Genetic analysis of tolerance to infections using random regressions: a simulation study

A N T T I   K A U S E* †

Wageningen University and Research Centre, Animal Breeding and Genomics Centre, Zodiac Building, Marijkeweg 40, 6709 PG, Wageningen, The Netherlands

Summary

Tolerance to infections is the ability of a host to limit the impact of a given pathogen burden on host performance. This simulation study demonstrated the merit of using random regressions to estimate unbiased genetic variances for tolerance slope and its genetic correlations with other traits, which could not be obtained using the previously implemented statistical methods. Genetic variance in tolerance was estimated as genetic variance in regression slopes of host performance along an increasing pathogen burden level. Random regressions combined with covariance functions allowed genetic variance for host performance to be estimated at any point along the pathogen burden trajectory, providing a novel means to analyse infection-induced changes in genetic variation of host performance. Yet, the results implied that decreasing family size as well as a non-zero environmental or genetic correlation between initial host performance before infection and pathogen burden led to biased estimates for tolerance genetic variance. In both cases, genetic correlation between tolerance slope and host performance in a pathogen-free environment became artificially negative, implying a genetic trade-off when it did not exist. Moreover, recording a normally distributed pathogen burden as a threshold trait is not a realistic way of obtaining unbiased estimates for tolerance genetic variance. The results show that random regressions are suitable for the genetic analysis of tolerance, given suitable data structure collected either under field or experimental conditions.

1. Introduction

The defence mechanisms against pathogens and parasites can be divided into two categories: resistance and tolerance. Resistance is the host trait that reduces the probability of infection or reduces the growth of the pathogen population within a host, both factors reducing pathogen burden within a host individual. Tolerance is defined as the ability of the host to limit the impact of a given pathogen burden on host health, performance and ultimately on fitness (Painter, 1958; Simms & Triplett, 1994). In farm animal science, tolerance is sometimes termed resilience (Clunies-Ross, 1932; Albers *et al.*, 1987; Bisset & Morris, 1996). Resistance and tolerance together

define a host's defensive capacity, yet they have very different influences on host − pathogen interactions. Increased host resistance typically induces selection pressure on a pathogen to overcome the resistance mechanisms, causing a continuous arms race between the host and the pathogen. In contrast, increased tolerance makes the pathogen harmless to the host reducing selection imposed on the pathogen. This reduces the arms race co-evolution and may provide an efficient alternative in the fight against pathogens (Mauricio *et al.*, 1997; Rausher, 2001; Bishop & MacKenzie, 2003; Best *et al.*, 2008). Accordingly, tolerance–resistance genetics is a fundamental component contributing to host − pathogen co-evolution and to practical plant and farm animal production.

Tolerance may be defined as a change in host performance in response to increasing pathogen burden (Simms, 2000). Thus, measuring tolerance follows the methodology of the analysis of reaction norms (Via & Lande, 1985). Analysing tolerance as a change in host performance from before to after pathogen attack,

* Corresponding author: Wageningen University and Research Centre, Animal Breeding and Genomics Centre, Marijkeweg 40, 6709 PG, Wageningen, The Netherlands. Tel: +31 317 484622. E-mail: antti.kause@mtt.fi
† Current address: MTT Agrifood Research Finland, Biotechnology and Food Research, Biometrical Genetics, FI-31600 Jokioinen, Finland.

an approach sometimes applied for 'resilience' analysis (Bisset & Morris, 1996), is unjustified because it includes both natural temporal variation in host performance (e.g. in growth curves) and impact of pathogen burden on host performance. Genetic variance in tolerance can be estimated as genetic variance in regression slopes of host performance along a gradient of increasing pathogen burden (Simms & Triplett, 1994; Mauricio *et al.*, 1997; Simms, 2000). Furthermore, a degree of genetic cost of tolerance can be measured as a genetic correlation between tolerance slope and host performance in a pathogen-free environment (i.e. the intercept of the pathogen burden–host performance relationship). A genetic trade-off between the slope and intercept has been suggested to be one potential reason for the maintenance of genetic variation in tolerance (Rausher, 1996; Agrawal *et al.*, 1999; Tiffin & Rausher, 1999).

Studies on tolerance genetics typically use an analysis of covariance (ANCOVA) to test for genetic variation in tolerance, where a significant *F*-test for family-by-pathogen burden interaction implies genetic variation for tolerance slopes (Simms & Triplett, 1994; Mauricio *et al.*, 1997; Koskela *et al.*, 2002; Råberg *et al.*, 2007). Furthermore, genetic relationships between tolerance and other traits are often estimated using correlations between family means for each trait. Family–mean correlations are typically downward-biased estimates of the true genetic correlations (Astles *et al.*, 2006). Thus, these approaches tend to produce biased estimates of genetic variances and genetic correlations for tolerance.

Random regressions allow sophisticated genetic analysis of traits defined as functions (Henderson, 1982; Kirkpatrick *et al.*, 1990; Meyer & Hill, 1997; Schaeffer, 2004). Random regressions have been applied to genetic analysis of growth curves, lactation curves and reaction norms in farm animal breeding (Henderson, 1982; Schaeffer, 2004), and also more recently in evolutionary genetics (e.g. Kingsolver *et al.*, 2004). When combined with pedigree information, random regression models allow the estimation of genetic variances for the regression coefficients and of genetic correlations with other traits. Moreover, regression coefficients combined with covariance functions (Kirkpatrick *et al.*, 1990) provide a potential novel means to estimate the degree to which pathogen infection induces changes in genetic variance of host performance.

Here I assessed the potential of random regressions for the genetic analysis of tolerance when a host population is under natural pathogen infection. This was done because a significant part of datasets from animal breeding programmes and wild populations come from field studies where natural infections with pathogens, parasitoids or production diseases occur. A simulation was conducted to study the effect of alternative data structures on bias and precision of (1) genetic variance of tolerance slope, (2) estimated breeding values (BVs), (3) genetic correlation between tolerance slope and intercept and (4) genetic variance in host performance along a pathogen burden gradient.

Two challenges likely to be encountered when analysing data from natural infection were considered. Firstly, genetic analysis of tolerance uses within-family information (full-sib and half-sibs, and more distant relatives) to estimate a family's tolerance slope. Thus, family size is a crucial parameter influencing estimation accuracy. Secondly, with natural infection individuals are not randomly allocated to control and infection treatments, as in a controlled experiment. For instance, under natural infection the individuals with the lowest performance (fitness, growth and reproduction) due to environmental or genetic factors may innately also have either the lowest pathogen burden (i.e. the highest resistance) or the highest pathogen burden. Accordingly, estimates of tolerance variance can become biased if pathogen burden within an individual and measures of host performance are innately related.

