,70] thus a stepwise approach was adopted. (Step 1) multi-trait animal models were used to estimate the genetic relationship between resistance and growth under infection; (Step 2) multi-trait models were used to provide evidence for genetic variation in tolerance of pigs to PRRS based on the genetic correlation between growth associated with low and high VL, respectively; (Step 3) a univariate random regression model was applied to obtain estimates for genetic variance in tolerance; and (Step 4) data were augmented using simulated performance in the absence of infection (ADG^0_{21} or ADG^0_{42}), with increasing simulated genetic correlation from weak to strong between ADG^0_{21} and ADG_{21} or ADG^0_{42} and ADG_{42}, respectively. The random regression models from step 3 were adapted to include variation in ADG^0_{21} or ADG^0_{42}.

2.3.4 Step 1: Animal models to estimate the genetic relationship between resistance and performance prior to and post infection

Our first step in analyzing variation in growth under infection was to estimate heritabilities and correlations between VL and growth in absence of and post infection with PRRSV using the following trivariate animal model:

$$\begin{bmatrix}\mathbf{y_1}\\ \mathbf{y_2}\\ \mathbf{y_3}\end{bmatrix} = \begin{bmatrix}\mathbf{X_1} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{X_2} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{X_3}\end{bmatrix}\begin{bmatrix}\mathbf{b_1}\\ \mathbf{b_2}\\ \mathbf{b_3}\end{bmatrix} + \begin{bmatrix}\mathbf{Z_1} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z_2} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{Z_3}\end{bmatrix}\begin{bmatrix}\mathbf{a_1}\\ \mathbf{a_2}\\ \mathbf{a_3}\end{bmatrix} + \begin{bmatrix}\mathbf{U_1} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{U_2} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{U_3}\end{bmatrix}\begin{bmatrix}\mathbf{p_1}\\ \mathbf{p_2}\\ \mathbf{p_3}\end{bmatrix} +$$

$$\begin{bmatrix}\mathbf{M_1} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{M_2} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{M_3}\end{bmatrix}\begin{bmatrix}\mathbf{l_1}\\ \mathbf{l_2}\\ \mathbf{l_3}\end{bmatrix} + \begin{bmatrix}\mathbf{e_1}\\ \mathbf{e_2}\\ \mathbf{e_3}\end{bmatrix}, \text{[1]}$$

where $\mathbf{y_1}$, $\mathbf{y_2}$ and $\mathbf{y_3}$ are phenotypes for body weight at the start of infection (BW_0) ($\mathbf{y_1}$), ADG_{21} or ADG_{42} ($\mathbf{y_2}$), and VL_{21} or VL_{42} ($\mathbf{y_3}$), respectively; $\mathbf{b_1}$, $\mathbf{b_2}$and $\mathbf{b_3}$ are the vectors of the fixed effects for the interaction of experimental trial and parity of the dam when offspring were born (trial-by-parity), sex of the offspring, and age at start of experimental infection, which was fitted as fixed covariates. To account for differences between viremia profiles and the two mathematical functions used to fit these, a binary variable associated with the viremia profile class (uni- or bi-modal) was also fitted as fixed effect; $\mathbf{a_1}$, $\mathbf{a_2}$ and $\mathbf{a_3}$ are vectors of additive genetic effects for each trait, with $\text{Var}\begin{bmatrix}\mathbf{a_1}\\ \mathbf{a_2}\\ \mathbf{a_3}\end{bmatrix} = \mathbf{G} \otimes \mathbf{A}$, where $\mathbf{G}$ is the genetic variance-covariance matrix and $\mathbf{A}$ the pedigree relationship matrix; $\mathbf{p_1}$, $\mathbf{p_2}$ and $\mathbf{p_3}$ are vectors of pen effects nested within a trial for each trait, with $\text{Var}\begin{bmatrix}\mathbf{p_1}\\ \mathbf{p_2}\\ \mathbf{p_3}\end{bmatrix} = \mathbf{I} \otimes \mathbf{K}$, where $\mathbf{I}$ is the identity matrix and $\mathbf{K}$ is the corresponding variance-covariance matrix of pen effects for the different traits; $\mathbf{l_1}$, $\mathbf{l_2}$ and $\mathbf{l_3}$ are the vectors of litter effects for each trait, with $\text{Var}\begin{bmatrix}\mathbf{l_1}\\ \mathbf{l_2}\\ \mathbf{l_3}\end{bmatrix} = \mathbf{I} \otimes \mathbf{L}$,

with corresponding variance-covariance matrix $\mathbf{L}$; $\mathbf{e_1}$, $\mathbf{e_2}$ and $\mathbf{e_3}$ are the vectors of error terms for each trait, with Var $\begin{bmatrix}\mathbf{e_1}\\ \mathbf{e_2}\\ \mathbf{e_3}\end{bmatrix} = \mathbf{I} \otimes \mathbf{R}$, where $\mathbf{R}$ is the variance-covariance matrix for the residual effects for each trait; and $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$, $\mathbf{Z}_1$, $\mathbf{Z}_2$ and $\mathbf{Z}_3$, $\mathbf{U}_1$, $\mathbf{U}_2$ and $\mathbf{U}_3$, and $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$ are the incidence matrices for the fixed, animal, pen and litter effects, respectively. In addition to the trivariate animal model, corresponding bivariate and univariate models were also used to check the robustness of variance components. As heritability estimates differed between models, heritability estimates were presented from the corresponding univariate models.

2.3.5 Step 2: Multi-trait models to examine evidence for genetic variation in tolerance: growth associated with low versus high VL

The trivariate model [1] from step 1 does not show how growth changes with respect to viral load, and, therefore, does not account for genetic variance in tolerance. A multi-trait sire model for ADG of progeny with categorized VL was used to assess sire-by-VL interactions to get a first indication whether sires varied genetically in tolerance to infection. If these genetic correlations are less than unity, this is indicative of sire rank changes when offspring are faced with low and high VL respectively, providing evidence for genetic variation in tolerance slope. Hence, individuals were sorted according to their VL from 0-21 dpi or 0-42 dpi, and partitioned into VL groups, where the low and high VL groups (n=330 each) consisted of individuals with VL values in the lower and upper quartiles, respectively, and the mid-range group consisted of the middle half of the data (n=660). A multi-trait sire model was then fitted to measures of ADG associated with low, mid and high VL from 0-21/0-42 dpi

(ADG_{low}, ADG_{mid} and ADG_{high}), respectively. The fixed and random effects of this model were identical to those used in model [1], with exception of **a**, which now refers to sire effects on performance and explains ¼ of the additive genetic variance, and of **e**, where residual covariance was fixed at 0. Furthermore, the pedigree relationship A-matrix was replaced with the genomic relationship matrix (G-matrix) to improve convergence.

2.3.6 Step 3: univariate random regression models for estimating genetic variance in tolerance

The multi-trait models in the previous steps provide evidence for genetic variation in tolerance but do not yield direct estimates of genetic variance in tolerance. A random regression reaction norm model was applied, whereby the origin of the reaction-norms was centered at the mean viral load values, thus providing only variance component estimates for level (ADG at mean VL) rather than vigor (ADG at zero VL). The following linear random regression sire model (RRM) for ADG on centered values of VL, which will be referred to as the level-slope model (as shown in Figure 1D), was used:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{X_{VL}b_s} + \mathbf{Za_i} + \mathbf{Z_{VL}a_s} + \mathbf{Up} + \mathbf{Ml} + \mathbf{e}, \quad [2]$$

where **y** is the vector of ADG_{21} or ADG_{42}, respectively; **b** is the vector of fixed effects outlined in model [1], with age and BW_0 included as additional fixed covariates to account for variation in age and body weight at the start of infection; and $\mathbf{b_s}$ is the population average tolerance slope; $\mathbf{a_i}$ and $\mathbf{a_s}$ are the sire effects on level and on tolerance slope, respectively, assumed to follow a multi-variate normal distribution with mean zero and $\text{Var}\begin{bmatrix}\mathbf{a_i}\\ \mathbf{a_s}\end{bmatrix} = \frac{1}{4}\mathbf{G_{RN}} \otimes \mathbf{A}$, with $\mathbf{G_{RN}} = \begin{bmatrix}\boldsymbol{\sigma}^2_{\mathbf{a_i}} & \boldsymbol{\sigma}_{\mathbf{a_i a_s}}\\ \boldsymbol{\sigma}_{\mathbf{a_i a_s}} & \boldsymbol{\sigma}^2_{\mathbf{a_s}}\end{bmatrix}$, where $\boldsymbol{\sigma}^2_{\mathbf{a_i}}$ and

$\boldsymbol{\sigma^2_{a_s}}$ are the variance of $\mathbf{a_i}$, and $\mathbf{a_s}$, respectively, $\boldsymbol{\sigma_{a_i a_s}}$ is the covariance between sire effects for level and slope; other random effects **p**, **l**, and **e** were fitted as described in model [1]; $\mathbf{X_{VL}}$ and $\mathbf{Z_{VL}}$ are the incidence matrices for population average tolerance slope and those associated with each sire, respectively, consisting of individual VL measures, and **X** is the incidence matrix for the fixed effects (including VL as fixed covariate) and **Z** is the incidence matrix for the random sire effect on level (**Z**).

To test the significance of sire effects on level and slope and to determine which of the models illustrated in Figure 1 best described the data, the model fit of the level-slope model [2] was compared with that of hierarchical models: (a) without any additive genetic effects (Figure 1A), (b) with only sire effects for level (Figure 1B), and (c) containing only sire effects on slope (Figure 1C). Significance of each random effect was assessed using the likelihood ratio test (LRT) [71], with the LRT test statistics below assumed to follow a χ^2 distribution, with 1 degree of freedom for inclusion of an additional sire effect (e.g. null to level model, including sire effect) and a mixture of 1 and 2 degrees of freedom for additional sire slope effects and covariance (for example, from level to level-slope model), [72,73].

2.3.7 Step 4: Random regression model using simulated performance in absence of infection for estimating genetic variance in tolerance

The random regression models fitted in step 3 generated potential confounding between level and tolerance slope variance estimates i.e. genetic variance in slope was absorbed by genetic variance in level due to the limited distribution of VL around average VL required to estimate the genetic variance in level. To assess whether confounding could be resolved by inclusion of performance measures of non-infected

relatives in the statistical models, growth in the absence of infection (ADG^0_{21} or ADG^0_{42}) was simulated for one hypothetical paternal half-sib for each individual with ADG_{21} and ADG_{42} records, respectively, thus doubling the size of the dataset. Data were simulated assuming a heritability of 0.4 for both ADG^0_{21} and ADG^0_{42} [74]. With the expectation that a higher r_g between the traits would imply less genetic variance in tolerance, low (0.05), moderate (0.30), strong (0.60) or high (0.90) genetic correlations (r_g) between ADG^0_{21} and $ADG_{21,}$ or ADG^0_{42} and ADG_{42}, respectively, were simulated (see **Appendix 2.1** for a detailed description of the simulations). Note that no assumptions were made with regards to genetic variance in tolerance. 10,000 replicates of simulated half-sib records for ADG^0_{21} and ADG^0_{42} were generated.

The random regression models [2] were then applied to the extended datasets for each replicate, where the response vector **y** now comprised either simulated ADG^0_{21} and measured ADG_{21}, or ADG^0_{42} and $ADG_{42.}$ VL was no longer centered at mean VL, but comprised VL equal to zero for the non-infected pigs and VL_{21} or VL_{42} for the infected pigs. The remaining fixed and random effects were identical to those in model [2], except no fixed effects or random pen or litter effects were fitted for the simulated half-sibs. Thus, by including simulated data of non-infected pigs model [2] was replaced by an intercept-slope model, with genetic variance estimated for growth in the absence of infection, and for tolerance slope.

As in step 3, hierarchical models were fitted (a) without any additive genetic effects for intercept or slope (Figure 1A, null model), and (b) with additive genetic effects for intercept only (Figure 1B, intercept-only model) and (c) with additive genetic effects for intercept and slope (Figure 1D, intercept-slope model). The model fit was assessed

using the loglikelihood ratio test outlined in step 3 above. Results were evaluated based on the mean and standard deviation of the estimates over replicates.

2.4 Results

2.4.1 Step 1: Relationship between resistance and performance prior to and post infection

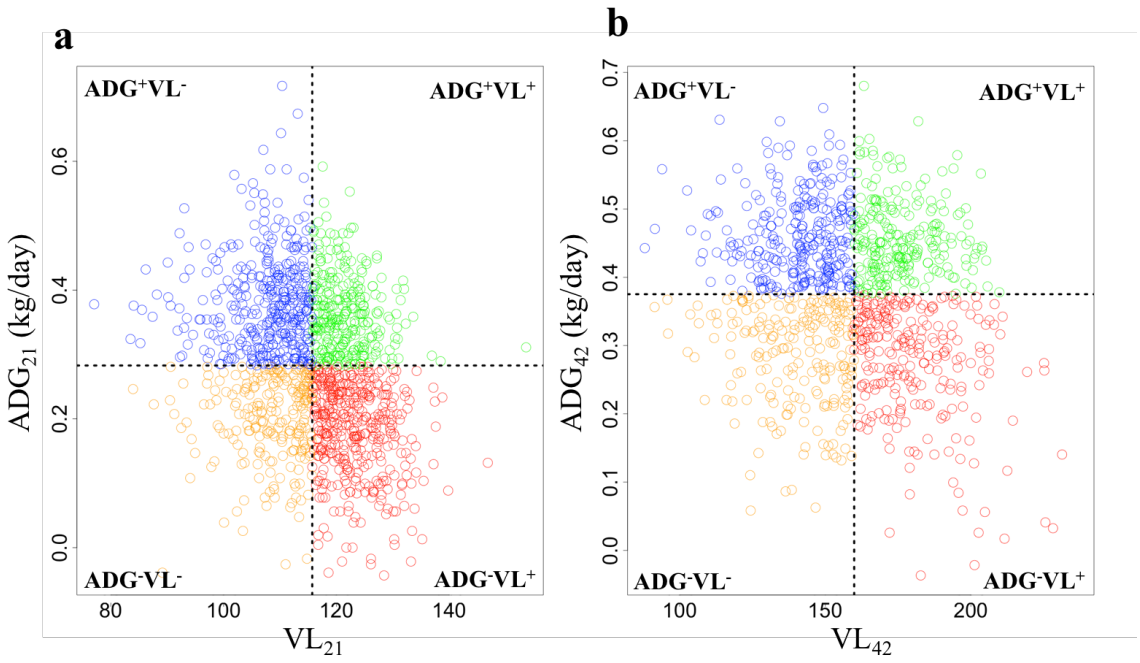
ADG_{21} and ADG_{42} ranged from a weight loss of 40 g/day to a weight gain of 720 and 680 g/day, respectively, with corresponding mean daily weight gains of 280 and 380 g/day (Table 2).

Table 2. Summary statistics of resistance and growth traits. Body weight at 0 dpi (BW_0), average daily gain and viral load from 0-21 and 0-42 dpi (ADG_{21}, ADG_{42}, VL_{21} and VL_{42}), respectively. AUC is the area under the curve for the log-transformed estimates for viral load in blood as measured by rtPCR.

Trait	Mean	Standard deviation	Minimum	Maximum	Number of records
BW_0 [kg]	7.30	1.39	3.45	12.88	1320
ADG_{21} [kg/day]	0.28	0.12	-0.04	0.72	1319
ADG_{42} [kg/day]	0.38	0.11	-0.04	0.68	1001
VL_{21} [AUC]*	115.69	9.37	77.04	153.62	1320
VL_{42} [AUC]*	159.90	23.42	88.00	236.35	1001

Figure 2 depicts the distributions of growth and VL for the two observation periods between 0-21 dpi and 0-42 dpi. The wide distribution of individuals with above average growth rate despite high VL ($ADG^{+}VL^{+}$), and with low growth rate despite low VL ($ADG^{-}VL^{-}$) may be indicative of phenotypic variation in tolerance.

Figure 2. **Scatter plots of data for ADG and VL from A) 0-21 and B) 0-42 dpi.** ADG and VL from 0-21 and 0-42 dpi (n=1320 and 1001, respectively) were distributed into one of four quadrants according to their growth and VL after infection with PRRS virus (n=330 and 250 in each quadrant for 0-21 and 0-42 dpi, respectively). The quadrants ($ADG^{+}VL^{-}$ blue, $ADG^{+}VL^{+}$ green, $ADG^{-}VL^{-}$ orange, $ADG^{-}VL^{+}$ red) refer to high growth rate and high resistance (low VL), high growth rate and low resistance (high VL), low growth rate and high resistance and low growth rate and high low, respectively. Quadrants were centered at mean VL and at mean ADG.



Growth rate under infection and resistance were found moderately heritable and had large standard errors (Table 3). Heritability estimates were similar for the two time periods considered.

Table 3. Estimates of heritability and correlations between resistance and growth traits. Heritability estimates (diagonal) and phenotypic (upper triangle) and genetic correlations (lower triangle) with standard errors (SE) from the trivariate animal model for Body Weight at 0 dpi (BW_0), Average Daily Gain and Viral Load from 0-21 and 0-42 dpi (ADG_{21}, ADG_{42}, VL_{21} and VL_{42}), respectively. Correlations between ADG_{21} and VL_{42} were not calculated, as VL is expected to impact ADG and not the other way around.

	Trait				
Trait	**BW_0**	**ADG_{21}**	**ADG_{42}**	**VL_{21}**	**VL_{42}**
BW_0	**0.11 (0.10)**	0.35 (0.03)	0.40 (0.03)	-0.21 (0.07)	-0.20 (0.03)
ADG_{21}	0.48 (0.30)	**0.29 (0.11)**	0.80 (0.01)	-0.29 (0.03)	.
ADG_{42}	0.24 (0.45)	1.00 (0.04)	**0.34 (0.14)**	-0.33 (0.03)	-0.36 (0.03)
VL_{21}	-0.33 (0.45)	-0.53 (0.27)	-0.64 (0.26)	**0.19 (0.11)**	0.80 (0.01)
VL_{42}	-0.54 (0.37)	.	-0.82 (0.16)	0.79 (0.14)	**0.18 (0.10)**

Although standard errors were high, genetic correlations between VL and growth under infection were statistically significantly different from one ($p<0.001$, based on the LRT comparing models with and without genetic correlations fixed to unity), indicating that not all genetic variation of growth under infection was explained by genetic differences in resistance (inverse of VL) (Table 3). Furthermore, genetic correlations between growth under infection and BW_0 were also significantly different

from one, implying that growth prior to and post infection were not under identical genetic regulation. Genetic correlations between growth under infection and VL were moderate to strong and negative whereas genetic correlations between growth under infection and BW_0 were moderately positive. Phenotypic correlations were of the same sign but generally weaker than the genetic correlations (Table 3). Phenotypically and genetically, these results indicate that pigs with greater resistance tend to grow faster.

2.4.2 Step 2: multi-trait models to examine evidence for genetic variation in tolerance. Trivariate models for growth at low, mid and high VL failed to converge for both time periods of infection. Using bivariate models for the upper and lower quartiles for VL, high genetic correlations of 0.94 (0.18) and 0.91 (0.13) between growth associated with low to high VL were identified for ADG_{21} and ADG_{42}, respectively. Genetic correlations significantly less than one would imply that growth rates associated with different degrees of infection severity, as indicated by low versus high VL, are genetically distinct traits and would thus be indicative of genetic variation in tolerance (Figure 1). Genetic correlations close to one indicate limited re-ranking among sires between high and low levels of VL and, thus limited genetic variance in tolerance. Furthermore, there was no significant difference between genetic variances of ADG associated with low and high VL, for either the 0-21 and 0-42 day period (where genetic variances for ADG associated with low and high VL were $2.10E^{-03}$($1.22E^{-03}$) and $4.56E^{-03}$($1.81E^{03}$) for 0-21 dpi, and $3.46E^{-03}$($1.24E^{-03}$) and $6.89E^{-03}$($2.18E^{-03}$) for 0-42 dpi, respectively). Referring to the expectations outlined in Figure 1, the results of this multi-trait model imply that random regression models of Step 3

with the same tolerance slope for each sire would provide a better fit to the data than models with different slopes for each sire (Figure 1C & D).

2.4.3 Step 3: Estimation of genetic variance in tolerance using univariate random regression models

Univariate random regression models without genetic effects, but including VL as fixed linear (and higher order polynomial) covariate were used to test the average association between growth and VL (Null model in Table 4). These identified a statistically significant linear association between growth and VL ($p<0.0001$), with a population average tolerance slope estimate of $-2.78E^{-03}$ ($3.32E^{-04}$) and $-1.28E^{-03}$ ($1.51E^{-04}$) kg/day per unit of VL increase for ADG_{21} regressed on VL_{21} and ADG_{42} on VL_{42}, respectively. This corresponds to an average growth rate difference of 213 g/day and 190 g/day between pigs with lowest and highest observed VL for the 21- and 42-day observation period, respectively, or differences in body weight of 4.5 and 8.0 kg over the 21- and 42-day observation periods, respectively. Similarly, body weight prior to infection had a significant association with ADG post infection (BW_0 $p<.0001$), with a positive regression coefficient of 0.025 (0.002) at 21 dpi and of 0.029 (0.003) at 42 dpi.

The log-likelihood of the model improved significantly when genetic effects (random sire effects) were included in the model (level model) ($p<.0001$) (Table 4). This indicates significant genetic variance in growth performance of pigs infected with PRRSV. However, including sire effects of slope only (Figure 1c) did not improve model fit over the null model ($p>0.60$) and resulted in negligibly small slope variance estimates.

Table 4. **Variance components for ADG (kg/d) from 0-21 dpi (4A) and 0-42 dpi (4B).** Variance components estimated from random regression models: null model, containing no genetic effect; level-only model, containing only the overall sire effect on growth under infection; slope-only model, containing only sire effect on the slope of the regression line of growth over VL; and level-slope model, containing sire effects on level and slope, respectively. All other fixed effects/covariates and random effects were identical between models. Results for ADG_{42} on VL_{21} were similar to those for ADG_{42} on VL_{42} and are therefore not shown.

4A

	Null model	**Level-only model**	**Slope-only model**	**Level-Slope model**
	Estimate (SE)	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**
Level	.	2.01E-03 (7.68E-04)	.	2.01E-03 (7.68E-04)
Covariance	.	.	.	2.21E-13 (1.04E-14)
Slope	.	.	4.37E-06 (2.06E-07)	1.00E-10 (4.71E-12)
Pen(Trial)	4.12E-04 (1.45E-04)	3.97E-04 (1.42E-04)	4.12E-04 (1.45E-04)	3.97E-04 (1.42E-04)
Litter	9.25E-04 (2.26E-04)	4.72E-04 (2.04E-04)	9.25E-04 (2.26E-04)	4.72E-04 (2.04E-04)
Residual	6.18E-03 (2.91E-04)	6.18E-03 (2.91E-04)	6.18E-03 (2.91E-04)	6.18E-03 (2.91E-04)
LogLikelihood	**2482.98**	**2495.03**	**2482.98**	**2495.20**

4B

	Null model	Level-only model	Slope-only model	Level-Slope model
	Estimate (SE)	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**
Level	.	2.32E-03 (1.02E-03)	.	2.33E-03 (1.03E-03)
Covariance	.	.	.	-3.95E-15 (2.04E-16)
Slope	.	.	8.60E-08 (1.47E-07)	3.41E-07 (5.94E-07)
Pen(Trial)	2.43E-04 (1.30E-04)	2.82E-04 (1.36E-04)	2.39E-04 (1.29E-04)	2.78E-04 (1.35E-04)
Litter	1.76E-03 (3.27E-04)	1.23E-03 (2.98E-04)	1.75E-03 (3.27E-04)	1.22E-03 (2.98E-04)
Residual	5.39E-03 (3.03E-04)	5.33E-03 (2.99E-04)	5.36E-03 (3.05E-04)	5.30E-03 (3.02E-04)
LogLikelihood	**1889.55**	**1911.18**	**1899.72**	**1911.35**

Models with sire effects on both level and slope, as well as a genetic covariance between them, yielded a significantly better model fit than the null model ($p<0.0001$). However, the level-slope model did not provide a significantly better fit than the level-only model for either 0 to 21 and 42 dpi ($p=1.00$ and 0.66, respectively) (Table 4).

All four models provided similar estimates of variance components for non-genetic random effects (Table 4). Estimates of the sire variance for level were very similar between the level-only model and the level-slope model and very low, whereas estimates for sire variance in tolerance slope differed slightly between the slope-only and the level-slope model (Table 4). The fact that addition of the slope did not affect the variance estimate for level suggests potential confounding of level and slope (see ***2.5.2 Statistical Considerations***) The estimate of the covariance between level and slope was close to zero, and constrained at the boundary for both time periods,

indicating numerical difficulties in accurately estimating these variance components. However, shifting the covariate VL to ensure a zero covariance between the new level and slope has no effect on the model likelihoods, suggesting that the results are robust. In conclusion, the random regression models did not allow estimation of genetic variance in tolerance of pigs to PRRSV infection. Based on statistical model fit alone, the level-only model accounting for genetic variance in growth rate at mean VL only constitutes a more appropriate model to describe genetic variation in growth response of infected pigs than the level-slope model accounting for genetic variance in both, growth rate at mean VL and tolerance. However, as outlined in more detail in the ***2.5.2 statistical considerations*** section below, it cannot be excluded that any genetic variance in tolerance that may exist is absorbed in the genetic variance for level because of the confounding between level and slope.

2.4.4 Step 4: Random regression models including simulated performance in absence of infection for estimating genetic variance in tolerance

Models with genetic effects on both intercept and slope, as well as a genetic covariance between them, consistently yielded a significantly better model fit than the null model ($p<0.0001$ for both 0-21 and 0-42 dpi), regardless of simulated genetic correlation between ADG^{0}_{21} and ADG_{21} or ADG^{0}_{42} and ADG_{42}. However, the intercept-slope model consistently provided a significantly superior fit over the intercept-only model only when the simulated genetic correlation between growth in absence of infection and growth under infection was low to moderate (Table 5). Generally, the ability to identify genetic variance in tolerance decreased with an increase in the simulated genetic correlation, as indicated by reduced improvement in log-likelihoods and a

lower proportion of replicates with significant genetic variation in tolerance slope ($p<0.05$) (Table 5). Somewhat surprisingly, for the 0-21dpi observation period, the majority of replicates indicated significant genetic variation in tolerance, even for strong genetic correlations between ADG^{0}_{21} and ADG_{21} (Table 5). In contrast, only low to moderate genetic correlations between ADG^{0}_{42} and ADG_{42} resulted in significant genetic variance in tolerance for the majority of replicates for the 42 day observation period (Table 5).

Table 5. Effect of the genetic correlation (r_g) between simulated ADG in the absence of infection with observed ADG under infection on evidence for genetic variance in tolerance. Effect of the genetic correlation (r_g) of simulated ADG^{0}_{21} with ADG_{21} and ADG^{0}_{42} with ADG_{42} on the average change in log-likelihood of the intercept-slope model over the intercept-only model (ΔLogLikelihood), the average p-value of log likelihood improvement, provided by a log-likelihood ratio test, and the proportion of the 10,000 replicates with significant genetic variance in tolerance (i.e. p-value of LRT was <0.05). Standard deviations over 10,000 replicates are shown in brackets.

ADG period (dpi)	r_g	ΔLogLikelihood	p-value	Proportion with significant genetic variance for tolerance (p<0.05)
0 to 21	**0.05**	10.96 (4.19)	0.000	1.00
	0.30	6.18 (3.39)	0.005	0.98
	0.60	2.32 (1.49)	0.041	0.76
	0.90	1.00 (2.12)	0.067	0.55
0 to 42	**0.05**	8.67 (4.40)	0.003	0.99
	0.30	4.34 (3.16)	0.023	0.87
	0.60	1.31 (1.56)	0.107	0.41
	0.90	-0.80 (2.43)	0.187	0.06

Table 6 shows that random regression sire models when including records from both non-infected and infected siblings can generate robust genetic variance estimates for both intercept and slope. As expected, genetic variance estimates for tolerance slope tended to decrease with increasing genetic correlations between ADG^{0}_{21} and ADG_{21} or ADG^{0}_{42} and ADG_{42}, whereas the genetic variance estimates in the intercept tended to increase (see **Appendix 2.1**). Genetic correlations between ADG in absence of infection and ADG under infection also affected the estimated genetic correlations between intercept and tolerance slope. Low (simulated) genetic correlations between ADG^{0}_{21} and ADG_{21} or ADG^{0}_{42} and ADG_{42}, respectively, led to negative genetic correlations between performance in the absence of infection and tolerance, whereas strong positive genetic correlations between the growth traits suggested that pigs with greater genetic growth in the absence of infection were also genetically more tolerant to infection.

Table 6. **Variance components of intercept, slope and covariances from random regression models.** Variance components estimated from random regression models based on simulated ADG^{0}_{21} and measured ADG_{21} (kg/d) or ADG^{0}_{42} and ADG_{42}. Fitted models were the intercept-only model, containing only the overall sire effect on intercept; and the intercept-slope model, containing sire effect on intercept and slope for ADG_{21} or ADG_{42}, respectively. All other fixed effects/covariates and random effects were identical between models. Standard errors (in brackets) were calculated as the standard deviation over 10,000 replicates. r_g is the simulated genetic correlation between ADG^{0}_{21} or ADG^{0}_{42} and ADG_{21} or ADG_{42}.

		Intercept-only model	**Intercept – slope model**		
ADG period (dpi)	**r_g**	**Intercept**	**Intercept**	**Covariance**	**Slope**
0 to 21	**0.05**	7.65E-04 (2.24E-04)	9.93E-04 (2.98E-04)	-7.57E-06 (3.97E-06)	2.24E-07 (5.80E-08)
	0.3	9.20E-04 (2.53E-04)	9.95E-04 (2.99E-04)	-3.03E-06 (3.27E-06)	1.44E-07 (5.09E-08)
	0.6	1.13E-03 (2.71E-04)	1.03E-03 (2.84E-04)	1.54E-06 (2.13E-06)	5.54E-08 (3.24E-08)
	0.9	1.34E-03 (2.31E-04)	1.19E-03 (1.95E-04)	3.35E-06 (1.06E-06)	1.25E-08 (6.33E-09)
0 to 42	**0.05**	9.20E-04 (2.83E-04)	1.10E-03 (3.47E-04)	-5.80E-06 (3.59E-06)	1.18E-07 (3.87E-08)
	0.3	1.09E-03 (3.05E-04)	1.12E-03 (3.47E-04)	-1.90E-06 (2.85E-06)	6.85E-08 (3.23E-08)
	0.6	1.28E-03 (3.15E-04)	1.14E-03 (3.19E-04)	-5.57E-07 (1.75E-06)	2.37E-08 (1.65E-08)
	0.9	1.47E-03 (2.57E-04)	1.21E-03 (4.10E-04)	1.95E-06 (1.21E-06)	1.09E-08 (7.13E-09)

2.5 Discussion

2.5.1 Summary of findings

Performance of an infected individual is likely to depend on its ability to restrict pathogen load (resistance) and its ability to limit the impact of infection (tolerance). The extensive PHGC dataset has identified substantial genetic variation in resistance of growing pigs to PRRS and led to the discovery of a major quantitative trait locus associated with both resistance and growth of pigs under infection [22,55,57,58]. Surprisingly, the dataset provided little evidence that pigs also vary genetically in tolerance to this virus. However, the simulations revealed that genetic variation in tolerance to PRRS may exist, depending on performance in the absence of infection (vigor). Furthermore, this analysis raised numerous statistical difficulties associated with genetic improvement of host tolerance, which could be overcome by including measures of performance of infected and non-infected relatives in the analysis.

Focusing on data from infected pigs alone, genetic correlations between body weight prior to infection, resistance (inverse of VL) and growth under infection were found to be moderately strong and positive, in line with previous studies [22,55]. This indicates that heavier individuals prior to infection counteract an increase in pathogen load, and thus tend to have lower VL, and therefore lower infection-induced reductions in growth rate. Genetic correlations between VL and growth were strongly negative, implying that animals that were genetically more resistant also tended to grow faster under infection. However, correlations were significantly different from one, indicating that genetic variation in growth of PRRSV infected pigs is not fully explained by heterogeneity in growth prior to infection and resistance. Therefore, genetic variation in tolerance may also play a part in host response to PRRSV

infection. However, the multi-trait model provided little evidence of genetic variation in tolerance. This was further supported by the random regression models. These showed that, although growth rate declined, on average, linearly with increasing VL, there was no statistically significant difference in tolerance between the sires of the infected piglets.

However, closer inspection of the underlying data structure raised suspicion that genetic variance in the reaction norm level absorbed genetic variance in tolerance due to confounding between level and slope in these data (see ***2.5.2 statistical considerations***). To disentangle the genetic variance in reaction-norm intercepts (i.e. growth rate in the absence of infection) and slopes (i.e. tolerance), the experimental data were augmented with simulated growth rates of non-infected relatives. The resulting data structure thus mimicked that of 'sib challenge tests' that are common practice in aquaculture and other livestock species [75–77]. The simulations demonstrated that inclusion of these additional data in the random regression models resolved the confounding between level and slope and resulted in more reliable genetic parameter estimates for tolerance. Crucially, the simulations revealed that it would be wrong to conclude that pigs in this study lacked substantial genetic variation in tolerance to PRRS, as was suggested by the models based on the collected data alone. As demonstrated by the simulations, genetic variance estimates for tolerance strongly depend on the genetic correlations between growth in the absence and growth under infection. Low to moderately strong genetic correlations between these two traits implied significant genetic variance in tolerance of the pigs in this study. Interestingly, estimated genetic correlations between body weight of pigs prior to infection and growth under infection were found moderately strong. Thus, if body weight prior to

infection was a reliable predictor for growth rate in the absence of infection, evidence for genetic variance in tolerance would emerge directly from the data.

2.5.2 Statistical considerations

Here, the conventional reaction-norm approach was adopted to model genetic variation in tolerance to infections [5,8]. Using both simulated and real data, we demonstrated that random regression models embedded in the mixed model machinery are a powerful tool to estimate genetic variance in tolerance for outbred populations if the data structure is appropriate [4,35,37]. Random regression models are also known to be highly sensitive to the underlying data structure and prone to generate inaccurate variance estimates for slope, in particular if sample size is limited or information on relatedness is poor, as was the case for the data in this study [35,36,70]. To prevent bias in the slope variance estimates [35,70], only sires that had more than 10 offspring were included in this study. However, the associated reduction of the data to records from only 54 mostly unrelated sires may have caused a trade-off between reducing bias and reducing statistical power, as indicated by lower heritabilities for ADG and VL than found in previous analyses on the same data [22,55]. Furthermore, to alleviate the potential impact of limited information of relatedness (as only sires and dams were known for the majority of pigs), the analyses were repeated including the genomic relationship matrix rather than the pedigree relationship matrix, which is not able to capture the difference between siblings due to Mendelian sampling. However, this had a negligible impact on the variance estimates and on the log-likelihoods of the reaction-norm models (results not shown).

