Edinburgh Research Explorer

Impact of genetic diversity on the prevalence and dynamics of infectious disease in livestock

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Early version, also known as pre-print

Implications of Host Genetic Variation on the Risk and Prevalence of Infectious Diseases Transmitted Through the Environment

Andrea B. Doeschl-Wilson,*,†,1 R. Davidson,‡ J. Conington,* T. Roughsedge,* M. R. Hutchings,‡ and B. Villanueva*,§

*Sustainable Livestock Systems Group and ‡Animal Health Group, Scottish Agricultural College, Edinburgh, EH9 3JG, United Kingdom, †Genetics and Genomics, The Roslin Institute, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, United Kingdom, and §Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Ministerio de Ciencia e Innovación, 28040 Madrid, Spain

ABSTRACT Previous studies have shown that host genetic heterogeneity in the response to infectious challenge can affect the emergence risk and the severity of diseases transmitted through direct contact between individuals. However, there is substantial uncertainty about the degree and direction of influence owing to different definitions of genetic variation, most of which are not in line with the current understanding of the genetic architecture of disease traits. Also, the relevance of previous results for diseases transmitted through environmental sources is unclear. In this article a compartmental genetic–epidemiological model was developed to quantify the impact of host genetic diversity on epidemiological characteristics of diseases transmitted through a contaminated environment. The model was parameterized for footrot in sheep. Genetic variation was defined through continuous distributions with varying shape and degree of dispersion for different disease traits. The model predicts a strong impact of genetic heterogeneity on the disease risk and its progression and severity, as well as on observable host phenotypes, when dispersion in key epidemiological parameters is high. The impact of host variation depends on the disease trait for which variation occurs and on environmental conditions affecting pathogen survival. In particular, compared to homogeneous populations with the same average susceptibility, disease risk and severity are substantially higher in populations containing a large proportion of highly susceptible individuals, and the differences are strongest when environmental contamination is low. The implications of our results for the recording and analysis of disease data and for predicting response to selection are discussed.

A detailed understanding of host genetic variation underlying infectious disease has become increasingly relevant in the face of newly emerging infectious diseases and the frequently observed reduced efficacies of conventional control strategies (Segal and Hill 2003). Over recent years, quantitative genetics applied to livestock diseases has provided abundant evidence that individuals vary genetically in their response to infectious challenge (Bishop *et al.* 2003; Nicholas 2005). Complementary to this, recent advances in genomics (*e.g.*, genome-wide association studies or whole-genome sequencing) provide deeper insights into the genetic architecture underlying an individual's response to infectious challenge (Lohmueller *et al.* 2003; Hirschhorn and Daly 2005; Daetwyler *et al.* 2008). Both quantitative genetics and genomic approaches combined thus offer the opportunity of a refined definition of host genetic variation in disease-dependent parameters.

A critical question arising from the current genetic advances is what impact the identified genetic heterogeneity in a population has on the risk of disease outbreak and subsequent disease severity. In particular, how are the epidemiological characteristics of a disease affected by the underlying genetic architecture of disease traits and by the population structure?

The importance of host heterogeneity on disease risk and prevalence has long been recognized by epidemiologists (*e.g.*, Hethcote 1978; Anderson and May 1984; Dye and Hasibeder 1986; Dushoff and Levin 1995; Levin and Durrett 1996; Dwyer *et al.* 1997; Keeling 1999). More recently,

interest in the impact of individual differences has been sparked by the accumulating evidence of the important role of "superspreaders" in the outbreak of epidemics (*e.g.*, Lloyd-Smith *et al.* 2005; Yates *et al.* 2006), as recently reported for severe acute respiratory syndrome (SARS) in humans (Shen *et al.* 2004) and various livestock diseases (Woolhouse *et al.* 1997; Matthews *et al.* 2006). Superspreading events are characterized by a small proportion of individuals contributing a disproportionally large amount to the transmission of infectious agents (*e.g.*, by unusually large numbers of contact or by supershedding with unusually effective excretion of infectious agents into the environment) and can be considered as extreme cases of host heterogeneity.