2. Materials and methods

(i) *Terminology used*

From hereon, pathogen burden is used as a general term to refer to a pathogen load of an individual, for instance, number or biomass of ecto- or endo-parasites, number of pathogens in a blood sample and in plants the number of herbivores or percentage leaf area lost to herbivores. Thus, pathogen burden is the inverse of resistance against infections, parasitoids or production diseases. In a random regression model, pathogen burden of individuals is on the *x*-axis and varies from zero upwards. Host performance refers to a host's trait value such as growth, reproduction performance or ultimately fitness. Tolerance is the slope of host performance when regressed against individuals' own pathogen burden phenotypes (Simms, 2000). The intercept refers to host performance in a pathogen-free environment (pathogen burden and thus *x*-axis value is zero). Alternatively, the intercept can be interpreted as initial host performance before infection (e.g. initial growth rate), and host performance under infection (corresponding to non-zero *x*-value) would be final host performance.

(ii) *Simulation of data*

Full-sib data with 100 families were simulated. Generation 1 consisted of 200 unrelated individuals, 100 sires and 100 dams. To estimate a family's tolerance slope, only within-family information was used.
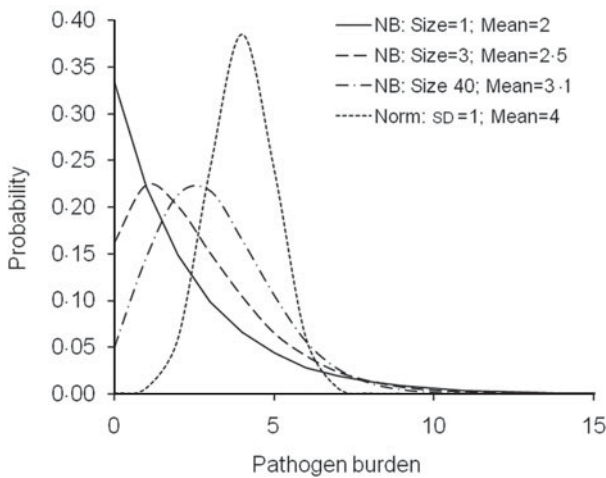
Fig. 1. Probability mass functions of the simulated pathogen burden distributions. For negative binomial (NB) and normal (Norm) distributions, alternative distributions were obtained by varying dispersion parameter (size), trait mean (Mean) and standard deviation (SD).

Similar designs are used e.g. in plant sciences, and sire models are common in animal breeding.

In generation 2, phenotypic values (p) of an individual for pathogen burden (PB), tolerance slope (b_1) and intercept (b_0) were simulated as the sum of genetic (a) and environmental effects (e) randomly sampled from a multi-trait distribution. A genetic value of an individual was simulated as the average of parents' genetic values (i.e. BVs) plus a Mendelian sampling term (ms):

$$a = \begin{bmatrix} a_{b_0} \\ a_{b_1} \\ a_{PB} \end{bmatrix}_{\text{offspring}} = \frac{1}{2}$$

$$\times \left(\begin{bmatrix} a_{b_0} \\ a_{b_1} \\ a_{PB} \end{bmatrix}_{\text{sire}} + \begin{bmatrix} a_{b_0} \\ a_{b_1} \\ a_{PB} \end{bmatrix}_{\text{dam}} \right) + \begin{bmatrix} ms_{b_0} \\ ms_{b_1} \\ ms_{PB} \end{bmatrix},$$

where

$$\text{Var}(a) = \begin{bmatrix} \sigma^2_{a(b_0)} & \sigma_{a(b_0 b_1)} & \sigma_{a(b_0 PB)} \\ & \sigma^2_{a(b_1)} & \sigma_{a(b_1 PB)} \\ \text{sym} & & \sigma^2_{a(PB)} \end{bmatrix}, \quad ms = \begin{bmatrix} ms_{b_0} \\ ms_{b1} \\ ms_{PB} \end{bmatrix},$$

var(ms) = 1/2 Var(a), and σ^2_a and σ_a are genetic variances and covariances, respectively. Environmental effects and their variances (σ^2_e) and covariances (σ_e) were defined as

$$e = \begin{bmatrix} e_{b_0} \\ e_{b_1} \\ e_{PB} \end{bmatrix} \text{ with } \text{Var}(e) = \begin{bmatrix} \sigma^2_{e(b_0)} & \sigma_{e(b_0 b_1)} & \sigma_{e(b_0 PB)} \\ & \sigma^2_{e(b_1)} & \sigma_{e(b_1 PB)} \\ \text{sym} & & \sigma^2_{e(PB)} \end{bmatrix}.$$

A phenotypic value of host performance (y_i) of an individual i at its own pathogen burden phenotype was calculated as $y_i = p_{b_{0i}} + p_{b_{1i}} p_{PB_i}$.

The simulated data were based on observations from studies on tolerance and resistance genetics in plants and animals (Mauricio *et al.*, 1997; Simms, 2000; Koskela *et al.*, 2002; Råberg *et al.*, 2007; Read *et al.*, 2008; Schneider & Ayres, 2008; Kuukka-Anttila *et al.*, 2010). In these papers, the relation between pathogen burden and host performance is typically observed to be linear, but non-linear responses are also realistic (see Discussion for the implications of the model assumptions). Heritability (h^2) and phenotypic variance (σ^2_p) of slope were varied because of a lack of data in the literature.

For pathogen burden, both normal and negative binomial distribution scenarios were simulated (Fig. 1). The normal distribution covers burden traits that are normally distributed at least after a transformation, such as log-normally and normally distributed pathogen/parasite counts, production diseases such as mastitis in dairy cows or ascites measured as a heart ratio in broilers (Green *et al.*, 2004; Zerehdaran *et al.*, 2006; Kuukka-Anttila *et al.*, 2010). The negative binomial distribution, in turn, is perhaps the most widely used distribution to describe empirical parasite/parasitoid burdens (Stear *et al.*, 1995).

For the normally distributed pathogen burden, genetic and environmental effects were randomly sampled from $\sim N(4, 0{\cdot}3)$ and $\sim N(4, 0{\cdot}7)$, respectively. Therefore, h^2 value of $0{\cdot}3$ was simulated. Pathogen burden with h^2 of $0{\cdot}05$ was also simulated but not reported here because the conclusions did not differ from the $h^2 = 0{\cdot}3$ alternative.

Negative binomial distribution of pathogen burden was simulated following the logic of simulating a threshold trait with an underlying normally distributed liability scale. Normally distributed pathogen burden was first simulated, as described above, and then the resulting phenotypic values were transformed to a negative binomial distribution using a link function (R Development Core Team, 2008). Alternative negative binomial distributions including very skewed ones were simulated by changing the pathogen burden mean and the parameter 'size' (Fig. 1). Size is a dispersion parameter for the shape of the gamma-mixing distribution (R Development Core Team, 2008). The alternative parameters used for negative binomial distributions were: size = 1 and mean = 2 (with SD = 2·45); size = 3 and mean = 2·5 (SD = 2·14); size = 40 and mean = 3·1 (SD = 1·83) (Fig. 1).

For intercept, genetic and environmental effects were randomly sampled from $\sim N(100, 120)$ and $\sim N(100, 280)$, respectively, corresponding to a phenotypic coefficient of variation (CV_P) of 20% and a h^2 of 0·3 for host performance in a pathogen-free environment. This represents a normal growth-related trait.