As is common practice for quantitative genetics models using REML, the likelihood ratio test (LRT) was used to test the significance of random effects such as the sire tolerance slope estimates [78]. For variance components constrained to the positive parameter space, the conventional LRT that assumes the test-statistics to follow a chi-square distribution with degrees of freedom equal to the number of additional parameters to be estimated in the more complex model has been described to be overly conservative [63]. For this reason the widely used adjustment of Stram & Lee [72] based on mixture distributions was applied. However, in this proposed adjustment individual subjects (in this case sires) were assumed independent. Due to lack of detailed pedigree information in the present study, the majority of sires were indeed assumed unrelated, with the exception of sires from trials 1-3. Repeating the analysis with the assumption that all sires were unrelated provided almost identical model results to those reported here. We therefore believe that the LRT is a valid method for testing the null hypotheses of zero genetic variance in tolerance and genetic correlations equal to zero or one in this study. Nevertheless, sires and sire by VL interactions were also fitted as fixed effects in the statistical models of step 3. In accordance with the results of modelling sires as random effect, there were no significant differences between the tolerance slopes associated with different sires according to the Wald test ($p=0.981$ and 0.081 for the 0 to 21 and 0 to 42 dpi time periods, respectively).

Perhaps most importantly, reaction-norms require considerable variation in the independent variable to generate unbiased tolerance slope estimates [7]. However, this study, in line with other infection challenge experiments, used an identical infection route, pathogen strain and dose for all individuals. Consequently, it provided a

relatively narrow value range for pathogen load (VL_{42} values ranged between 88 and 236 AUC in our study), with no values close to zero. To better accommodate the distribution of the data in the models, the VL was centered at the mean VL value, in line with common practice in the animal breeding literature [36,67,70]. However, the relatively narrow range of the VL of offspring, combined with the relatively low numbers of offspring for some sires, may have hampered the ability of these models to disentangle sire effects on level and slope. This confounding is likely further aggravated by genetic variation in resistance to PRRS, which implies that VL is not homogeneously distributed amongst sires, with more resistant sires predominantly having progeny with low VL, and less resistant sires predominantly progeny with high VL.

Considering all these effects combined, the weak evidence for significant genetic variation in tolerance to PRRS from the random regression models in this study may simply reflect a lack of statistical power to disentangle sire effects on regression slope and level. The complementary simulation studies presented here, which assumed that additional performance measures of related uninfected individuals were available, demonstrated one way of increasing statistical power. Similarly, it might be possible to increase statistical power by harnessing information from repeated measures of growth and pathogen load for each individual over the course of infection in the statistical models. By increasing the range of distribution of VL for each individual, a more robust slope may be fitted through the centre of the data, alluding to an "overall" picture of tolerance across multiple time-points in infection.

2.5.3 Implications for genetic improvement of tolerance of pigs to PRRS and other diseases

Genetic improvement of tolerance may have several advantages over improving resistance. Firstly, host resistance limits pathogen replication within the host and, as a consequence, selection for host resistance may impose selection advantages on pathogen strains that can overcome host resistance mechanisms and eventually result in a loss of selection advantage of the host [32,79]. Given the high mutation rate of RNA viruses such as PRRSV [80], this is a potential pitfall for a long-term breeding strategy focused on resistance. It has been proposed that, theoretically, tolerance might not impose such selection pressure on the pathogen [32].

Secondly, it has been suggested that improving host tolerance may offer cross-protection against other strains of the virus, or other prevalent infectious agents, as tolerance mechanisms primarily target host-intrinsic damage prevention or repair mechanisms, compared to resistance mechanisms, which interfere directly with the pathogen life-cycle [5,6,11]. This is particularly relevant for PRRS, which is often associated with co-infection with other respiratory viruses, such as PCV2 or the influenza virus, which can mimic the respiratory clinical signs associated with PRRS [81]. Furthermore, in a globalized animal breeding market, where PRRS is endemic and highly prevalent in farms, (estimated 60-80% in the U.S, and up to 79% in mainland Europe), and where environmental conditions are difficult to improve, eradication of the virus has proven to be challenging [82–84]. Selective breeding for tolerance is considered desirable when pathogen prevalence is high, when pathogen elimination has proven difficult and when pathogens can evolve rapidly to evade control measures that aim to interfere with the pathogen life-cycle [24]. All these cases

apply to PRRS. Therefore, improvement of tolerance of pigs to this ubiquitous virus may constitute a viable alternative to eradication programs, as it would allow pigs to maintain homeostasis despite infection [81]. However, tolerance would result in continued presence of the virus which could rebound and result in further pathogenesis in the host and threats to the herd. Thus, distinction between resistance and tolerance in genetic improvement programs is imperative if they have different effects on pathogen prevalence and evolution, as implied by theory [26,32].

Obtaining reliable tolerance estimates from natural disease outbreaks is extremely difficult due to the myriad of confounding factors (e.g. difference in exposure and onset of infection, differences in the individual immune status, co-infections), which can severely bias tolerance estimates and mask the underlying genetic signal [4,37]. For this reason, empirical evidence for genetic variation in host tolerance to infections stems primarily from challenge experiments in inbred lines of model species [5,30,85]. The PHGC challenge data constitute a unique data source for investigating the genetic basis and relative importance of host resistance and tolerance in outbred pigs' responses to virus infections, as it provides the required measures of both pathogen load and performance for large sample sizes, without the confounding factors inherent to field data. The analyses of these data however demonstrated that the limited data range produced in challenge experiments, together with other factors that affect the distribution of the data, such as genetic variance in host resistance, can easily blur the tolerance signal in multi-trait and reaction-norm models, and highlight the importance of performance records of non-infected relatives for obtaining accurate tolerance estimates.

Collecting equivalent performance records of non-infected relatives of the challenged individuals would be extremely valuable for establishing the relationship between tolerance and performance in the absence of infection, and for identifying shared or distinct genomic regions associated with these traits. A strong genetic correlation between these traits would imply that one could select for high performance at the nucleus to improve tolerance and performance in the more infectious commercial farms. In the current pig breeding structure, a direct data pipeline of performance measures between pigs in commercial farms experiencing disease outbreaks and those of related selection candidates in the almost pathogen free nucleus may be useful. Obtaining unbiased and comparable measures of within-host pathogen load from natural disease outbreaks constitutes the main challenge for producing reliable tolerance estimates from natural disease outbreaks [4]. A practically more feasible approach is to estimate genetic correlations between performance in clean and infectious environments and to include performance during disease outbreaks in the selection criterion [15,86], although this approach does not allow distinction between resistance and tolerance.

Based on resource-allocation theory and earlier findings, resistance and tolerance are conventionally considered as alternative host defense mechanisms to infections, leading to the notion of a trade-off between improving resistance and tolerance. A companion genome-wide association study on the same PHGC data indeed found different regions that were associated with tolerance and with resistance [87]. Emerging evidence from different studies suggests that both resistance and tolerance mechanisms may be required for effective host protection to infection and that the optimal host response to infection likely depends on a carefully timed interaction

between pathogen elimination (i.e. resistance) mechanisms and host mechanisms that promote tissue damage control and increase disease tolerance [85,88]. The aforementioned companion study identified several overlapping genomic regions associated with resistance and tolerance of pigs to PRRS and found that the WUR10000125 SNP, previously associated to confer greater resistance to PRRS (lower VL_{21}), also confers greater tolerance. Valuable insights about these interactions could be harnessed from the available longitudinal measures of pathogen burden and growth, e.g. by following the infection trajectories of individuals and target entire trajectories rather than resistance or tolerance for genetic improvement [41,85].

In order to target both resistance and tolerance in a sustainable breeding program, the epidemiological and evolutionary consequences of genetic selection in either or both traits combined must be studied in more detail. In particular, it needs to be determined whether evolutionary theory predicting a lower risk of pathogen evolution from selection for improved host tolerance rather than resistance hold in the case of PRRS; and to what extent genetically more resistant or tolerant pigs are also less infectious [60,89,90]. It is probable that control of PRRS and other infectious diseases by genetic selection is a "balancing act" [3], involving mechanisms associated with resistance and tolerance to provide the fittest pigs.

2.5.4 Conclusions

Using evidence from the available data alone suggests that growing piglets differ genetically in resistance but did not explicitly show evidence for genetic differences in tolerance to PRRSV infection. However, statistical constraints may have masked genetic variation in tolerance. Currently unknown genetic correlations between

performance under and in absence of PRRSV infection could reveal significant genetic variance in tolerance. Future studies are warranted to validate the results in this study for infections with the same and different strains of the PRRS virus, including vaccine strains. This study shows that genetics of tolerance is more difficult to analyze than genetics of resistance, and is therefore more difficult to target in genetic improvement.

Appendix 2.1

2.1 Simulating ADG in absence of infection

To assess the impact of additional measures of performance in the absence of infection (i.e. true intercept) on genetic parameter estimates for tolerance, average daily gain in absence of infection (ADG_0) was simulated for paternal half-sibs of the infected individuals in the dataset. ADG^0_{21} and ADG^0_{42} were simulated for both observation periods from 0-21dpi and from 0-42dpi, respectively, assuming a heritability of 0.4 [74] and a genetic variance half of that of the genetic variance for average daily gain under infection ADG_I (where ADG_I stands for ADG_{21} and ADG_{42}, respectively). A smaller variance was assumed for ADG_0 since part of the variation in ADG_I is due to different levels of VL. Since it was expected that genetic variance estimates for tolerance depend on the genetic relationship between ADG_0 and ADG_I (see Figure 1 in main manuscript), different values of genetic correlations (r_g) between ADG_0 and ADG_I were simulated representing weak, moderately low, moderately strong and strong genetic relationships between the traits with corresponding values of 0.05, 0.30, 0.60 and 0.90, respectively.

The simulated values for ADG^0_{21} and ADG^0_{42} were produced as follows:

2.1.1 Step 1: Calculate sire breeding values for growth under infection (ADG_I)

Firstly, sire breeding values were calculated for ADG under infection (ADG_I) from 0-21 or 0-42 dpi (ADG_{21} and ADG_{42}, respectively) using sire estimated breeding values (EBV_S) from the level model, described in main text, using the following equation:

$$BV_s^{ADG_I} = EBV_S^{ADG_I} + PE, \quad [1]$$

Where the prediction error (PE) was sampled using:

$$\text{PEV } = \sigma^2_{A_{ADG_I}} - var\left(EBV(ADG_I)\right),$$

where $\sigma^2_{A_{ADG_I}}$ is the additive genetic variance of ADG under infection, and sire EBVs were acquired from the level model.

2.1.2 Step 2: Calculate sire breeding values for growth in the absence of infection (ADG_0)

Breeding values were calculated for ADG in the absence of infection (ADG_0) for 0-21 or 0-42 dpi (ADG^0_{21} or ADG^0_{42}, respectively) using the following equation:

$$BV_s^{ADG_0} = b_{BV^{ADG_0},BV^{ADG_I}} BV_s^{ADG_I} + RES_1, \quad [2]$$

Where the regression coefficient $b_{BV^{ADG_0},BV^{ADG_I}}$ was calculated as:

$$r_g \frac{\sigma_{A_{ADG_0}}}{\sigma_{A_{ADG_I}}}$$

where r_g is the simulated genetic correlation between ADG_0 and ADG_I (simulated as either 0.05, 0.30, 0.60 or 0.90), and $\sigma_{A_{ADG_0}}$ and $\sigma_{A_{ADG_I}}$ are the square roots of the additive genetic variance in ADG_0 and ADG_I, respectively, and the residual (RES_1) was sampled from:

$$N(0, \left(1 - r_g\right)^2 \sigma^2_{A_{ADG_0}})$$

2.1.3 Step 3: Calculate phenotypic values for growth without infection for progeny (half sibs) (P_{Prog})

Finally, phenotypic values of growth without infection were simulated for one paternal half-sib of each infected individual in the data set, as follows:

$$ADG_0 = \frac{1}{2} BV_s^{ADG_0} + RES_2, \quad [3]$$

where RES_2 was sampled from:

$$N(0, (1 - \frac{1}{4} h^2) \sigma^2_{p_{ADG_0}})$$

where phenotypic variance of ADG_0 ($\sigma^2_{p_{ADG_0}}$) was calculated as:

$$\frac{\sigma^2_{A_{ADG_0}}}{h^2_{ADG_0}}$$

Following this approach, phenotypes for ADG^0_{21} and ADG^0_{42} were generated for one half-sib per infected individual in the dataset, thus doubling the size of the dataset. For each simulated population, 10,000 replicates were generated. Table 1 below shows the means (and standard errors) for the simulated true ADG_0 breeding values, the genetic variances and correlations between true breeding values of ADG_0 and ADG_I.

Table 1. Mean sire breeding value and genetic variance of simulated performance in absence of infection (ADG_0), where r_g is simulated genetic correlation between ADG_0 and ADG under infection (ADG_I) for the 21- or 42-day observation periods. Correlations between breeding values of ADG_0 and ADG_I are also shown. Standard error over 10,000 replicates is shown in brackets.

r_g	Mean ($BV_s^{ADG_0}$)	$\sigma^2_{A_{ADG_0}}$	Correlations between $BV_s^{ADG_0}$ and $BV_s^{ADG_I}$
	21 day observation period		
0.05	2.34E-04 (1.20E-04)	9.36E-04 (1.55E-05)	0.05 (0.14)
0.30	1.04E-03 (1.15E-04)	1.26E-03 (1.62E-05)	0.30 (0.12)
0.60	2.07E-03 (9.41E-05)	3.39E-03 (1.61E-05)	0.60 (0.08)
0.90	3.01E-03 (5.23E-05)	3.44E-03 (1.75E-05)	0.90 (0.02)
	42 day observation period		
0.05	1.81E-04 (1.47E-04)	1.32E-03 (2.33E-05)	0.05 (0.14)
0.30	1.46E-03 (1.40E-04)	1.35E-03 (2.32E-05)	0.30 (0.12)
0.60	2.83E-03 (1.16E-04)	1.51E-03 (2.65E-05)	0.60 (0.08)
0.90	4.34E-03 (6.42E-05)	1.44E-03 (2.76E-05)	0.90 (0.02)

Chapter 3

Harnessing longitudinal information to identify genetic variation in tolerance of pigs to Porcine Reproductive and Respiratory Syndrome virus infection

3.1 Introduction

In chapter 2, evidence for genetic variation in tolerance of pigs to PRRS was inconclusive due to limited statistical power to distinguish between parameters related to overall growth response under infection and those related to tolerance. It was concluded that more measurements, in particularly a wider spread of measurements, would be required to resolve this confounding and to accurately estimate genetic variance in tolerance. Estimation of genetic variance in tolerance in chapter 2 was based on the cumulative host response over specific periods of infection, such as from 0 to 21 days post-infection (dpi) representing the acute phase of infection, or from 0 to 42 dpi representing the prolonged period until the end of the experiment. This resulted in single measurements for cumulative virus load and growth for each animal included in the random regression sire models.

The hypothesis of this chapter is that genetic variance in tolerance may be identified by considering individual stages of infections and by utilizing information from the repeated individual measures of viremia and growth in the statistical models. This is built on previous research, which indicated that the genetic expression of host response traits is dynamic, changing throughout time-course of infection [37,85]. Similarly, tolerance has been identified as likely being time-dependent [31,37,39,85].

Recent studies of the PHGC data have fitted mathematical Woods functions through the repeated viremia measurements of individuals to create a viremia curve profile for each individual that describes within-host viremia changes over the duration of infection (figure 1) [62]. These Woods function viremia profiles allow partitioning of host response into different stages of infection, such as the phase of viremia increase towards peak viremia or phases related to viremia clearance, respectively [55,62].

The aims of this chapter were to estimate genetic variation in tolerance at different stages of infection, defined by viremia curve characteristics. Secondly, by using repeated measures for each individual at each stage of infection, this research aimed to increase the statistical power to detect genetic variation in tolerance to PRRS over the 42-day observation period. Thirdly, this study aimed to identify the genetic relationship between tolerance and resistance, and whether this changed over different stages of infection. Finally, this study aimed to investigate the association between a previously identified quantitative trait locus (QTL), shown to have significant effects on cumulative VL and growth, with tolerance, and whether the strength of this association changed through stages of infection.

3.2 Methods

3.2.1 Data

The data analysed in this chapter came from the same pigs of the nine PRRS Host Genetics Consortium (PHGC) PRRSV challenge trials as used in chapter 2, with the exception that pigs that had experienced a rebound in viremia were omitted from the analyses presented in this chapter. Rebounders were omitted to remove noise associated with the later stage of infection; this way all pigs considered here

experienced a gradual viremia decline after peak viremia had been reached. From the resulting 1011 pigs (offspring of 49 sires), repeated growth and viremia measures were used in the tolerance analysis of this chapter. Please see chapter 2 for more information on the infection protocols and trials.

3.2.2 The WUR10000125 single nucleotide polymorphism (SNP) and associated resistance and growth genotypes

All PHGC pigs were genotyped with 60K Beadchips [46]. Previous genome wide association studies performed on the same dataset as used here had identified a quantitative trait locus (QTL) on chromosome 4, where the single nucleotide polymorphism (SNP) WUR10000125 (WUR) was strongly associated with differences in growth under infection and in resistance. More specifically, the favourable B allele corresponded to both higher resistance to PRRSV and faster growth over the 21-day infection period (indicated by lower within-host viral load (VL21)). Individuals with AB genotype had 4.5% lower VL and were 2kg heavier than individuals with the AA genotype, with the genomic window containing the WUR genotype explaining 13.2% and 9.14% of the genetic variance for VL and growth, respectively [55,57,58]. Thus, using the Illumina A/B genotype reference system, each pig of this study was assigned to one of the three WUR genotypes; the number of pigs with AA, AB and BB were n=689, 286 and 36, respectively.

3.2.3 Defining stages of infection

The experimental observation period (0-35 dpi for trials 7 and 8, and 0-42 dpi for all other trials) was partitioned into three stages, i.e. early, mid and late stage of infection.

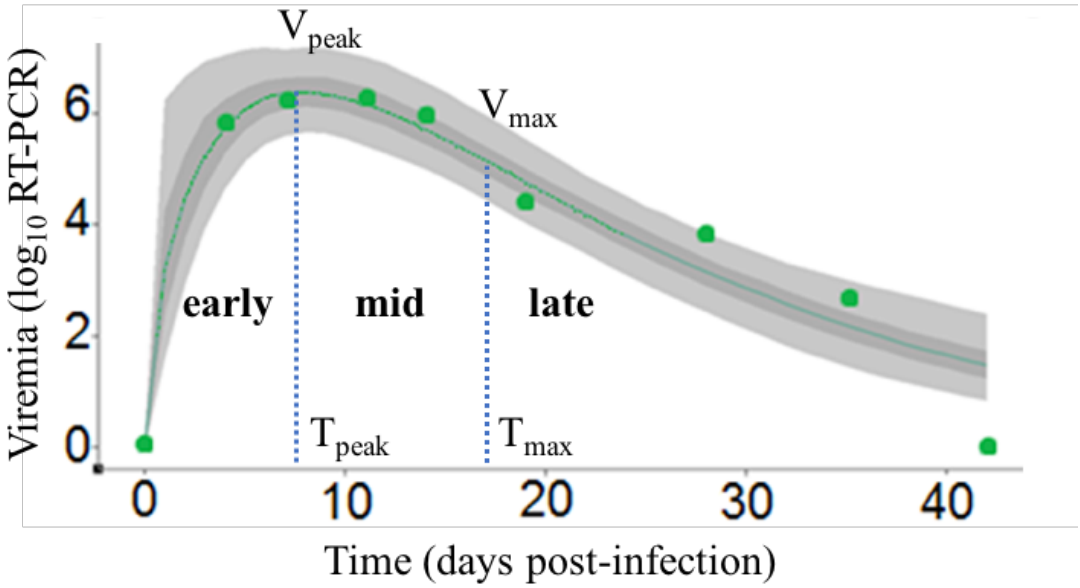
These stages were defined according to viremia profile characteristics, so that the early stage corresponds to the phase of rapid viremia increase towards peak viremia, the mid stage corresponds to the initial phase of rapid post-peak viremia decline, and the late stage corresponds to the phase where viremia continues to decline but at a decreasing rate. The mathematical definition of these stages was achieved by fitting individual viremia profiles, defined by the Woods function described below, to the repeated viremia measures of each individual pig, as illustrated in Figure 1 [55,62]. The Woods function, which had been previously shown to be a good fit to the viremia data of each individual pig that did not experience viremia rebound [62], is given by:

$$V(t) = a_1 t^{b_1} e^{-c_1 t}, \quad [1]$$

where $V(t)$ is $\log_{10}$ scaled serum viremia at t dpi; a_1 affects the magnitude of points on the curve; b_1 is indicator of the rate of increase to peak viremia, c_1 is an indicator of the rate of decline after peak viremia. Using the analytical expression [1], two critical time points can be derived by differentiation: $\mathrm{T}_{\mathrm{peak}_i} = \frac{\widehat{b_{1_i}}}{\widehat{c_{1_i}}}$ for each individual i, the time when viremia reaches its maximum value and $T_{max_i} = \frac{\widehat{b_{1_i}} + \sqrt{\widehat{b_{1_i}}}}{\widehat{c_{1_i}}}$, the time post peak viremia when the rate of viremia decline reaches its maximum (i.e. the inflection point of the Woods function).

These critical time points (T_{peak} and T_{max}, respectively) were used to define lower and upper boundaries of each stage of infection in this study. Stages of infection were thus defined from time of initial infection (0 days post-infection (dpi)) to T_{peak} (*early*); from T_{peak} to T_{max} (*mid*), and from T_{max} to end of observed infection period, which was at 35 dpi for trials 7 and 8, and 42 dpi for all other trials (*late*).

Figure 1. Illustration of an individual viremia profile used to define stages of infection, where the green dots represents log-transformed viremia measures, and the green line represents the fitted Woods function over the time-course of infection (up to 42 dpi) (with 90% and 95% credibility intervals in grey) for one individual. *Early* stage of infection is defined from initial infection (0 dpi) to time where peak viremia (V_{peak}) is reached (T_{peak}, in this example, approximately 9 dpi); *mid* stage of infection is defined from T_{peak} to T_{max} (t_{max} in this example, approximately 16 dpi); and *late* stage of infection, defined from T_{max} to end of experiment (42 dpi, or 35 dpi for trials 7 & 8). Figure modified from Islam et al 2013 [62].



As the function was fit to each individual separately, separate values for the parameters T_{peak} and T_{max} were obtained for each individual (Table 1). Table 1 shows the variation in the corresponding time periods for each stage of infection. Average duration of each stage of infection was 6.90, 9.19 and 14.26 dpi for early, mid and late stages of infection, respectively, corresponding to average T_{peak} and T_{max} of 6.89 and 16.16, respectively.

Table 1. Descriptive statistics of duration of each stage of infection, where *early* stage is defined as initial infection (0 dpi) to T_{peak}; *mid* stage is defined as T_{peak} to T_{max}, and *late* stage of infection is defined from T_{max} to end of the observation period (35 dpi for trials 7 an 8 and 42 dpi for all other trials). Descriptive statistics of T_{peak} and T_{max} are also shown. All measures below are in dpi.

Stage of infection	Mean	Standard deviation	Minimum	Maximum
Early	6.90	1.32	3.01	10.60
Mid	9.19	2.11	4.04	14.87
Late	14.26	0.60	13.53	14.92
$\mathbf{T_{peak}}$	6.89	1.32	3.01	10.60
$\mathbf{T_{max}}$	16.16	2.38	10.01	26.97

Using these definitions of infection stage, resistance was then quantified, as previously, as the inverse of the cumulative log-transformed viral load in blood (VL) over a given stage of infection [5,4,55,6]. This was calculated as the area under the curve of the log-transformed viremia estimates throughout the observation period, obtained by numerical integration the Woods function, described in model [1] above over the corresponding time periods, yielding estimates for VL_{early}, VL_{mid} and VL_{late}. To calculate the corresponding ADG for each stage of infection, a linear spline curve was fitted through weekly body weights (BW) for each individual using the *smooth.spline* function in R (Dipley, BD and Maechler, M). BW was interpolated at time of peak viremia (T_{peak}) and time of maximal viremia decay (T_{max}) using the *predict* function in R. From this, ADG_{early}, ADG_{mid} and ADG_{late} were calculated by dividing the difference in body weight at the start and end of the stage in consideration by the corresponding duration.

3.2.4 Estimating genetic variance in tolerance slope at each stage of infection

To estimate genetic variance in tolerance, a random regression reaction norm model was applied to the data, whereby the origin of the reaction-norms was centred at the mean viral load for each stage of infection, thus providing only variance component estimates for level (ADG at mean VL) rather than vigour (ADG at zero VL) for each infection stage. As in chapter 2, the linear random regression sire model (RRM) for ADG on centred values of VL, which will be referred to as the level-slope model, was used to identify genetic variance in tolerance at each stage of infection, where ADG at early, mid and late stages of infection were regressed on VL at the corresponding stages of infection. Additionally, to account for a potential time lag between changes in viremia and changes in growth, ADG at mid and late stages of infection were also regressed on VL at earlier stages of infection.

As in chapter 2, to test the significance of sire effects on level and slope and to determine which of the models best described the data, the model fit of the level-slope model was compared with that of hierarchical models without any additive genetic effects, and with only sire effects for level. Significance of each random effect was assessed using the likelihood ratio test (LRT) [71], with the LRT test statistics below assumed to follow a χ^2 distribution, with 1 degree of freedom for inclusion of an additional sire effect (e.g. null to level model, including sire effect) and a mixture of 1 and 2 degrees of freedom for additional sire slope effects and covariance (for example, from level to level-slope model) [72,73].

3.2.5 Estimating genetic variance in tolerance slope over all stages of infection using a repeated measurement model

To increase statistical power to identify genetic variance in tolerance across all stages of infection, repeated measures of ADG and VL at each stage of infection were used (i.e. to triple the size of the dataset) in the univariate following random regression sire model (referred to as the level-slope repeated measure model):

$$\mathbf{y} = \mathbf{Xb} + \mathbf{X_{VL}b_s} + \mathbf{Za_i} + \mathbf{Z_{VL}a_s} + \mathbf{Wc} + \mathbf{Up} + \mathbf{Ml} + \mathbf{e} \quad [2]$$

where $\mathbf{y}$ is the vector of repeated ADG measures of early, mid and late stages of infection; $\mathbf{b}$ is the vector of fixed effects outlined in model [2], with BW at each stage of infection and age included as additional fixed covariates to account for body weight at the different stages of infection and variation in age at start of infection, respectively; and $\mathbf{b_s}$ is the population average tolerance slope; $\mathbf{a_i}$ and $\mathbf{a_s}$ are the sire effects on level and on tolerance slope, respectively, assumed to follow a multi-variate normal distribution with mean zero and $\text{Var}\begin{bmatrix}\mathbf{a_i}\\ \mathbf{a_s}\end{bmatrix} = \frac{1}{4}\mathbf{G_{RN}} \otimes \mathbf{A}$, with $\mathbf{G_{RN}} = \begin{bmatrix}\boldsymbol{\sigma}^2_{\mathbf{a_i}} & \boldsymbol{\sigma}_{\mathbf{a_i a_s}}\\ \boldsymbol{\sigma}_{\mathbf{a_i a_s}} & \boldsymbol{\sigma}^2_{\mathbf{a_s}}\end{bmatrix}$, where $\boldsymbol{\sigma}^2_{\mathbf{a_i}}$ and $\boldsymbol{\sigma}^2_{\mathbf{a_s}}$ are the variance of $\mathbf{a_i}$, and $\mathbf{a_s}$, respectively, $\boldsymbol{\sigma}_{\mathbf{a_i a_s}}$ is the covariance between sire effects for level and slope; **c** is the variance of permanent environmental effect; the random pen and litter effects **p** and **l** were fitted as described in chapter 2; **e** is the vector of error terms, with $\text{var}(\mathbf{e}) = \mathbf{I} \otimes \mathbf{R}$, where **R** is the variance-covariance matrix. To prevent upward bias genetic variance in slope caused by heterogeneous variance error [35,91], 2 heterogeneous residual structures based on data ordered from low to high VL (n=1517 and 1516, respectively).

$\mathbf{X_{VL}}$ and $\mathbf{Z_{VL}}$ are the incidence matrices for population average tolerance slope and those associated with each sire, respectively, consisting of individual VL measures,

and **X** is the incidence matrix for the fixed effects (including VL as fixed covariate) and **Z** is the incidence matrix for the random sire effect on level (**Z**); and **W**, **U** and **M** are the incidence matrices for permanent environmental, pen and litter effects, respectively. Model fit was compared to null and level repeated measure models using LRTs.

3.2.6 Estimating the relationship between resistance and tolerance

To determine the relationship between resistance and tolerance at each stage of infection, phenotypic and genetic correlations of, and between, traits were estimated using the following bivariate mixed sire model in ASReml 3.0 [63]:

$$\begin{bmatrix}\mathbf{y_1}\\ \mathbf{y_2}\end{bmatrix} = \begin{bmatrix}X_1 & 0\\ 0 & X_2\end{bmatrix}\begin{bmatrix}\mathbf{b_1}\\ \mathbf{b_2}\end{bmatrix} + \begin{bmatrix}Z_1 & 0\\ 0 & Z_2\end{bmatrix}\begin{bmatrix}\mathbf{a_1}\\ \mathbf{a_2}\end{bmatrix} + \begin{bmatrix}U_1 & 0\\ 0 & U_2\end{bmatrix}\begin{bmatrix}\mathbf{p_1}\\ \mathbf{p_2}\end{bmatrix} + \begin{bmatrix}M_1 & 0\\ 0 & M_2\end{bmatrix}\begin{bmatrix}\mathbf{l_1}\\ \mathbf{l_2}\end{bmatrix} + \begin{bmatrix}\mathbf{e_1}\\ \mathbf{e_2}\end{bmatrix} \quad \textbf{[3]}$$

where $\mathbf{y_1}$ and $\mathbf{y_2}$ are ADG and VL for different stages of infection (early, mid or late), respectively; $\mathbf{b_1}$ and $\mathbf{b_2}$ are a vectors of fixed effects/covariates for ADG or VL, containing the interaction of experimental trial and parity of the dam, when offspring were born (trial-by-parity), sex of the offspring, and BW_0 (for ADG only) and age, which were fitted as fixed covariates, VL at stage of infection was also fitted as a fixed covariate for ADG to fit an average slope; $\mathbf{a_1}$ and $\mathbf{a_2}$ are the sire effects for ADG and VL, respectively, where $\mathbf{a_1}$ includes the sire effect on slope and level, and were assumed to follow a multi-variate normal distribution with mean zero and $\text{Var}\begin{bmatrix}\mathbf{a_1}\\ \mathbf{a_2}\end{bmatrix} =$ $G \otimes A$, with $G = \begin{bmatrix}\sigma^2_{a_l} & \sigma_{a_l a_s} & \sigma_{a_l a_{VL}}\\ \sigma_{a_l a_s} & \sigma^2_{a_s} & \sigma_{a_s a_{VL}}\\ \sigma_{a_l a_{VL}} & \sigma_{a_s a_{VL}} & \sigma^2_{Va_{VL}}\end{bmatrix}$, where $\sigma^2_{a_l}$, $\sigma^2_{a_S}$, and $\sigma^2_{a_{VL}}$ are the sire variances of level, slope and VL, respectively, respectively, $\sigma_{a_l a_s}$ is the covariance between the sire effects on level and slope, $\sigma_{a_l a_{VL}}$ is the covariance between sire effects

on level and VL, $\sigma_{a_s a_{VL}}$ is the covariance between sire effects on slope and VL; and A is the numerator relationship matrix among sires; $\mathbf{p_1}$ and $\mathbf{p_2}$ are vectors of pen effects nested within a trial for each trait, with Var $\begin{bmatrix}\mathbf{p_1}\\ \mathbf{p_2}\end{bmatrix} = \mathrm{I} \otimes \mathrm{K}$, where I is the identity matrix and K is the corresponding variance-covariance matrix of pen effects for the different traits; $\mathbf{l_1}$ and $\mathbf{l_2}$ are the vectors of litter effects for each trait, with Var $\begin{bmatrix}\mathbf{l_1}\\ \mathbf{l_2}\end{bmatrix} = \mathrm{I} \otimes \mathrm{L}$, with the corresponding variance-covariance matrix L; $\mathbf{e_1}$ and $\mathbf{e_2}$ are the vectors of error terms for each trait, with Var $\begin{bmatrix}\mathbf{e_1}\\ \mathbf{e_2}\end{bmatrix} = \mathrm{I} \otimes \mathrm{R}$, where R is the variance-covariance matrix for the residual effects for each trait; and X_1 and X_2, Z_1 and Z_2, U_1 and U_2, and M_1 and M_2 are the incidence matrices for the fixed, animal, pen and litter effects, respectively, for each trait. For the repeated measurement model, model [3] was applied to the vectors of repeated measurements of ADG and VL at each stage of infection. Additionally, BW at each stage of infection was fitted as a fixed covariate, and permanent environmental effects were also included as an additional random effect (as in model [2]).

3.2.7 Association of the WUR resistance genotype with tolerance

Association of the WUR genotype with tolerance was estimated by including genotype and genotype-by-VL as fixed covariates in models [2] and [3], providing estimates of least-square means for the genotype effect on growth and tolerance slope, respectively. Significance of the associations of the genotype with traits were assessed using an f-test, where $p<0.05$ was the significance threshold.

3.3 Results

3.3.1 Single measures model for different stages of infection

3.3.1.1 Evidence for genetic variation in tolerance at different stages of infection

The null model including VL as a fixed covariate identified a statistically significant linear association between growth and VL at only mid stage VL ($p=0.02$, $p<.0001$ for ADG_{mid} and ADG_{late}, respectively), and at late stage VL $p<.0001$ for ADG_{late}. Growth at any stage of infection was not associated with early stage VL. Only models including mid and late stage VL are therefore considered from hereon. Population average tolerance slope estimates were negative but generally very flat: for ADG_{mid} and ADG_{late} on VL_{mid} growth rate only decreased on average by -1.42E-03 (4.65E-04) and -2.43E-03 (4.85E-04) kg/day per unit of VL increase, respectively, and for ADG_{late} on VL_{late} the decrease in growth rate was -1.91E-03 (4.39E-04) kg/day per unit VL increase. The log-likelihood of the model was significantly improved when genetic effects (random sire effects) were included in the model (level model) at all stages of infection ($p<.0001$), indicating significant genetic variance in growth of pigs infected with PRRSV. Models with sire effects on both level and tolerance slope, as well as a genetic covariance between them, yielded a significantly better fit than the null model at all stages of infection ($p<.0001$). However, a statistically significant improvement of the model fit of the level-slope model over the level model was only observed when ADG at late stage of infection was regressed on VL at mid stage of infection ($p<0.05$) (Table 2, bold). All other models constrained covariance between level and slope at zero, similar to what was seen in the level slope models investigated in chapter 2.