In contrast to the well-established modeling approaches that deal with variability in disease-independent parameters such as age, sex, and contact rate (*e.g.*, Ferguson and Garnett 2000; Keeling and Eames 2005; Bansal *et al.* 2007), models dealing with host heterogeneity in disease-dependent parameters (*e.g.*, susceptibility, infectivity, and recovery rate) vary greatly in their approaches and conclusions. Further, as outlined in detail below, a review of these models reveals that the assumptions and methods applied are generally not in line with the current understanding of host genetic variation and may not hold for a number of important diseases. Many epidemiological models describe host heterogeneity in disease-dependent parameters in terms of variation in the basic reproductive number R_0, which is the average number of infections caused by an infected individual over its lifetime in an otherwise susceptible population (Anderson and May 1991). However, models differ in their specific definition of heterogeneity and this leads to contradicting conclusions (*e.g.*, Woolhouse *et al.* 1998; Springbett *et al.* 2003; Nath *et al.* 2008). For example, Lloyd-Smith *et al.* (2005) defined heterogeneity as variation in a continuously distributed "individual reproductive number" and predicted that the probability of stochastic disease extinction in heterogeneous populations increases with an increasing degree of heterogeneity. In contrast, Springbett *et al.* (2003) arrived at opposite predictions by defining heterogeneity as a number of distinct genotypes with different values of R_0 and associating greater heterogeneity with a greater number of genotypes.

Although the basic reproductive number R_0 is very important from an epidemiological point of view, describing host genetic variation by differences in R_0 does not lend itself easily to genetic analyses and interpretation as R_0 encompasses a variety of traits associated with the host as well as characteristics of the pathogen and the environment (Anderson and May 1991). Also, given that different genes are likely to control different components of the host response to infection and that genetic control strategies may target specific loci on the genome (Nath *et al.* 2008), it may be important to distinguish between heterogeneity occurring in different host-specific traits (*e.g.*, susceptibility, infectivity, and recovery rate) when assessing their impact on epidemiological characteristics. This argument is supported by recent studies (Yates *et al.* 2006; Nath *et al.* 2008), which found that predictions for epidemic characteristics are sensitive to the source of host variation.

Regardless of the trait for which heterogeneity is assumed, genetic heterogeneity has been defined in most models by a small number of homogeneous subgroups with different trait values. In genetic terms, this corresponds to a limited number of genetic loci controlling the epidemiological traits of interest. However, traits such as susceptibility and infectivity are complex quantitative traits and are thus likely to be affected by many genes and by the environment (Hayes and Goddard 2001; Lohmueller *et al.* 2003; Hirschhorn and Daly 2005). Hence, host variation is more realistically described by continuous distributions than by few homogeneous subgroups representing distinct genotypes.

Finally, up to now the influence of individual variation on the emergence and spread of infectious diseases has been studied for cases where infectious pathogens are transmitted through direct contact between susceptible and infectious individuals (*e.g.*, Anderson and May 1984; Hethcote and Van Ark 1987; Springbett *et al.* 2003; Lloyd-Smith *et al.* 2005; Yates *et al.* 2006; Nath *et al.* 2008) or for vector-borne infections (*e.g.*, Dye and Hasibeder 1996; Lloyd *et al.* 2007). However, many important diseases, particularly livestock diseases, are transmitted through environmental contamination rather than through direct contact with infectious or intermittent hosts (*e.g.*, some cases of mastitis in cattle; footrot in sheep; and parasitic gastroenteritis in humans, sheep, and cattle). Results of existing models may not directly apply to these types of diseases, as environmental factors influencing pathogen decay may alter the effect of host genetic heterogeneity on disease establishment and progression. Furthermore, the majority of existing epidemiological studies focus on the risk of epidemic outbreaks. Many livestock diseases are, however, endemic, and control strategies aim at minimizing disease severity rather than at reducing the chance of large-scale epidemic outbreaks. For these cases, characteristics such as disease prevalence and severity and their changes over time are important characteristics that need to be assessed.