For the tolerance slope, a fixed mean of zero was always used, and a phenotypic variance (σ^2_p) of 30

Table 1. *A summary of simulated scenarios*

Population structure	h^2 for pathogen burden and intercept	Slope σ^2_p/h^2	r_E	r_G	Family size	Distribution parameters
Negative binomial distribution						
Varied family size	0·3	30/0·3	0	0	10, 30, 50, 100, 200	Size=1, mean=2, SD=2·45; Size=3, mean=2·5, SD=2·14; Size=40, mean=3·1, SD=1·83
Normal distribution						
Varied family size	0·3	5, 15, 30/0·05 or 0·3	0	0	10, 30, 50, 100, 200	Mean=4, SD=1
Varied r_E between intercept and pathogen burden	0·3	30/0·3	−0·5, −0·3, −0·1, 0, 0·1, 0·3, 0·5	0	100	Mean=4, SD=1
Varied r_G between intercept and pathogen burden	0·3	30/0·3	0	−0·5, −0·3, −0·1, 0, 0·1, 0·3, 0·5	100	Mean=4, SD=1

σ^2_p, phenotypic variance; h^2, heritability; r_E, environmental correlation; r_G, genetic correlation.

was simulated, but alternative h^2 value scenarios were applied (see below). In the base scenario, environmental and genetic correlations between pathogen burden, slope and intercept were all simulated to be zero.

The simulated values for slope and pathogen burden (both normal and negative binomial distribution) correspond to a CV_P of 30% for host performance under pathogen burden, but phenotypic variance increased with increasing pathogen burden. The increased variance is natural because with increasing pathogen burden the tolerance slopes diverge more and more from each other. For slope h^2 of 0·3, host performance h^2 remains constant at 0·3 along the parasite burden gradient. For slope h^2 of 0·05, host performance h^2 is reduced from 0·30 at pathogen-free environment to 0·10 at pathogen burden of 8. For the normally distributed pathogen burden, scenarios with CV_P of 25% or 22% for host performance under pathogen burden were also simulated to assess the effect of the increased variance assumption on the results. These were obtained by reducing the slope σ^2_p from the original 30 to 15 and 5, respectively.

The simulation was repeated 500 times for each alternative scenario.

(iii) *Alternative scenarios studied*

The alternative scenarios simulated are summarized in Table 1. Firstly, the effect of family size was assessed. Family size was set to 10, 30, 50, 100 or 200. Given a fixed number of 100 families, population size hence ranged from 1000 up to 20 000. For the normally distributed pathogen burden, scenarios with low and high heritabilities for slope ($h^2 = 0·05$ and 0·3, respectively) were simulated by randomly sampling genetic and environment effects of tolerance slope from $\sim N(0, \; 30 \times h^2)$ and $\sim N(0, \; 30 \times (1 - h^2))$, respectively. For the negative binomially distributed pathogen burden, only slope h^2 of 0·3 was simulated.

For the remaining scenarios, normally distributed pathogen burden, family size of 100, population size of 10 000 and tolerance slope h^2 of 0·3 were simulated.

Secondly, the effect of non-zero environmental (r_E) and genetic correlation (r_G) between intercept and pathogen burden was assessed. The correlations were simulated, separately for r_E and r_G, to be either −0·5, −0·3, −0·1, 0, 0·1, 0·3 or 0·5 (Table 1). A negative correlation means that low host performance in a pathogen-free environment is related to high pathogen burden.

(iv) *Genetic analysis of simulated data*

In the genetic analysis, phenotypic data from generation 2 individuals were used. The phenotypic values of pathogen burden and performance value at that burden were analysed. Thus, each individual had only

one host performance observation. ASReml software (Gilmour *et al.*, 2006) was used to fit a full-sib random regression model:

$$y_{ij} = \mu + b_{0j} + b_{1j}PB_{ij} + \varepsilon_{ij}, \tag{1}$$

where y_{ij} is host performance of an individual i from family j at its pathogen burden, μ is the population mean, b_{0j} is the random genetic effect of intercept for a family j, b_{1j} is the random genetic effect of regression slope for a family j, PB_{ij} is pathogen burden of an individual i from family j and ε_{ij} is the random error term. This is a mixed model similar to ANCOVA, the only difference being that the regression terms are now random factors. To avoid heterogeneous error variance inflating genetic variance in slope (Lillehammer *et al.*, 2009), residual variance was estimated within six pathogen burden classes along the *x*-axis. The six classes were defined as: pathogen burden <2, $2–3$, $3–4$, $4–5$, $5–6$ and >6. Genetic correlation between slope and intercept was constrained to be between 0.99 and -0.99.

When non-zero environmental or genetic correlation between intercept and pathogen burden was simulated, an additional genetic model was used to correct for the potential bias. In this case, the phenotypic values of host performance in a pathogen-free environment were added as a regression term (covariate) in the model (1).

These models resulted in estimated genetic variances and EBVs for slope and intercept, genetic correlation between slope and intercept, and six separate error variances along the *x*-axis. Genetic variance was calculated as two times the full-sib family variance. Because each individual has one observation for the reaction norm, residual variance of slope cannot be estimated, although it was originally simulated into the data. Accuracy (r_{IH}) of EBVs was calculated as a Pearson correlation coefficient between true BVs and EBVs of the families ($n = 100$ families).

The genetic variance in host performance as a function of pathogen burden was calculated as $x'_{PB} \mathbf{G} x_{PB}$ (Kolmodin & Bijma, 2004), where

$$G = \begin{bmatrix} \sigma^2_{a(b_0)} & \sigma_{a(b_0, b_1)} \\ \sigma_{a(b_0, b_1)} & \sigma^2_{a(b_1)} \end{bmatrix}.$$

The term x_{PB} is a vector [1 burden]' in which burden refers to a pathogen burden value on the *x*-axis. The use of covariance function is demonstrated only for the normally distributed pathogen burden scenario with varied family size and slope h^2 of 0.05.

For the varied family size scenario, a traditional normally distributed trait was also simulated as a control trait whose behaviour is well understood. Genetic and environmental effects were randomly sampled from $\sim N(100, 120)$ and $\sim N(100, 280)$, creating a trait similar to body weight with h^2 of 0.3 and CV_P of 20%. The genetic model used to analyse the simulated data was: $y_{ij} = \mu + fam_j + \varepsilon_{ij}$, where fam_j is a random full-sib family effect used to estimate genetic variance and EBVs.

3. Results

(i) *Effect of family size on estimates of slope genetic variance*

For the normally distributed pathogen burden, small family size resulted in elevated estimates for tolerance genetic variance (Table 2). For instance, with family size of 10, estimated genetic variances for slope were 7.9 and 2.0 times higher than the simulated value in low- and high-slope h^2 scenarios, respectively. To obtain unbiased estimates for tolerance genetic variance, family sizes of 100 and 200 were needed for low- and high-slope h^2 scenarios, respectively (Table 2). For the negative binomial distribution of pathogen burden, the same pattern was observed but the upward bias was stronger at low family size (Table 3).

For the normally distributed pathogen burden, accuracy of slope EBVs remained low in the low-slope h^2 scenario ($r_{IH} = 0.13–0.41$) but remained moderate to high in the high-slope h^2 scenario ($r_{IH} = 0.47–0.78$); accuracy increased with increasing family size in both cases (Table 2). For the negative binomial distribution of pathogen burden, the accuracies were of similar magnitude or higher (for the larger family sizes) than for the normally distributed pathogen burden scenarios (Table 3).