Genetic variances of level were relatively stable across all stages of infection for both level-only model and level-slope model (Table 2). Genetic variance estimates for

tolerance slope varied greatly depending on stage of infection, by up to four orders of magnitude (Table 2). As would be expected, genetic variance for slope was largest when ADG_{late} was regressed on VL_{mid} (1.77E-05 (1.83E-05)). Similarly, covariances between level and slope varied greatly in magnitude depending on stage of infection, but all covariances were negative. The genetic covariance estimate for the ADG_{late} - VL_{mid} model was noticeably stronger than for the other models (Table 2).

Table 2. **Genetic variance components for ADG (kg/d) at mid and late stage of infection regressed on VL at mid and late stage of infection, respectively**. The last column denotes the p-value of the LRT used to test whether the level-slope model significantly improves the model fit over the level model. For definition of stages of infection, see text.

		Level-only model	Level-Slope model			Model fit
VL	ADG	Level	Level	Covariance	Slope	P-value
Mid	Mid	3.47E-03 (1.56E-03)	3.39E-03 (1.56E-03)	-1.02E-11 (5.73E-13)	3.40E-06 (6.24E-06)	0.743
	Late	2.33E-03 (1.24E-03)	2.52E-03 (1.30E-03)	-6.90E-05 (9.61E-05)	**1.64E-05 (1.02E-05)**	**0.024**
Late	Late	1.96E-03 (1.15E-03)	1.97E-03 (1.15E-03)	-8.47E-13 (4.77E-14)	1.13E-09 (6.35E-11)	1.000

3.3.1.2 Relationship between level, slope and resistance at different stages of infection

For ADG_{late} regressed on VL_{mid}, genetic correlation between level and tolerance slope was -0.38 (0.45); between level and VL, the genetic correlation was -0.40 (0.37); and between VL and slope, the genetic correlation was 0.31 (0.50). Due to high standard errors, these genetic correlations were not significantly different from zero. Phenotypic

correlation was generally lower than genetic correlation: between level and VL was 0.27 (0.05), indicating that individuals that had higher VL were likely to grow faster growth under infection (i.e. less resistant individuals grew faster); and between slope and VL was 0.23 (0.06), indicating that individuals with higher VL tended to be more tolerant (i.e. that more resistant individuals were less tolerant and vice versa). Phenotypic correlations between level and slope could not be calculated as there was no covariance between level and slope for residual or additional random effects.

3.2.1.3 Association of WUR genotype with tolerance across individual stages of infection

The WUR genotype did not have an association with tolerance slope at any stage of infection ($p>0.05$ for each stage).

3.3.2 Repeated measures model:

3.3.2.1 Genetic variation in tolerance across all stages of infection

A very weak, but statistically significant association between growth and VL was detected by the repeated measurement model ($p=0.01$), with an average slope of 8.17E-04 (4.10E-04) kg/day per unit VL increase. Similarly, a weak but significant association between BW at each stage of infection with growth under infection was found ($p<.0001$) with a positive regression coefficient of 2.74E-02 (1.08E-03). In other words, pigs that are 1kg heavier at a specific stage of infection tended to grow on average by 20g/d faster.

Table 3. **Variance components (with standard errors) for ADG (kg/d)**, obtained by the repeated measures model including ADG and VL measures for all three defined stages of infection (early, mid and late), together with the LogLikelihood value associated with the different models. Variance components estimated from random regression models: null model, containing no genetic effect; level-only model, containing only the overall sire effect on growth under infection; and level-slope model, containing sire effects on level and slope, respectively. All other fixed effects/covariates and random effects were identical between models. Variance component estimates for the other random effects (e.g. pen, litter, common environmental, residuals) were identical between the models and not shown.

	Null repeated measure model	**Level repeated measure model**	**Level-Slope repeated measure model**
Level	.	1.33E-03 (5.76E-04)	1.80E-03 (7.39E-04)
covariance	.	.	8.10E-05 (5.05E-05)
Slope	.	.	1.61E-05 (4.36E-06)
LogLikelihood	**4717.39**	**4725.31**	**4767.28**

The log-likelihood of the repeated measures model improved significantly when genetic effects (i.e. random sire effects) were included in the model (level model) ($p<.0001$) (Table 3). Furthermore, a statistically significant increase in the log-likelihood of the level-model was observed when genetic effects for level, slope and covariance between the latter were included ($p<.0001$), indicating statistically significant genetic variation in tolerance. Similar to the single measures model, variance estimates were relatively robust across the different models.

3.3.2.2 Relationship between level, tolerance slope and resistance across all stages of infection

Genetic correlation between level and tolerance slope was 0.51 (0.23), indicating that individuals that genetically had faster growth rate under infection were likely to be genetically more tolerant. Between level and VL, the genetic correlation was -0.92 (0.27), indicating that individuals with higher genetic resistance tended to grow faster under infection. Between VL and tolerance slope, the genetic correlation was -0.79 (0.31), indicating a favourable genetic relationship between resistance and tolerance. Phenotypic correlation was generally lower than genetic correlation: between level and VL it was -0.13 (0.03), and between VL and tolerance slope it was -0.08 (0.02).

3.3.2.3 Association of WUR genotype with tolerance across all stages of infection

The WUR genotype was suggestively associated with tolerance over all stages of infection, (p=0.045). As with resistance, the B allele also conferred higher tolerance, where individuals with the AB genotype had a tolerance slope 1.6% ± 0.88% higher than individuals with the AA genotype. As there were only 36 individuals with the BB genotype, these results are not reported.

3.4 Discussion

3.4.1 Summary of findings

In contrast to the inconclusive evidence for genetic variation in tolerance of pigs to PRRS obtained in the precious chapter, this study found strong evidence for significant genetic variation in tolerance of pigs to PRRS. This was achieved by partitioning the entire infection period into three different stages of infection based on individual

viremia profile characteristics (early, mid and late), and by either assessing tolerance genetics at each particular stage, or by utilizing the information from longitudinal data in a repeated measurement model.

According to the single stage models, significant genetic variance in tolerance could only be detected when late stage ADG regressed against mid stage VL. This would indicate that genetic variation in tolerance is sensitive to the timing of measurements, and that a unit reduction in VL at the rapid phase of post-peak viremia decline is associated with genetic differences in pigs' growth responses at the late stage of infection: whereas some pigs may be genetically predisposed to experience compensatory growth, other pigs are genetically predisposed to suffer a prolonged growth depression. At all other stages of infection, lack of evidence for genetic variance in tolerance was accompanied by similar statistical constraints as observed in chapter 2 (e.g. confounding between level and slope, convergence issues).

The statistical constraints associated with the single measures models for tolerance could be overcome by adopting a repeated measurement model that included viremia and growth measures at the different stages of infection. This model provided strong evidence for significant genetic variation in tolerance, as indicated by the considerably better model fit of the level-slope model over level-only model. Furthermore, it provided plausible non-zero estimates with low standard errors for the "overall" genetic variance in tolerance across all stages of infection, as well as for the genetic correlations between tolerance and resistance.

3.4.2 Association of the WUR genotype with tolerance

The WUR genotype was not statistically associated with tolerance slope at any stage of infection ($p>0.05$). This was in line with Rashidi et al. who, using a random regression model regressing ADG from 0 to 28 dpi on VL from 0 to 14 dpi, using a slope-only model (see chapter 2) with the same PHGC dataset, found that WUR was not associated with tolerance slope. Instead, only a suggestive association with a genomic region on chromosome 1 [87]. However, the WUR genotype was suggestively associated with overall tolerance slope through all stages of infection. It is likely that WUR, which has been shown to be significantly associated with resistance, is associated with tolerance when genetic correlation between resistance and tolerance is strong and positive, as was the case with the repeated measurement level-slope model. This indicates that resistance and tolerance may be pleiotropic, with the WUR SNP affecting both traits. The association of WUR with resistance and growth at different stages of infection is further discussed in chapter 4.

3.4.3 The importance of stages of infection to identify genetic variation in tolerance

Partitioning the 42-day observation period into different stages of infection helped to resolve the previously encountered statistical constraints for detecting genetic variation in tolerance of pigs to PRRS in two ways: firstly, it provided repeated measurements of virus load and growth to boost the statistical power of the models, and secondly it helped to focus the analysis to a stage of infection where genetic differences in tolerance are most pronounced.

In this study, the individual stages of infection were defined with help of the mathematical Woods function, which had been previously shown to provide a close fit

to the viremia data of the majority of pigs in the PHGC trials that had not experienced a viremia rebound. Adopting this function allowed the definition of infection phases based on viremia profile properties, such as the phase associated with viremia increase towards peak viremia, the phase of rapid post-peak decline, and the later phase corresponding to more gradual viremia decline. The question arises how sensitive the results are to the definition of infection stages. For this purpose, the analyses of this chapter were repeated for an alternative definition of stages of infection, in which the stages were simply defined according to the following three fixed time periods: from 0-7 dpi (early), from 7-21 dpi (mid) and from 21-42dpi (late). As shown in the appendix to this chapter, the results were similar: The repeated measurement model provided strong evidence for significant genetic variation in tolerance to infection and the single measures models applied to VL and ADG at different stages of infection provided evidence for genetic variance in tolerance only for the late stage of infection (although in this case when ADG_{late} was regressed on VL_{late}, rather than on VL_{mid} as was the case when stages of infection were defined based on viremia characteristics). The results thus imply that the results are robust to different definitions of infection stages. Nevertheless, it should be mentioned that the statistical models with fixed time-period infection stages were more prone to numerical constraints during the fitting process than infection stages based on viremia characteristics. For example, estimates for genetic correlations between resistance and tolerance could not be obtained from the repeated measures model with infection stages defined by fixed time intervals.

Conventionally, host resistance and tolerance are considered as two alternative host defence mechanisms to fight infections. Based on this, one may expect a negative genetic correlation between both traits. However, in this study, the single measurement

model applied to individual infection stages and the repeated measurement model provided ambiguous results with regards to the genetic correlation between resistance and tolerance: whereas the single stage analysis was suggestive for an antagonistic genetic correlation between resistance and tolerance (estimated genetic correlation between VL and tolerance slope at mid stage of infection was 0.31 (0.5)), the multi-stage analysis indicated a strong favourable genetic relationship between both traits (estimated genetic correlation between VL and tolerance slope was -0.79 (0.31). The results would thus suggest that the relationship between resistance and tolerance is time dependent. More effective reduction in pathogen load at a particular stage of infection may temporarily decrease host tolerance, but the compromise between the traits may disappear when traits are considered over a longer period of infection.

3.4.4 Implications for breeding programmes and conclusions

The results of this study suggest that, in principle, genetic selection for greater tolerance of pigs to PRRS is possible. However, in practical terms, tolerance may still be difficult to target for genetic selection, given that multiple repeated measures are required for each individual, and information of performance and within-host pathogen load are required for multiple offspring of each sire. The strong favourable genetic relationship between resistance and tolerance, and the fact that the WUR resistance allele also seems to infer greater tolerance observed in this study would imply that both mechanisms likely work together to improve host resilience and that selection for greater resistance will simultaneously improve tolerance. Thus, genetic selection for resistance may be a sufficient strategy for improving host response to infection.

Appendix 3.1

3.1 Stage of infection defined by duration

In the main text, we defined stages of infection by individual viremia profile characteristics to identify genetic variation in tolerance at each stage of infection, and across all stages of infection using a repeated measures model. In this appendix, we used fixed time intervals as a proxy for different stages of infection. Stages of infection were defined from 0 to 7 dpi (early), 7 to 21 dpi (mid) and 21 to 42 dpi (35 dpi for trials 7 and 8) (late), as these durations similar to the average viremia peak, and maximal rate of viremia clearance used to define stages of infection in the main text. It should be noted, that although 14 dpi was closer to maximal rate of viremia clearance, previous research had used up to 21 dpi to define acute stage of infection, so 21 dpi onwards was used to define stages here. These duration-defined stages of infection were implemented using model [2] to identify genetic variation at each stage of infection and in repeated measures model [3] to identify genetic variation in tolerance across all stages (see main text).

3.2 Genetic variation in tolerance at different stages of infection

The null model including VL as a fixed covariate identified a statistically significant linear association between growth and VL ($p<.0001$). The log-likelihood of the model was significantly improved when genetic effects (random sire effects) were included in the model (level model) at all stages of infection ($p<.0001$), indicating significant genetic variance in growth of pigs infected with PRRSV. However, as indicated in

Table A1, significant improvement of the level-slope model fit over the level-only model only occurred at "late" stage of infection i.e. in the 21-42 dpi infection period.

Table A1. Genetic variance components for ADG (kg/d) at each stage of infection defined by fixed time periods from 0 to 7, 7 to 21 and 21 to 42 dpi. The last column denotes the p-value of the LRT used to test whether the level-slope model significantly improves the model fit over the level model. For definition of stages of infection, see main text.

		Level-only model	**Level-Slope model**			**Model fit**
VL	**ADG**	**Level**	**Level**	**Covariance**	**Slope**	**P-value**
0-7	**0-7**	2.82E-03 (1.41E-03)	2.84E-03 (1.42E-03)	4.55E-05 (1.64E-04)	5.39E-06 (3.29E-05)	0.871
7-21	**7-21**	3.08E-03 (1.34E-03)	3.08E-03 (1.34E-03)	2.55E-13 (1.40E-14)	1.02E-10 (5.59E-12)	1.000
	21-42	2.89E-03 (1.59E-03)	2.91E-03 (1.60E-03)	-1.17E-05 (1.01E-04)	6.86E-08 (3.76E-09)	0.966
21-42	**21-42**	2.28E-03 (1.40E-03)	2.90E-03 (1.64E-03)	7.33E-05 (9.55E-05)	**2.90E-03 (1.64E-03)**	**0.011**

3.3 Genetic variation in tolerance across all stages of infection

When the repeated measurement model was applied, the level-slope model was a significantly better fit than the level-only model ($p<.0001$) (Table A2), indicating genetic variation in tolerance across all stages of infection could also be detected when stages of infection were defined by fixed time periods rather than by time periods defined by viremia curve characteristics.

Table A2. Variance components for ADG (kg/d) over all duration-defined stages of infection (0 to 7, 7 to 21 and 21 to 42 dpi)

	Null repeatability model	**Level repeatability model**	**Level-Slope repeatability model**
Level	.	2.70E-03 (1.03E-03)	2.81E-03 (1.09E-03)
covariance	.	.	-2.20E-05 (4.91E-05)
Slope	.	.	9.80E-06 (2.86E-06)
Residual 1	3.06E-02 (1.15E-03)	3.06E-02 (1.15E-03)	2.96E-02 (1.11E-03)
Residual 2	1.62E-02 (6.44E-04)	1.62E-02 (6.39E-04)	1.52E-02 (6.03E-04)
LogLikelihood	**4059.50**	**4071.44**	**4106.00**

3.4 Relationship between level, slope and resistance from 21-42 dpi

Correlations between level, tolerance slope and resistance were estimated for the infection duration 21-42 dpi. Genetic correlation between level and slope was 0.25 (0.43). Between level and VL, the genetic correlation was -0.61 (0.37), indicating that, overall, individuals who were more resistant had faster growth; and between VL and slope, the genetic correlation was 0.13 (0.38). Phenotypic correlation was generally lower than genetic correlation: between level and VL was 0.37 (0.08), indicating that those that less resistant individuals grew faster under infection. Between VL and slope, the phenotypic correlation was 0.16 (0.05), indicating that more resistant individuals tended to be less tolerant and vice versa, however correlations were low, so this association was not strong. Phenotypic correlations between level and slope could not be calculated as there was no covariance between level and slope for residual or additional random effects.

3.5 Association of WUR genotype with tolerance slope

The WUR was not associated with tolerance slope at any stage of infection or across all stages of infection ($p>0.05$).

3.6 Relationship between level, slope and resistance across all duration stages of infection

The bivariate model ran into convergence issues, where covariances were constrained to boundary.

Chapter 4

Genome-wide association studies for resistance and growth at different stages of infection

4.1 Introduction

Genome-wide association studies (GWAS) have played an important part in genetic analyses of host response to Porcine Reproductive and Respiratory Syndrome (PRRS) virus infections. To date, various host genomic regions have been identified which were associated with growth and resistance of pigs to PRRSV infection [55,57,58]. In particular, using the same PHGC data as used in this thesis, a single nucleotide polymorphism (SNP) was identified that explains a considerable amount of the total genetic variance in resistance and weight gain of weaned piglets following experimental infection. This WUR10000125 (WUR) SNP covers a ~1MB region at *Sus Scrofa* chromosome (SSC) 4, where the favourable B allele corresponded to higher resistance (indicated by lower within-host viral load (VL)) and faster growth rate. Furthermore, genetic correlations between growth and resistance were shown to be positive and strong (ranging from 0.57 to 0.75 for different virus strains [22,55,57]). As such, selection for improved resistance would be expected to simultaneously improve growth under infection and vice versa.

In existing publications, the relationship between resistance and growth has, thus far, been exclusively assessed based on cumulative response over 0 to 21 or 0 to 42 days post-infection (dpi). However, as indicated in chapter 3, partitioning the infection

period into separate stages of infection, e.g. based on individual viremia curve characteristics, may provide novel insights about how the relationship between resistance and growth over time and the effect of the WUR genotype on it. This study aimed to define and assess the genetic basis of, and relationship between, growth and resistance across different stages of infection, defined by viremia curve characteristics. Secondly, this study investigated the association of the WUR genotype with these traits, and whether the strength of this association changed through stages of infection. Finally, by conducting a genome-wide association study (GWAS), this study aimed to identify genomic regions associated with each trait, and how this association differed across stages of infection. Note that it was not possible to carry at a GWAS for tolerance, as this was part of an alternative PhD project, and thus, due to legal reasons, outside the scope of this thesis.

4.2 Methods

4.2.1 Data

The same PHGC data as used in chapter 3 were used in this chapter. In this study, data from 1183 pigs were used [3]. Using the Illumina A/B genotype reference system (see chapter 3), the number of animals in each WUR genotype category were where n=689, 286 and 36 for AA, AB and BB genotypes, respectively. The genotypes were approximately equally distributed across trials and pens within trial.

4.2.2 Statistical models

4.2.2.1 Defining stages of infection

The stages of infection were defined as outlined in chapter 3. Briefly, based on individual viremia profile characteristics, stages were defined as *early* (from start of experimental infection to time of peak viremia); *mid* (from time of peak viremia to time of maximal viremia clearance rate); and *late* (from time of maximal viremia clearance rate to the end infection). Using these stages of infection, resistance was defined for each individual as the area under the viremia curve at early, mid and late stages of infection (VL_{early}, VL_{mid}, and VL_{late}, respectively); and growth was defined as average daily gain for the respective stages of infection (ADG_{early}, ADG_{mid} and ADG_{late}).

4.2.2.2 Genetic models

To determine the underlying genetic basis of ADG and VL at each stage of infection, heritabilities, phenotypic and genetic correlations of, and between, traits were estimated using bivariate mixed animal model in ASReml 3.0 [63]:

$$\begin{bmatrix}\mathbf{y_1}\\ \mathbf{y_2}\end{bmatrix} = \begin{bmatrix}\mathbf{X_1} & \mathbf{0}\\ \mathbf{0} & \mathbf{X_2}\end{bmatrix}\begin{bmatrix}\mathbf{b_1}\\ \mathbf{b_2}\end{bmatrix} + \begin{bmatrix}\mathbf{Z_1} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z_2}\end{bmatrix}\begin{bmatrix}\mathbf{a_1}\\ \mathbf{a_2}\end{bmatrix} + \begin{bmatrix}\mathbf{U_1} & \mathbf{0}\\ \mathbf{0} & \mathbf{U_2}\end{bmatrix}\begin{bmatrix}\mathbf{p_1}\\ \mathbf{p_2}\end{bmatrix} + \begin{bmatrix}\mathbf{M_1} & \mathbf{0}\\ \mathbf{0} & \mathbf{M_2}\end{bmatrix}\begin{bmatrix}\mathbf{l_1}\\ \mathbf{l_2}\end{bmatrix} + \begin{bmatrix}\mathbf{e_1}\\ \mathbf{e_2}\end{bmatrix} \quad \mathbf{[1]}$$

where $\mathbf{y_1}$ and $\mathbf{y_2}$ are ADG and VL for different stages of infection (early, mid or late), respectively; $\mathbf{b_1}$ and $\mathbf{b_2}$ are a vectors of fixed effects/covariates for ADG or VL, containing the interaction of experimental trial and parity of the dam, when offspring were born (trial-by-parity), sex of the offspring, and BW_0 (for ADG only) and age, which were fitted as fixed covariates; $\mathbf{a_1}$ and $\mathbf{a_2}$ are the additive genetic effects for ADG and VL, respectively, assumed to follow a multi-variate normal distribution with

mean zero and Var $\begin{bmatrix}\mathbf{a_1}\\ \mathbf{a_2}\end{bmatrix} = \mathbf{G} \otimes \mathbf{A}$, with $\mathbf{G} = \begin{bmatrix}\boldsymbol{\sigma}^2_{\mathbf{a_1}} & \boldsymbol{\sigma}_{\mathbf{a_1a_2}}\\ \boldsymbol{\sigma}_{\mathbf{a_1a_2}} & \boldsymbol{\sigma}^2_{\mathbf{a_2}}\end{bmatrix}$, where $\boldsymbol{\sigma}^2_{\mathbf{a_1}}$ and $\boldsymbol{\sigma}^2_{\mathbf{a_2}}$ are the variance of $\mathbf{a_1}$, and $\mathbf{a_2}$, respectively, $\boldsymbol{\sigma}_{\mathbf{a_1a_2}}$ is the covariance between animal effects for each trait, and **A** is the numerator relationship matrix among animals; $\mathbf{p_1}$ and $\mathbf{p_2}$ are vectors of pen effects nested within a trial for each trait, with Var $\begin{bmatrix}\mathbf{p_1}\\ \mathbf{p_2}\end{bmatrix} = \mathbf{I} \otimes \mathbf{K}$, where **I** is the identity matrix and **K** is the corresponding variance-covariance matrix of pen effects for the different traits; $\mathbf{l_1}$ and $\mathbf{l_2}$ are the vectors of litter effects for each trait, with Var $\begin{bmatrix}\mathbf{l_1}\\ \mathbf{l_2}\end{bmatrix} = \mathbf{I} \otimes \mathbf{L}$, with corresponding variance-covariance matrix **L**; $\mathbf{e_1}$, and $\mathbf{e_2}$ are the vectors of error terms for each trait, with Var $\begin{bmatrix}\mathbf{e_1}\\ \mathbf{e_2}\end{bmatrix} = \mathbf{I} \otimes \mathbf{R}$, where **R** is the variance-covariance matrix for the residual effects for each trait; and $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{Z}_1$, $\mathbf{Z}_2$, $\mathbf{U}_1$, $\mathbf{U}_2$, $\mathbf{M}_1$, and $\mathbf{M}_2$ are the incidence matrices for the fixed, animal, pen and litter effects, respectively. A bivariate model was used to provide more robust estimates than a multivariate model.

4.2.2.3 Association of WUR genotype with stages of infection

To ensure highest accuracy, associations of the WUR genotype for ADG and VL at early, mid and late stages of infection were estimated for each trait using the corresponding univariate animal models to model [2] (i.e. same fixed and random effects), but also including the WUR genotype as a fixed effect, providing estimates of least-square means for the genotype effect on ADG and VL. Statistical associations of the WUR genotype were assessed using an f-test, where $p<0.05$ was the significance threshold.

4.2.2.4 Genome Wide Association Study

Associations of SNP genotypes with individual traits were analysed fitting all SNP simultaneously in the genome-wide association study implemented using GenSel software [92]. In this approach, all SNPs were fitted simultaneously as random effects in an iterative manner. GenSel incorporates Bayesian methods, which generate posterior distributions for the parameter of interest (here effect size) based on information from both the phenotypic data and the prior information to identify relevant SNPs [47,49,50]. A burn-in period was specified, where 1000 initial iterations are removed from analysis [49,50]. Iterations then sample from the posterior distribution of previous iterations. If a SNP was fitted in the model ($\boldsymbol{\delta_i}$=1) and explained a larger proportion of the genetic variance than other fitted SNPs, then that particular marker was more likely to be included in the model in subsequent iterations. For this, a parameter π was specified, where π is the proportion of markers expected to have no significant effect on the trait. Bayesian variable selection assumes that a SNP has a zero effect on the trait of consideration, with probability equal to π, or has an effect sampled from a normal distribution with probability equal to 1-π. The parameter π can be calculated as (1 – number of animals genotyped/ number of SNPs), which in this study gave π=0.974. By specifying this value, the model was informed that 0.926% of markers will be associated with the phenotype. In this study, where 50,000 SNPs were included in the analysis, the model thus contained 463 markers fitted in each iteration of Bayesian variable selection.

As nearby genetic markers are in high LD, they will be in high LD with a causative mutation. Therefore, association of all SNPs within a genomic region were considered using a window-based approach, where effects of all SNPs were summed within each

window. Each non-overlapping window was defined by 1 Megabase (Mb) segments on the chromosome [93]. The GenSel software allowed only fixed effects and covariates to be included in the model. Thus previously fitted random effects (e.g. pen or litter) were included as fixed effects so that the only random effects were the SNP effects, leading to model [2] below:

$$\mathbf{y} = \mathbf{Xb} + \sum_{\mathbf{i=1}}^{\mathbf{k}} \mathbf{z_i}\, \alpha_i \delta_i + \mathbf{e}\ [\mathbf{2}]$$

where **y** is ADG or VL for each stage of infection (early, mid or late), respectively; **X** is the incidence matrix relating ADG or VL to fixed effects; **b** is a vectors of fixed effects/covariates for ADG or VL, containing the interaction of experimental trial and parity of the dam when offspring were born (trial-by-parity), sex of the offspring, pen, and BW0 (for ADG only) and age, which were fitted as fixed covariates (due to limitations of software in including additional random effects); $\mathbf{z_i}$ is a vector of the genotype covariates for SNP *i* for each trait, α_i is the allele substitution effect for SNP *i* for each trait, and δ_i is an indicator for whether SNP *i* was included ($\delta i = 1$) or excluded ($\delta i = 0$) in the model for a given iteration of the Monte Carlo Markov Chain. The prior probability of $\delta i = 0$ was set equal to $\pi = 0.974$, and implemented using Bayes-B [94]; and **e** was a vector containing residual errors.

All SNPs were fitted simultaneously as random effects in an iterative manner, over 50,000 iterations.

The posterior distribution of the proportion of genetic variance explained by each of these windows was derived as such: The total genomic estimated breeding value (GEBV) for each individual was computed by summing the product of the genotype covariate and the estimate of SNP effects across all evaluated SNPs for each

individual. For each window, the posterior means of the relevant SNP effects were multiplied by their corresponding genotype covariates and summed to calculate the window GEBV of each individual. The variance of these window GEBV across individuals, expressed as a proportion of the variance of the total GEBV across individuals, was used to identify genomic regions that were most strongly associated with the phenotype. To control for false positives, a threshold to define significant windows was set using a posterior probability of inclusion over 0.80 [49].

4.3 Results

4.3.1 Genetic basis and relationship between ADG and VL across stages of infection

ADG across all stages of infection ranged from -0.11 to 0.80kg/day (table 1). Mean ADG increased with stage of infection from 0.25 to 0.43kg/day at early and late stages of infection, respectively, indicating that individuals were, generally, able to grow faster at later stages of infection. Mean VL increased from 37.07 to 67.58 from early to late stage of infection, respectively. The VL increase is mainly attributed to the longer time period associated with the late stage of infection (on average 14.3 days compared to 6.9 and 9.2 days associated with the early and mid stage of infection, respectively). Whereas the phenotypic variance for growth is stable across the different stages, phenotypic variance in VL increases over time (table 1).

Table 1. Descriptive statistics of growth (ADG (kg/day) and viral load (VL (log_{10}RT-PCR) at early, mid and late stages of infection.

Trait	Mean	Standard deviation	Minimum	Maximum
ADG_{early}	0.25	0.13	-0.11	0.64
ADG_{mid}	0.29	0.13	-0.09	0.68
ADG_{late}	0.43	0.14	0.00	0.80
VL_{early}	37.07	6.49	16.36	57.21
VL_{mid}	58.37	9.21	35.64	87.11
VL_{late}	67.58	10.39	39.05	99.31

Both ADG and VL were found to be moderately heritable across all stages of infection (table 2). VL had lowest heritability at early stage of infection (from 0 dpi to peak viremia (0.17)) and reached highest heritability at the mid stage (between peak viremia and max viremia clearance (0.31)). Conversely, ADG was most heritable at earliest stage of infection (0.28) and least heritable by late (0.18).

Table 2. **Heritabilites, of growth (ADG) and viral load (VL) at early, mid and late stages of infection, and genetic and phenotypic correlations between the traits (lower and upper diagonal, respectively).** Standard errors are in brackets.

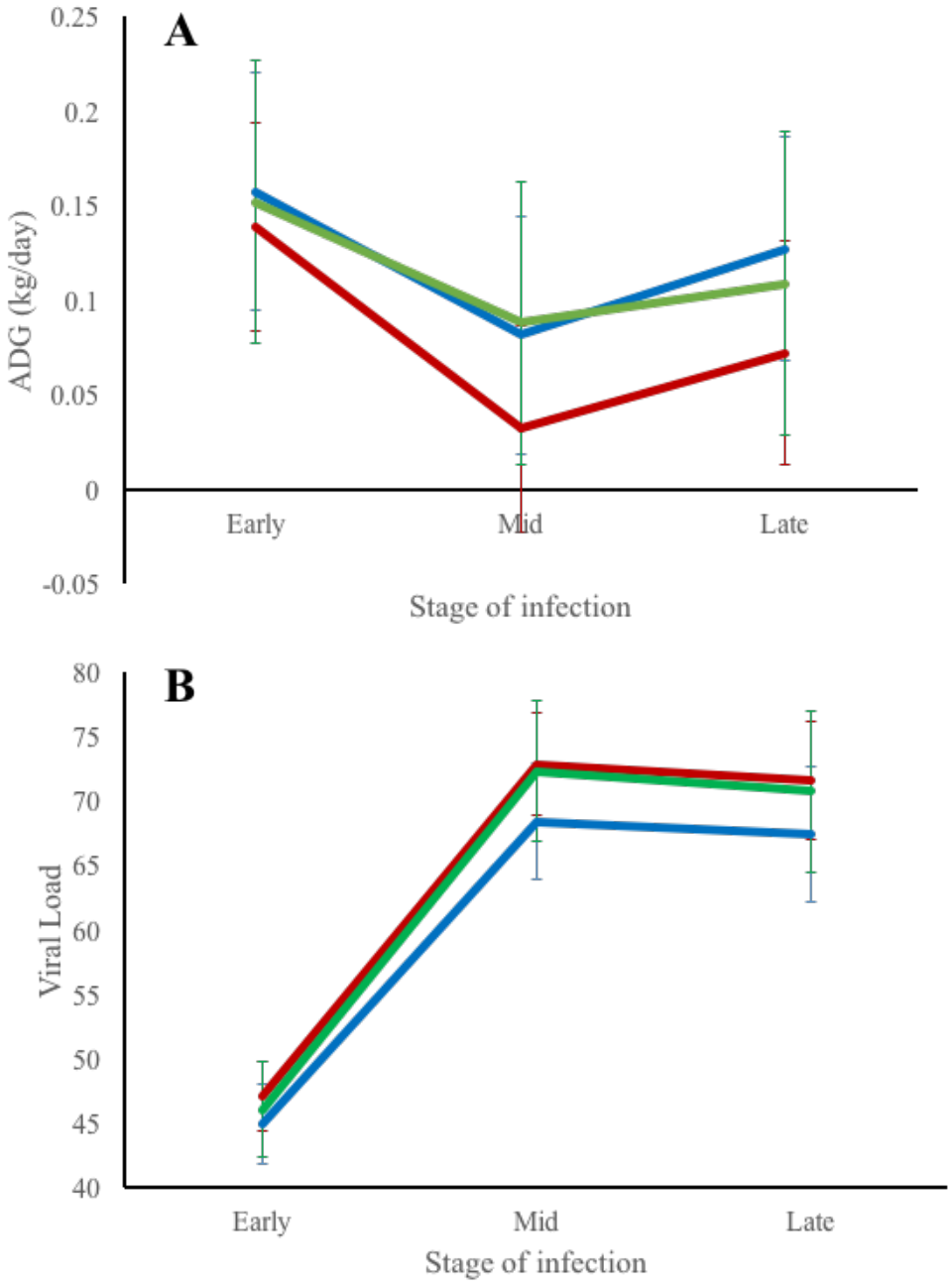
Trait	ADG_{early}	ADG_{mid}	ADG_{late}	VL_{early}	VL_{mid}	VL_{late}
ADG_{early}	**0.28 (0.07)**	0.35 (0.03)	0.24 (0.03)	0.03 (0.03)	-0.06 (0.03)	-0.12 (0.03)
ADG_{mid}	0.74 (0.13)	**0.23 (0.07)**	0.35 (0.03)	0.04 (0.03)	-0.07 (0.03)	-0.12 (0.03)
ADG_{late}	0.81 (0.20)	0.84 (0.18)	**0.18 (0.08)**	-0.04 (0.03)	-0.17 (0.03)	-0.18 (0.03)
VL_{early}	0.22 (0.36)	0.83 (0.26)	0.07 (0.38)	**0.17 (0.10)**	0.43 (0.03)	-0.19 (0.03)
VL_{mid}	-0.21 (0.28)	0.45 (0.28)	-0.38 (0.29)	0.78 (0.15)	**0.31 (0.13)**	0.73 (0.02)
VL_{late}	-0.47 (0.28)	-0.23 (0.32)	-0.61 (0.27)	0.01 (0.37)	0.82 (0.12)	**0.20 (0.10)**

Genetic correlations between all stages of infection were generally high and positive for ADG (in the range of 0.74-0.84). Phenotypic correlations were also positive, but weaker. Genetic correlations between stages for VL were strong between consecutive stages of infection (0.73 to 0.82), but very weak between early and late stage (-0.01), indicating that individuals with lower VL early in infection did not necessarily imply lower VL at late stage. Phenotypic correlations were moderate to strong between consecutive stages, but weak and negative between early and late stages. Estimates for genetic correlations between ADG and VL varied depending on stage of infection, and had high standard errors. At the early and mid stage of infection, genetic correlations between ADG and VL were weak to moderate (0.22 and 0.45 for early and mid stages, respectively), indicating that individuals with may have had greater genetic resistance at these stages of infection tended to grow faster during these stages. However, by late stage of infection, genetic correlation between ADG and VL was strong and negative (-0.61). The strongest genetic relationship between VL and ADG (0.83 (0.26)) was observed between ADG at mid stage of infection and VL at early stage of infection, implying that individuals with lower genetic resistance at early stage of infection tended to grow faster at the mid stage, and vice versa (see discussion). However, standard errors were high for genetic correlations between ADG and VL, with exception of ADG_{mid} with VL_{early} and ADG_{late} with VL_{late}. Phenotypic correlations between ADG and VL were typically weak across all stages of infection.