This study aimed to address the above-described shortcomings. For this purpose a stochastic genetic–epidemiological model for infectious diseases transmitted through the contaminated environment was developed to investigate the influence of host genetic heterogeneity, in combination with environmental conditions, on disease emergence and progression in the population. Genetic heterogeneity was defined in a way that is consistent with the current understanding of the genetic structure underlying variation in the response to infectious challenge. The model was parameterized for footrot in sheep, an important bacterial disease in sheep-producing countries worldwide and for which there is much evidence for host genetic variation in diverse epidemiological traits (Emery *et al.* 1984; Raadsma *et al.* 1994, 1995; Nieuwhof *et al.* 2008).

Materials and Methods

The epidemiological model

The model developed in this study is a stochastic compartmental "SLDCRS" model, in which animals of a closed population of constant size (*i.e.*, no birth, removal, or death) may progress through different disease states [susceptible (S), latent infected (L), diseased (D), asymptomatic carrier (C), recovered (R), susceptible (S)] over the time course of infection (Anderson and May 1991). For animal i in disease state X, a value $T_X(i)$ is defined, describing the expected time that animal i spends in category X given progression to X at a previous time step. Infection occurs through environmental contamination, which is quantified by the density of infectious pathogen shed by infected animals in categories D and C into the environment. Each individual has its predefined shedding rates $\kappa_D(i)$ and $\kappa_C(i)$ associated with the diseased and the carrier state, respectively, which are defined as the number of bacteria shed into the environment per unit area per day. It is assumed that the total amount of infectious pathogens shed by individual animals is cumulative and uniformly dispersed, so that all animals have equal exposure to the pathogens in the contaminated environment. Infectious bacteria are assumed to have a mean survival time in the environment of T_B days.

The epidemic process is simulated as a Poisson process, *i.e.*, as a series of random events in continuous time that occur independently of one another. The possible event types are the progression of an animal from one disease state to the next (*i.e.*, S → L, L → D, D → C, C → R, and R → S), as well as pathogen shedding by an animal of category D or C, and the decay of environmental pathogens. Both the time between successive events (*i.e.*, the interevent time) and the probability of a specific event to occur are determined by the set of individual transition rates $r_X(i) = 1/T_X(i)$ from a state X ($X = S, L, D, C,$ or R) to the next state in the SLDCRS sequence, by the pathogen shedding rates $\kappa_D(i)$ and $\kappa_C(i)$ and by bacterial decay rates $r_B = 1/T_B$ as follows: in the population as a whole we define the average rate of infection of susceptible individuals as $r_1 = \sum_{i \in S} r_S(i)E$, where E is the environmental bacterial load, $r_S(i) = 1/T_S(i)$, and the sum is taken over all susceptible individuals. The transition events corresponding to other states (*i.e.*, $X = L, D, C,$ or R) occur at average rates $r_k = \sum_{i \in X} r_X(i)$, for $k = 2, 3, 4, 5$, where $r_X(i) = 1/T_X(i)$ was defined above. Similarly, shedding occurs at average rates $r_6 = \sum_{i \in D} k_D(i)$ and $r_7 = \sum_{i \in C} k_C(i)$, respectively, depending on whether bacteria are shed from diseased individuals (*D*) or asymptomatic carriers (*C*), and bacterial decay occurs at an average rate $r_8 = r_B E$. Note that event rates change during the time course of the disease outbreak as individuals move between compartments and environmental contamination changes.

The simulations of the stochastic model were carried out using Gillespie's direct algorithm (Gillespie 1977). In short, the interevent times are sampled from an exponential distribution with parameter $R = \sum_{k=1}^{8} r_k$ (Renshaw 1991). The specific event type that then occurs is obtained by drawing a random variate from a discrete distribution with probability $p(k) = r_k/R$ corresponding to event type k, $k = 1$–8, as defined above. Once the event type has been chosen, a similar sampling process is applied to determine the affected individual on the basis of its relative rate compared to that of other individuals belonging to the same category.