When CV_P of 25% or 22% for host performance under the normally distributed pathogen burden was simulated, the bias for slope variance was even greater and accuracy for slope EBVs lower (results not shown) compared to the scenario with CV_P of 30% (Table 2).

In contrast to slope variance, heritability for body weight, which served as a control trait, was estimated without bias even with the smallest family size of 10 (Table 2). Similarly, accuracy for body weight remained high at all family sizes, ranging from 0.80 at family size of 10 to 0.99 at family size of 200. This demonstrates that the estimation of genetic parameters for a slope is a challenge compared to a normal growth trait.

(ii) *Effect of family size on estimates of intercept genetic variance*

For the normally distributed pathogen burden, the intercept showed a pattern similar to the slope, but with higher average accuracies ($r_{IH} = 0.27–0.90$; Table 2). With family size of 10, estimated genetic variance for intercept was 1.9 times higher than

Table 2. *Estimated genetic parameters ($\pm$ SD) for a scenario with normally distributed pathogen burden, varied family size (10–200) and two simulated tolerance slope heritabilities (0·05 and 0·3)*

Slope h^2 scenario	Parameter estimated/ simulated values	Family size				
		10	30	50	100	200
		Estimated values				
	Slope V_G					
0·05	1·5	11·8 ± 16·2	4·96 ± 5·25	3·47 ± 4·00	2·19 ± 2·22	1·66 ± 1·42
0·3	9·0	17·7 ± 17·9	11·5 ± 7·80	10·2 ± 5·57	9·22 ± 3·79	9·13 ± 2·47
	Slope r_{IH}					
0·05		0·13 ± 0·22	0·20 ± 0·26	0·26 ± 0·23	0·31 ± 0·22	0·41 ± 0·16
0·3		0·47 ± 0·27	0·65 ± 0·11	0·69 ± 0·11	0·74 ± 0·10	0·78 ± 0·14
	Intercept V_G					
0·05	120	230 ± 255	147 ± 104	143 ± 72·6	138 ± 47·0	126 ± 30·6
0·3	120	228 ± 247	148 ± 110	130 ± 83·3	116 ± 50·7	120 ± 34·4
	Intercept r_{IH}					
0·05		0·39 ± 0·42	0·73 ± 0·21	0·80 ± 0·11	0·87 ± 0·03	0·90 ± 0·02
0·3		0·27 ± 0·41	0·60 ± 0·13	0·65 ± 0·10	0·71 ± 0·06	0·79 ± 0·04
	Body weight h^2					
0·05	0·3	0·30 ± 0·06	0·30 ± 0·05	0·30 ± 0·04	0·30 ± 0·04	0·30 ± 0·04
0·3	0·3	0·30 ± 0·06	0·30 ± 0·05	0·30 ± 0·04	0·30 ± 0·04	0·30 ± 0·04
	Body weight r_{IH}					
0·05		0·80 ± 0·04	0·92 ± 0·02	0·95 ± 0·01	0·97 ± 0·006	0·99 ± 0·003
0·3		0·80 ± 0·04	0·92 ± 0·02	0·95 ± 0·01	0·97 ± 0·006	0·99 ± 0·004
	$r_{Slope–Intercept}$					
0·05	0·0	−0·66 ± 0·44	−0·25 ± 0·62	−0·22 ± 0·52	−0·06 ± 0·52	−0·00 ± 0·40
0·3	0·0	−0·51 ± 0·47	0·06 ± 0·60	0·04 ± 0·53	0·08 ± 0·42	0·03 ± 0·22

Heritability of pathogen burden, intercept and body weight were simulated to be 0·3.
Slope h^2 and Slope V_G, slope heritability and slope genetic variance, respectively. Intercept V_G, intercept genetic variance; Body weight h^2, heritability of body weight; r_{IH}, correlation between the true and EBVs (i.e. accuracy); $r_{Slope–Intercept}$, genetic correlation between slope and intercept.

simulated in both low- and high-slope h^2 scenarios. The accuracy of the intercept was higher for the low-slope h^2 scenario. This result can be explained by the fact that increased slope genetic variation makes intercept estimation more prone to errors. This can be demonstrated with mean square errors (MSE) of slope EBVs which quantify the overall estimation error of the EBVs. For slope h^2 of 0·05, MSE for slope EBVs was 2·07, 1·34, 1·07, 0·80 and 0·70 for family sizes of 10, 30, 50, 100 and 200. For slope h^2 of 0·3, MSE for slope EBVs were more than two times higher: 4·75, 3·17, 2·73, 2·18 and 1·82. When slopes were estimated with large absolute error (the scenario with slope h^2 of 0·3), the accuracy of the intercept estimates also declined.

For the negative binomial distribution of pathogen burden, accuracy of intercept estimates was higher ($r_{IH} = 0·62–0·97$; Table 3) than for the respective normally distributed pathogen burden scenarios ($r_{IH} = 0·27–0·79$; Table 2). This occurred because as the distribution is pushed to the left, more observations are close to the intercept (0, x-axis). Thus, with low family size intercept variance was only little or not at all biased (Table 3).

(iii) *Effect of family size on estimates of host performance genetic variance*

The bias in slope and intercept variances due to low family size led to upward-biased estimates for genetic variance in host performance along the pathogen burden gradient, especially at both ends of the burden spectrum (Fig. 2). Using the six separate environmental variances estimated along the pathogen burden, heritability of host performance can also be calculated. For instance, with slope h^2 of 0·05 and family size of 100, the estimated (simulated) host performance h^2 at pathogen burdens of 1·5, 2·5, 3·5, 4·5, 5·5 and 6·5 were 0·25 (0·26), 0·21 (0·22), 0·18 (0·18), 0·16 (0·15), 0·14 (0·13) and 0·12 (0·11), respectively, implying accurate estimation for large family sizes.

(iv) *Environmental correlation between pathogen burden and host performance*

When environmental correlation between pathogen burden and intercept was changed to be either negative or positive, estimated genetic variance for slope

Table 3. *Estimated genetic parameters* ($\pm$ SD) *for a scenario with negative binomial pathogen burden distribution, varied family size* (10–200), *and tolerance slope heritability of 0·3*