4.3.2 Association of WUR genotype with ADG and VL across different stages of infection

The WUR genotype was associated with VL across at stages of infection ($p<.0001$), but only significantly associated with ADG at mid and late stages of infection ($p<.0001$). Although standard errors were high individuals with the favourable B allele generally had higher ADG than those with genotype AA (Figure 1a). Similarly, individuals with heterozygous AB genotype had lower VL across all stages of infection than those with the AA genotype (Figure 1b). As there were few individuals with BB genotype, only differences between AA and AB are considered from here-on. The largest difference in ADG between genotypes was at the late stage of infection, where individuals with AB genotype grew, on average, 55g/day faster than those with AA. However, largest difference in VL occurred at the mid stage of infection, where AB animals were had 6.10% lower VL than AA animals.

Figure 1. **Effect of WUR genotype of a) ADG and b) Viral load (VL) across early, mid and late stages of infection**, where B allele is the favourable allele. The red, blue and green lines correspond to AA, AB and BB genotypes, respectively. Lease-square means of ADG (kg/day) with standard errors and VL (area under viremia curve) are on the y-axis.

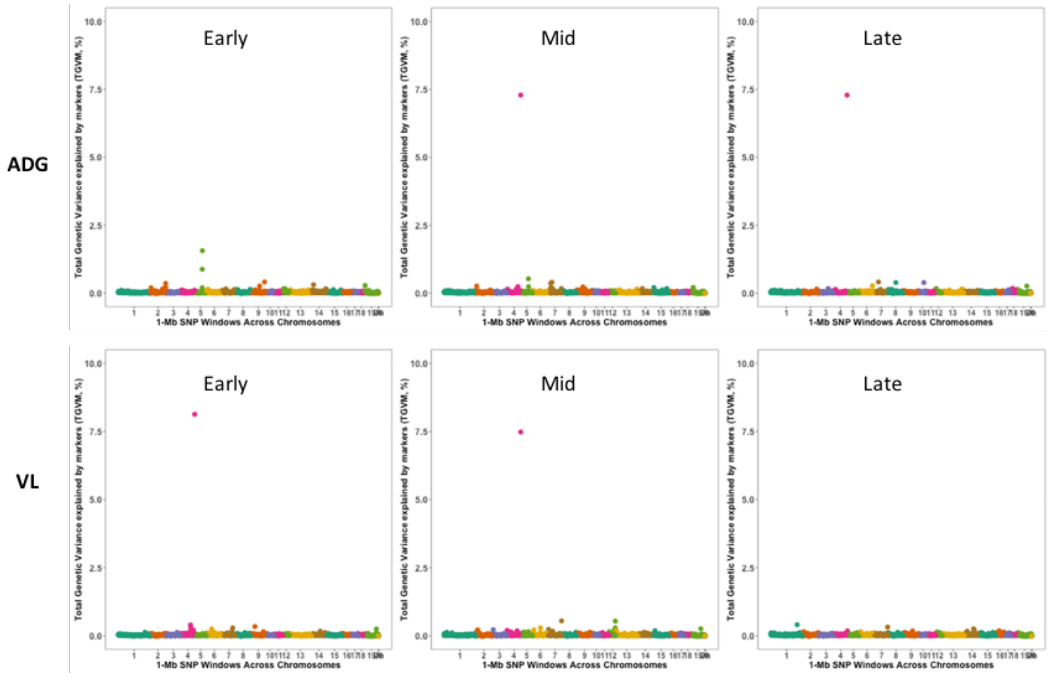


4.3.3 Genome Wide Association Study

This study did not find any genomic regions explaining a significant percentage of the genetic variance of ADG at early stage of infection (Figure 2), although a small percentage of the total genetic variation was explained by two regions on chromosome 5 (1.56% and 0.88% by each window). Additionally, there were no significant genomic regions for VL at late stage of infection identified by the GWAS.

The only genomic region which explained a significant percentage of the total genetic variance (i.e. had a posterior greater than 0.8) was on chromosome 4 containing 40 SNP (window 793). This genomic region explained 7.29% of the total genetic variance for ADG at both mid and late stage of infection, and 8.13% and 7.48% of total genetic variance for VL at early and mid-stage of infection, respectively. This window contains the previously identified WUR genotype, in which SNP have been shown to be in strong LD [23]. As such, no further investigation was carried out on the SNP within this window (see discussion).

Figure 2. Manhattan plots of the genome-wide association study of ADG (upper row) and VL (lower row) at early, mid and late stage of infection showing percentage of total genetic variance explained by non-overlapping 1Mb genomic regions, arranged by chromosome 1 to 18 and X. SNPs of unknown region were removed.



4.4 Discussion

4.4.1 Summary of findings

Partitioning infection into stages of infection informed by viremia profile characteristics provided new insight into how the genetic regulation of pigs' response to experimental PRRSV infection changed over the time-course of infection. ADG and VL were heritable across all stages of infection, which implies that genetic selection for resistance to PRRS is expected to affect growth and viremia patterns at all stages of infection. Heritability for resistance was greatest for VL during the highly immune-active time-period between peak viremia and maximal rate of viremia clearance. This

would suggest that the phase associated with the most rapid viremia decline would be easiest to manipulate by genetic selection. For ADG, heritability was greatest at the early stage of infection. It is not clear whether this reflects heritable genetic differences in growth per se, or captures already different growth responses to infection. Indeed, genetic correlation between VL and ADG were strongest between pre-peak VL and post-peak ADG, indicating that growth response lags behind viremia response.

Genomic regions containing the WUR SNP explained a significant percentage of total genetic variance at early-mid stage of VL and mid-late stage ADG, with the AB genotype conferring both higher resistance and higher growth under infection. This is consistent with previous studies, which found similar heritabilities and associations of the WUR genotype with these traits at 21 and 42 dpi [57,55]. No previously unidentified genomic regions associated only with a particular phase of infection were found.

4.4.2 Why define stages of infection?

Defining stages of infection based on underlying pathogen load could provide deeper and novel insights into the genetics underlying the mechanisms influencing an individual's infection profile. Cumulative viral load for a prolonged period of time (e.g. 21 or 42 days) provides a summary of how an individual copes with infection, but does not capture the dynamic changes of the individual's viremia curve resulting from different sets of immune functions at different stages of infection. For example, two individuals may have equal VL over the 21-day infection period, but one of the individuals may experience higher viremia over a short time, while the other may have low viremia over a prolonged time. This may reflect different immune response

patterns, but the genetic signal is blurred by using cumulative VL as phenotype. By defining stages of infection based on underlying viremia curve characteristics, the corresponding VL measures are more likely to reflect different sets of immune response parameters to PRRSV, and thus possibly different sets of genomic regions associated with these. For example, the time of fastest rate of viremia clearance, which defined the boundary between mid and late phase in this study, coincides with the time at which neutralizing antibodies are produced at highest rates [95]. Thus, the late stage of infection is likely determined by the ability of the individual to mount an effective adaptive immune response [96]. There has been evidence of genetic variation in antibody response to PRRS. Indeed, Hess et al. estimated heritability of 0.13 for serum levels of IgG during PRRS infection [95], and Serão et al identified genomic regions associated with IgG in serum of PRRS infected animals at 46 dpi [97].

This study found indeed that heritabilities for resistance and growth, and their genetic correlation changed considerably over the time-course of infection. Furthermore, the WUR SNP that had been previously identified to explain a significant part of genetic variation in VL and growth over the 21 and 42-day infection period, was found to be only associated with particular phases of infection in this study. However, no novel QTL associated with only a particular stage of infection, was detected.

4.4.3 Relationship between resistance and growth

The WUR genotype was associated with growth only at mid to late stages of infection, but associated with VL at early to mid stages of infection. In line with this, genetic correlation was strongly positive between ADG at mid stage of infection and VL at early stage of infection (0.83 (0.26)), implying that genetically resistant individuals

with lower VL earlier in infection would subsequently have slower growth rate at later stage of infection. This is likely due to individuals with higher genetic resistance needing to divert more resources away from growth to combat infection, where greater investment in early immune response to infection would result in a less severe infection, but a temporary reduction in growth. These results are consistent with resource allocation theory, where a trade-off between fighting infection and growth would occur due to limited energy resources available for both biological processes [13].

At the late stage of infection, the genetic correlation between ADG and VL shifts from strongly positive to negative (e.g. genetic correlation between VL_{late} and ADG_{late} is -0.61 (0.27)). This is likely due to return to homeostasis or even compensatory growth, where those animals that mange to clear the virus faster can allocate resources to growth [98]. In summary, these results combined imply that selection for lower VL at the post-peak viremia phase may lead to reduction in growth, but faster growth at the later stages of the infection period and overall. It is only by examining the dynamic interplay between growth and resistance across different stages of infection, that this relationship becomes clear and gives biological meaning.

It should be noted that in chapter 3, genetic variation in tolerance was only identified when late ADG was regressed on mid VL. This is surprising, given the strong genetic correlation between ADG and VL at these stages of infection.

4.4.4 Association of WUR with growth and resistance to PRRS

By analysing the host response to infection based on underlying biological properties, we were able to identify with which stages of infection the WUR genotype was

associated. The GWAS showed that the only window which explained a significant percentage of the total genetic variance was window 793, which contains the previously identified WUR SNP [23,57,58]. Window 793 was significant only at early-mid stage for VL, and mid-late stage for ADG. The effect of the WUR SNP was significant at these stages also. This SNP has been shown to be in high linkage disequilibrium (LD) with other SNPs in this window [23,57]. Thus, it is likely that this window was significant due to the effect of the WUR SNP. Furthermore, WUR has previously been shown to be in high LD with a causative mutation in the guanylate binding protein 5 (GBP5 gene), which plays a role in the inflammatory immune response [99,100]. In line with Boddicker et al and Hess et al, this study found a beneficial effect of the WUR genotype for both ADG and VL, where individuals with the favourable AB genotype demonstrated higher ADG and lower VL [55,57]. Additionally, Hess et al found that the genotype at the WUR SNP was associated with all viremia curve characteristics, where individuals with the AB genotype generally had lower V_{peak}, T_{peak} was reached 0.2 days earlier, V_{max} was 3.8% faster and T_{peak} was reached 0.68 days earlier. In this study, the genotype had strongest association with VL at early stage of infection, explaining 8.13% of the total genetic variance.

Thus, selection to increase the frequency of the B allele of the WUR genotype would be expected have the greatest influence on early viremia patterns. However, the selection to response is likely to be limited, as only a small percentage of the total genetic variance is explained by genomic region containing the WUR SNP. Resistance is a highly polygenic trait, with multiple pathways involved in mounting an effecting immune response to PRRS. As there are no windows that explain a significant percentage of total genetic variance of VL at late stage of infection, it is likely that

there is no single genomic region that has a large effect at this stage of infection, given that it was moderately heritable. Indeed, the AB genotype may increase resistance at early-mid stages of infection, with a consequent positive effect on growth at the subsequent stages. In chapter 3, we found that the WUR genotype had no association with tolerance at any stage of infection. This, together with the results of this chapter, would thus imply that the association between WUR and ADG at mid to late stage of infection is a direct consequence of the effect of the genotype on resistance.

4.4.5 Conclusions

Using stages of infection defined by viremia curve characteristics is suitable for analysing the genetic change and interplay between growth and resistance over the time-course of infection. This study found that the genetic relationship between resistance and growth change considerably over different stages of infection. The results of this study support the resource allocation theory, suggesting a trade-off between resistance and growth i.e. that early investment in resistance has a short-term effect on growth, but leads to faster growth at late stage of infection. This study established that the WUR SNP would confer resistance at early stage-mid stage of infection, which would have a beneficial impact on growth at later stages of infection. Furthermore, no further genomic regions were associated with resistance or growth.

Chapter 5

Health trajectories reveal the dynamic contributions of host genetic resistance and tolerance to infection outcome

Graham Lough, Ilias Kyriazakis, Silke Bergmann, Andreas Lengeling, and Andrea B. Doeschl-Wilson

Authors' contributions

Design and conduct of infection experiment: AL, SB; Development of statistical methods: GL, ADW, IK, AL; Data analysis: GL; Interpretation and writing of the manuscript: GL.

5.1 Abstract

In chapters 2 and 3, we found limitations with adopting random regression models to identify genetic variation in tolerance in outbred populations. Through utilising repeated measures, we were able to identify a positive relationship between resistance and tolerance over all stages of infection. However, in chapter 4, we found that the relationship between growth and resistance changed over the time-course of infection. In this chapter, we propose a novel methodology that enables interpretation of the dynamics between resistance and tolerance over the time-course of infection.

Resistance and tolerance are two alternative strategies hosts can adopt to survive infections. Both strategies may be genetically controlled. To date, the relative contribution of resistance and tolerance to infection outcome is poorly understood. Here, we use a bioluminescent *Listeria monocytogenes* (*Lm*) infection challenge model

to study the genetic determination and dynamic contributions of host resistance and tolerance to listeriosis in four genetically diverse mouse strains. Using conventional statistical analyses, we detect significant genetic variation in both resistance and tolerance, but cannot capture the time-dependent relative importance of either host strategy. We overcome these limitations through the development of novel statistical tools to analyse individual infection trajectories portraying simultaneous changes in infection severity and health. Based on these tools, early expression of resistance, followed by expression of tolerance, emerge as important hallmarks for surviving *Lm* infections. Our trajectory analysis further reveals that survivors and non-survivors follow distinct infection paths, which are also genetically determined, and provides new survival thresholds as objective endpoints in infection experiments. Future studies may use trajectories as novel traits for mapping and identifying genes that control infection dynamics and outcome.

5.2 Introduction

Two alternative host response strategies to pathogen challenge contribute to survival: *resistance*, defined as the ability of a host to limit or inhibit pathogen replication, thus reducing infection severity [19]; and *tolerance*, defined as the ability of an infected host to limit the impact of infection on fitness or health. Tolerance mechanisms reduce or prevent damage associated with pathogen challenge, but have no direct impact on the pathogen itself [5,6,19,101]. In addition, tolerance is also an important mechanism for the co-evolution of symbiotic interactions between beneficial commensal microbes and the host, which has been long recognised in both plants and animals [102,103]. As host strategies, both resistance and tolerance may be genetically determined [5,30,34].

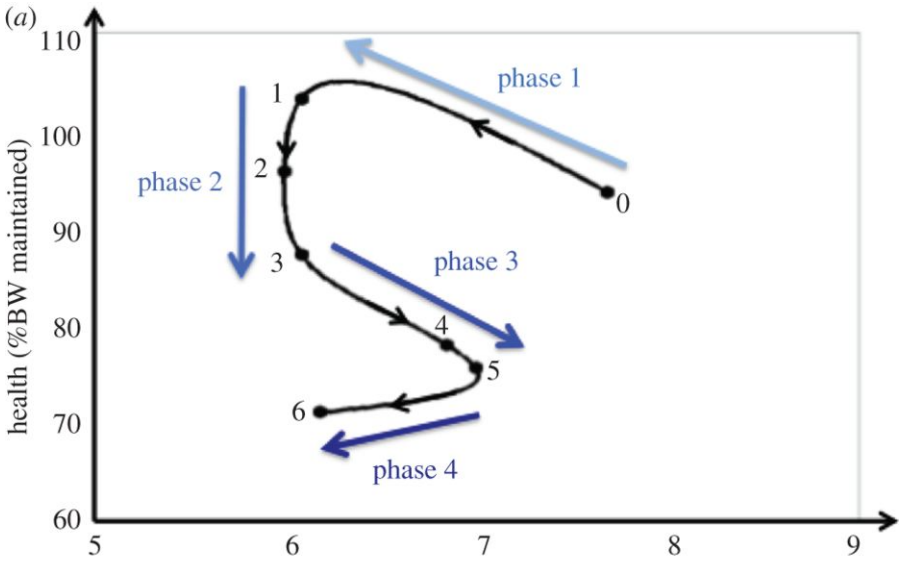
Assessment of their relative contribution to survival requires quantitative estimates of resistance and tolerance based on empirical evidence. Resistance may be defined as the inverse of infection severity, conventionally quantified by measures of within-host pathogen burden. Obtaining quantitative estimates of tolerance has proven difficult in practice, owing to its statistical definition as reaction norm of health with respect to changes in pathogen burden [5,9,35] and the high frequency of measurements associated with constructing and analysing reaction-norms [4,7,35,37]. Although conceptually defined at the individual level, quantitative tolerance estimates can usually only be obtained at the level of groups of (related) individuals, which constitutes a major limitation to unravelling the host genetic regulation of tolerance [4,35,37].

These static definitions of resistance and tolerance, and the limitation of estimating tolerance at the group level, cannot further our understanding of the relative contribution of resistance and tolerance to individual survival, which is likely to change over the time course of infection [41]. Recently, individual health trajectories have been introduced as potentially powerful tools to capture the dynamic nature of infection and its impact on health in individual hosts [38,41]. Trajectories are constructed by plotting individual measurements of infection severity (e.g. pathogen burden) against health in 2-dimensional space at different stages of the infection. Following this pair-wise progression over time produces a trajectory that illustrates the dynamic interplay of resistance and tolerance mechanisms by describing how changes in within-host pathogen burden are associated with changes in health throughout the infection period, not currently captured by static definitions of resistance and tolerance (Figure 1A). We suggest that using trajectories as an alternative to conventional

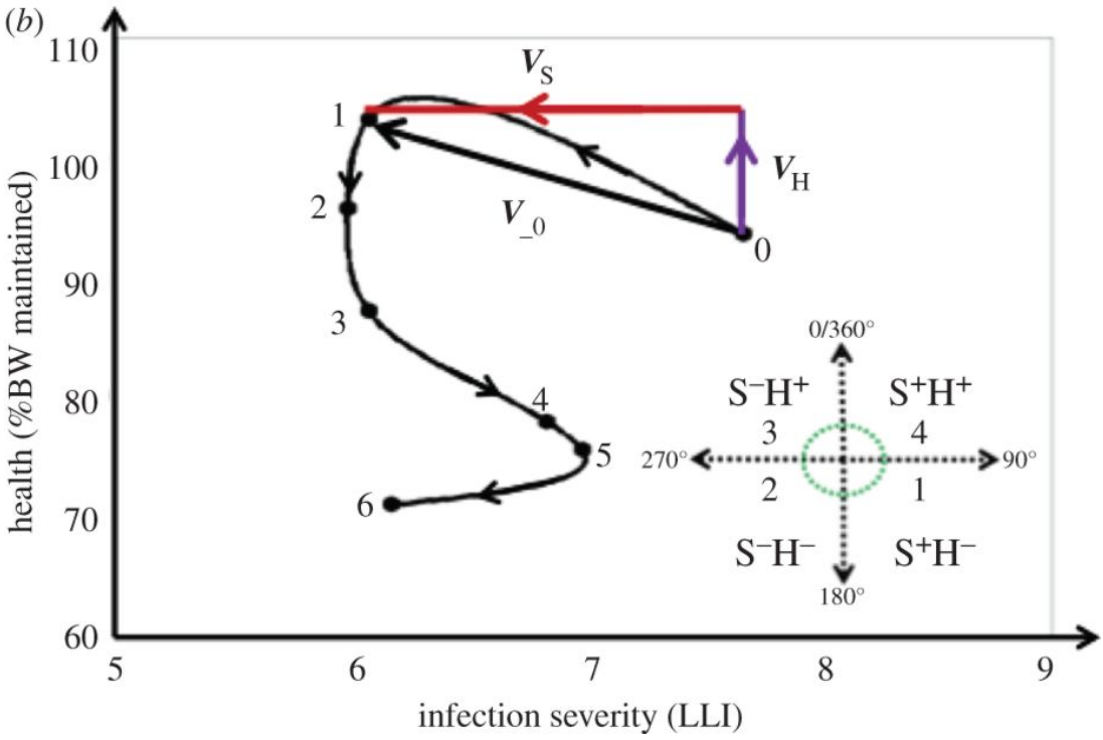
statistical analysis of resistance and tolerance will help describe individuals' infection paths towards a specific outcome (e.g. death or survival), and reveal critical stages of the infection associated with the greatest impact on health or fitness. It has been postulated that infection trajectories can be classified into distinct trajectory types [38,41] that may be linked to genetic background of the host [41]. It may thus be possible to map host genotypes to specific trajectory types, and target these for genetic improvement of host response to infections [38].

Figure 1. An infection severity - health trajectory for an individual mouse.

Graph 1A. Illustration of trajectory phases from start of infection to death. A trajectory for an individual mouse was produced by plotting longitudinal pairwise measurements of body weight (BW) and infection severity (bacterial load measured by log-transformed light intensity plus one [LLI]), in a 2-dimensional space, and following their progression over time. The graph also shows the four characteristic phases of infection associated with distinct changes in infection severity and health (indicated by arrows), as described in the text.



Graph 1B. Illustration of trajectory vectors and resulting sequences. The trajectory vector $V_{\_0}$ with components V_H (change in %BW) and V_s (change in infection severity), represents simultaneous change in infection severity (decrease, S-) and health (increase, H+) between 0-1 dpi. The bottom right panel shows the 4 quadrants (S^+H^-, S^-H^-, S^-H^+, S^+H^+) in the Severity-Health (SH) plane specifying the direction of a trajectory vector, together with the associated sequence numbers (1-4). Each trajectory is mapped to a sequence comprising 14 numbers, representing the directions of the trajectory vectors at 14 consecutive daily intervals, and with the sequence number 0 indicating death. The trajectory of the individual depicted in this figure corresponds to the sequence {32211200000000}, i.e. it died after dpi 6.



Although proven powerful on conceptual grounds, the use of trajectories to study host response to infection has not previously been supported by experimental data [38,41].

Their wider application in infectious disease research has been hampered by the lack of statistical methods for quantitative trajectory analyses [37]. Trajectories often display loops (Figure 1A), which implies that they cannot be represented by mathematical functions, and are thus not amenable to conventional statistical models. In this study, we develop a novel statistical framework for quantitative trajectory analysis, making use of non-invasive bioluminescent imaging tools to analyse the time course of listeria infection in four inbred mouse strains. *Listeria monocytogenes* (*Lm*) is a Gram-positive, facultative intracellular bacterium that causes food-borne infections in animals and humans. *Lm* is responsible for the life-threatening disease listeriosis in elderly and immunocompromised individuals [104,105]. In healthy individuals, *Lm* infections are usually self-limited but can cause acute, febrile gastroenteritis [106]. Inbred mouse strains differ substantially in their apparent susceptibility to listeriosis, through contributions of multiple genetic loci [107–109], but genetic variation in tolerance to *Lm*, and the relative contributions of resistance and tolerance to survival are currently unknown. We use this model system to (i) determine whether there is genetic variation in tolerance to *Lm,* and whether mouse strains rank similarly in terms of resistance and tolerance, (ii) study the kinetic infection severity-health relationships using trajectory analysis, and their association with survival and (iii) assess whether different host genotypes map on to distinct trajectory types.

5.3 Methods

5.3.1 Mice

The data were obtained from *Lm* infection challenge experiments of 108 mice from four genetically diverse inbred mouse strains as outlined in Bergmann et al [110]. Briefly, female mice aged between 9-10 weeks from the strains A/J, BALB/cJ (BALB) and C57BL/6J (B6J), and C3HeB/FeJ (C3H) were orally infected with bioluminescent *Lm* as described below. The inbred mouse strains were selected due to known differences in resistance to listeriosis development, similar mature body weights, and for their suitability for *in vivo* bioluminescence imaging (BLI). All mice were subjected to BLI or analysed for bacterial organ loads. On 1, 3, 5 and 7 days post infection (dpi), 18 to 30 mice from each strain (27, 18, 30 and 30 mice for A/J, C3H, BL6 and BALB, respectively) were sacrificed to measure colony forming units (CFU) of *Lm* from organ homogenates [110]. This enabled assessment of the spread of *Lm* to different internal organs, and to calibrate the infection severity measures obtained by the BLI analysis (see below). Ten mice per strain were maintained after inoculation until 14 dpi, or until they had to be euthanized due to reaching humane endpoints of infection severity. This was the case for all mice from strain C3H, 80% of A/J mice and 40% of BALB mice, which were all euthanized between 5-7 dpi due to onset of clinical signs of advanced listeriosis, according to established protocols and approved animal welfare regulations [110]. All mice were housed under specific-pathogen-free conditions. At the start of the experiment, all mice had reached mature body weights. Thus, any changes in body weight post infection were assumed to be a direct consequence of the infection challenge.

5.3.2 Infection protocol

Prior to challenge with *Lm*, the mice were acclimatised for 1 to 2 weeks in the facility. On the day prior to infection, the mice were starved overnight, with drinking water replaced with carbonate buffered water. The next day, mice were intragastrically challenged with $5x10^9$ CFU *Lm* EGDe-InlA-mur-lux, an internalin A (*inlA*) modified strain of *Lm* as previously described [110,111]. After infection challenge mice had *ad libitum* access to both food and water.

5.3.3 Measurement of infection severity and health

In line with the literature, resistance was quantified as an inverse measure of infection severity [5], defined here in terms of log-transformed measures of light-intensity (LLI) obtained daily from bioluminescent *in vivo* imaging (See Appendix 5.1). Higher LLI values correspond to higher *Lm* loads, which is indicative of higher infection severity [19].

As *Lm* infection in adult mice causes a significant drop in body weight (BW), BW was used as an indicator of impact of the infection on health. BW was recorded for each individual mouse immediately prior to infection, and daily post infection over the 14-day duration of the experiment. The impact of infection on health at a particular dpi was then represented as percentage loss of BW at that day from the initial BW at 0 dpi, and percentage of BW maintained at that day was considered as the daily indicator for health.

5.3.4 Conventional statistical analysis of resistance and tolerance

The statistical analysis used data only from the 40 mice that had not been analysed prior to 14 dpi for CFU counts, as only these provided information about the association of resistance and tolerance to survival. Data were analysed with the SAS statistical package (2010, version 9.3) using procedure proc MIXED.

5.3.4.1 Estimating resistance and tolerance based on peak infection severity and minimum health

In accordance with Raberg et al [5], we defined resistance in terms of maximum infection severity, here represented by peak LLI levels over the 2 week observation period. Tolerance estimates were obtained accordingly based on maximum infection severity (peak LLI) and minimum health (maximum %BW loss) achieved during the observation period.

To assess genetic variation in resistance, a linear mixed model was used with peak LLI as the dependent variable, and mouse strain, binary survival outcome (succumbed to infection within 14 dpi (surviving)/did not succumb to infection (non-surviving)), and their interactions as fixed effects.

In line with existing studies of tolerance genetics [5,7], an analysis of covariance (ANCOVA) was used to assess genetic variation in tolerance. The ANCOVA was performed using maximum %BW loss as dependent variable and peak LLI as independent variable, and mouse strain, survival outcome and the corresponding interactions as fixed effects. The intercept was fixed at zero, corresponding to zero percent BW loss in the absence of infection. The ANCOVA slope coefficients resulting from regressing individual health measures against infection severity provide

group estimates of tolerance, where steeper negative slopes correspond to less tolerant groups. Differences between these slope estimates based on the F-test statistics for strain-by-infection severity interaction thus provide evidence for genetic variation in tolerance.

5.3.4.2 Assessing the sensitivity of resistance and tolerance estimates to time of measurement

To determine the sensitivity of resistance and tolerance estimates to the timing of measurement, we replaced the extreme measures of peak LLI and maximum %BW loss with daily measures of LLI and %BW loss to obtain daily least square means (LSM) for infection severity (inverse of resistance) and tolerance slope for every mouse strain by survival outcome using the repeated measurement models as outlined in the Appendix 5.2.

5.3.5 Infection severity - health trajectories

Infection severity-health (SH) trajectories were generated by plotting the 14 daily health measurements (represented by %BW maintained) against the corresponding infection severity measures (LLI) recorded until 14 dpi for each individual, or until time of death if infection-dependent euthanasia occurred prior to 14 dpi. Successive scatter points were connected using the spline curve and, for illustrative purposes, smoothed using the SM30 smoothing procedure in SAS. Figure 1A shows an example of a trajectory; individual trajectories of all mice in consideration are presented in Appendix 5.3.

5.3.5.1 Trajectory comparison and numerical representation

Trajectories were first visually inspected to determine common features and differences related to levels and timing of simultaneous changes in infection severity and health. "Bad neighbourhoods", associated with subsequent death due to infection, were identified in the 2D phase plane by simply overlaying trajectories of surviving mice and those that succumbed to infection.

For statistical comparison of trajectories associated with different individuals, trajectories were mapped to numerical sequences, which were constructed as follows: first, for each individual trajectory, daily 2D vectors $\boldsymbol{V}_{\boldsymbol{k}} = (V_{(S_k)}, V_{(H_k)})$ were produced as shown in Figure 1B, where $V_{(S_k)}$ represents the change in infection severity (in LLI units) from day *k*-1 to day *k*, and $V_{(H_k)}$ represents the corresponding change in health (%BW change). The magnitude of $\boldsymbol{V}_{\boldsymbol{k}}$, i.e. $\boldsymbol{V}_{\boldsymbol{k}}$ given by

$$\boldsymbol{V}_{\boldsymbol{k}} = \sqrt{(V_{(S_k)}^2 + V_{(H_k)}^2)}$$

describes the rate of change in the 2D host state between days *k-1* and *k* (Figure 1B). The direction of $\boldsymbol{V}_{\boldsymbol{k}}$ (determined by the signs of both vector components) indicates whether an increase or decrease in infection severity (S^+ / S^-) between days *k-1* and *k* is associated with a simultaneous improvement or deterioration in health (H^+ / H^-). Four possible sign combinations give rise to four SH categories (1=S^+H^-, 2=S^-H^-, 3 = S^-H^+ and 4 = S^+H^+) according to which quadrant in the SH plane the trajectory vector faces (Figure 1B). To reduce the impact of measurement noise in the statistical analysis, S^+ or H^- were only assigned if infection intensity had increased by more than 0.1 LLI units, and BW had dropped by more than 2% compared to the last measurements, respectively. Otherwise, changes in either direction were assigned to

S^- and H^+, respectively. Finally, stringing the 14 daily consecutive SH combinations together generated an SH - time series for each individual represented by a sequence of numbers between 0 and 4, where 0 indicates death of the host and 1-4 refer to the different combinations of simultaneous changes in infection severity and health as specified above (Figure 1B).

5.3.5.2 Statistical analysis of trajectory sequences

Representing trajectories by numerical sequences allowed quantitative comparison of trajectories associated with different individuals. For this purpose, Hamming distances between all pairs of individual trajectory sequences, describing the proportions of non-zero sequence elements that differed between two sequences, were calculated. To prevent sequences associated with individuals who succumbed to infection within the 14 days observation period being assigned shorter distances, sequences were truncated to the last time point where both individuals of the pair in consideration were still alive. In order to determine whether the 40 trajectories could be classified into few distinct types depending on their patterns as had been proposed previously [38,41], cluster analysis was carried out using an agglomerative clustering method, '*clusterdata'* function in Matlab (version R2013b), with the truncated Hamming distances as measure of similarity, and the weighted average distance as distance metric between clusters. The number of maximum clusters specified was 2, 4 and 6. Resulting clusters were visualized using BioLayout Express 3D [112].

Furthermore, a permutation test (in which trajectory sequences were randomised) was applied to test statistically significant differences between (truncated) trajectory sequences belonging to different clusters, mouse strains or survival groups. The

permutation test assessed whether truncated Hamming distances between any two groups were on average significantly larger than the corresponding within-group distances.

5.4 Results

Infection established in all mice, as indicated by high levels of LLI, and all mice experienced a drop in BW at a certain stage of infection, although at varying levels and duration. None of the mice were able to clear the infection within the 14-day experimental infection period. All B6J mice survived the infection period, whereas all C3H mice succumbed to infection within 6 dpi. There was within-strain variation in survival outcome for A/J and BALB mice: two A/J mice and six BALB mice survived until the end of the experimental observation period. This led to the following six mouse strains by survival outcome groups: C3H, B6J, A/J non-survivors, A/J survivors, BALB non-survivors and BALB survivors.