Parameterization of the model for footrot in sheep (benchmark model for homogeneous populations)

The model was parameterized for footrot in sheep, with *Dichelobacter nodosus* bacteria as the primary infectious agent (Egerton 2000). Various studies have demonstrated that the progress of infection follows that described by a compartmental SLDRCS model (Sinclair 1957; Abott 2000). In the benchmark model there was no variation between individuals in any of the host-specific epidemiological parameters ($T_S, T_L, T_D, T_C, T_R, \kappa_D,$ and κ_C) defined above. Parameter values for the benchmark model were adopted, where possible, from a recent deterministic epidemiological model of footrot (Nieuwhof *et al.* 2009), as listed in Table 1.

There are no estimates of the duration of the asymptotic carrier state in the literature. However, evidence exists that only a small proportion of animals per flock become asymptomatic carriers (Depiazzi *et al.* 1998). Since the benchmark model allows only all or no animals to become carriers, a short duration of 1 day was assumed for this state ($T_C = 1$ day). This conservative approach was chosen to prevent overestimation of the influence of asymptomatic carriers on epidemiological characteristics.

Footrot prevalence and severity are strongly dependent on environmental conditions affecting bacterial survival (Abbot and Egerton 2003; Conington *et al.*, 2010). Different environmental conditions were modeled using different values for the bacterial survival time T_B (Table 1).

Modeling heterogeneous populations and assessing the influence of heterogeneity

Field studies have shown that host genetic factors influence an animal's susceptibility to footrot as well as the speed of recovery and duration of immunity and that these traits are of complex inheritance; *i.e.*, genetic variation is best described by continuous distributions (Raadsma *et al.* 1994, 1995; Nieuwhof *et al.* 2008). For ease of interpretation, variation between hosts was considered separately for each host-specific epidemiological trait listed in Table 1 (*i.e.*, T_S, T_L, T_D, T_C, T_R, κ_C, or κ_D). A heritability of 1 was assumed for all traits; *i.e.*, it was assumed that environmental variation is fully represented by the stochastic nature of the model. Values for the host-specific epidemiological trait in question were drawn from a two-parameter continuous gamma distribution with shape parameter α and scale parameter θ [*i.e.*, probability density function is $x^{\alpha-1}(\exp(-x/\theta)/\Gamma(\alpha)\theta^\alpha)$]. The individual transition rates between successive disease states in the epidemiological model are then the inverse of the corresponding gamma-distributed time periods. Hence,

Table 1 Parameter values for the benchmark model assuming no variation between hosts

Parameter	Definition	Benchmark value
T_S	Duration of susceptible period per unit of environmental pathogen density	20,000 (corresponds to a rate of 0.00005)[a]
T_L	Duration of latent period	10 days[b]
T_D	Duration of infectious period with clinical symptoms (diseased period)	20 days[b]
T_C	Duration of asymptomatic infectious carrier state	1 day[a]
T_R	Duration of immunity after recovery	30 days[b]
κ_D, κ_C	Shedding rates of animals in categories D (infectious with clinical symptoms) and C (asymptomatic carrier), respectively	1.0 *D. nodosus* bacteria per unit area per day for animals in categories D and C[a]
T_B	Survival time for infectious bacteria in the environment	1, 1.5, or 12.0 days[b]
N	Population size	1000[a]

[a] Arbitrary values. The value for T_S was chosen to yield footrot prevalences similar to those reported in field studies (*e.g.*, Nieuwhof *et al.* 2008) for unit shedding rates κ_D and κ_C, respectively. For more information on the value for T_C, see text.
[b] Benchmark values adopted from Nieuwhof *et al.* (2009).