Slope h^2 scenario	Parameter estimated/ simulated values	Dispersion parameters: size/mean	Family size				
			10	30	50	100	200
			Estimated values				
	Slope V_G						
0·3	9·0	1/2	21·5+5·43	14·7+3·22	9·80+2·48	9·35+1·85	9·05+1·59
0·3	9·0	3/2·5	19·6+5·95	13·1+3·34	9·18+2·39	9·03+1·74	8·91+1·64
0·3	9·0	40/3·1	16·6+5·46	12·6+3·24	9·60+2·70	8·92+1·94	9·04+1·61
	Slope r_{IH}						
0·3		1/2	0·41+0·12	0·62+0·07	0·72+0·05	0·82+0·03	0·89+0·02
0·3		3/2·5	0·44+0·10	0·65+0·07	0·73+0·04	0·83+0·03	0·90+0·02
0·3		40/3·1	0·46+0·16	0·65+0·07	0·72+0·05	0·82+0·04	0·89+0·02
	Intercept V_G						
0·3	120	1/2	120+39·4	120+25·7	121+20·6	121+19·4	119+18·1
0·3	120	3/2·5	117+48·7	117+28·4	121+23·2	122+20·4	120+18·2
0·3	120	40/3·1	131+44·3	115+31·0	117+25·5	119+21·7	120+18·7
	Intercept r_{IH}						
0·3		1/2	0·72+0·05	0·87+0·03	0·91+0·02	0·95+0·01	0·97+0·01
0·3		3/2·5	0·67+0·07	0·83+0·03	0·88+0·02	0·93+0·01	0·96+0·01
0·3		40/3·1	0·62+0·09	0·77+0·04	0·82+0·03	0·88+0·02	0·93+0·02
	$r_{Slope-Intercept}$						
0·3	0·0	1/2	−0·05+0·33	0·02+0·22	0·01+0·16	−0·02+0·12	0·01+0·10
0·3	0·0	3/2·5	−0·15+0·36	−0·06+0·16	0·01+0·17	−0·02+0·11	0·01+0·10
0·3	0·0	40/3·1	−0·19+0·55	−0·03+0·26	−0·01+0·15	−0·02+0·11	−0·01+0·10

Heritability of pathogen burden and intercept were simulated to be 0·3.
Slope h^2 and Slope V_G, slope heritability and slope genetic variance, respectively; Intercept V_G, intercept genetic variance; r_{IH}, correlation between the true and EBVs (i.e. accuracy); $r_{Slope-Intercept}$, genetic correlation between slope and intercept.
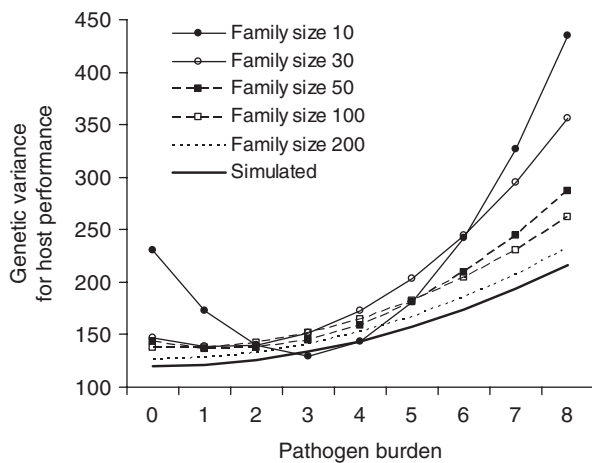

Fig. 2. Genetic variance in host performance as a function of pathogen burden. The bold line shows the simulated values, and the other lines are estimates for alternative family sizes predicted by a covariance function using the estimated genetic (co)variances of tolerance slope and intercept (Table 2). Simulated slope h^2 was 0·05.

increased symmetrically (Table 4). For example, when environmental correlation was −0·5 (or 0·5), estimated genetic variance for slope was 15·6 times higher than the simulated value. This occurred because pathogen burden was non-randomly distributed within a family, and a part of the initial variation in host performance was translated to tolerance variation.

The accuracy of the tolerance slope remained high, decreasing only slightly from $r_{IH} = 0·74$ at simulated r_E of 0·0 to $r_{IH} = 0·70$ at r_E of 0·5 (Table 4). Thus, the family differences in slopes were increased (slope V_G increased) but the consistency of the family slope EBVs with the true BVs remained unaffected. In contrast, the accuracy of the intercept decreased from 0·71 at simulated r_E of 0·0 to $r_{IH} = 0·46$ at simulated r_E of 0·5 (−0·5). Consequently, when family slopes were biased, the intercept EBVs became less consistent with the true BVs. Similar to the slope, increasing or decreasing environmental correlation elevated the estimated genetic variance for intercept (Table 4).

Genetic variance of slope was estimated without bias when host performance in a pathogen-free environment was included as a covariate in the statistical model (Table 4). Moreover, standard deviation of slope variance was reduced (SD = 1·60–1·93) compared to the model without the covariate (for $r_E = 0·0$, SD = 3·59). Similarly, the inclusion of host performance in a pathogen-free environment increased slope accuracy to 0·96 compared to the original 0·71 (Table 4). Accuracy increased because variation in initial host performance made an individual's

Table 4. *Genetic parameters ($\pm$ SD) for a scenario with varied environmental correlation (r_E) between pathogen burden and intercept estimated with a statistical model either including or excluding host performance in a pathogen-free environment as a covariate*

Parameter estimated	Simulated value	r_E						
		−0·5	−0·3	−0·1	0	0·1	0·3	0·5
		Estimated value						
Statistical model without the covariate								
Slope V_G	9·0	140 ± 12·2	56·7 ± 8·35	14·2 ± 4·34	9·40 ± 3·59	14·4 ± 4·23	57·2 ± 7·97	142 ± 11·7
Slope r_{IH}		0·71 ± 0·07	0·72 ± 0·07	0·74 ± 0·09	0·74 ± 0·12	0·75 ± 0·08	0·72 ± 0·08	0·70 ± 0·08
Intercept V_G	120	905 ± 245	409 ± 127	150 ± 58·8	117 ± 51·6	149 ± 59·9	413 ± 121	921 ± 251
Intercept r_{IH}		0·46 ± 0·11	0·51 ± 0·11	0·70 ± 0·07	0·71 ± 0·05	0·71 ± 0·07	0·51 ± 0·10	0·46 ± 0·12
$r_{Slope–Intercept}$	0·0	−0·89 ± 0·03	−0·76 ± 0·07	−0·20 ± 0·30	0·07 ± 0·41	−0·21 ± 0·30	−0·76 ± 0·06	−0·89 ± 0·03
Statistical model with the covariate								
Slope V_G	9·0	9·35 ± 1·93	9·32 ± 1·66	9·20 ± 1·60	9·48 ± 1·78	9·40 ± 1·81	9·33 ± 1·82	9·44 ± 1·82
Slope r_{IH}		0·96 ± 0·10	0·96 ± 0·09	0·96 ± 0·09	0·96 ± 0·09	0·96 ± 0·09	0·96 ± 0·07	0·96 ± 0·08
$r_{Slope–Intercept}$	0·0	−0·36 ± 0·50	−0·34 ± 0·49	−0·28 ± 0·50	−0·33 ± 0·53	−0·39 ± 0·50	−0·40 ± 0·49	−0·40 ± 0·50

Family size is 100, heritability of pathogen burden, slope and intercept were simulated to be 0·3, genetic correlation between slope and intercept simulated to be 0.

Slope V_G, slope genetic variance; Intercept V_G, intercept genetic variance; r_{IH}, correlation between the true and EBVs (i.e. accuracy); $r_{Slope–Intercept}$, genetic correlation between slope and intercept.

performance along a pathogen burden axis deviate from its family's average regression line. Thus, including initial host performance did not just reduce bias but also decreased error variance around the family's slope.