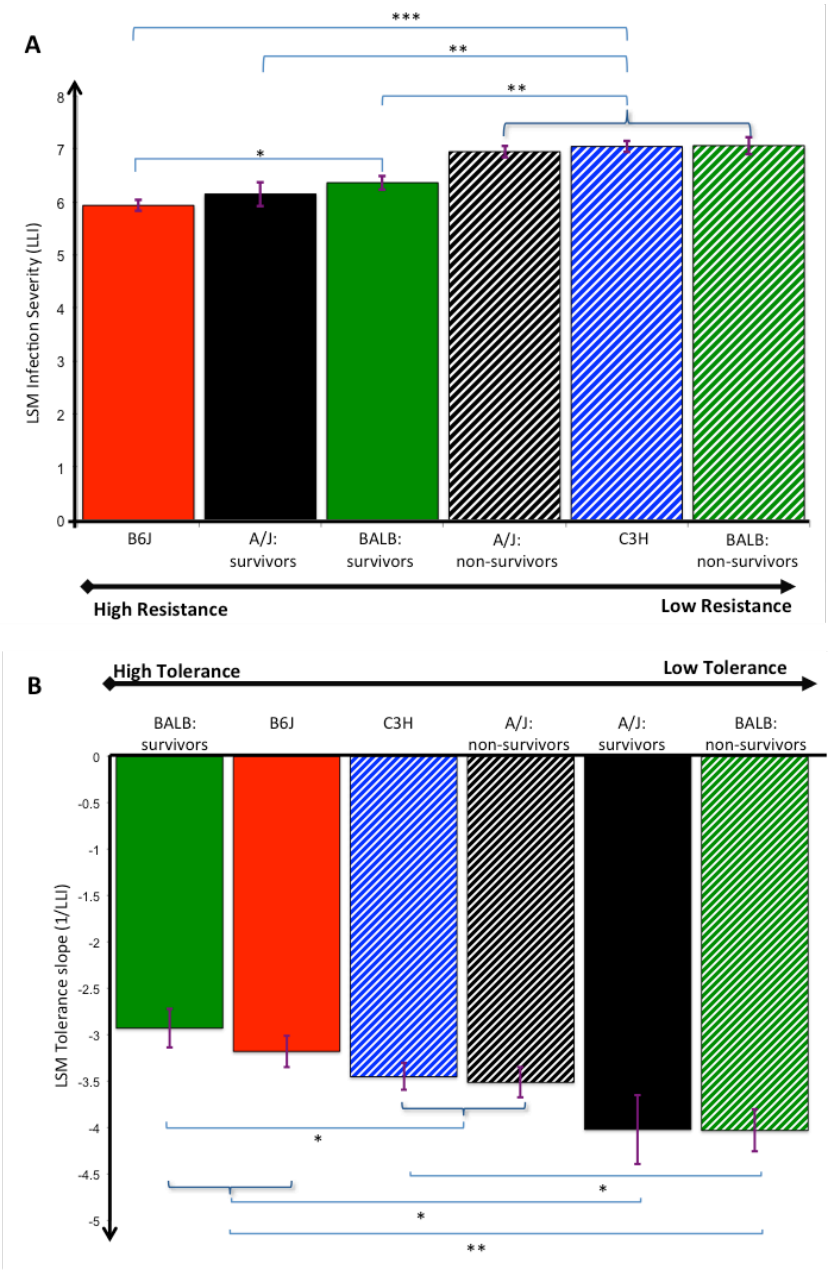
5.4.1 Estimates of resistance and tolerance

5.4.1.1 Estimates of resistance and tolerance based on peak infection severity and minimum health

Analysis of resistance revealed statistically significant strain and survival effects, as well as strain-by-survival interactions ($p<0.05$) (Figure 2A). There was no statistically significant difference in resistance amongst the non-survivors, but all non-survivors (non-surviving BALB or A/J mice and C3H) ranked significantly lower in terms of resistance than any survivor (Figure 2A). The mouse strains also differed significantly in tolerance to *Lm* infection, and tolerance varied between survival groups within the

mouse strains (Figure 2B). However, the ranking of the strains differed for the two traits (Figure 2). In particular, non-survivors did not rank consistently lower in tolerance than survivors. C3H and B6J strains characterised by 0% and 100% survival, respectively, were at the opposite ends of the resistance spectrum, but had similar tolerance estimates.

Figure 2. Least square mean (LSM) estimates for resistance and tolerance based on measures of maximum infection severity and minimum health. **(A)** LSM Resistance, (quantified as the inverse of peak infection severity (LLI)) **(B)** LSM Tolerance (defined as regression slope based on regressing maximum %BW loss against peak infection severity (LLI)).



Different mouse inbred strains are indicated by different colours and mouse strains or defined subgroups that succumbed to infection are indicated by dotted lines. LSM

*estimates (standard errors, SE) for peak infection severity (LLI) ordered from least to most resistant were: BALB non-survivors: 7.06(0.16); C3H: 7.04(0.10); A/J non-survivors: 6.95(0.11); BALB survivors: 6.36(0.13); A/J survivors: 6.14(0.23); B6J: 5.93(0.10). LSM tolerance estimates (SE) ordered from least to most tolerant were: BALB non-survivors: -4.03(0.23); A/J survivors: -4.02(0.37); A/J non-survivors: -3.51(0.17); C3H: -3.45(0.15); B6J: -3.18(0.17); BALB survivors: -2.93(0.21). Stars indicate statistically significant differences from pair-wise comparison (*** $p<0.001$; ** $p<0.01$, * $p<0.05$).*

5.4.1.2 Dynamic trends in resistance and tolerance estimates

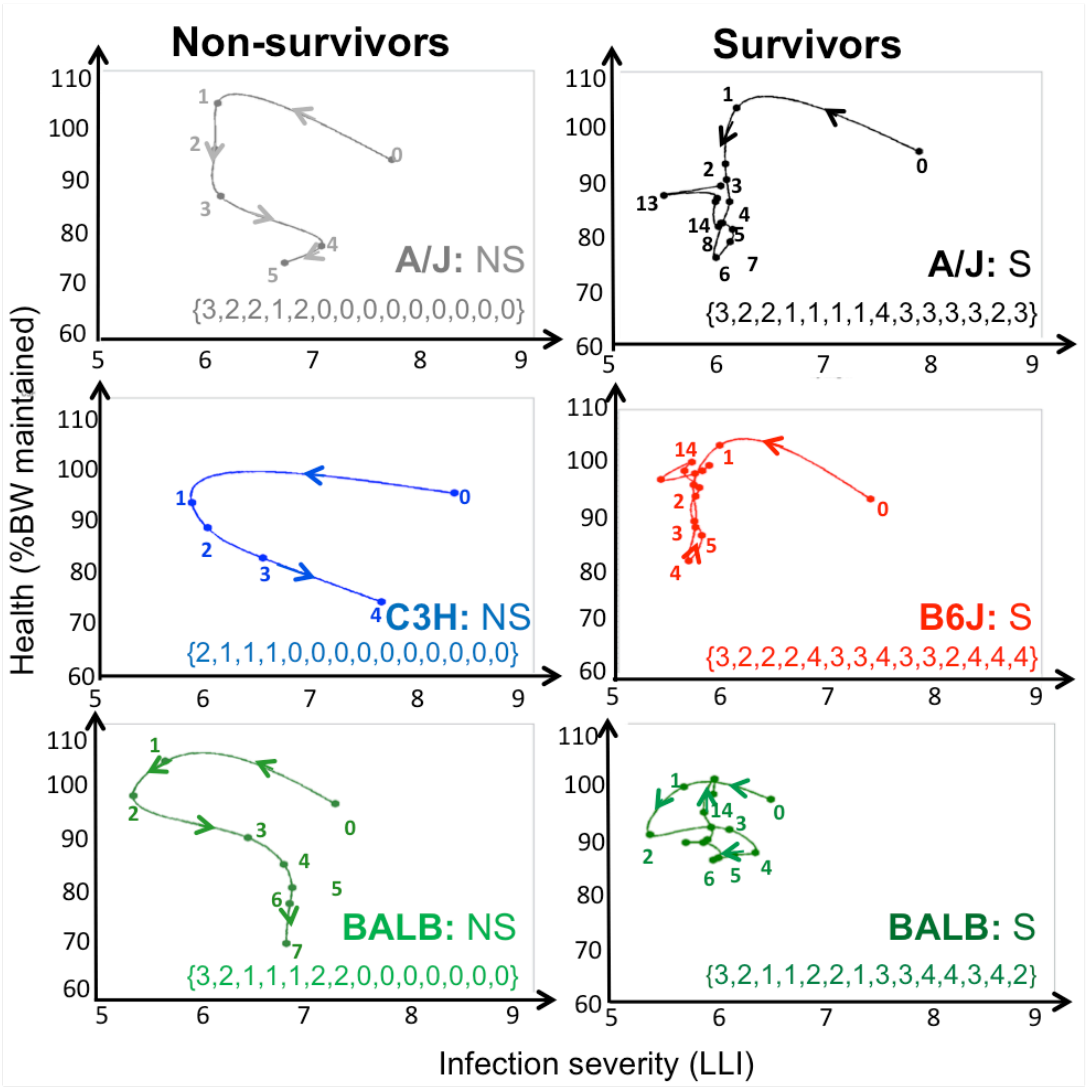
Although the actual resistance and tolerance estimates were sensitive to the timing of measurements, genetic variation in both resistance and tolerance could be detected throughout the entire 14-day infection period (Appendix 5.2 Figure and Appendix 5.2). Ranking in both traits was relatively time stable, except for a reverse in ranking of A/J mice, which started the experiment as the most tolerant strain, and emerged as the least tolerant out of the three remaining mouse strains. B6J emerged as the most resistant mouse strain already after 2 dpi, and eventually also as the most tolerant strain. In accordance with the results above, non-survivors differed significantly from survivors in resistance only, indicating that resistance may be more important than tolerance for survival of *Lm* infection.

5.4.2 Trajectory Analysis

5.4.2.1 Trajectory characteristics and determinants of survival

Visual inspection of individual trajectories (Figure 3 and Appendix 5.3) revealed common patterns in individual's routes of infection and distinct survival characteristics. Four distinct phases over the course of infection were identified with characteristic changes in infection severity and health, as represented by different infection severity and health (SH) combinations (Figure 3).

Figure 3. Representative trajectories for each mouse-strain by survival outcome, together with the corresponding numerical sequence representing daily changes in infection severity and health (SH). We refer to Fig. 1 and the main text for explanation of the latter. The numbers in the trajectories denote the day at which the measurements were taken. 'S' and 'NS' denote survivors and non-survivors, respectively.

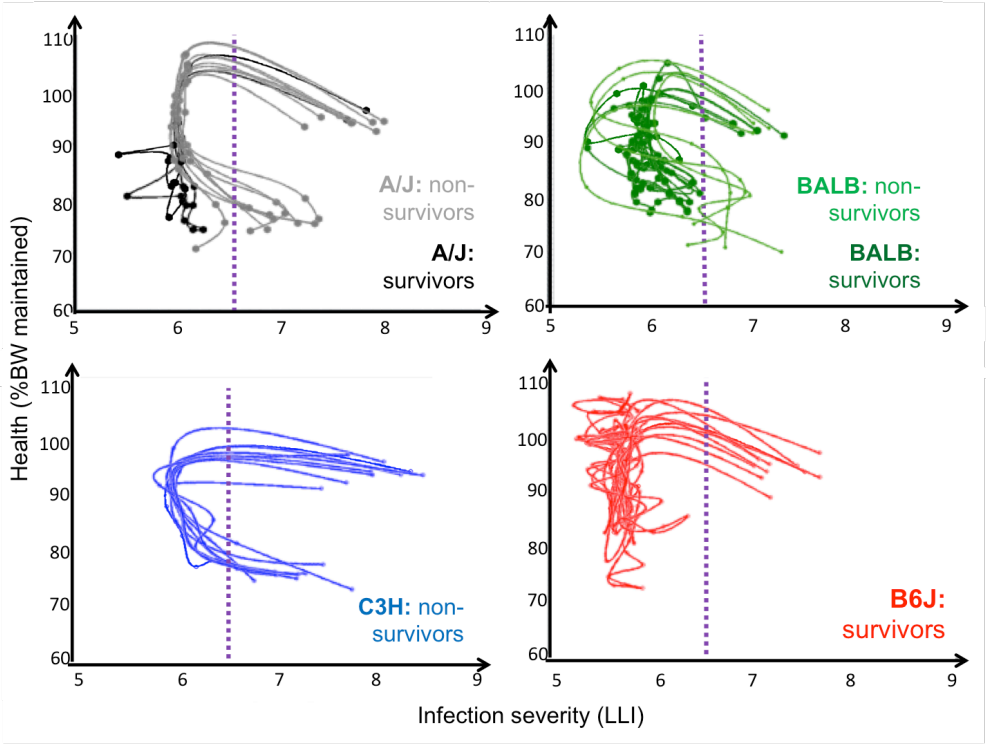


Phase 1 described the establishment of infection during 0-1 dpi, and is related to the initial reduction in infection severity due to partial clearance of the inoculated

pathogen accompanied by partial recovery in body weight (i.e. S^-H^+). This phase was seen in all mice except C3H mice, most of which experienced weight loss (i.e. S^-H^-). Phase 2 corresponded to a period during which infection severity was stable but BW continued to drop (S^-H^-). Phase 3 was associated with resurgence in pathogen load and continued weight loss (i.e. S^+H^-). The final phase 4 differed between survivors and non-survivors. All survivors regained weight and controlled infection severity. Non-survivors, in contrast, continued to lose body weight, although some were able to limit pathogen load. Mice in this phase generally fluctuated between expression of S^-H^+ and S^+H^+. With the exception of phase 1, which lasted one day for all mice, the duration of the individual phases varied between mice. Only mice that survived the infection experienced an improvement in health (H+) at some stage after 4 dpi.

By overlaying trajectories, an infection severity threshold of approximately 6.5 LLI units for pathogen resurgence could be identified that discriminated between survival and death (Figure 4). All mice that had crossed this threshold after 1 dpi succumbed to infection, regardless of their genotype, the exact day when the threshold was crossed (which occurred between 4 and 5 dpi), or whether infection severity temporarily decreased thereafter. All mice that suppressed pathogen replication below this threshold survived. Interestingly, there was no such discriminating threshold for BW loss within the limits of animal welfare regulations (Figure 4). The results suggest that trajectories can provide more predictive thresholds for the definition of ethical experiment terminating endpoints than arbitrary set cut-off values for BW losses.

Figure 4. Spaghetti plots of individual trajectories of each mouse strain. The purple line shows the infection severity threshold that discriminates between survivors and non-survivors, independent of mouse inbred strain. All non-survivors, except for one A/J mouse had crossed this threshold after 1 dpi, whereas none of the survivors had crossed this threshold after 1dpi.



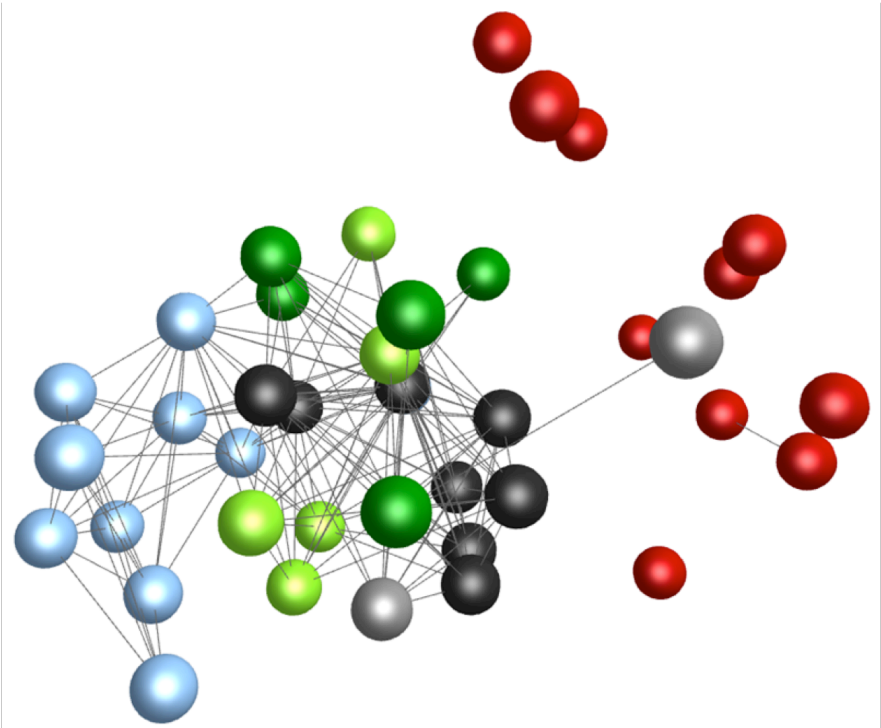
5.4.2.2 Statistical analysis and genetic footprint of infection severity-health trajectories

Cluster analysis, combined with a permutation test, applied to the corresponding (truncated) trajectory sequences indicated that individual trajectories group into distinct trajectory types (Appendix 5.4). When the stipulated maximum number of clusters was two, the resulting clusters comprised of either non-survivors or survivors. The permutation test confirmed a highly significant difference in the truncated SH sequences associated with both survival groups ($p<0.0001$). Hence, infection paths of

mice that succumbed to infection were significantly different to those of surviving mice at the early stage of infection.

When the stipulated maximum number of clusters was gradually increased, four distinct trajectory clusters ($p<0.02$ for all cluster pairs) emerged, with the greatest sequence differences found between clusters comprising exclusively C3H and B6J mice, respectively (Fig 5). The different clusters correspond to different survival outcomes and SH patterns within survivors / non-survivors, respectively, rather than to the four different inbred mouse strains (Figure 5). However, each trajectory cluster is dominated by a specific mouse strain, suggesting that trajectories are partly genetically determined (Figure 5).

Figure 5. Graphical representation of similarities between truncated trajectory sequences associated with different mouse inbred strains and survival groups within the mouse strains.



Each node (ball) represents an individual mouse, and each edge (connecting line) represents the degree of similarity between trajectory sequences of two mice, represented by 1-H, where H is the pairwise Hamming distance. Only similarities 1-H > 0.8 are depicted in the graph. Colours represent different mouse strains and survival groups, respectively. The graph was produced with the BioLayout Express 3D software [112], which spatially distributes the nodes according to the similarity measure 1-H, so that individuals with similar trajectory sequences (i.e. 1-H close to 1) are placed in close proximity to each other, whereas individuals with different trajectory sequences (i.e. 1-H close to 0) are placed far apart. The graph illustrates that differences between trajectory sequences were on average smallest within each mouse strains and greatest between C3H mice (light blue; all succumbing to infection) and B6J mice (red; all surviving infection). Within the A/J (black) and BALB (green) mice, sequences associated with survivors and non-survivors survival (survivors: light, non-survivors: dark) did not fall into different visual clusters. The graphical results were confirmed by statistical cluster analysis (see text).

The permutation test applied to different mouse strains confirmed a statistically significant difference ($p<0.05$) between trajectory sequences associated with different mouse strains, except for A/J and BALB mice (Table 1). Interestingly, the trajectory sequences of surviving and non-surviving BALB or A/J mice were statistically indistinguishable ($p=0.53$ and $p=0.06$, respectively), implying that within a mouse strain trajectory sequences alone are insufficient for predicting survival outcome.

Table 1. P-Values of permutation test used to assess whether mouse strains differ significantly in their SH trajectory sequences.

Between-strain	**A/J:**	**C3H**	**B6J**	**BALB**
A/J:	.	$<10^{-6}$	0.0003	0.31
C3H		.	$<10^{-6}$	0.0002
B6J			.	0.0002
BALB				.

5.5 Discussion

Host genetic variation in both resistance and tolerance can account for a substantial part of the observed variation in host response to infection [5,30,35,113]. Many studies have provided conclusive evidence for genetic variation in host resistance to *Listeria monocytogene*s [5,114–117]. We have demonstrated that mice from genetically distinct inbred strains, previously found to differ significantly in resistance to the bacteria, also differ in tolerance. Resistance and tolerance are often considered as static traits that constitute alternative host defence strategies against invading pathogens [5,6]. The data show clearly that expression of resistance (reduction in pathogen load) and tolerance (damage prevention and repair) and their relative contribution to survival varies over the time course of infection. We therefore propose a paradigm shift in considering resistance and tolerance as dynamic, rather than static traits. In practice, this can only be achieved through time series measurements in individual infected hosts, which in turn depend upon non-invasive technologies, such as imaging. Recently, there have been rapid advances in the development of such imaging

technologies [118–120], and this has led to increasing demands for advanced statistical tools to analyse infection dynamics, such as the trajectory methods proposed here.

The novelty of this research lies in the development of simple and versatile mathematical tools for capturing the dynamic development of resistance and tolerance in each individual, and their relative importance on the outcome of infection. The conventional reaction-norm approach to tolerance has severe limitations that have hampered progress in tolerance studies [4]. Firstly, it usually restricts tolerance estimates to group level, which is not helpful for improving tolerance of individuals or identifying tolerance genes. Secondly, the high data demand associated with estimating tolerance parameters from this approach limits reaction norms to linear models, thus ignoring all biological understanding of the highly nonlinear and time-dependent relationship between pathogen burden and health [7]. In contrast, individual trajectories, which can be easily constructed if longitudinal measurements are available, illustrate how changes in infection severity are related to health change within each individual throughout infection. Although not synonymous with resistance and tolerance, the 2D trajectory vectors crudely reflect how resistance and tolerance are co-expressed at different stages of infection. For example, simultaneous decrease in infection severity and health (S^-H^-) reflects expression of resistance at the cost of deterioration in health, indicating incomplete tolerance. As demonstrated in this study, trajectories can reveal distinct phases of infection associated with different patterns of co-expression of resistance and tolerance, and illustrate for each individual the 2D path towards death or survival. Previous studies of trajectories have defined 'bad neighbourhoods' in the infection-severity health plane that appear predictive for fatal infection outcome [5,37,38]. In the mouse data we identified an infection severity

threshold that discriminated between death and survival (Figure 4). All mice that succumbed to the infection experienced a drastic increase in infection severity between 3-7 dpi, whereas all survivors managed to restrict pathogen resurgence below this threshold during this critical phase. The critical infection severity threshold was independent of host genotypes and timing. Any mouse that crossed this threshold eventually succumbed to infection, even if thereafter it managed to reduce infection severity below the threshold. Interestingly, there was no apparent health threshold that discriminated between death and survival. All mice experienced body weight loss as a consequence of infection, but all of the survivors and none of the non-survivors managed to recover some of the lost weight, in some cases despite continued increase in pathogen load. Our results thus indicate that both early expression of resistance and tolerance at the later stages of infection are important determinants of survival to *Lm* infections. Discriminatory thresholds, such as those identified here, can provide more informative criteria than body weight for defining humane endpoints for termination of animal infection.

Individual trajectories have been used previously to classify and predict host responses to infection but their assessment was limited to qualitative analysis [38,41]. By transforming visual trajectories into numerical sequences that preserve the key topological trajectory features, we were able to subject trajectories to rigorous statistical analysis. Our statistical analysis confirmed that individual trajectories cluster into a limited number of genetically regulated distinct trajectory types [38,41]. Trajectories thus open new avenues for genetic studies of host response to infections. Future studies may focus on genetic dissection of different trajectory types to identify novel genetic variants that control infection dynamics at a molecular level.

Despite a clear genetic footprint in trajectory patterns, trajectory sequences could not capture within-strain differences in survival outcome. This could be due to several reasons: firstly, body weight may only be a crude indicator of health [5]. Alternatively, since mice were euthanized due to welfare considerations based upon weight loss, some of the mice classified as non-survivors may actually have survived the infection. Furthermore, survival outcome may be partly determined by individual differences in the gastro-intestinal flora, which have been found to show substantial inter-strain variation even in similarly highly controlled environments as used in our study [121]. Higher-dimensional trajectories comprising other types of measurements (e.g. related to the immune response or microbiota) in addition to measures of health and pathogen load may shed light on relevant host response mechanisms controlling an individual's infection path and its outcome. Note that, although more difficult to visualize, multi-dimensional trajectories can still be represented as a series of vectors defined by their direction and length and are thus amenable to similar statistical analyses as those presented here.

Previous studies have estimated an antagonistic relationship between resistance and tolerance at the phenotypic and genetic level [30,122]. These results led to the notion of a trade-off between resistance and tolerance mechanisms and their consideration as alternative host defence strategies to fight infections, and shaped predictions of evolutionary consequences for both hosts and pathogens [32,79,123,124]. These studies do not take account of the dynamic relationships that emerge from our study. For example, expression of resistance at the early stages of infection is likely to affect the expression of tolerance at the later stages as fast pathogen clearance may prevent tissue damage and obviate any requirement for damage prevention or repair

mechanisms associated with tolerance. Conventional statistical models that do not account for this kind of interdependence between traits may produce a spurious antagonistic relationship between traits, both on the phenotypic, as well as on the genetic level [125], even if resistance and tolerance are controlled by different sets of genes or genetic pathways as suggested by immunological evidence [19,101]. Trajectories enable us to bypass the complex relationship between resistance and tolerance, and may give rise to novel phenotypes for future genetic analyses that may lead to the discovery of genes that control an individual's infection path.

All of the infected mice experienced a substantial body weight loss after the initial reduction in pathogen burden, but of itself, body weight loss did not influence survival. From a resource allocation perspective, this would suggest that resistance mechanisms are costly [13,126]. By reducing resources allocated to other functions, such as searching for and digesting food, the host may be able to direct resources to the immune response, resulting in temporary body weight loss. This has been put forward as the evolutionary basis of pathogen-induced anorexia and 'sickness' behaviours [12,126–128]. Our previous applications of the resource allocation theory to assess the effect of genetic resistance on the long-term effects of infection, showed that hosts with greater genetic resistance may suffer greater performance loss (e.g. growth or body weight loss) in the short-term, but are able to revert to original levels of performance faster than non-resistant genotypes [129]. This is consistent with our study, where the B6J mice emerged as most resistant genotype after 3 dpi and were the only mouse strain that managed to fully restore the original body weight within 14 dpi.

To study the dynamic co-expression patterns of resistance and tolerance to *Listeria monocytogenes* in different mouse inbred strains we have taken advantage of a bioluminescent *Lm* infection model in which the listerial strain EGDe-InlA-mur-lux recognises the host receptor E-cadherin and intestinal expressed N-cadherin [110,111,130,131] Other *Lm* infection models have been shown to elicit different host responses [131,132]. It would be interesting to apply our novel methods to data from these models to determine whether the dynamic contributions of resistance and tolerance to survival are preserved across different pathogen and host strains. However, this would require the introgression of humanised alleles of *CDH1* (encoding E-cadherin) into the different mouse genetic backgrounds followed by repeated backcrossing to make mice permissive to oral *Lm* challenge [110,114].

In conclusion, our study complements existing evidence for genetic variation in both host resistance and tolerance, and for the importance of both host strategies in fighting infections [5,6,9,19,30,34,101–103,109,129]. However, our study also highlights the potential benefits that may arise from considering the dynamic patterns of co-expression of genetic resistance and tolerance over the time course of infection. Trajectories capture the dynamic signature and genetic footprint of both mechanisms on the level of individuals, together with their impact on survival.

Appendix 5.1

Quantifying infection severity by bioluminescence

In this study resistance was quantified as an inverse measure of infection severity, which was derived from bioluminescence as follows: For detection of bioluminescence, mice were anesthetized with isoflurane and monitored using the IVIS 200 imaging system (CaliperLS), as outlined in Bergman et al. [22]. Increased light intensity (LI) (measured in photons p/sec/cm2/sr (RAD) units) corresponded to increased Lm loads, which was indicative of higher infection severity [22]. Light intensity was high for all mice at 0 dpi, indicating a successful administration of the Lm inoculum, and subsequently varied over time between the different mouse strains. A Pearson correlation of r=0.81 between log transformed colony forming units (CFU) tissue counts and log-transformed light-intensity plus one (LLI) measures indicated a high correlation between the two measures of infection severity. Therefore, LLI was found to be a valid measure of infection severity in this study.

Appendix 5.2

Sensitivity of resistance and tolerance estimates to the timing of measurement

5.2.1 Statistical models

To assess genetic variation in resistance based on longitudinal measures of infection severity, a linear mixed model was used with log-transformed light intensity (LLI) as the dependent variable, and mouse strain, binary survival outcome (succumbed to infection within 14 dpi / did not succumb to infection within 14 dpi), day of measure (dpi), and all statistically significant interactions as fixed effects, and individual mouse as random effect. To account for potential auto-correlations between repeated measurements a heterogeneous first-order autoregressive covariance structure (ARH(1)) was fitted. This covariance structure assumes that variances in the resistance measures can differ between time points and covariances between measures at different time points decrease exponentially over time. This structure was selected amongst the range of available correlation structure based on the lowest Akaike's Information Criterion (AIC) used as model fit statistics. Statistical comparison between individual strain/survival differences and different time points were conducted using the CONTRAST statement in SAS (SAS Institute Inc. 2011. SAS/STAT® 9.3 User's Guide. Cary, NC: SAS Institute Inc.). The analysis thus provided least square means (LSM) and standard errors for infection severity for every mouse strain by survival outcome at each day post infection.

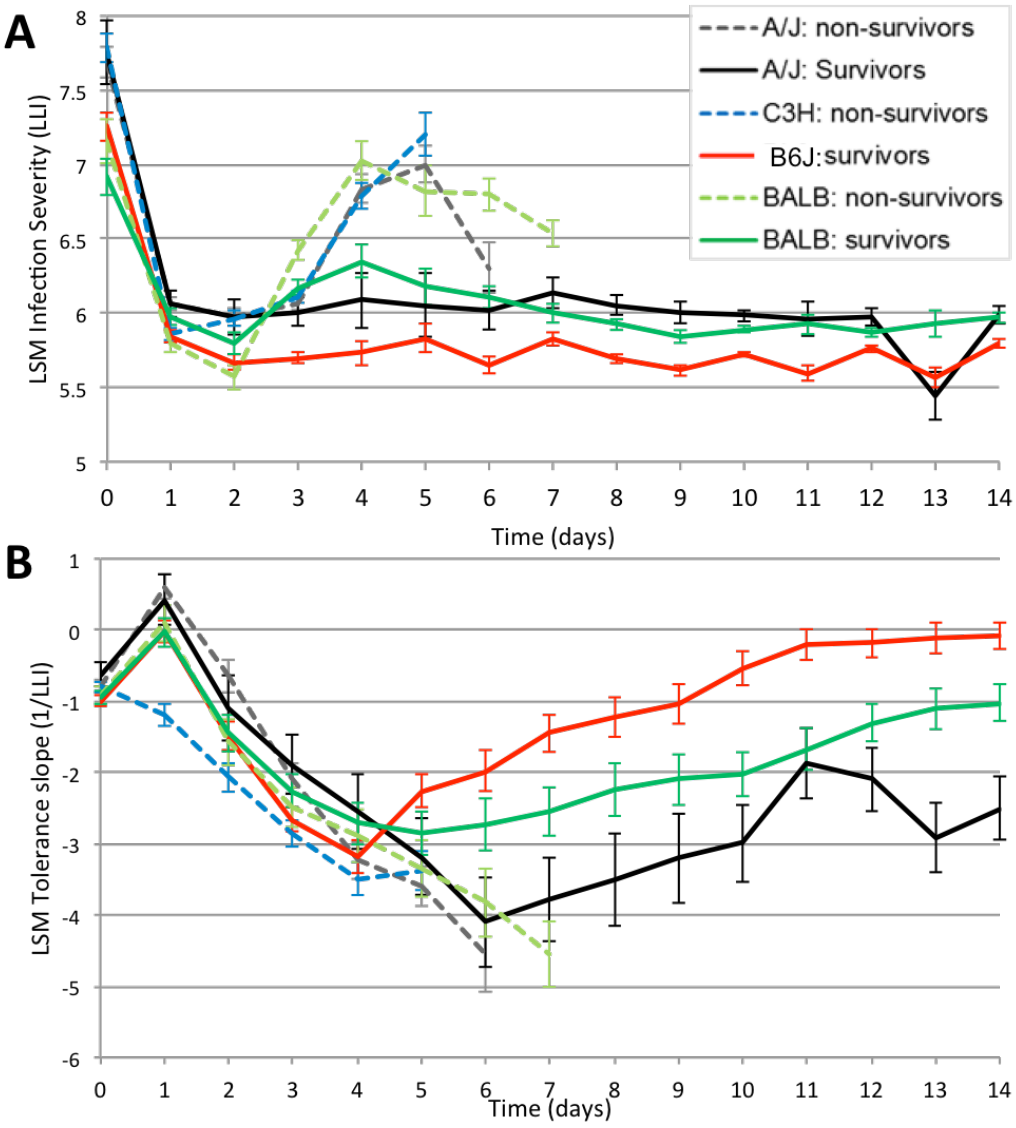
Similarly, the linear mixed model for carrying out the repeated measures ANCOVA for tolerance used daily %BW loss as dependent variable, daily LLI as covariate, and

mouse strain, binary survival outcome, day of measure (dpi), and all statistically significant interactions as fixed effects, and individual mouse as random effect. The intercept was fixed at zero, corresponding to zero percent BW loss in the absence of infection. The analysis provided least square means (LSM) tolerance slope estimates with standard errors for every mouse strain by survival outcome at each day post infection. Significant differences between these slope estimates based on the F-test statistics for mouse strain-by-infection severity-by-dpi interaction provided evidence for genetic variation in estimated tolerance trends over time. The analyses provided least square means (LSM) and standard errors for infection severity (inverse of resistance) and tolerance slope estimates for every mouse strain by survival outcome at each dpi.

5.2.2 Results

Appendix 5.2 Figure shows the daily **Least square mean (LSM) estimates (and standard errors) for resistance and tolerance.** Initial reduction in infection severity occurred in all strains between 0 and 1 dpi, reflecting partial clearance of the inoculated bacteria (Appendix 5.2 Figure).

Appendix 5.2 Figure. Least square mean (LSM) estimates (and standard errors) for resistance and tolerance estimates over the time-course of infection based on daily measures of infection severity and health. **(A)** LSM Resistance, **(B)** LSM Tolerance. Different mouse inbred strains are indicated by different colours and mouse strains or defined subgroups that succumbed to infection are indicated by dotted lines. Caution is advised when interpreting the results for the surviving A/J mice, as there were only 2 survivors.