for $T_X \sim$ Gamma (α, θ), $r_X = 1/T_X \sim$ InverseGamma $(\alpha, 1/\theta)$. Gamma distributions can adopt a variety of shapes, such as concave $(\alpha < 1)$, nearly symmetric $(\alpha$ large$)$, or unimodal and strongly right skewed $(\alpha \approx 1)$. The shape of the probability density function is determined by the shape parameter α, and the coefficients of variation, dispersion, skewness, and kurtosis, are $1/\sqrt{\alpha}$, $1/\alpha$, $2/\sqrt{\alpha}$, and $1/\alpha$, respectively. Smaller values of α thus indicate greater heterogeneity as well as a greater degree of asymmetry and peakedness. As an example, supershedders in a disease to which the "20/80 rule" (20% of the population contributes 80% of the entire transmission potential) applies (Woolhouse *et al.* 1997; Matthews *et al.* 2006) can be represented as the top 20% from the right-hand tail of a gamma distribution for κ_D or κ_C with shape parameter $\alpha = 0.25$. From a genetic point of view, the shape of the gamma distribution reflects the genetic architecture of the trait. For example, an infinitesimal genetic model assuming that the trait in question is affected by a large number of genes each with an infinitesimally small effect, would correspond to a symmetric distribution, whereas the presence of few loci with large effects on the trait of interest and an unbalanced allele frequency at these loci in the population, or dominance, could give rise to a skewed distribution.

The central question addressed in this study is how the source of genetic variation and the underlying genetic architecture affect disease risk and severity in the population. For this purpose, distributions for each epidemiological trait for which variation was assumed were generated with the same trait mean $(\alpha\theta)$ as that of the benchmark model for homogeneous populations (Table 1). For each of these epidemiological traits, distributions with a variety of shapes were generated by choosing $\alpha = 0.25$, 1.25, 2.5, and 12.5 (Figure 1). Hence, to assess the influence of the source of genetic variation, outputs corresponding to distributions for different epidemiological traits of the same shape were compared (*i.e.*, the same value of α and a different mean value $\alpha\theta$). For assessing the role of different genetic architectures underlying a particular epidemiological trait, outputs corresponding to distributions of different shapes for the particular epidemiological trait in question were compared (*i.e.*, different α and the same mean value $\alpha\theta$).

### Simulation process and assessment of model results

In the simulations, infection was introduced into an initially fully susceptible population through one latent infected individual. The simulations were carried out for closed populations with no birth or death or removal processes and for which disease progression was not influenced by treatment effects. The model was run for a simulated time period of 3 years, after which the infection either died out or, as indicated by preliminary results, converged to equilibrium prevalence in all simulated scenarios. Ten thousand replicates were generated for each parameter combination.

The impact of host heterogeneity (in terms of both the source of variation and the underlying genetic architecture) was assessed by four different epidemic characteristics: first, we calculated the probability of the disease to exist in the population at different time points (calculated as the proportion of replicates with at least one infected individual
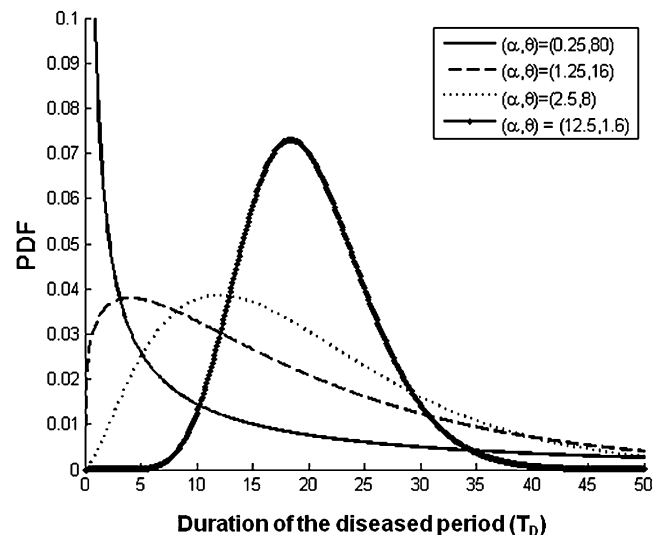
Figure 1 Probability density function (PDF) for the gamma distribution with distribution parameters (α, θ) used to describe host variation in the duration of the diseased period (T_D). The distributions have the same mean value $T_D = 20$ days.