The intercept results from the analysis where the covariate was included are not comparable to the simulated intercept values because the covariate removes a major part of the variance from the original simulated intercept. Thus, intercept variance ranged from 2·40 to 3·61 (results not shown) although the simulated value was 120.

(v) *Genetic correlation between pathogen burden and host performance*

The effect of genetic correlation between pathogen burden and intercept on estimated genetic variance for slope was similar but weaker (Table 5) than the effect of the respective environmental correlation (Table 4). For instance, with a genetic correlation of −0·5 (or 0·5), slope variance was overestimated by a factor of 1·8. Slope accuracy remained high for all r_Gs tested, and intercept accuracy was only slightly reduced when r_G deviated from zero. Again, using initial host performance as a covariate in the analysis corrected for the fact that pathogen burden was non-randomly distributed within families (Table 5).

(vi) *Cost of tolerance*

In all the above scenarios, zero genetic correlation between intercept and slope was simulated. For the normally distributed pathogen burden, a decrease in

family size led to strongly negative genetic correlation between slope and intercept (Table 2). For the negative binomial distribution of pathogen burden, the bias was smaller or almost non-existent when the distribution was much skewed to the right (size = 1, mean = 2; Table 3). Having observations close to the intercept prevented the intercept from becoming automatically biased by a bias in the slope.

Similarly, with increasing or decreasing environmental (Table 4) or genetic correlation (Table 5) between pathogen burden and performance in a pathogen-free environment, the genetic correlation between slope and intercept became highly negative. The negative correlation occurred because when the slope for a family was overestimated, the intercept for the family decreased (and vice versa). These results imply that the analyses falsely indicated a presence of a genetic trade-off (i.e. cost of tolerance).

When initial host performance was used as a covariance in the scenarios with either environmental or genetic correlation between pathogen burden and intercept, the genetic correlation between the slope and intercept was always moderately negative (range from −0·28 to −0·48; Tables 4 and 5). This downward bias is unfortunate because, on the one hand, genetic variance for the slope was correctly estimated, but on the other hand, the genetic correlation was not.

4. Discussion

(i) *Random regression models*

The results showed that random regression models can be successfully applied to estimate genetic

Table 5. *Genetic parameters ($\pm$ SD) for a scenario with varied genetic correlation (r_G) between pathogen burden and intercept estimated with a statistical model either including or excluding host performance in pathogen-free environment as a covariate*

Parameter estimated	Simulated value	r_G						
		−0·5	−0·3	−0·1	0	0·1	0·3	0·5
		Estimated value						
Statistical model without the covariate								
Slope V_G	9·0	16·2 ± 4·8	11·5 ± 4·0	9·6 ± 3·7	9·2 ± 3·6	9·4 ± 3·8	11·6 ± 4·0	16·0 ± 4·5
Slope r_{IH}		0·74 ± 0·11	0·74 ± 0·10	0·74 ± 0·11	0·73 ± 0·13	0·73 ± 0·11	0·75 ± 0·09	0·75 ± 0·08
Intercept V_G	120	161 ± 67	130 ± 54	120 ± 55	119 ± 53	122 ± 52	131 ± 56	163 ± 57
Intercept r_{IH}		0·68 ± 0·08	0·71 ± 0·06	0·72 ± 0·06	0·72 ± 0·06	0·72 ± 0·06	0·72 ± 0·06	0·68 ± 0·08
$r_{Slope–Intercept}$	0·0	−0·29 ± 0·28	−0·06 ± 0·35	0·09 ± 0·41	0·05 ± 0·42	0·09 ± 0·40	−0·04 ± 0·37	−0·32 ± 0·23
Statistical model with the covariate								
Slope V_G	9·0	9·39 ± 1·73	9·59 ± 1·65	9·27 ± 1·78	9·48 ± 1·76	9·61 ± 1·82	9·42 ± 1·74	9·33 ± 1·78
Slope r_{IH}		0·96 ± 0·09	0·95 ± 0·11	0·96 ± 0·10	0·95 ± 0·10	0·96 ± 0·09	0·96 ± 0·05	0·96 ± 0·06
$r_{Slope–Intercept}$	0·0	−0·37 ± 0·52	−0·42 ± 0·52	−0·35 ± 0·49	−0·42 ± 0·48	−0·48 ± 0·48	−0·35 ± 0·47	−0·37 ± 0·53

Family size is 100, heritability of pathogen burden, slope and intercept were simulated to be 0·3, environment correlation between slope and intercept simulated to be 0.

Slope V_G, slope genetic variance; Intercept V_G, intercept genetic variance; r_{IH}, correlation between the true and EBVs (i.e. accuracy); $r_{Slope–Intercept}$, genetic correlation between slope and intercept.

parameters for tolerance, given suitable data structure. This permits unbiased estimation of genetic (co)variance for both slope (tolerance) and intercept (performance in a pathogen-free environment), which is not possible using the traditional ANCOVA approach. When each animal has several observations, e.g. before and after an infection, the model can include genetic, permanent environmental and environmental effects for the slope, allowing the estimation of heritability for tolerance. Furthermore, random regressions can be combined with covariance functions to calculate genetic variance and heritability for host performance at any point along the pathogen burden x-axis trajectory (Fig. 2). This provides a novel way to increase our understanding of infection-induced changes in genetic variance and evolutionary responses in life histories. Pathogen infections can indeed change heritabilities of performance traits, which then impact genetic responses to selection (e.g. Vehviläinen *et al.*, 2008; Lewis *et al.*, 2009). In contrast to pathogen-induced changes, changes in heritabilities due to other biotic (e.g. diet) and abiotic factors have been a focus of active research (Kause & Morin, 2001; Charmantier & Garant, 2005).

The focus here was on the analysis of field data, but the statistical method presented is applicable to experimental treatments as well. An experimental manipulation of pathogen burden level is preferred over a natural infection, because of the better control on experimental design.

A linear relationship between pathogen burden and reduction in host performance was simulated here. So far, the genetic analyses of tolerance/resistance to infections have assumed such a linear relationship (e.g. Mauricio *et al.*, 1997; Simms, 2000; Koskela *et al.*, 2002; Råberg *et al.*, 2007; Kuukka-Anttila *et al.*, 2010). However, in reality, the relationship may be non-linear, for example, sigmoidal or plateau-linear. Such non-linear relationships can be analysed with random regressions using the same principles as for a linear regression (Kirkpatrick *et al.*, 1990; Schaeffer, 2004). Yet, non-linearity makes the tolerance analysis more challenging. Typically, large sample sizes are needed to prove that non-linear regressions fit the data better than the linear ones. Moreover, when reaction norms are non-linear, the position of a family along the pathogen burden x-axis impacts the estimated regression parameters of the family.

(ii) *Experimental designs: family size*

Tolerance slope is estimated within each family and thus family size is a crucial parameter influencing estimation accuracy. The results implied that decreasing family size leads to upward-biased genetic variance estimates for tolerance slope. With decreasing family size, it is increasingly difficult to accurately estimate the true slope of a family. When a small number of individuals are sampled for each family, the sample is no longer representative of the true distribution and single observations have strong impact on the slope estimate. For some families the slope is underestimated, for others overestimated, and thus genetic variance estimate for slope is artificially increased and selection accuracy decreased.