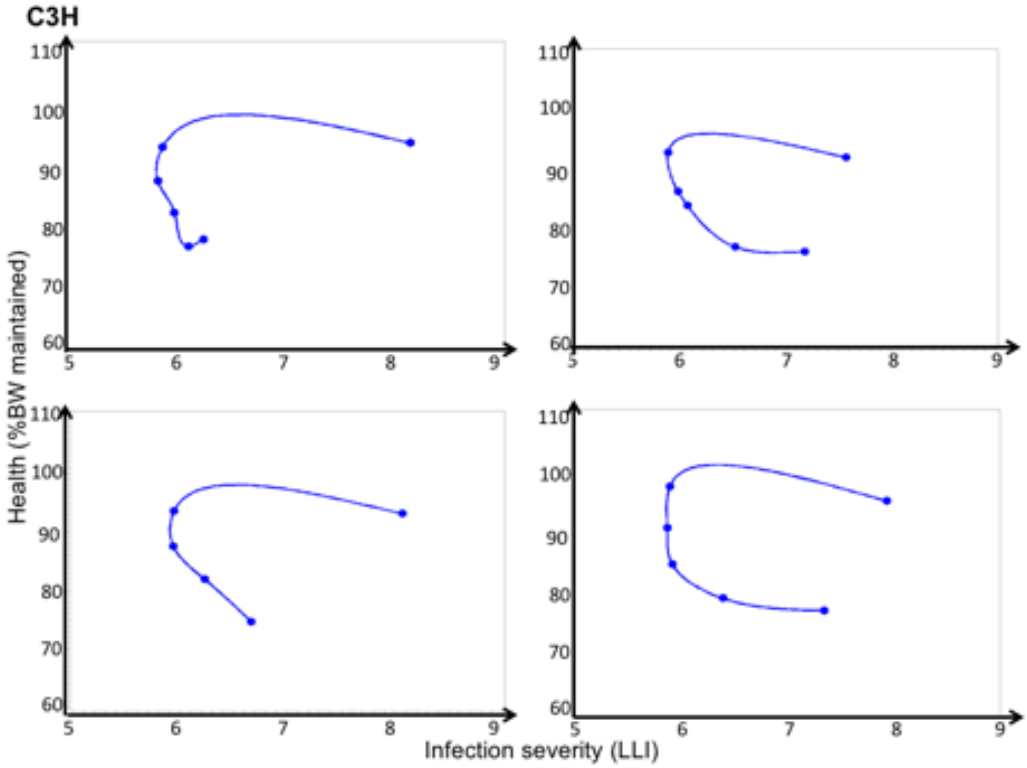


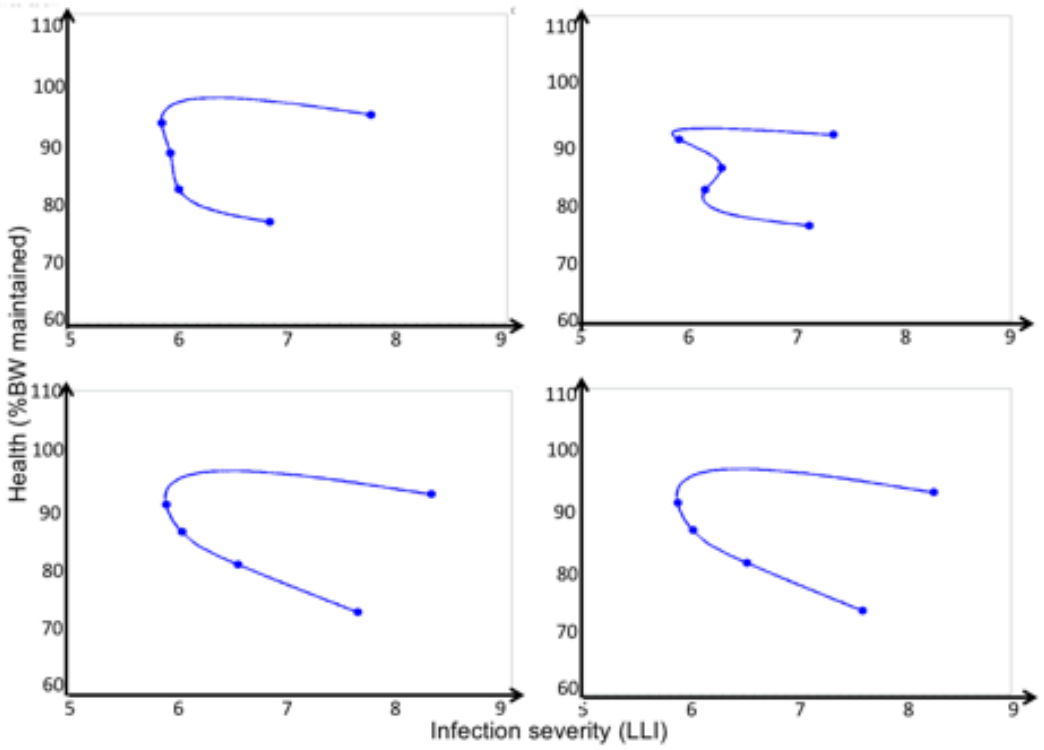
Bacterial replication was then moderately controlled in all surviving mouse strains after 1 dpi, whereas all mice that succumbed to infection experienced resurgence in infection severity from 3 dpi onwards (Appendix 5.2 Figure A). The statistical analysis for resistance revealed statistically significant strain, survival and day effects, as well as strain by survival by day interactions ($p<0.05$). Differences in resistance between mouse strains were greatest at 5 dpi, with the most striking difference between survivors and non-survivors (Appendix 5.2 Figure A). For the surviving mice, the ranking between strains was relatively stable over time, with B6J maintaining highest resistance estimates for the majority of time points.

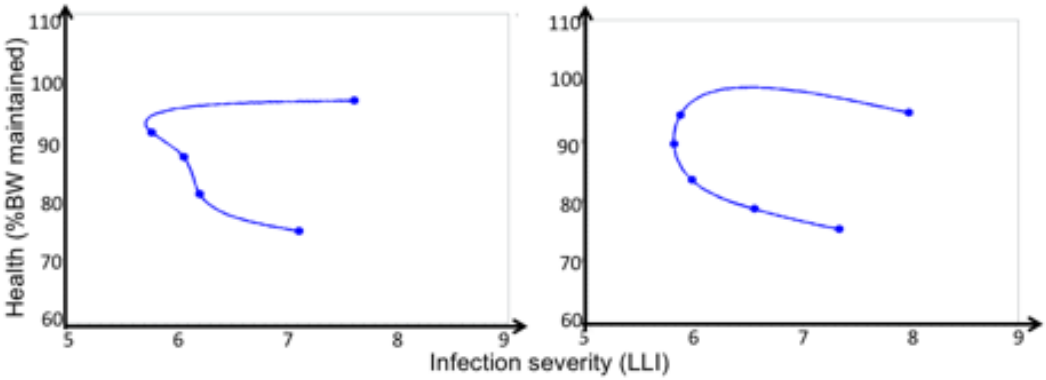
Similar to resistance, each of the fixed effects and all interactions fitted in the statistical models for tolerance proved statistically significant ($p<0.05$), indicating that the trends in daily tolerance estimates vary substantially between the mouse strains and survival groups. In contrast to resistance, survivors do not necessarily rank above non-survivors in tolerance estimates at any day post infection (Appendix 5.2 Figure B). Appendix 5.2 Figure B also reveals a common time trend in tolerance estimates: With the exception of C3H, all mouse strains experienced initial increased levels of estimated tolerance between 0 and 1 dpi. After 1 dpi, tolerance estimates decrease to negative values for all mouse strains and remain negative throughout the two weeks observation period. Depending on the mouse strain and survival group, minimum levels of tolerance were reached between 4 to 7 dpi, after which tolerance estimates gradually increase for all surviving mouse strains. B6J is the only mouse strain to fully recover its body weight as illustrated by a LSM tolerance estimate close to zero at 14 dpi.

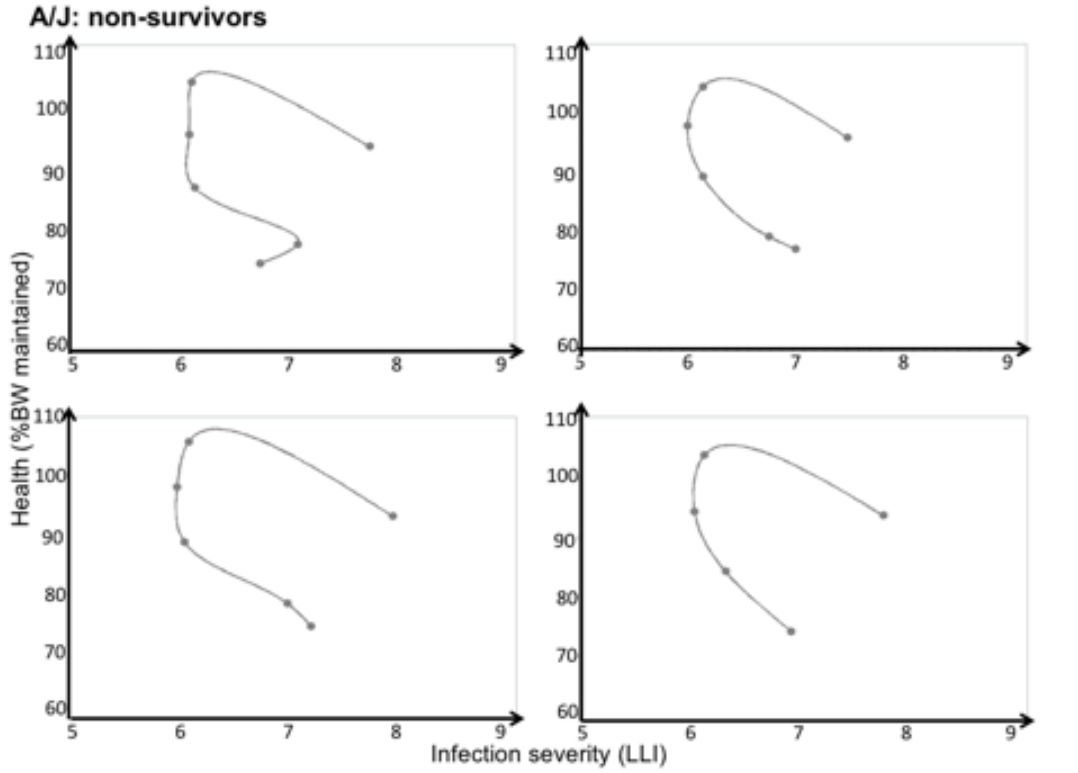
Appendix 5.3

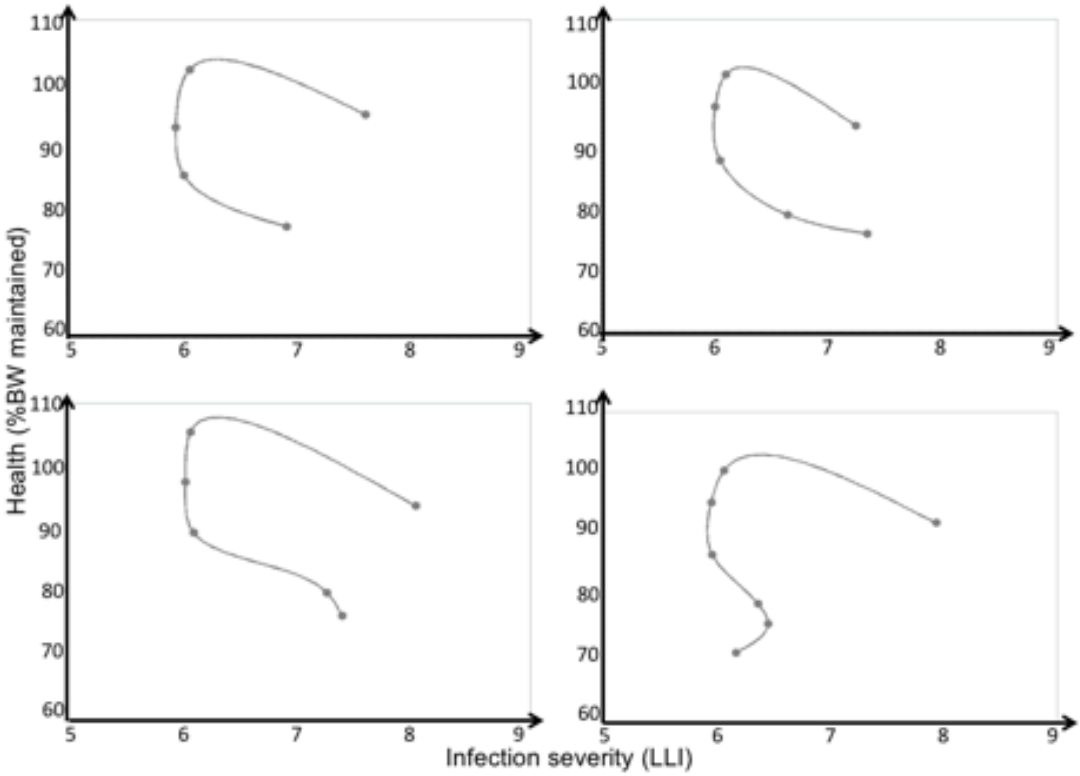
Individual trajectories for each mouse strain

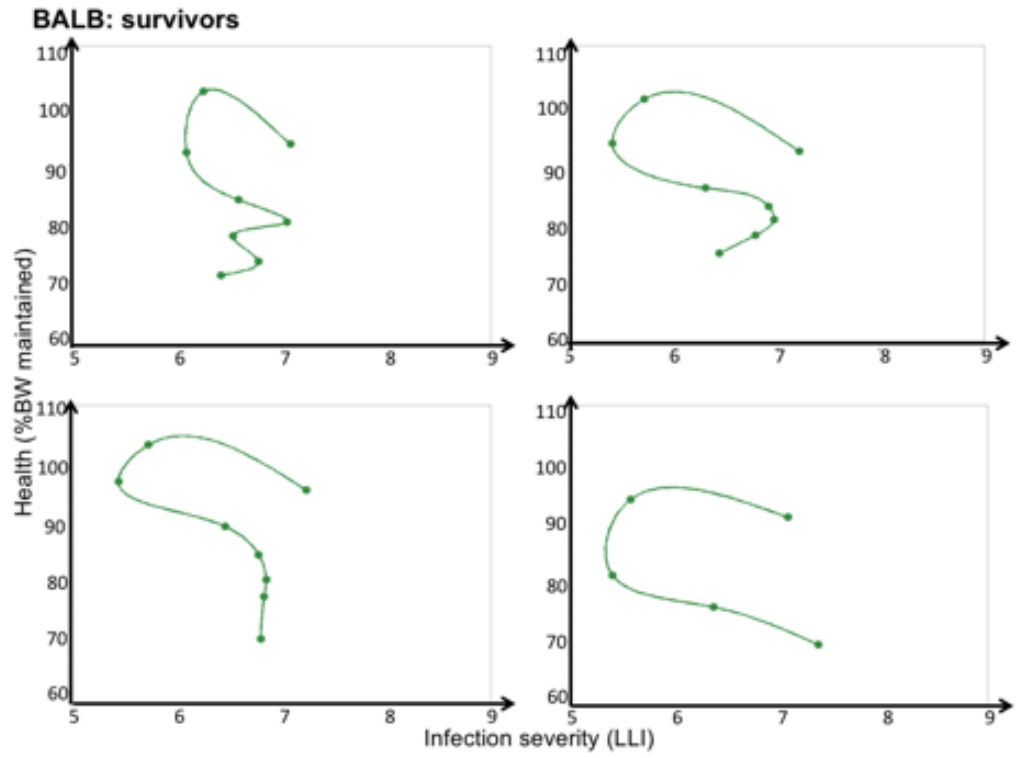


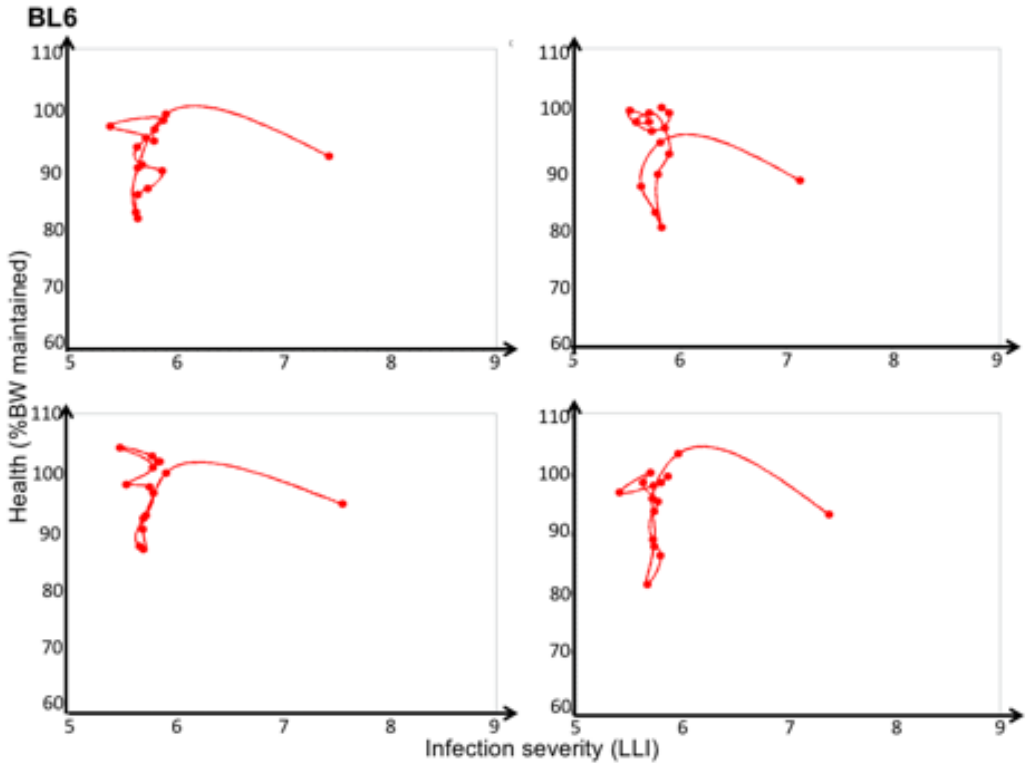

Health (%BW maintained)
110
100
90
80
70
60
5
6
7
8
9
Infection severity (LLI)

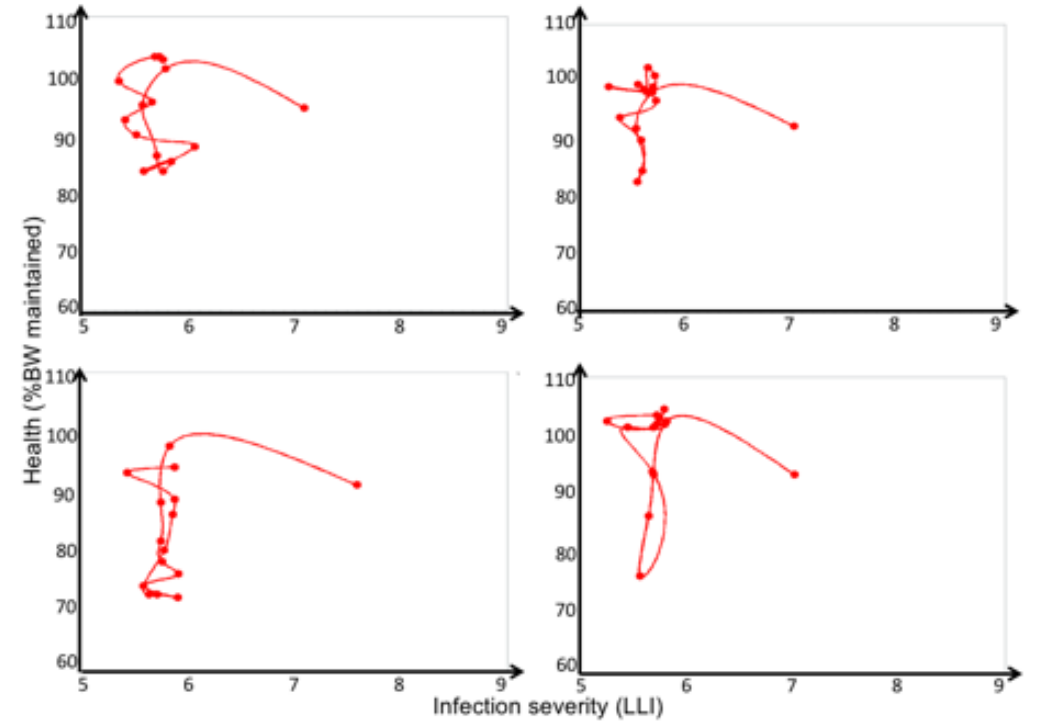

Health (%BW maintained)
110
100
90
80
70
60
5
6
7
8
9
Infection severity (LLI)

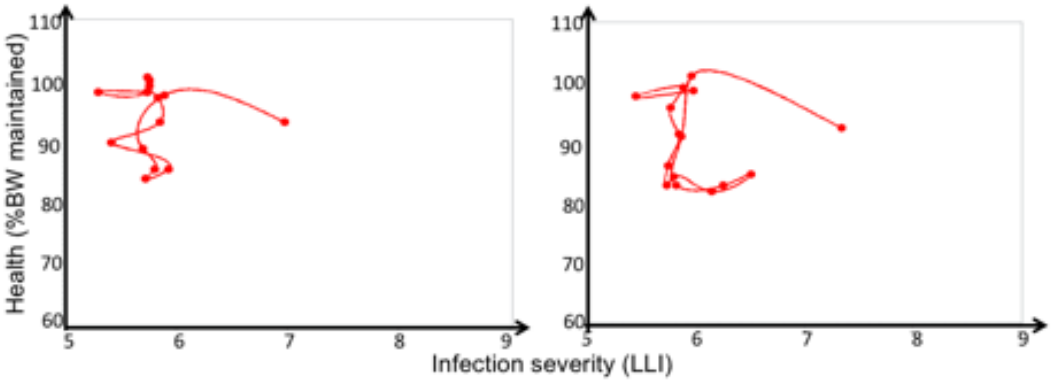

A/J: non-survivors
Health (%BW maintained)
Infection severity (LLI)

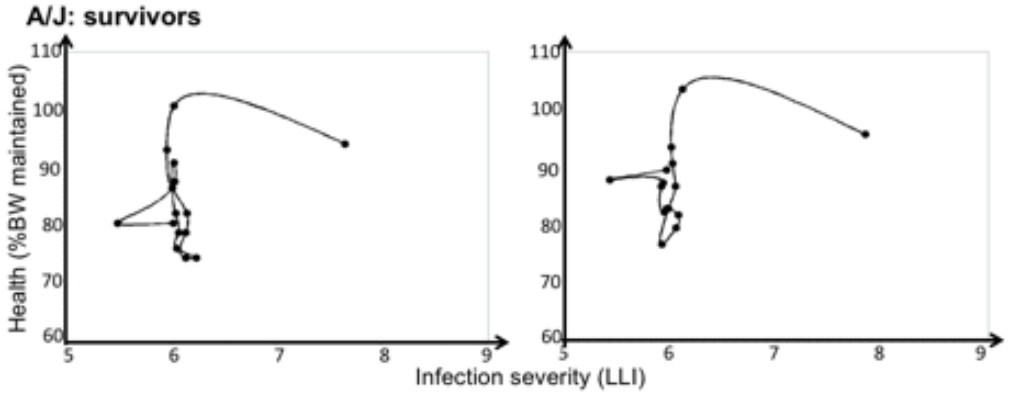

Health (%BW maintained)
Infection severity (LLI)

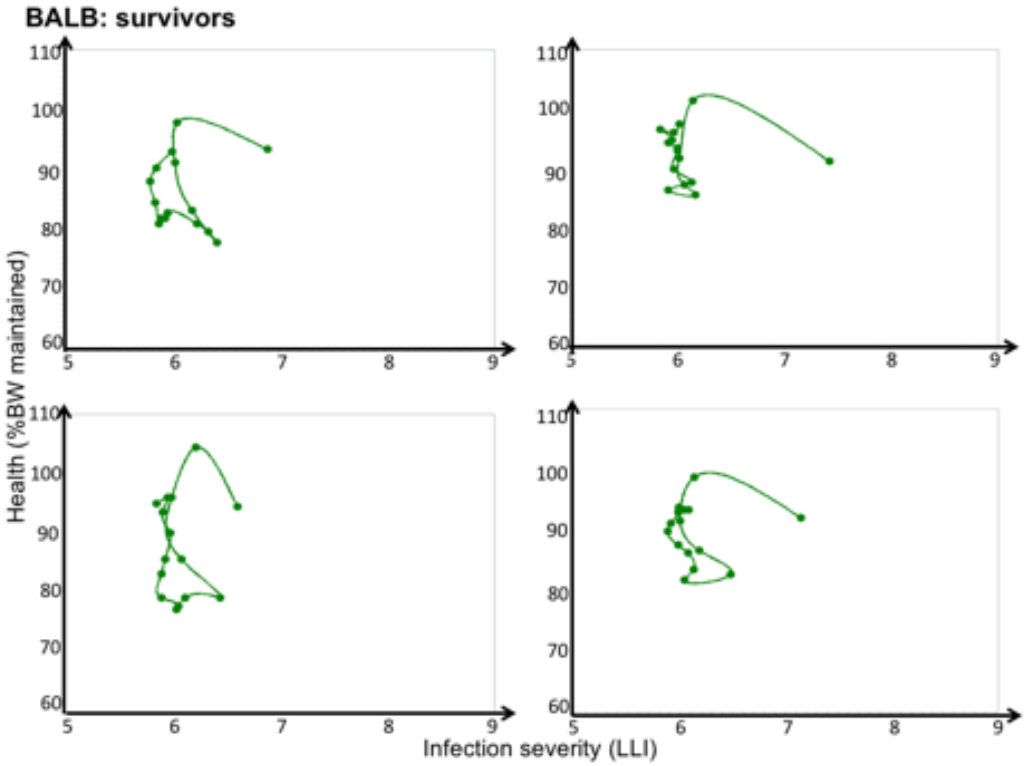

BALB: survivors
Health (%BW maintained)
Infection severity (LLI)

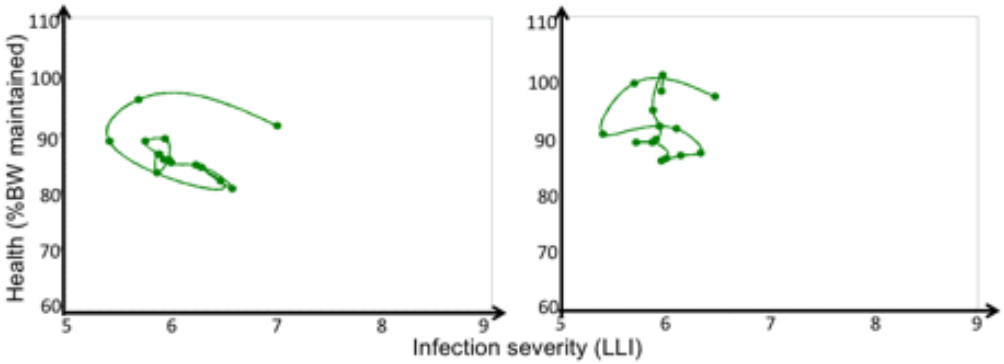

BL6
Health (%BW maintained)
Infection severity (LLI)


Health (%BW maintained)
110
100
90
80
70
60
5
6
7
8
9
Infection severity (LLI)


Health (%BW maintained)
110
100
90
80
70
60
5
6
7
8
9
Infection severity (LLI)


A/J: survivors
Health (%BW maintained)
110
100
90
80
70
60
5
6
7
8
9
Infection severity (LLI)


BALB: survivors
Health (%BW maintained)
110
100
90
80
70
60
5
6
7
8
9
Infection severity (LLI)


Health (%BW maintained)
110
100
90
80
70
60
5
6
7
8
9
Infection severity (LLI)

Appendix 5.4

Numerical trajectory sequences

Denoting simultaneous changes in infection severity (S) and health (H) over time for 40 mice from four different inbred strains, and corresponding trajectory types resulting from the cluster analysis with 4 clusters. Sequence numbers 1,2,3,4 and 0 refer to $S^{+}H^{-}$, $S^{-}H^{-}$, $S^{-}H^{+}$, $S^{+}H^{+}$, and death, respectively. Day i (top row) corresponds to the time period between day i-1 and day i.

Cluster	Mouse strain	Day														
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	C3H: non-survivors	1	2	1	1	1	1	1	0	0	0	0	0	0	0	0
1	C3H: non-survivors	1	2	1	1	1	1	1	0	0	0	0	0	0	0	0
1	C3H: non-survivors	1	2	1	1	1	1	0	0	0	0	0	0	0	0	0
1	C3H: non-survivors	1	2	1	1	1	1	0	0	0	0	0	0	0	0	0
1	C3H: non-survivors	1	2	1	1	1	1	0	0	0	0	0	0	0	0	0
1	C3H: non-survivors	1	2	1	2	1	1	0	0	0	0	0	0	0	0	0
2	A/J: non-survivors	1	3	2	2	1	1	2	0	0	0	0	0	0	0	0
2	A/J: non-survivors	1	3	2	1	1	1	0	0	0	0	0	0	0	0	0
2	A/J: non-survivors	1	3	2	1	1	1	1	0	0	0	0	0	0	0	0
2	A/J: non-survivors	1	3	2	2	1	1	1	0	0	0	0	0	0	0	0
2	A/J: non-survivors	1	3	2	2	1	1	0	0	0	0	0	0	0	0	0
2	A/J: non-survivors	1	3	2	2	1	1	1	0	0	0	0	0	0	0	0
2	A/J: non-survivors	1	3	2	1	1	1	1	0	0	0	0	0	0	0	0
2	A/J: non-survivors	1	3	2	2	1	1	1	2	0	0	0	0	0	0	0
2	BALB: non-survivors	1	3	2	1	1	1	2	1	2	0	0	0	0	0	0
2	BALB: non-survivors	1	3	2	1	1	1	1	2	2	0	0	0	0	0	0
2	BALB: non-survivors	1	3	2	1	1	1	2	2	2	0	0	0	0	0	0
2	BALB: non-survivors	1	3	2	1	1	1	0	0	0	0	0	0	0	0	0
2	C3H: non-survivors	1	2	2	1	1	1	1	0	0	0	0	0	0	0	0
2	C3H: non-survivors	1	2	2	1	1	1	0	0	0	0	0	0	0	0	0
2	C3H: non-survivors	1	3	2	1	1	1	0	0	0	0	0	0	0	0	0
2	C3H: non-survivors	1	2	2	1	1	1	1	0	0	0	0	0	0	0	0
3	BL6: survivors	1	3	2	2	2	3	3	4	4	3	3	4	3	4	2
3	BL6: survivors	1	3	2	1	2	2	3	4	3	3	3	4	3	1	4
3	BL6: survivors	1	3	2	2	2	3	3	4	4	3	3	4	3	4	4
3	BL6: survivors	1	3	2	2	2	2	4	3	3	4	3	4	2	4	4
3	BL6: survivors	1	3	2	1	2	4	4	3	4	3	3	4	3	4	4
3	BL6: survivors	1	3	2	1	4	2	4	3	4	3	3	3	3	3	3
3	BL6: survivors	1	3	2	2	4	4	4	4	3	3	4	4	3	4	3
4	A/J: survivors	1	3	2	2	1	1	1	1	4	3	3	3	3	2	2
4	A/J: survivors	1	3	2	2	2	2	2	2	4	3	3	3	3	3	3
4	BALB: survivors	1	3	2	1	1	1	3	3	3	3	3	3	3	3	3
4	BALB: survivors	1	3	2	2	1	1	3	4	3	3	3	3	3	3	4
4	BALB: survivors	1	3	2	2	1	1	2	2	3	3	3	3	3	3	3
4	BALB: survivors	1	3	2	1	1	1	2	3	3	3	3	3	4	3	4
4	BALB: survivors	1	3	2	1	1	3	1	3	3	3	3	4	4	3	2
4	BALB: survivors	1	3	2	1	3	1	2	2	4	3	3	4	4	3	4
4	BL6: survivors	1	3	2	2	3	2	1	2	1	3	3	4	4	4	3
4	BL6: survivors	1	3	2	2	3	3	4	3	4	3	4	4	4	3	3
4	BL6: survivors	1	3	2	2	3	2	4	3	4	3	3	4	4	2	3

S+H-
S-H-
S-H+
S+H+

Chapter 6

Alternative approaches to tolerance: infection trajectories of pigs infected with Porcine Reproductive and Respiratory Syndrome virus

6.1 Introduction

Using random regression models applied to single cumulative measures of pathogen load and growth for paternal half-sibs we did not find conclusive evidence for genetic variation in tolerance of pigs to PRRS (chapter 2). Only when the 21 or 42 day observation period was broken down into individual stages of infection, or when longitudinal information was harnessed with a repeated measures model, could genetic variation in tolerance of pigs to PRRS be identified (see chapter 3). However, the random regression models were sire models, implying that estimated breeding values could only be obtained per sire rather than for each individual. As such, tolerance maybe of limited use for genetic selection for improvement of host response to infection.

Studies, including our own study presented in chapter 5, have shown that within an individual, both expression of resistance and tolerance change over the time-course of infection [31,37,39,85]. In chapter 5 we have shown that infection trajectories that follow the pairwise progression of host performance and within-host pathogen load of each individual over time offer a way to capture these dynamics on the individual level.

We have also demonstrated that trajectories are partly determined by underlying host genetics (chapter 5, [85]). However, the mouse study was limited to four genetically distinct highly inbred mouse strains. The PRRS Host Genetics Consortium (PHGC) dataset, which provides repeated measures of viremia and growth and was used to examine genetic variation in tolerance of pigs to PRRS in chapters 2 to 4, offers a unique dataset to construct and examine infection trajectories of outbred animal populations.

In chapter 5, genetic differences between the inbred mouse strains could be captured by applying permutation tests to numerical trajectory sequences of each individual from the four distinct strains. This methodology does not lend itself to study genetic variation in infection trajectories of outbred populations. For this reason, a different approach was adopted to statistically analyse infection trajectories of PRRSV infected pigs. Rather than performing quantitative genetic analysis directly on the trajectories, the approach in this chapter was to derive novel phenotypes based on trajectory characteristics, and then to apply standard quantitative genetic mixed models to these phenotypes. This study thus aimed to a) construct infection trajectories of pigs infected with PRRSV in order to derive novel phenotypes from trajectories for subsequent quantitative genetic analysis; b) identify genetic variation in these trajectory-derived phenotypes; c) assess whether the previously identified WUR genotype was associated with the novel phenotypes.

6.2 Methods

6.2.1 Infection experiment and data

The PRRS Host Genetics Consortium (PHGC) data from the 9 PRRSV challenge trials described in chapter 1 were used (see chapter 1 for details). As the novel phenotypes derived from trajectories required unidirectional movement, i.e. no loops in the trajectories, only measurements from 14 dpi onwards were used in this study, as according to previous analyses, all individuals had reached peak viremia by 10.6 dpi and started to gradually reduce viremia [62]. Similarly, only individuals that did not have a rebound in viremia were used, as rebounders also generate trajectory loops. The data analysed in this study thus comprised of 1183 pigs from 104 sires.

6.2.2 Construction of PRRSV infection trajectories and definition of trajectory phenotypes

Trajectories were constructed using weekly repeated measures of growth (Average Daily Gain (ADG)) for each individual. ADG was calculated as growth rate between day *k-1* and *k*. For example, ADG between 21 and 28 dpi (ADG_{28}) was the difference in BW between these two days divided by 7. Weekly changes in viremia were also used to construct trajectories using predictions from the Woods function (see chapter 3). For example, $viremia_{28}$ was predicted using Woods viremia predictions at 28 dpi. In this way, these weekly ADG and viremia measures were used to construct trajectories for each individual.

To derive trajectory components, weekly two-dimensional vectors were calculated for each trajectory:

$$\boldsymbol{V_k} = (V_{(PL_k)}, V_{(ADG_k)})$$

where $V_{(PL_k)}$ represents change in pathogen load (PL = viremia) from week k-1 to k, and $V_{(ADG_k)}$ represents corresponding change in performance (ADG) for the same duration. The weekly trajectory vectors gave rise to two phenotypes: The first, θ_k, is calculated by:

$$\theta_k = \tan^{-1}(\frac{V_{(ADG_k)}}{V_{(PL_k)}})$$

where θ_k is the angle between the trajectory vector $\boldsymbol{V_k}$ and the viremia-axis. Thus θ_k assumes values between -90º and 90º, where the sign of the angle corresponds to negative or positive changes in $V_{(ADG_k)}$. Thus, high positive values of θ_k represent a fast increase in weekly growth rate as viremia decreases from week k-1 to k.