The family size needed for accurate slope estimation is specific for each dataset. Thus, general recommendations for appropriate sample sizes are difficult to provide.

Knap & Su (2008) applied random regressions on pig litter size data to estimate genetic variance for slopes of reaction norms across North and Latin America, Europe, Asia and Australia. Consistent with our results, they observed that with decreasing sire family size from 146 to 43 (with the number of observations decreasing from 2 97 518 to 33 641), the sire reaction norms became less and less parallel increasing the estimated genetic variance for slopes. A similar result was observed in the simulation by Sae-Lim *et al.* (2010) who examined optimal population structures for estimating a genetic correlation between two environments, a measure of genotype-by-environment interaction. The smaller the family size in each environment, the less accurately were family means in each environment estimated, and thus the greater the artificially introduced variance for the genotype-by-environment interaction estimate (Sae-Lim *et al.*, 2010).

Our results also showed that for the normally distributed pathogen burden, genetic correlation between tolerance slope and intercept was biased downwards when family size was low. An upward (downward) bias in the slope of a family pushes the intercept downwards (upwards), creating an artificial negative genetic correlation. One solution for getting accurate intercept estimates is to have a larger group of uninfected individuals or to maintain a pathogen-free control treatment. Accordingly, the genetic correlation between tolerance slope and intercept was much less biased, or not biased at all, when the negative binomial distribution with pathogen burden observations aggregated around zero was used.

(iii) *Experimental designs: correlation between pathogen burden and host performance*

The results showed that when initial host performance was related to pathogen burden, indicated either as environmental or genetic correlation between intercept and pathogen burden, estimates for genetic variance for tolerance were upward biased, and an artificial negative genetic correlation between tolerance slope and intercept was easily obtained. It is well established that individuals with initially different growth or life-history trait levels may be differently exposed to infections, parasites and production diseases (Arendt, 1997; Rauw *et al.*, 1998; Moghadam *et al.*, 2001; Kause *et al.*, 2005), confounding the cause-and-effect relation between pathogen burden and reduction in host performance in field datasets.

In full-sib and sire models, within-family variation that includes both Mendelian sampling term and environmental (co)variance is not specifically modelled. In an outbred population, the Mendelian sampling term is equal to half of the genetic (co)variance. Thus, half of the genetic covariance between pathogen burden and initial host performance is translated to within-family covariance, and the effect is similar for the respective environmental covariance; upward-biased genetic variance estimates for tolerance slopes appear.

One way of addressing this issue is to include initial host performance as a covariate in the statistical model. This is convenient for traits that can be repeatedly recorded from the same individuals, and for experimental data where timing and amount of pathogen challenge are known. However, in field data, initial host performance under conditions of no infection may be unknown, and individuals may not have experienced a pathogen infection for the same duration.

An eye fluke *Diplostomum* spp. (Helminth, Trematoda) infecting farmed rainbow trout *Oncorhynchus mykiss* (Walbaum) is a potential target for applying the method even when individuals are under field conditions. *Diplostomum* spp. impairs vision by causing cataracts in the lens of the fish eye, reducing fish growth (Kuukka-Anttila *et al.*, 2010). Number of flukes within an eye (resistance) can be repeatedly counted on live fish using a portable slit-lamp microscope, and fish performance is easily repeatedly recorded from thousands of individuals (Kuukka-Anttila *et al.*, 2010). The method can be potentially applied also in sheep-breeding programmes, in which the degree of intestinal nematode infection is recorded from faecal worm-egg counts. This can be done under natural infections or after experimentally infecting animals in the field (Albers *et al.*, 1987; Woolaston & Windon, 2001).

In plant–herbivore interaction studies with naturally occurring leaf damage, the concern has been that microenvironment that affects plant fitness can also affect herbivore density and thus amount of leaf damage experienced by the plant (Tiffin & Inouye, 2000). For example, if both plant fitness and herbivore abundance are greater in more sunny microclimates, then estimates of tolerance may be artificially inflated. Such microclimate effects are challenging to correct for in the statistical analysis.

However, in the present study, selection accuracy remained moderate for all the simulated correlation scenarios, confirming that artificial selection based on family slope EBVs would be effective means for genetic improvement. This is beneficial for practical breeding attempts. Currently the major pig, poultry, dairy cattle and aquaculture breeding programmes are not selecting for tolerance to infections or production diseases.

An additional recording challenge is that many pathogen-induced infections in animals are most easily recorded as binary traits (e.g. 0 = healthy individual, 1 = infected). Yet, it is likely that infected individuals differ in the degree of pathogen burden, or that some seemingly healthy individuals are in fact mildly infected, i.e. they have a subclinical infection. Coding a normally distributed burden as a binary trait can easily lead to over- or underestimation of slope genetic variance. This artefact occurs because in case of binary scale pathogen burden the x-axis scale for tolerance regression ranges from 0 to 1, and burden variance is always $p(1-p)$, where p is prevalence of infected individuals. However, the real unrecorded continuous scale pathogen burden can have low or high variance depending on a disease, time and environment. Changing the x-axis scale while holding host performance observations unchanged artificially changes the steepness of tolerance slopes and thus alters slope variance.

Genetic variance for slopes becomes artificially underestimated when the liability scale variance in burden is low compared to the binary scale variance. This makes family slopes more parallel, reducing slope variance. In contrast, coding a liability scale burden with high variance to be a binary trait with lower variance makes family slopes artificially more diverged, creating overestimated slope genetic variance. For instance, x-axis scale can be changed from a range of '0 to 1000' pathogens to '0 to 1'. In this way, tolerance slopes become artificially steeper with high genetic variance. Consequently, recording a normally distributed pathogen burden as a threshold trait is not a realistic way to obtain unbiased estimates for tolerance genetic variance.

In conclusion, sufficiently large family sizes, accurate pathogen burden recording and control over the relation between initial host performance and pathogen burden are needed for unbiased and precise genetic analysis of tolerance. These are challenges especially for field studies, but realistic designs can be obtained by means of carefully planned experimentation and data collection.

References

Agrawal, A. A., Strauss, S. Y. & Stout, M. J. (1999). Costs of induced responses and tolerance to herbivory in male and female fitness components of wild radish. *Evolution* **53**, 1093–1104.

Albers, G. A., Gray, G. D., Piper, L. R., Barker, J. S., Le Jambre, L. F. & Barger, I. A. (1987). The genetics resistance and resilience to *Haemonchus contortus* infection in young merino sheep. *International Journal for Parasitology* **17**, 1355–1363.

Arendt, J. D. (1997). Adaptive intrinsic growth rates: an integration across taxa. *Quarterly Review of Biology* **72**, 149–177.

Astles, P. A., Moore, A. J. & Preziosi, R. F. (2006). A comparison of methods to estimate cross-environment genetic correlations. *Journal of Evolutionary Biology* **19**, 114–122.

Best, A., White, A. & Boots, M. (2008). Maintenance of host variation in tolerance to pathogens and parasites. *Proceedings of the National Academy of Sciences USA* **105**, 20786–20791.