The second trajectory derived component was the *magnitude* of the weekly trajectory vector $||V_k||$, and was given by:

$$||V_k|| = \sqrt{(V_{PL_k})^2 + (V_{ADG_k})^2})$$

which describes the magnitude of the combined weekly change in growth rate and pathogen load (viremia).

Figure 1. Trajectories from two individuals (A and B), where each vector connects weekly pairwise measures of viremia and ADG (starting at 7 dpi, through weekly to 42 dpi). Each trajectory thus depicts dynamic changes in ADG with respect to changes in viremia. Individual A increases in ADG while steadily reducing viremia from 7 dpi to the end of the experiment, whereas the growth rate of individual B decreases overall. Also illustrated in Figure 1A is the trajectory vector between 21 and 28 dpi: $\mathbf{V}_{28}$, with components V_{ADG} (change in ADG) and V_{PL} (change in pathogen load (viremia)), which represents change in pathogen load (viremia) and ADG over the 1 week period.

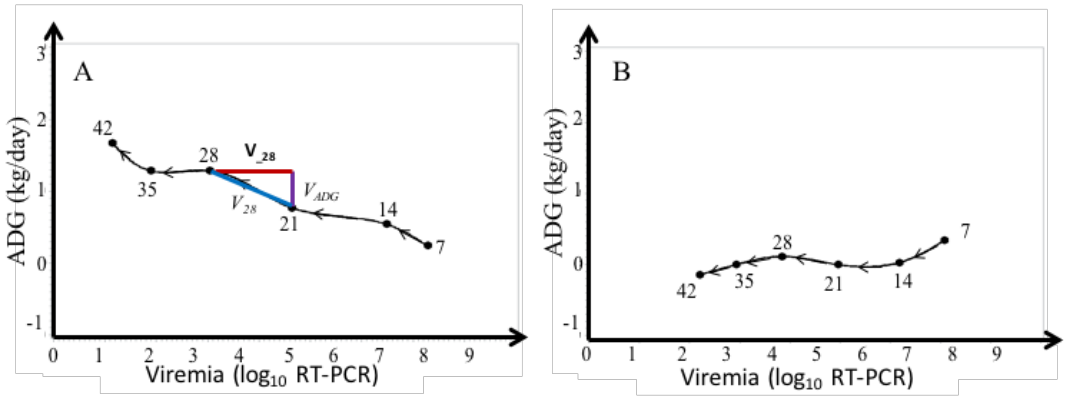
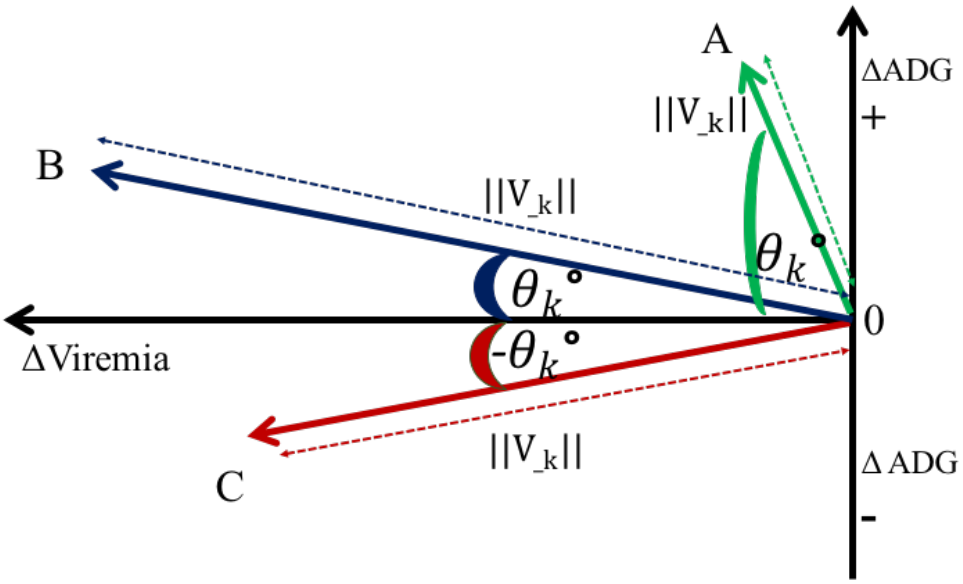


Figure 2. **Illustration of two novel phenotypes derived from trajectory vector components for 3 individuals** (A, B and C, coloured green, blue and red, respectively), where each line represents a trajectory vector from week k-1 to k. $\theta_{\_k}$ represents the angle from origin of each vector with the viremia axis, where higher value of $\pm\theta_{\_k}$ corresponds to higher gain/loss in ADG with respect to a unit change in viremia. $||V_{\_k}||$ represents the magnitude of each vector from week k-1 to k, where the higher the value of $||V_{\_k}||$, the greater the rate of combined change in viremia and growth. Individual A has the highest $\theta_{\_k}$, but the lowest $||V_{\_k}||$; in contrast to individuals A and B, individual C has negative $\theta_{\_k}$, so experiences a reduced ADG with decreasing viremia. Individual B experiences the fastest rate of change ($||V_{\_k}||$), and individual A has the slowest rate of change.



6.2.3 Statistical analyses

All statistical analyses were carried out using ASReml 3.0 (Gilmour et al., 2009). To identify genetic variation in trajectory components, the following univariate animal model was adopted:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Up} + \mathbf{Ml} + \mathbf{e}, \qquad [1]$$

where $\mathbf{y}$ is a vector comprising individual records for $\theta_{\_k}$ or $||V_{\_k}||$ from week k-1 to k; $\mathbf{b}$ is a vector of fixed effects, including for the interaction of experimental trial and parity of the dam when offspring were born (trial-by-parity), sex of the offspring, and age, which was fitted as a fixed covariate; $\mathbf{a}$ is a vector of additive genetic effects, with $\mathrm{Var}(\mathbf{a}) = \mathbf{G}\otimes\mathbf{A}$ is the genetic variance-covariance matrix where $\mathbf{G}$ is a vector of genetic variances and $\mathbf{A}$ the pedigree relationship matrix; $\mathbf{p}$ is a vector of pen effects nested within trial, with $\mathrm{Var}(\mathbf{p}) = \mathbf{I}\otimes\mathbf{K}$, where is the identity matrix, and $\mathbf{K}$ is the corresponding vector of variances for pen effects; $\mathbf{l}$ is a vector of litter effects, where $\mathrm{Var}(\mathbf{l}) = \mathbf{I}\otimes\mathbf{L}$, where $\mathbf{L}$ is the corresponding vector of pen effect variances; and $\mathbf{e}$ is a vector of residuals, with $\mathrm{Var}(\mathbf{e}) = \mathbf{I}\otimes\mathbf{R}$ where $\mathbf{R}$ is the variance-covariance matrix for the residual effects for each trait; and $\mathbf{X}$, $\mathbf{Z}$, $\mathbf{U}$ and $\mathbf{M}$ are incidence matrices for fixed, animal, pen and litter effects, respectively. Traits and model residuals were normally distributed, as indicated by a Shapiro-Wilk test ($p<.0001$).

6.4 Results

Mean $\theta_{\_k}$ was lowest from 21-28 dpi (-4.11) and highest by end of infection (2.08), indicating that recovery in growth with respect to unit changes in viremia was higher towards the end of infection (Table 1). There was a huge variation in $\theta_{\_k}$ associated with each weekly trajectory vector, in particular, at the later infection stages.

Magnitude of vectors ($||V_{\_k}||$) decreased throughout infection from 21 dpi onwards, with highest at 14-21 dpi (1.79) and lowest at 35-42 dpi (0.73), indicating that weekly changes in ADG and viremia reduced in magnitude after 3 weeks post infection. Standard deviation also decreased with infection time.

Table 1. Summary statistics of weekly $||\mathbf{V}_{\mathbf{\_k}}||$ and $\boldsymbol{\theta}_{\boldsymbol{\_k}}$, where $||V_{\_k}||$ is rate of change between two weekly measurements (indicated by days post infection (dpi)), where high $||V_{\_k}||$ means higher magnitude of vector between two measurements; and $\theta_{\_k}$ (between $-90°$ and $90°$) is the degree change in ADG with respect to change in viremia between consecutive weeks, respectively, where higher positive or negative angle means higher gain/loss of performance with respect to change in pathogen load.

	dpi	mean	sd	min	max
$\boldsymbol{\theta}_{\boldsymbol{\_k}}$ (°)	14-21	-2.55	7.94	-55.83	49.93
	21-28	-4.11	12	-61.41	56.2
	28-35	-2.13	22.4	-76.8	69.28
	35-42	2.08	31.04	-73.55	83.25
$\|\|\mathbf{V}_{\mathbf{\_k}}\|\|$	14-21	1.79	0.48	0.34	3.37
	21-28	1.42	0.26	0.27	2.94
	28-35	1.02	0.23	0.44	2.49
	35-42	0.73	0.24	0.1	1.99

Heritability of $\theta_{\_k}$ was not significantly different from zero across all time-points of infection, indicating that there was no identifiable genetic variation (Table 2). Rate of change ($||V_{\_k}||$) was of low to moderate heritability at earlier time of infection (0.20 and 0.12 for 14-21 and 21-28 dpi, respectively), but was not significantly different from zero later in infection. Furthermore, inclusion of WUR in the model was not statistically significant at any time of infection for either $\theta_{\_k}$ or $||V_{\_k}||$, indicating that WUR was not associated with either of the traits.

Table 2. Heritability (h^2), standard error (in brackets) of increase in growth, with respect to change in viremia ($\theta_{\_k}$) and rate of change in growth and viremia ($||\mathbf{V}_{\mathbf{k}}||$). Standard errors are in brackets. p-values for inclusion of WUR in the model are also given.

Trait	dpi	h2	WUR (p-value)
$\theta_{\_k}$ (°)	14-21	0.00	0.110
	21-28	0.00	0.600
	28-35	0.02 (0.04)	0.890
	35-42	0.00 (0.04)	0.300
$\|\|\mathbf{V}_{\mathbf{k}}\|\|$	14-21	0.20 (0.09)	0.129
	21-28	0.12 (0.08)	0.278
	28-35	0.00	0.159
	35-42	0.00	0.161

Given the low heritabilities of each trait at all time points of infection, genetic correlations between traits were not estimated. Phenotypic correlations between the traits, without adjusting for additional fixed and random effects, were not significantly different from zero.

6.5 Discussion

6.5.1 Summary of findings

This study aimed to derive novel phenotypes from trajectories which explain simultaneous changes in host performance with changes in pathogen load under PRRS infection. Two phenotypes were derived which explained gain/loss in weekly ADG in relation to change in viremia ($\theta_{\_k}$); and the magnitude of combined change in ADG and viremia between two consecutive weeks ($||\mathrm{V}_{\mathrm{k}}||$). However, with exception of

$||V_{\_k}||$ from 14-21 dpi, the traits were not found to have a significant genetic component. Given this, the WUR genotype was not associated with either trait.

6.5.2 Alternative approaches to analyse trajectories of outbred populations

Previous studies have highlighted the potential usefulness of individual trajectories for classifying and predicting host responses to infection [38,40,41]. However, no study to date has attempted to estimate genetic variance of trajectories of infected farm animals. Indeed, it is not obvious how to estimate genetic variance and breeding values for entire infection trajectories.

In chapter 5, we quantified differences in individual trajectories based on sequence construction followed by a permutation test [85]. However, a permutation test can only be applied to identify statistical differences between a limited number of groups (such as the 4 inbred mouse strains used to identify genetic variation in trajectories between strains in chapter 4), and was thus an inappropriate methodology to quantify genetic variation in infection dynamics of an outbred population. Defining novel phenotypes derived from trajectory characteristics, as performed in this chapter, constituted a straight-forward approach to carry out quantitative genetic analyses on trajectories.

An alternative quantitative genetics approach to estimate genetic variation of unidirectional (i.e. no loops) trajectories would be by adopting random regression models with spline functions fitted through the data trajectories [66,133]. These could also be exploited to estimate covariance functions affecting trajectory shape. This random regression approach was indeed attempted, but genetic variation could not be identified due to convergence issues. This is somewhat not surprising, since the linear random regression models were already difficult to fit to the PHGC data (see chapters

3 and 4). As such, this approach was not pursued further and the derivation of trajectory phenotypes was done instead.

6.5.3 Potential use of trajectory phenotypes for genetic improvement of host response to infection

In chapters 3 and 4, we identified a changing relationship between growth and resistance over the time-course of infection. Trajectories complement these analyses as they describe how growth rate changes with respect to changes in within-host VL over the time course of infection. Although don't lend themselves easily to quantitative genetic analysis, trajectory-derived phenotypes may provide valuable phenotypes which could be used for genetic selection of response trajectories. For example, a trajectory where rapid increase in viremia is not accompanied by drastic reductions in growth may be desirable. This could be achieved by increasing the trajectory angle $\theta_{\_k}$ at the time when viremia increase is fastest. $\theta_{\_k}$ could be loosely interpreted as a proxy for tolerance (higher growth at higher within-host viral load) measured at the individual level. If a pathogen is prevalent in a population and tolerance is the preferred trait to mitigate the impact of infection on the host, these trajectory-derived phenotypes could allow genetic selection for "tolerance" without the statistical requirements and complications of random regression methodology identified in chapters 2 and 3. Furthermore, genetic correlations between trajectory angles would provide useful insights whether genotypes experiencing low reduction in growth at an earlier stage of infection would be expected to also grow faster at the later stages of infection. These measures would thus provide useful insight about which trajectories are feasible to achieve through genetic selection.

The requirement for repeated measurements is a major bottleneck for implementing trajectories into breeding programmes. In a farm setting, it may not be possible to record repeated measures for each animal. As such, it may be possible to construct crude trajectory vectors and phenotypes $\theta_{\_k}$ and $||V_{\_k}||$ between two consecutive measurements to interpret overall change in growth with relation to long-term reduction in VL. Clearly, further studies are needed before infection trajectories can be implemented in breeding programmes.

6.5.4 Future research

The primary interest in adopting trajectory analysis is to identify the interaction between tolerance and resistance and how they change over the time-course of infection. Trajectory-derived phenotypes allow analysis of simultaneous changes in growth and VL that can be applied to an outbred population. Given that this study did not detect genetic variation in trajectory-derived phenotypes, alternative phenotypes or approaches which capture host infection dynamics and allow for genetic selection on the individual level may need to be explored. However, these phenotypes may be applicable to other data, if repeated measures of performance and pathogen load are available.

Chapter 7

General Discussion

7.1 Aims of study

This research aimed to: 1) Identify genetic variation in tolerance of pigs to PRRS using conventional random regression methodology; 2) Examine the relationship between resilience, resistance and tolerance over different stages of infection; 3) Assess genomic regions associated with growth and resistance at different stages of infection 4) Identify genetic variation in tolerance and resistance to listeria infection in mice using conventional linear mixed model methodology; 5) Develop a novel trajectory analysis methodology to analyse changes in resistance and tolerance over time, and determine genetic variation in trajectories; 6) Apply trajectory analysis to PRRS by deriving novel phenotypes from trajectories.

In chapter 2, we could not conclusively identify genetic variation in tolerance of pigs to PRRS based on random regression methods using the same time periods as previous studies [23,57,58]. However, this highlighted the potential limitations of adopting a reaction norm approach to identify genetic variation in tolerance in outbred populations (described on page 104). To overcome these limitations, performance in absence of infection (*vigour*) was simulated for half-sibs of individuals in the experiment. Genetic variation in tolerance to PRRS was thus demonstrated under the assumptions made in simulations, but only when genetic correlation between vigour and growth under infection was low to moderate. In chapter 3, stages of infection (early, mid and late) were defined based on underlying viremia profile characteristics for each individual. Using these stages of infection, we identified genetic variation in

tolerance at mid to late stage of infection. Additionally, by adopting a more robust repeated measurement model, we identified genetic variation in tolerance over the 21 and 42 days infection period, and a positive genetic correlation between resistance and tolerance, indicating that individuals who are more genetically resistant were also likely to be more genetically tolerant to PRRS. Furthermore, we identified that the WUR genotype previously associated with resistance was also suggestively associated with tolerance slope, highlighting that, over all stages of infection, resistance and tolerance may be pleiotropic. However, the WUR genotype was not associated with tolerance slope at any stage of infection; and resistance and tolerance were possibly negatively correlated. This indicated that care must be taken as to when genetic selection is used to improve host response to PRRS. Indeed, in chapter 4, the relationship between growth under infection (resilience) and resistance changed at each stage of infection. Individuals who had high resistance were likely to lose weight at early stage, but clear the virus by late stage, thus recovering in growth. However, at mid-stage of infection, less resistant individuals had higher growth, demonstrating that genetic selection for improved growth during this stage of infection could have undesirable consequences. Finally, this study showed that the WUR genotype, previously associated with resistance and growth at 21 and 42 dpi, respectively, was significant only at early-mid stages for resistance, but mid-late stages for growth, indicating a lag effect of VL on growth. No other genomic regions were associated with growth or resistance at any stage of infection.

In chapter 5, a diversion from virus infected pigs to bacterial infected mice was performed, as the mice data provided a suitable dataset for developing and validating a novel trajectory analysis without the inherent noise affecting disease phenotypes of

outbred populations. By adopting novel trajectory analysis, this study demonstrated that genetic variation in resistance and tolerance to listeria in mice existed. Furthermore, we found that early expression of resistance determined survival of infection. Survivors and non-survivors followed different paths of infection, which were partially determined by genetics. The results were encouraging for applying trajectory analysis also on the outbred pig population of this thesis. In chapter 6, we derived novel phenotypes from PRRS trajectories of pigs that explained changes in growth in relation to change in pathogen load as a proxy for tolerance, and applied them to the PRRS data. However, there was limited evidence to suggest genetic variation in these phenotypes.

This chapter will discuss 1) whether tolerance is a feasible trait for genetic improvement of livestock response to infectious disease; 2) whether genetic selection for resilience is a suitable infectious disease control strategy; 3) the adoption of novel methodology to analyse changes in resistance and tolerance over time, and the potential for genetic selection for host response at different stages of infection, and finally 4) implications of this research in a wider context.

7.2 Is tolerance a feasible trait for genetic improvement of livestock response to infectious disease?

Robust phenotypes are required for identification of genetic variation in traits, and to permit genetic selection for improvement of host response to infection. Conceptually, tolerance is a powerful and useful descriptive phenotype that has attractive qualities and outcomes which should appeal to animal breeders. However, to date, there is still a large gap in knowledge of whether animals in outbred populations vary genetically

in tolerance. It is likely that this has been partly due to difficulty in meeting the statistical requirements to accurately estimate genetic variance in tolerance to infection.

Tolerance is a phenotype which describes change in host performance with respect to pathogen load; a trait which could provide excellent benefits in both maximising growth under difficult conditions, and that has potential evolutionary benefits if implemented into a breeding program, as highlighted in the general introduction. However, conventional methodology to estimate tolerance may be difficult to implement in practice, where a reaction norm approach is adopted to regress health or performance on within-host pathogen load [8].

In chapters 2 and 4, we identified genetic variation in tolerance of pigs to PRRS, and in mice to listeria using a reaction norm approach where maximum body weight (BW) loss was regressed on maximum pathogen load [85]. As such, this approach was adequate to identify genetic variation in tolerance. However, the reaction-norm approach has been relatively simple to implement in estimating genetic variation in tolerance where groups are few. Raberg et al used 5 inbred mouse strains [5], Blanchet et al. used 8 sites of measurement [29], and, in chapter 5, four highly inbred mouse strains were used [85]. This allows adoption of a regression approach in which group-by-pathogen load interaction is included as a fixed covariate. By adopting this approach, genetic variation in tolerance could be identified by a statistically significant interaction between family (e.g. genotype or strain) and pathogen load in an ANCOVA [5].

If a reaction-norm approach is required for a more complex population structure, random regression models (RRMs) must be adopted to estimate genetic variance in

tolerance [4,35,37].

While RRMs have previously been shown to be powerful tools that require few parameters to estimate genotype-by-environment interaction [134], there are complications and potential confounding when applied to tolerance weakens their application to identify genetic variation in tolerance. Firstly, RRMs are highly senstitive to the underlying data structure and may generate biased estimates, particularly if sample size is limited or relatedness is poor. [35,36,70]. Bias in slope can be caused if family size is small [35]. As such multiple measures per group and a large amount of statistical power is required. In chapter 3, we found difficulty in implementing RRMs to identify genetic variation in tolerance of pigs to PRRS for these reasons. Genetic variation could only be identified under certain assumptions made in simulations. However, the requirement to double the size of the dataset was likely helpful in increasing statistical power required to identify genetic variation in tolerance, by reducing confounding between level and slope, and increasing the distribution of data around centred VL. Indeed, in chapter 4, we used repeated measures for each individual to triple the size of the dataset, which provided adequate statistical power to estimate genetic variance in tolerance slope. This requirement for large amount of data to overcome limits the potential of random regression when applied to animal breeding. Furthermore, tolerance is conceptually defined at the individual level, where change in host performance with regards to change in within-host pathogen load. However, tolerance can only be defined at the group level using a reaction norm approach, which limit their application in animal breeding [4,5].

Reaction-norms estimate tolerance as a static trait, which, as found in chapters 2 and 3, is unlikely to be the case, given the changing relationship between resistance and

tolerance and resistance and growth. Again, repeated measures for each individual required to correctly estimate potential changes in tolerance further limits the application of a reaction norm for improvement of host response to infection.

Accurate estimation of tolerance on performance in absence of infection (vigour) provides an additional challenge. Assuming a linear relationship between pathogen load and performance, heterogeneity in vigour will affect the slope (see figure 1 in general introduction). In this research, estimates of vigour were not available for the PHGC data. The weak evidence for significant genetic variation in tolerance to PRRS from the univariate random regression models in this study may simply reflect a lack in statistical power for disentangling sire effects on regression slope and level. Genetic variation in tolerance was only evident where low genetic correlations with vigour were simulated (chapter 2). This may be partly overcome by setting vigour to 0, or 100% for all individuals, as we did in chapter 5, and modelling subsequent losses in performance. However, estimation of tolerance is still dependent on vigour, which is an additional complication of application of RRMs to animal breeding.

It is important to note that genetic variation in resistance to PRRS could cause potential caveats for estimating tolerance, as it implies that VL values are not homogeneously distributed amongst sires, with more resistant sires having predominantly low VL measures, and less resistant sires predominantly high VL values. This dependence of tolerance on resistance is another limitation to their application. A further complication is that dependence of tolerance on resistance may be also depend on the type of immune response mechanism at time of measurement, i.e. tolerance may be a due to direct (damage caused by pathogen) or indirect (damage caused by immune response) tolerance mechanisms, and this may change throughout infection. Indeed, this

highlights another possible problem with tolerance as a phenotype. Firstly, at what time and how to measure within-host pathogen load is a decision that will greatly affect tolerance. It is likely that the most robust measure is the area under the curve of pathogen load [4]. However, this requires repeated measures, and may be difficult to implement (see chapters 1 to 3). If repeated measures are available however, a possible proxy for tolerance may be found with trajectories and their derived phenotypes which address the various potential problems with tolerance applied to outbred populations outlined above.

7.3 Would selecting for resilience be sufficient?

Resilience is commonly defined as the ability of the host to maintain performance when exposed to infection, i.e. the genetic effect on performance during an outbreak of disease [25,134]. Resilience differs from tolerance in that the underlying pathogen load of the host is not considered. In chapters 1 and 2 we discussed the statistical constraints of estimating genetic variance in tolerance. Resilience is unlikely to have these statistical constraints. As such, resilience may be a more practical and attractive trait to implement into an animal breeding program. However, resistance and tolerance together contribute to overall resilience of an individual [4]. A careful balance between resistance and tolerance that would maximise resilience of the individual to infection is the ideal genetic selection program for control of infectious disease. It should be noted that where within-host pathogen load is low, both tolerant and non-tolerant individuals may be resilient to infection, given that an increase in pathogen load could lead to loss of tolerance. However, where pathogen load is high, it is only tolerant individuals that exhibit resilience. Thus, care must be taken to account for within-host

pathogen load when selecting for resilience. In chapter 3, we identified an overall positive strong relationship between resistance and tolerance i.e. that individuals who were more resistant were also more tolerant. Furthermore, we identified that the WUR genotype that was known to be associated with growth under infection and resistance to PRRS [23,57,58], was also suggestively associated with tolerance. This indicates that resistance and tolerance may be pleiotropic, indicating that selection of one trait may potentially have a positive effect on the other. If this is indeed the case, then selection for resilience would likely improve both resistance and tolerance to PRRS.

It could be concluded, then, that from a pure production point of view at least, genetic selection for resilience may indeed be the answer to improving pigs response to PRRSV infection. However, it should be noted that this may not be the case for other infectious diseases. As stated in the introduction, there are theoretical different epidemiological and evolutionary outcomes for selecting for resistance or tolerance, or indeed, a balance of both. Briefly, resistance could put selection pressure on the pathogen and drive pathogen co-evolution, where tolerance would not affect the pathogen, per se, and therefore would not drive pathogen co-evolution. Furthermore, only resistance could potentially eradicate a pathogen from a population, where tolerance would only mitigate the impact of infection on the host [32,79,124]. As such, selection for tolerance would be preferred where a pathogen is highly prevalent and difficult to eradicate from a population, such is the case with PRRS. Our finding that resistance and tolerance are positively correlated indicates that the epidemiological and evolutionary theory may be more convoluted, in that there may not be the trade-off between resistance and tolerance highlighted in the literature [30], at least when applied to PRRS.

Still, great care must be taken as to the timing of genetic selection. Having identified genetic variation in tolerance at mid-late stage of infection, we also analysed the relationship between resistance and tolerance at this stage. Although standard errors were high, there was a negative genetic correlation between resistance and tolerance at this stage, suggesting that there is a trade-off between resistance and tolerance at certain stages of infection. In which case, selecting for resilience may have undesired consequences on either tolerance or resistance at a certain stage of infection. Indeed, in chapter 4 we analysed the relationship between resistance and resilience at different stages of infection (growth under infection). Early in infection, resistant individuals had slower growth, in accordance with resource allocation theory, which would imply that energy resources were diverted from growth to combat infection [13]. Indeed, a strong positive correlation between VL and growth at mid stage infection suggested that those who were less resistant had consequently higher growth under infection. If this is the case, again, there would those selected for resilience at this stage would have higher VL, and would thus be more likely to spread the virus. This research also found a beneficial effect of the WUR genotype for growth, tolerance and resistance, where individuals with the favourable B allele demonstrated higher growth, less steep tolerance slope, and lower VL. It is likely then, that in this case, selecting for resilience would also improve resistance and tolerance to PRRS, and vice versa. Further research must be carried out into the underlying relationship between resistance and tolerance affecting host resilience to other infections.

7.4 The application of trajectories in animal breeding

A linear relationship between pathogen load and performance under infection cannot be assumed, as demonstrated in chapter 5, where trajectories revealed distinct phases of infection associated with different patterns of co-expression of resistance and tolerance. Trajectories are a conceptual tool which allow identification of 'bad neighbourhoods' in the infection-severity health plane, and illustration the 2D path towards death or survival for each individual [85]. Clinical evidence and mathematical dynamical systems theory stipulate that infection severity-health trajectories can be classified into few distinct types which may be useful for disentangling resistance and tolerance, which would be of great use for animal breeding [38,41]. Assuming resistance and tolerance are genetically regulated, implies that trajectories should also be under host genetic control, as was found in chapter 5, and may thus may be amenable to genetic selection. Immunological studies suggest that the outcome of infection may not necessarily depend on either resistance or tolerance alone, but is more likely to depend on genetically regulated dynamic interactions between the two mechanisms. Thus both mechanisms need to be evoked at the right time to ensure survival [5]. Considering resistance and tolerance as static traits that represent alternative host defense strategies may limit the power of genetic studies, not only because tolerance cannot be measured on the individual level, but also because resistance or tolerance may be expressed at different levels throughout the infection and depend on each other. Ignoring the dynamic co-expression of resistance and tolerance within a host may however lead to biased estimates of phenotypic and genetic correlations between both traits. For example, expression of resistance at the early stages of infection is likely to affect the expression of tolerance at the later stages

as fast pathogen clearance may prevent tissue damage and thus the onset of repair mechanisms associated with tolerance. Conventional statistical models that don't account for this kind of interdependence between traits would result in a spurious antagonistic relationship between the traits, both on the phenotypic as well as on the genetic level [125], even if resistance and tolerance were controlled by different sets of genes or genetic pathways as suggested by immunological evidence [5,101]. Trajectories capture the dynamic signature and genetic footprint of both mechanisms on the level of individuals, together with their impact on survival. They may thus constitute useful phenotypes to identify novel genes that control infection dynamics and outcome.

However, there are several drawbacks to trajectories that may limit their application to animal breeding. Firstly, they require longitudinal measures for each individual, which may be difficult to obtain in the field. Secondly, individuals may be at different stages of infection (see chapters 2 and 3), and thus measures taken in the field to define trajectories may be misleading. Thirdly, trajectories are a visual tool, and thus lend themselves more easily to qualitative rather than quantitative assessment. This allows identification of properties and that are shared and different between genetic background, for example expression of resistance and associated change in performance. However, this is not practical for genetic selection. Instead, trajectories need to be quantified.

In chapter 6, we tried to identify genetic variation in trajectories by fitting a spline through the PHGC data using RRMs. Unfortunately, we could not determine if individuals varied genetically in trajectories due to model convergence issues. Instead, we derived novel phenotypes, which could serve as a proxy for tolerance by analyzing

change in host performance with respect to change in pathogen load. These may serve as novel phenotypes that can be applied to animal breeding. However, we did could not identify genetic variation in the phenotypes with this data. This may have been due to complications of individuals being at different stages of infection and highlights a potential constraint of trajectory analysis applied to animal breeding. However, it may be possible that there is genetic variation identifiable using other data, and this should be fully explored in future research.

By utilising repeated measures for each individual, stages of infection were defined, in chapters 2 and 3, to explore the relationship between resistance, tolerance and growth under infection. Genetic variation was evident in both resistance and growth under infection, and this relationship was found to change throughout infection, with early resistance being associated with loss in growth, and less resistant individuals having faster growth under infection, implying potential tolerance. Indeed, if resistance and tolerance are overall under pleiotropic control, as suggested by results of chapter 3, it may be that genes conferring early resistance also confers later tolerance to infection. That the two are related means that it may be timing that is most important to select for an ideal balance between resistance tolerance. It may be possible, then, that defining stages of infection is another solution to exploring resistance-tolerance dynamics in outbred populations. Indeed, although the genetic signal was not as clear, it was possible to use duration to define stages of infection, making the definition of stages of infection easier to implement in animal breeding. Again, further research is required into different infection models to explore the possibilities of trajectory analysis and stages of infection.

7.5 Implications of research

This study identified genetic variation in tolerance of pigs to PRRS in an outbred population using repeated measures. Tolerance may thus be a viable trait which can be explored in different infection models and under different scenarios. We have already made aware the potential pitfalls and limitations of tolerance and a phenotype. Here we discuss the potential implications of our findings, and the possibilities of using tolerance to explore new research.

In chapters 2 and 3 we found that the WUR genotype was associated with resilience and resistance at different stages of infection, and with tolerance over all stages of infection. Genetic selection for individuals with the AB WUR genotype will likely simultaneously increase resilience, resistance and tolerance to PRRS, but this may affect other genomic regions associated with each trait. Whilst we did not identify other genomic regions statistically associated with these traits, that may not necessarily be the case for other infections. Furthermore, we found that WUR may be associated with growth at later stage of infection as a consequence of its association with resistance at early stage of infection. So, it may be that the WUR genotype was suggestively associated with tolerance over all stages of infection due to the positive relationship between resistance and tolerance. Further genomic analysis is required to analyse the relationship of genomic regions associated with these traits and between traits.

The effect of resistance on tolerance cannot be ignored in future research. All the evidence provided from our research suggests that the two are related. Selection for one is likely to have consequences on other traits not targeted directly in the selection criterion. Further work must be carried out using longitudinal methods in different

infection models as to the different dynamics between resistance and tolerance, their contribution to overall host resilience, and the impact of their association on the epidemiology and evolution of the pathogen of interest. As large amounts of data are required for such infection experiments, simulations may be of interest to predict the effect of this different relationship on breeding values. For example, in chapter 2, we simulated performance in absence of infection (vigour) for relatives of individuals under infection, and tested the strength of association between vigour and tolerance slope, and the consequent impact of this relationship of the significance on tolerance slope. Although these findings are under a given set of assumptions, they constitute a valuable starting point at which to disentangle resistance and tolerance, and their contribution to resilience. By using breeding values for resilience and resistance it may be possible to infer their effect on tolerance by simulating a changing genetic correlation between resistance and resilience. Consequent predictions of genetic selection could then be tested, and if they are of interest, epidemiological models could be adopted to assess the impact of selection on spread of the pathogen of interest.

In order to assess the impact of genetic selection for resistance, tolerance or trajectory characteristics on PRRS epidemiology and virus evolution, a genetic-epidemiological prediction model could be developed based on the results of this thesis. Such models already exist for other diseases, e.g. bacterial and parasite infections in small ruminants [135–137]. For PRRS, the genetic epidemiological model of MacKenzie et al. [138] or various epidemiological models for PRRS developed [139–142] may be a useful starting points. The results of this thesis would help to make these models more useful for breeding purposes. The theory of resistance driving pathogen co-evolution is also called into question if resistance and tolerance to PRRS are co-expressed at different

stages of infection. Pathogen co-evolution may be forced by individuals at early stage of infection, where those individuals express resistance. However, given the strong genetic relationship between resistance and tolerance, theoretically, those individuals may also be able to tolerate different strains of the pathogen, if mechanisms of tolerance reduce the damage caused by the pathogen or host immune response. That is, individuals will resilient to different strains of the pathogen, but may also drive pathogen co-evolution. This must be taken into account for heterogeneous populations, and any further evolutionary models should consider this if predicting mutations in PRRSV with relation to genetic selection programs.

7.6 Conclusions

With the PHGC dataset, it could be concluded that resilience may be a sufficient target for a genetic selection program to improve the response of pigs to PRRSV infection. Improving resilience in this case will likely both improve host resistance and tolerance to PRRS. However, there is potential unknown epidemiological and evolutionary implications of co-expression of resistance and tolerance at different stages of infection that must be accounted for before undertaking such a breeding program. Indeed, with the alternative Listeria dataset for inbred mouse populations studies within this thesis, we highlighted potential uses of trajectory analysis for subsequent genetic studies. Successful integration of trajectories in genetic evaluations of outbred animal populations would account for the relationship between resistance and tolerance, and may increase overall resilience. Finally, should we aim for genetic improvement of host resistance or tolerance to infectious disease? For PRRS, the answer is likely both.

Literature cited

1. Bennett R, IJpelaar J. Updated Estimates of the Costs Associated with Thirty Four Endemic Livestock Diseases in Great Britain: A Note. J. Agric. Econ. [Internet]. Blackwell Publishing Ltd; 2005 [cited 2016 Nov 18];56:135–44. Available from: http://doi.wiley.com/10.1111/j.1477-9552.2005.tb00126.x

2. Plastow GS, Plastow GS. Genomics to benefit livestock production: improving animal health. Rev. Bras. Zootec. [Internet]. Sociedade Brasileira de Zootecnia; 2016 [cited 2016 Nov 18];45:349–54. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-35982016000600349&lng=en&nrm=iso&tlng=en

3. Lunney JK, Chen H. Genetic control of host resistance to porcine reproductive and respiratory syndrome virus (PRRSV) infection. Virus Res. [Internet]. 2010 [cited 2014 Feb 25];154:161–9. Available from: http://www.sciencedirect.com/science/article/pii/S0168170210002819

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

5. Råberg L, Sim D, Read AF, Raberg L, Sim D, Read AF. Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. Science (80-.). 2007;318:812–4.

6. Medzhitov R, Schneider DS, Soares MP. Disease Tolerance as a Defense Strategy. Science (80-.). 2012;335:936–41.

7. Råberg L, Graham AL, Read AF, Raberg L, Graham AL, Read AF. Decomposing health: tolerance and resistance to parasites in animals. Philos. Trans. R. Soc. B-

Biological Sci. 2009;364:37–49.

8. Simms EL. Defining tolerance as a norm of reaction. Evol. Ecol. 2000;14:563–70.

9. Simms EL, Triplett J. COSTS AND BENEFITS OF PLANT-RESPONSES TO DISEASE - RESISTANCE AND TOLERANCE. Evolution (N. Y). 1994;48:1973–85.

10. Detilleux JC. A mathematical model to study resistance and tolerance to infection at the animal and population levels: application to E. coli mastitis. Front. Genet. [Internet]. Frontiers Media SA; 2012 [cited 2017 Aug 22];3:146. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23248637

11. Ayres JS, Schneider DS. Tolerance of Infections. In: Paul WE, editor. Annu. Rev. Immunol. Vol 30. Palo Alto: Annual Reviews; 2012. p. 271–94.

12. Coop RL, Kyriazakis I. Influence of host nutrition on the development and consequences of nematode parasitism in ruminants. Trends Parasitol. 2001;17:325–30.

13. Rauw WM. Immune response from a resource allocation perspective. Front. Genet. 2012;3:267.

14. Kemp SJ, Teale AJ. Genetic basis of trypanotolerance in cattle and mice. Parasitol. Today. 1998;14:450–4.

15. Bisset SA, Morris CA. Feasibility and implications of breeding sheep for resilience to nematode challenge. Int. J. Parasitol. 1996;26:857–68.

16. Schafer JF. Tolerance to Plant Disease. Annu. Rev. Phytopathol. [Internet]. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA ; 1971 [cited 2016 Nov 18];9:235–52. Available from: http://www.annualreviews.org/doi/10.1146/annurev.py.09.090171.001315

17. Strauss SY, Agrawal AA. The ecology and evolution of plant tolerance to herbivory. Trends Ecol. Evol. 1999;14:179–85.

18. Pilson D. The evolution of plant response to herbivory: simultaneously considering resistance and tolerance in Brassica rapa. Evol. Ecol. 2000;14:457–89.

19. Schneider DS, Ayres JS. Two ways to survive infection: what resistance and tolerance can teach us about treating infectious diseases. Nat. Rev. Immunol. 2008;8:889–95.

20. Davies G, Genini S, Bishop SC, Giuffra E, Adzhubei AA, Vlasova A V, et al. An assessment of opportunities to dissect host genetic variation in resistance to infectious diseases in livestock. animal [Internet]. Cambridge University Press; 2009 [cited 2016 Nov 18];3:415–36. Available from: http://www.journals.cambridge.org/abstract_S1751731108003522

21. Stear MJ, Bishop SC, Mallard BA, Raadsma H. The sustainability, feasibility and desirability of breeding livestock for disease resistance. Res. Vet. Sci. [Internet]. 2001 [cited 2015 Aug 16];71:1–7. Available from: http://www.sciencedirect.com/science/article/pii/S0034528801904965

22. Boddicker N, Waide EH, Rowland RRR, Lunney JK, Garrick DJ, Reecy JM, et al. Evidence for a major QTL associated with host response to porcine reproductive and respiratory syndrome virus challenge. J. Anim. … [Internet]. 2012;90:1733–46. Available from: http://www.journalofanimalscience.org/content/90/6/1733.full.pdf%5Cnhttp://www.journalofanimalscience.org/content/90/6/1733.short

23. Boddicker NJ, Bjorkquist A, Rowland RRR, Lunney JK, Reecy JM, Dekkers JCM. Genome-wide association and genomic prediction for host response to porcine

reproductive and respiratory syndrome virus infection. Genet. Sel. Evol. [Internet]. 2014 [cited 2016 Aug 19];46:18. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24592976

24. Bishop SC, Woolliams JA. Genomics and disease resistance studies in livestock. Livest. Sci. [Internet]. 2014 [cited 2016 Mar 22];166:190–8. Available from: http://www.research.ed.ac.uk/portal/en/publications/genomics-and-disease-resistance-studies-in-livestock(19b64ea0-225c-4c14-bf07-de11ad89db2e)/export.html

25. Morris C, Bisset S, Vlassoff a., Wheeler M, West C, Devantier B, et al. Selecting for resilience in Romney sheep under nematode parasite challenge, 1994-2007. New Zeal. J. Agric. Res. 2010;53:245–61.

26. Bishop SC. A consideration of resistance and tolerance for ruminant nematode infections. Front. Genet. 2012;3:168.

27. Bishop SC. Genetic resistance to infections in sheep. Vet. Microbiol. 2015;181:2–7.

28. Doeschl-Wilson AB, Kyriazakis I. Should we aim for genetic improvement in host resistance or tolerance to infectious pathogens? Front. Genet. 2012;3:272.

29. Blanchet S, Rey O, Loot G. Evidence for host variation in parasite tolerance in a wild fish population. Evol. Ecol. 2010;24:1129–39.

30. Mauricio R, Rausher MD, Burdick DS. Variation in the defense strategies of plants: Are resistance and tolerance mutually exclusive? Ecology. 1997;78:1301–11.

31. Howick VM, Lazzaro BP. Genotype and diet shape resistance and tolerance across distinct phases of bacterial infection. BMC Evol. Biol. [Internet]. 2014 [cited 2016 Nov 24];14:56. Available from:

http://bmcevolbiol.biomedcentral.com/articles/10.1186/1471-2148-14-56

32. Roy BA, Kirchner JW. Evolutionary dynamics of pathogen resistance and tolerance. Evolution (N. Y). 2000;54:51–63.

33. Rausher MD. Genetic analysis of coevolution between plants and their natural enemies. Trends Genet. 1996;12:212–7.

34. Lefevre T, Williams AJ, de Roode JC. Genetic variation in resistance, but not tolerance, to a protozoan parasite in the monarch butterfly. Proc. R. Soc. B-Biological Sci. 2011;278:751–9.

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

36. Knap PW, Su G. Genotype by environment interaction for litter size in pigs as quantified by reaction norms analysis. Animal. 2008;2:1742.

37. Hayward AD, Nussey DH, Wilson AJ, Berenos C, Pilkington JG, Watt KA, et al. Natural Selection on Individual Variation in Tolerance of Gastrointestinal Nematode Infection. Plos Biol. 2014;12.

38. Schneider DS. Tracing Personalized Health Curves during Infections. Plos Biol. 2011;9.

39. Kutzer MAM, Armitage SAO. The effect of diet and time after bacterial infection on fecundity, resistance, and tolerance in *Drosophila melanogaster*. Ecol. Evol. [Internet]. 2016 [cited 2016 Nov 24];6:4229–42. Available from: http://doi.wiley.com/10.1002/ece3.2185

40. Torres BY, Oliveira JHM, Thomas Tate A, Rath P, Cumnock K, Schneider DS, et al. Tracking Resilience to Infections by Mapping Disease Space. Dobson AP, editor. PLOS Biol. [Internet]. Public Library of Science; 2016 [cited 2016 Nov

24];14:e1002436. Available from: http://dx.plos.org/10.1371/journal.pbio.1002436

41. Doeschl-Wilson AB, Bishop SC, Kyriazakis I, Villanueva B. Novel methods for quantifying individual host response to infectious pathogens for genetic analyses. Front. Genet. 2012;3:266.

42. Dekkers JCM. Application of genomics tools to animal breeding. Curr. Genomics [Internet]. Bentham Science Publishers; 2012 [cited 2016 Nov 2];13:207–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23115522

43. Balding DJ. A tutorial on statistical methods for population association studies. Nat. Rev. Genet. [Internet]. 2006 [cited 2016 Nov 2];7:781–91. Available from: http://www.nature.com/doifinder/10.1038/nrg1916

44. Dean M. Approaches to identify genes for complex human diseases: Lessons from Mendelian disorders. Hum. Mutat. [Internet]. 2003 [cited 2016 Nov 2];22:261–74. Available from: http://doi.wiley.com/10.1002/humu.10259

45. Andersson L. Genome-wide association analysis in domestic animals: A powerful approach for genetic dissection of trait loci. Genetica. 2009;136:341–9.

46. Ramos AM, Crooijmans RPMA, Affara NA, Amaral AJ, Archibald AL, Beever JE, et al. Design of a High Density SNP Genotyping Assay in the Pig Using SNPs Identified and Characterized by Next Generation Sequencing Technology. Orban L, editor. PLoS One [Internet]. Public Library of Science; 2009 [cited 2016 Nov 2];4:e6524. Available from: http://dx.plos.org/10.1371/journal.pone.0006524

47. Fernando RL, Toosi A, Garrick DJ, Dekkers JCM. Application of Whole-Genome Prediction Methods for Genome-Wide Association Studies: a Bayesian Approach. Proceedings.

48. Meuwissen TH, Hayes BJ, Goddard ME. Prediction of total genetic value using

genome-wide dense marker maps. Genetics [Internet]. 2001 [cited 2016 Nov 2];157:1819–29. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11290733

49. Fernando RL, Garrick D. Bayesian methods applied to GWAS. Methods Mol. Biol. [Internet]. 2013 [cited 2016 Nov 2];1019:237–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23756894

50. Garrick DJ, Fernando RL. Implementing a QTL detection study (GWAS) using genomic prediction methodology. Methods Mol. Biol. [Internet]. 2013 [cited 2016 Nov 2];1019:275–98. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23756895

51. Doeschl-Wilson AB, Kyriazakis I, Vincent A, Rothschild MF, Thacker E, Galina-Pantoja L. Clinical and pathological responses of pigs from two genetically diverse commercial lines to porcine reproductive and respiratory syndrome virus infection. J. Anim. Sci. [Internet]. 2009 [cited 2016 Mar 1];87:1638–47. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19181772

52. Holtkamp DJ. Assessment of the economic impact of porcine reproductive and respiratory syndrome virus on United States pork producers. J. Swine Heal. Prod. 2013;21:72–84.

53. Fang Y, Schneider P, Zhang WP, Faaberg KS, Nelson EA, Rowland RRR. Diversity and evolution of a newly emerged North American Type 1 porcine arterivirus: analysis of isolates collected between 1999 and 2004. Arch. Virol. [Internet]. 2007 [cited 2016 Aug 22];152:1009–17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17323198

54. Kimman TG, Cornelissen LA, Moormann RJ, Rebel JMJ, Stockhofe-Zurwieden N. Challenges for porcine reproductive and respiratory syndrome virus (PRRSV)

vaccinology. Vaccine. 2009;27:3704–18.

55. Hess AS, Islam ZZ, Hess MK, Rowland RRRR, Lunney JJK, Doeschl-Wilson AA, et al. Comparison of host genetic factors influencing pig response to infection with two North American isolates of porcine reproductive and respiratory syndrome virus. Genet. Sel. Evol. [Internet]. BioMed Central; 2016 [cited 2016 Jul 1];48:43. Available from: http://gsejournal.biomedcentral.com/articles/10.1186/s12711-016-0222-0

56. Lewis CRG, Ait-Ali T, Clapperton M, Archibald AL, Bishop S. Genetic Perspectives on Host Responses to Porcine Reproductive and Respiratory Syndrome (PRRS). Viral Immunol. [Internet]. 2007 [cited 2016 Jan 11];20:343–58. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17931105

57. Boddicker NJ, Garrick DJ, Rowland RRR, Lunney JK, Reecy JM, Dekkers JCM. 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 [Internet]. 2013;45:48–58. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23914972%5Cnhttp://onlinelibrary.wiley.com/store/10.1111/age.12079/asset/age12079.pdf?v=1&t=hv3zspsd&s=3c4408fb229d53874e7ea1053c7835de64accb0a%5Cn%3CGo to ISI%3E://WOS:000329456900006%5Cnhttp://onlinelibrary.wiley.com/store/10.11

58. Boddicker NJ. The genetic basis of host response to experimental infection with the porcine reproductive and respiratory syndrome virus in pigs [dissertation]. 2013;198.

59. Orndorff PE, Hamrick TS, Smoak IW, Havell EA. Host and bacterial factors in listeriosis pathogenesis. Vet. Microbiol. 2006;114:1–15.

60. Vale PF, McNally L, Doeschl-Wilson A, King KC, Popat R, Domingo-Sananes MR, et al. Beyond killing. Evol. Med. Public Heal. [Internet]. Oxford University Press; 2016 [cited 2016 Aug 5];2016:148–57. Available from: http://emph.oxfordjournals.org/lookup/doi/10.1093/emph/eow012

61. Truong HM, Lu Z, Kutish GF, Galeota J, Osorio FA, Pattnaik AK. A highly pathogenic porcine reproductive and respiratory syndrome virus generated from an infectious cDNA clone retains the in vivo virulence and transmissibility properties of the parental virus. Virology [Internet]. 2004 [cited 2015 Sep 7];325:308–19. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15246270

62. Islam ZU, Bishop SC, Savill NJ, Rowland RRR, Lunney JK, Trible B, et al. Quantitative analysis of porcine reproductive and respiratory syndrome (PRRS) viremia profiles from experimental infection: A statistical modelling approach. PLoS One. 2013;8.

63. Gilmour AR, Gogel BJ, Cullis BR, Thompson R. ASReml User Guide.

64. VanRaden PM. Efficient Methods to Compute Genomic Predictions. J. Dairy Sci. [Internet]. 2008 [cited 2016 Dec 6];91:4414–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18946147

65. Stowe K a., Marquis RJ, Hochwender CG, Simms EL. the Evolutionary Ecology of Tolerance To Consumer Damage. Annu. Rev. Ecol. Syst. [Internet]. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA ; 2000 [cited 2016 Nov 14];31:565–95. Available from: http://www.annualreviews.org/doi/10.1146/annurev.ecolsys.31.1.565

66. Schaeffer LR. Application of random regression models in animal breeding. Livest. Prod. Sci. 2004;86:35–45.

67. Kolmodin R, Bijma P. Response to mass selection when the genotype by environment interaction is modelled as a linear reaction norm. Genet. Sel. Evol. [Internet]. 2004;36:435. Available from: http://www.gsejournal.org/content/36/4/435

68. Van Tienderen PH, Koelewijn HP. Selection on reaction norms, genetic correlations and constraints. Genet. Res. [Internet]. 1994 [cited 2016 Jul 1];64:115–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7813902

69. Strandberg E. Analysis of genotype by environment interaction using random regression models. Proc. 8th World Congr. Genet. Appl. to Livest. Prod. Belo Horizonte, Minas Gerais, Brazil, 13-18 August, 2006. [Internet]. Instituto Prociência; 2006 [cited 2016 Jan 8]. p. 25–05. Available from: http://www.cabdirect.org/abstracts/20063170059.html;jsessionid=43ED913AB724C52AE70F07FEE821A063

70. Calus MPL, Bijma P, Veerkamp RF. Effects of data structure on the estimation of covariance functions to describe genotype by environment interactions in a reaction norm model. Genet. Sel. Evol. [Internet]. BioMed Central Ltd; 2004 [cited 2016 Jan 8];36:489–507. Available from: http://www.gsejournal.org/content/36/5/489

71. Lynch M, Walsh B. Genetics and Analysis of Quantitative Traits [Internet]. Sinauer; 1998 [cited 2016 Jan 11]. Available from: https://books.google.co.uk/books/about/Genetics_and_Analysis_of_Quantitative_Tr.html?id=UhCCQgAACAAJ&pgis=1

72. Stram DO, Lee JW. Variance components testing in the longitudinal mixed effects model. Biometrics [Internet]. 1994 [cited 2016 Jan 8];50:1171–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7786999

73. Visscher PM, Medland SE, Ferreira MAR, Morley KI, Zhu G, Cornes BK, et al. Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. PLoS Genet. [Internet]. 2006 [cited 2016 Jan 7];2:e41. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1413498&tool=pmcentrez&rendertype=abstract

74. Chen P. Genetic improvement of lean growth rate and reproductive traits in pigs. [cited 2016 Nov 14]; Available from: http://lib.dr.iastate.edu/rtd

75. Ødegård J, Meuwissen TH, Meuwissen T, Hayes B, Goddard M, Luan T, et al. Identity-by-descent genomic selection using selective and sparse genotyping. Genet. Sel. Evol. [Internet]. BioMed Central; 2014 [cited 2016 Aug 9];46:3. Available from: http://www.gsejournal.org/content/46/1/3

76. Kause a., van Dalen S, Bovenhuis H. Genetics of Ascites Resistance and Tolerance in Chicken: A Random Regression Approach. G3: Genes|Genomes|Genetics. 2012;2:527–35.

77. de Greef KH, Janss LL, Vereijken AL, Pit R, Gerritsen CL. Disease-induced variability of genetic correlations: ascites in broilers as a case study. J. Anim. Sci. [Internet]. 2001 [cited 2016 Dec 5];79:1723–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11465359

78. Wilson AJ, Réale D, Clements MN, Morrissey MM, Postma E, Walling C a., et al. An ecologist's guide to the animal model. J. Anim. Ecol. 2010;79:13–26.

79. Restif O, Koella JC. Concurrent evolution of resistance and tolerance to pathogens. Am. Nat. 2004;164:E90–102.

80. Drake JW, Holland JJ. Mutation rates among RNA viruses. Proc. Natl. Acad. Sci.

U. S. A. [Internet]. 1999 [cited 2016 Jan 11];96:13910–3. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=24164&tool=pmcentrez&rendertype=abstract

81. Rowland RRR, Lunney J, Dekkers J. Control of porcine reproductive and respiratory syndrome (PRRS) through genetic improvements in disease resistance and tolerance. Front. Genet. [Internet]. 2012 [cited 2015 Sep 21];3:1–6. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3565991&tool=pmcentrez&rendertype=abstract

82. Zimmerman JJ, Yoon KJ, Pirtle EC, Wills RW, Sanderson TJ, McGinley MJ. Studies of porcine reproductive and respiratory syndrome (PRRS) virus infection in avian species. Vet. Microbiol. [Internet]. 1997 [cited 2015 Sep 21];55:329–36. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9220630

83. Albina E. Epidemiology of porcine reproductive and respiratory syndrome (PRRS): an overview. Vet. Microbiol. [Internet]. 1997 [cited 2015 Sep 21];55:309–16. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9220627

84. de Pax X, Vega D, Duran C., Angulo J. PRRS prevalence in Europe: Perception of the pig veterinary practitioners - PRRS.com [Internet]. ESPHM 2015. 2015 [cited 2015 Sep 22]. p. 1. Available from: https://www.prrs.com/en/publications/abstracts/prrs-prevalence-europe-perception-the-pig-veterinary-practitioners/

85. Lough G, Kyriazakis I, Bergmann S, Lengeling A, Doeschl-Wilson AB. Health trajectories reveal the dynamic contributions of host genetic resistance and tolerance to infection outcome. Proc. Biol. Sci. [Internet]. 2015 [cited 2015 Nov

18];282:20152151–20152151. Available from: http://rspb.royalsocietypublishing.org/content/282/1819/20152151

86. Herrero-Medrano JM, Mathur PK, Napel J ten, Rashidi H, Alexandri P, Knol EF, et al. Estimation of genetic parameters and breeding values across challenged environments to select for robust pigs. J. Anim. Sci. [Internet]. American Society of Animal Science; 2015 [cited 2016 Oct 18];93:1494. Available from: https://dl.sciencesocieties.org/publications/jas/abstracts/93/4/1494

87. Rashidi H. Breeding against infectious diseases in animals [Internet]. Wageningen University; 2016. Available from: http://library.wur.nl/WebQuery/wurpubs/498758

88. Allen JE, Sutherland TE. Host protective roles of type 2 immunity: parasite killing and tissue repair, flip sides of the same coin. Semin. Immunol. [Internet]. 2014 [cited 2016 Feb 7];26:329–40. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4179909&tool=pmcentrez&rendertype=abstract

89. Gopinath S, Lichtman JS, Bouley DM, Elias JE, Monack DM. Role of disease-associated tolerance in infectious superspreaders. Proc. Natl. Acad. Sci. U. S. A. [Internet]. 2014 [cited 2016 Mar 8];111:15780–5. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4226084&tool=pmcentrez&rendertype=abstract

90. Anacleto O, Garcia-Cortés LA, Lipschutz-Powell D, Woolliams JA, Doeschl-Wilson AB. A Novel Statistical Model to Estimate Host Genetic Effects Affecting Disease Transmission. Genetics [Internet]. 2015 [cited 2016 Feb 15];201:871–84. Available from:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4649657&tool=pmcentrez&rendertype=abstract

91. Lillehammer M, Odegård J, Meuwissen THE. Reducing the bias of estimates of genotype by environment interactions in random regression sire models. Genet. Sel. Evol. 2009;41:30.

92. Fernando R, Garrick D. GenSel – User Manual for a portfolio of Genomic Selection related Analyses. 2009;

93. Hayes B, Goddard M. Genome-wide association and genomic selection in animal breeding. Genome [Internet]. 2010 [cited 2016 Nov 2];53:876–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21076503

94. Habier D, Fernando RL, Kizilkaya K, Garrick DJ, Meuwissen T, Hayes B, et al. Extension of the bayesian alphabet for genomic selection. BMC Bioinformatics [Internet]. BioMed Central; 2011 [cited 2016 Nov 2];12:186. Available from: http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-186

95. Hess AS. Genetic and biological factors influencing host response to porcine reproductive and respiratory syndrome virus in growing pigs. [cited 2016 Nov 17]; Available from: http://lib.dr.iastate.edu/etd

96. Lopez OJ, Osorio FA. Role of neutralizing antibodies in PRRSV protective immunity. Vet. Immunol. Immunopathol. [Internet]. 2004 [cited 2016 Nov 10];102:155–63. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0165242704002363

97. Serao NVL, Matika O, Kemp RA, Harding JCS, Bishop SC, Plastow GS, et al. Genetic analysis of reproductive traits and antibody response in a PRRS outbreak herd. J. Anim. Sci. [Internet]. American Society of Animal Science; 2014 [cited 2016

Aug 19];92:2905–21. Available from:
http://www.animalsciencepublications.org/publications/jas/abstracts/92/7/2905
98. EXTON MS. Infection-Induced Anorexia: Active Host Defence Strategy. Appetite [Internet]. 1997 [cited 2016 Nov 10];29:369–83. Available from:
http://linkinghub.elsevier.com/retrieve/pii/S0195666397901161
99. Koltes JE, Fritz-Waters E, Eisley CJ, Choi I, Bao H, Kommadath A, et al. Identification of a putative quantitative trait nucleotide in guanylate binding protein 5 for host response to PRRS virus infection. BMC Genomics [Internet]. BioMed Central; 2015 [cited 2016 Oct 27];16:412. Available from:
http://www.biomedcentral.com/1471-2164/16/412
100. Shenoy AR, Wellington DA, Kumar P, Kassa H, Booth CJ, Cresswell P, et al. GBP5 promotes NLRP3 inflammasome assembly and immunity in mammals. Science [Internet]. 2012 [cited 2016 Oct 27];336:481–5. Available from:
http://www.ncbi.nlm.nih.gov/pubmed/22461501
101. Glass EJ. The molecular pathways underlying host resistance and tolerance to pathogens. Front. Genet. 2012;3:263.
102. Rodriguez R, Redman R. More than 400 million years of evolution and some plants still can't make it on their own: plant stress tolerance via fungal symbiosis. J. Exp. Bot. [Internet]. 2008 [cited 2015 Aug 27];59:1109–14. Available from:
http://jxb.oxfordjournals.org/cgi/content/long/59/5/1109
103. Zelante T, Iannitti RG, Fallarino F, Gargaro M, De Luca A, Moretti S, et al. Tryptophan Feeding of the IDO1-AhR Axis in Host-Microbial Symbiosis. Front. Immunol. [Internet]. 2014 [cited 2015 Sep 3];5:640. Available from:
http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4266093&tool=pmcentre

z&rendertype=abstract

104. Ramaswamy V, Cresence VM, Rejitha JS, Lekshmi MU, Dharsana KS, Prasad SP, et al. Listeria - review of epidemiology and pathogenesis. J. Microbiol. Immunol. Infect. 2007;40:4–13.

105. Barbuddhe SB, Chakraborty T. Listeria as an Enteroinvasive Gastrointestinal Pathogen. In: Sasakawa C, editor. Mol. Mech. Bact. Infect. Via Gut. 2009. p. 173–95.

106. Ooi ST, Lorber B. Gastroenteritis due to Listeria monocytogenes. Clin. Infect. Dis. 2005;40:1327–32.

107. Gervais F, Stevenson M, Skamene E. GENETIC-CONTROL OF RESISTANCE TO LISTERIA-MONOCYTOGENES - REGULATION OF LEUKOCYTE INFLAMMATORY RESPONSES BY THE HC LOCUS. J. Immunol. 1984;132:2078–83.

108. Cheers C, McKenzie IFC. RESISTANCE AND SUSCEPTIBILITY OF MICE TO BACTERIAL-INFECTION - GENETICS OF LISTERIOSIS. Infect. Immun. 1978;19:755–62.

109. Boyartchuk VL, Broman KW, Mosher RE, D'Orazio SEF, Starnbach MN, Dietrich W. Multigenic control of Listeria monocytogenes susceptibility in mice. Nat. Genet. 2001;27:259–60.

110. Bergmann S, Beard PM, Pasche B, Lienenklaus S, Weiss S, Gahan CGM, et al. Influence of internalin a murinisation on host resistance to orally acquired listeriosis in mice. Bmc Microbiol. [Internet]. 2013;13:90. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3640945&tool=pmcentrez&rendertype=abstract

111. Wollert T, Pasche B, Rochon M, Deppenmeier S, van den Heuvel J, Gruber AD, et al. Extending the host range of Listeria monocytogenes by rational protein design. Cell. 2007;129:891–902.

112. Theocharidis A, van Dongen S, Enright AJ, Freeman TC. Network visualization and analysis of gene expression data using BioLayout Express(3D). Nat. Protoc. 2009;4:1535–50.

113. Vale PF, Fenton A, Brown SP. Limiting Damage during Infection: Lessons from Infection Tolerance for Novel Therapeutics. Plos Biol. 2014;12.

114. Lecuit M. Human listeriosis and animal models. Microbes Infect. 2007;9:1216–25.

115. Skamene E. GENETIC-REGULATION OF HOST-RESISTANCE TO BACTERIAL-INFECTION. Rev. Infect. Dis. 1983;5:S823–32.

116. Beutler B, Goodnow CC. How host defense is encoded in the mammalian genome INTRODUCTION. Mamm. Genome. 2011;22:1–5.

117. Chapman SJ, Hill AVS. Human genetic susceptibility to infectious disease. Nat. Rev. Genet. 2012;13:175–88.

118. Pittet MJ, Weissleder R. Intravital Imaging. Cell. 2011;147:983–91.

119. Hoerr V, Tuchscherr L, Hüve J, Nippe N, Loser K, Glyvuk N, et al. Bacteria tracking by in vivo magnetic resonance imaging. BMC Biol. [Internet]. 2013 [cited 2015 Feb 18];11:63. Available from: http://www.biomedcentral.com/1741-7007/11/63

120. Attia AS, Schroeder KA, Seeley EH, Wilson KJ, Hammer ND, Colvin DC, et al. Monitoring the inflammatory response to infection through the integration of MALDI IMS and MRI. Cell Host Microbe [Internet]. 2012 [cited 2015 Feb

18];11:664–73. Available from:
http://www.sciencedirect.com/science/article/pii/S1931312812001655
121. Friswell MK, Gika H, Stratford IJ, Theodoridis G, Telfer B, Wilson ID, et al. Site and strain-specific variation in gut microbiota profiles and metabolism in experimental mice. PLoS One [Internet]. Public Library of Science; 2010 [cited 2015 Sep 3];5:e8584. Available from:
http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2798964&tool=pmcentrez&rendertype=abstract
122. Athanasiadou S, Tolossa K, Debela E, Tolera A, Houdijk JGM. Tolerance and resistance to a nematode challenge are not always mutually exclusive. Int. J. Parasitol. [Internet]. 2015 [cited 2015 Feb 11]; Available from:
http://www.sciencedirect.com/science/article/pii/S0020751915000193
123. Maze-Guilmo E, Loot G, Paez DJ, Lefevre T, Blanchet S. Heritable variation in host tolerance and resistance inferred from a wild host-parasite system. Proc. R. Soc. B-Biological Sci. 2014;281.
124. Gandon S, Vale PF. The evolution of resistance against good and bad infections. J. Evol. Biol. 2014;27:303–12.
125. Kause A, Odegard J. The genetic analysis of tolerance to infections: a review. Front. Genet. 2012;3:262.
126. Coop RL, Kyriazakis I. Nutrition-parasite interaction. Vet. Parasitol. 1999;84:187–204.
127. Kyriazakis I, Tolkamp BJ, Hutchings MR. Towards a functional explanation for the occurrence of anorexia during parasitic infections. Anim. Behav. 1998;56:265–74.

128. Kyriazakis I. Is anorexia during infection in animals affected by food composition? Anim. Feed Sci. Technol. 2010;156:1–9.

129. Doeschl-Wilson AB, Brindle W, Emmans G, Kyriazakis I. Unravelling the Relationship between Animal Growth and Immune Response during Micro-Parasitic Infections. PLoS One. 2009;4.

130. Monk IR, Casey PG, Hill C, Gahan CGM. Directed evolution and targeted mutagenesis to murinize listeria monocytogenes internalin A for enhanced infectivity in the murine oral infection model. Bmc Microbiol. [Internet]. 2010 [cited 2015 Sep 3];10:318. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3016325&tool=pmcentrez&rendertype=abstract

131. Tsai Y-H, Disson O, Bierne H, Lecuit M. Murinization of internalin extends its receptor repertoire, altering Listeria monocytogenes cell tropism and host responses. PLoS Pathog. [Internet]. 2013 [cited 2015 Sep 3];9:e1003381. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3667765&tool=pmcentrez&rendertype=abstract

132. Disson O, Grayo S, Huillet E, Nikitas G, Langa-Vives F, Dussurget O, et al. Conjugated action of two species-specific invasion proteins for fetoplacental listeriosis. Nature [Internet]. 2008 [cited 2015 Sep 3];455:1114–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18806773

133. Meyer K, Kirkpatrick M. Up hill, down dale: quantitative genetics of curvaceous traits. Philos. Trans. R. Soc. Lond. B. Biol. Sci. [Internet]. 2005 [cited 2016 Feb 1];360:1443–55. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1569513&tool=pmcentre

z&rendertype=abstract

134. Rashidi H, Mulder H a., Mathur P, Van Arendonk J a M, Knol EF. Variation among sows in response to porcine reproductive and respiratory syndrome. J. Anim. Sci. 2014;92:95–105.

135. Bishop SC, Stear MJ, Barger IA, Barnes EH, Dobson RJ, Gettinby G, et al. Modelling responses to selection for resistance to gastro-intestinal parasites in sheep. Anim. Sci. [Internet]. Cambridge University Press; 1997 [cited 2017 Jan 24];64:469–78. Available from: http://www.journals.cambridge.org/abstract_S1357729800016088

136. Nieuwhof G, Conington J, Bishop SC. A genetic epidemiological model to describe resistance to an endemic bacterial disease in livestock: application to footrot in sheep. Genet. Sel. Evol. [Internet]. 2009 [cited 2017 Jan 24];41:19. Available from: http://www.gsejournal.org/content/41/1/19

137. Doeschl-Wilson AB, Davidson R, Conington J, Roughsedge T, Hutchings MR, Villanueva B. Implications of Host Genetic Variation on the Risk and Prevalence of Infectious Diseases Transmitted Through the Environment. Genetics [Internet]. 2011 [cited 2017 Jan 24];188:683–93. Available from: http://www.genetics.org/cgi/doi/10.1534/genetics.110.125625

138. MacKenzie K, Bishop SC. Utilizing stochastic genetic epidemiological models to quantify the impact of selection for resistance to infectious diseases in domestic livestock. J. Anim. Sci. [Internet]. 2001 [cited 2017 Jan 24];79:2057–65. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11518213

139. Evans CM, Medley GF, Creasey SJ, Green LE. A stochastic mathematical model of the within-herd transmission dynamics of porcine reproductive and

respiratory syndrome virus (PRRSV): Fade-out and persistence. Prev. Vet. Med. [Internet]. 2010 [cited 2017 Jan 24];93:248–57. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20004990

140. Jeong J, Aly SS, Cano JP, Polson D, Kass PH, Perez AM. Stochastic model of porcine reproductive and respiratory syndrome virus control strategies on a swine farm in the United States. Am. J. Vet. Res. [Internet]. 2014 [cited 2017 Jan 24];75:260–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24564311

141. Arruda AG, Friendship R, Carpenter J, Greer A, Poljak Z. Evaluation of Control Strategies for Porcine Reproductive and Respiratory Syndrome (PRRS) in Swine Breeding Herds Using a Discrete Event Agent-Based Model. Sanchez J, editor. PLoS One [Internet]. 2016 [cited 2017 Jan 24];11:e0166596. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27875546

142. Charpin C, Mahé S, Keranflec'h A, Belloc C, Cariolet R, Le Potier M-F, et al. Infectiousness of pigs infected by the Porcine Reproductive and Respiratory Syndrome virus (PRRSV) is time-dependent. Vet. Res. [Internet]. 2012 [cited 2017 Jan 24];43:69. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23061672