Bishop, S. C. & MacKenzie, K. M. (2003). Genetic management strategies for controlling infectious diseases in livestock populations. *Genetics Selection Evolution* **35**, S3–S17.

Bisset, S. A. & Morris, C. A. (1996). Feasibility and implications of breeding sheep for resilience to nematode challenge. *International Journal for Parasitology* **26**, 857–868.

Charmantier, A. & Garant, D. (2005). Environmental quality and evolutionary potential: lessons from wild populations. *Proceedings of the Royal Society B: Biological Sciences* **272**, 1415–1425.

Clunies-Ross, I. (1932). Observations on the resistance of sheep to infestations by the stomach worm *Haemonchus contortus*. *Journal of the Council for Scientific and Industrial Research* **5**, 73–80.

Gilmour, A. R., Gogel, B. J., Cullis, B. R. & Thompson, R. (2006). *ASReml User Guide Release 2.0*. Hemel Hempstead, UK: VSN International Ltd.

Green, M. J., Green, L. E., Schukken, Y. H., Bradley, A. J., Peeler, E. J., Barkema, H. W., de Haas, Y., Collis, V. J. & Medley, G. F. (2004). Somatic cell count distributions during lactation predict clinical mastitis. *Journal of Dairy Science* **87**, 1256–1264.

Henderson, C. R. (1982). Analysis of covariance in the mixed model: higher-level, non homogeneous, and random regressions. *Biometrics* **38**, 623–640.

Kause, A. & Morin, J. P. (2001). Seasonality and genetic architecture of development time and body size in the birch feeding sawfly *Priophorus pallipes*. *Genetical Research* **78**, 31–40.

Kause, A., Ritola, O., Paananen, T., Wahlroos, H. & Mäntysaari, E. A. (2005). Genetic trends in growth, sexual maturity and skeletal deformations, and rate of inbreeding in a breeding programme for rainbow trout. *Aquaculture* **247**, 177–187.

Kingsolver, J. G., Ragland, G. J. & Shlichta, J. G. (2004). Quantitative genetics of continuous reaction norms: thermal sensitivity of caterpillar growth rates. *Evolution* **58**, 1521–1529.

Kirkpatrick, M., Lofsvold, D. & Bulmer, M. (1990). Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* **124**, 979–993.

Knap, P. W. & Su, G. (2008). Genotype by environment interaction for litter size in pigs as quantified by reaction norms analysis. *Animal* **2**, 1742–1747.

Kolmodin, R. & Bijma, P. (2004). Response to mass selection when the genotype by environment interaction is modelled as a linear reaction norm. *Genetics Selection Evolution* **36**, 435–454.

Koskela, T., Puustinen, S., Salonen, V. & Mutikainen, P. (2002). Resistance and tolerance in a host plant–holoparasitic plant interaction: genetic variation and costs. *Evolution* **56**, 899–908.

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

Lewis, C. R. G., Torremorell, M., Galina-Pantoja, L. & Bishop, S. C. (2009). Genetic parameters for performance traits in commercial sows estimated before and after an outbreak of porcine reproductive and respiratory syndrome. *Journal of Animal Science* **87**, 876–884.

Lillehammer, M., Ødegård, J. & Meuwissen, T. H. E. (2009). Reducing the bias of estimates of genotype by environment interactions in random regression sire models. *Genetics Selection Evolution* **41**, 30.

Mauricio, R., Rausher, M. D. & Burdick, D. S. (1997). Variation in the defense strategies of plants: are resistance and tolerance mutually exclusive? *Ecology* **78**, 1301–1311.

Meyer, K. & Hill, W. G. (1997). Estimation of genetic and phenotypic covariance functions for longitudinal or 'repeated' records by restricted maximum likelihood. *Livestock Production Science* **47**, 185–200.

Moghadam, H. K., McMillan, I., Chambers, J. R. & Julian, R. J. (2001). Estimation of genetic parameters for ascites syndrome in broiler chickens. *Poultry Science* **80**, 844–848.

Painter, R. H. (1958). Resistance of plants to insects. *Annual Review of Entomology* **3**, 267–290.

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0, http://www.R-project.org

Råberg, L., Sim, D. & Read, A. F. (2007). Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. *Science* **318**, 812–814.

Rausher, M. D. (1996). Genetic analysis of coevolution between plants and their natural enemies. *Trends in Genetics* **12**, 212–217.

Rausher, M. D. (2001). Co-evolution and plant resistance to natural enemies. *Nature* **411**, 857–864.

Rauw, W. M., Kanis, E., Noordhuizen-Stassen, E. N. & Grommers, F. J. (1998). Undesirable side effects of selection for high production efficiency in farm animals: a review. *Livestock Production Science* **56**, 15–33.

Read, A. F., Graham, A. L. & Råberg, L. (2008). Animal defenses against infectious agents: is damage control more important than pathogen control? *PLoS Biology* **6**(12), e1000004. doi:10.1371/journal.pbio.1000004

Sae-Lim, P., Komen, H. & Kause, A. (2010). Bias and precision of estimates of genotype-by-environment interaction: a simulation study. *Aquaculture* **310**, 66–73.

Schaeffer, L. R. (2004). Application of random regression models in animal breeding. *Livestock Production Science* **86**, 35–45.

Schneider, D. S. & Ayres, J. S. (2008). Two ways to survive infection: what resistance and tolerance can teach us about treating infectious diseases. *Nature Reviews Immunology* **8**, 889–895.

Simms, E. L. (2000). Defining tolerance as a norm of reaction. *Evolutionary Ecology* **14**, 563–570.

Simms, E. L. & Triplett, J. (1994). Costs and benefits of plant responses to disease: resistance and tolerance. *Evolution* **48**, 1973–1985.

Stear, M. J., Bairden, K., Duncan, J. L., Gettinby, G., McKellar, Q. A., Murray, M. & Wallace, D. S. (1995). The distribution of faecal nematode egg counts in Scottish Blackface lambs following natural, predominantly *Ostertagia circumcincta* infection. *Parasitology* **110**, 573–581.

Tiffin, P. & Inouye, B. D. (2000). Measuring tolerance to herbivory: Accuracy and precision of estimates made using natural versus imposed damage. *Evolution* **54**, 1024–1029.

Tiffin, P. & Rausher, M. D. (1999). Genetic constraints and selection acting on tolerance to herbivory in the common morning glory *Ipomoea purpurea*. *American Naturalist* **154**, 700–716.

Vehviläinen, H., Kause, A., Quinton, C., Koskinen, H. & Paananen, T. (2008). Survival of the currently fittest – genetics of rainbow trout survival across time and space. *Genetics* **180**, 507–516.

Via, S. & Lande, R. (1985). Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* **39**, 505–522.

Woolaston, R. R. & Windon, R. G. (2001). Selection of sheep for response to *Trichostrongylus colubriformis* larvae: genetic parameters. *Animal Science* **73**, 41–48.

Zerehdaran, S., van Grevehof, E. M., van der Waaij, E. H. & Bovenhuis, H. (2006). A bivariate mixture model analysis of body weight and ascites traits in broilers. *Poultry Science* **85**, 32–38.