Slow bacterial decay *(T_B = 12 days)*   Fast bacterial decay *(T_B = 1.5 days)*

Variation in host susceptibility (T_S)
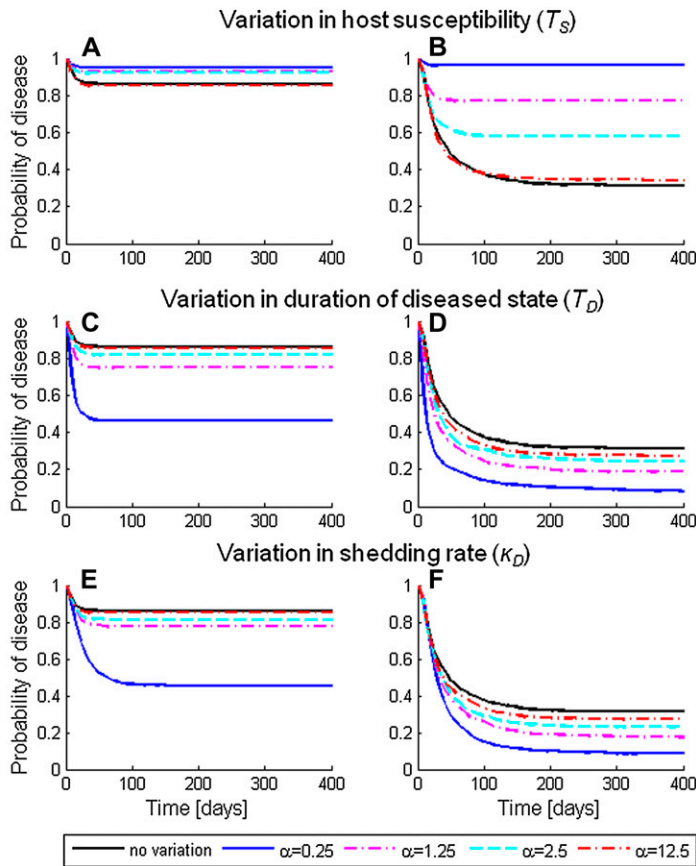
Figure 2 (A–F) Impact of variation in epidemiological traits on the probability of disease prevailing in the population over time. The probability of disease prevalence at time t was calculated as the proportion of replicates with at least one infected individual at time t. Please see File S1 for the impact of variation in all epidemiological traits for which variation was introduced.

Variation in duration of diseased state (T_D)

Variation in shedding rate (κ_D)

Legend: ── no variation   ── α=0.25   ─·─ α=1.25   ─ ─ α=2.5   ─··─ α=12.5

at the time in question, where infected individuals are individuals of categories L, D, or C), providing thus predictions for the risk of disease establishment and persistence over the duration of at least 3 years. Second, we assessed disease prevalence patterns over time. Third, we recorded disease incidence, *i.e.*, the total number of new infections over the simulated time period, as a measure of disease severity. Finally, as the stochastic epidemiological model provides continuous information of the disease state of every individual over time, observable phenotypic infection characteristics, such as the frequency of infections (*i.e.*, number of times that the individual moved from state S to L) and the number of days that an individual was infected (*i.e.*, spent in categories L, D, or C) throughout the simulated 3 years, were calculated for each individual. These characteristics may constitute the estimated or observed phenotypes in genetic analyses of field data and hence form the bases of selection or management strategies. We therefore assessed how population averages and variances of these observable infection characteristics are affected by variations in epidemiological input parameters. The averages and variances were calculated as averages over all individuals, over all replicates avoiding stochastic extinction, and refer to the last year in the simulations, *i.e.*, when the system reached a steady state.

Results

Impact of variation on the probability of disease establishment

For environmental conditions favoring bacterial survival, footrot either disappears from the population within a few weeks after introduction or becomes persistent (Figure 2, A, C, and E and supporting information, File S1). When bacterial decay is fast, the risk of footrot establishment and long-term persistence is considerably lower than when bacterial decay is slow, except for $\alpha = 0.25$. For instance, for homogeneous populations the probability of footrot persistence, estimated by the proportion of replicates predicting positive prevalence at the end of the simulated time period, was 0.86 when $T_B = 12$ days (Figure 2, A, C, and E), 0.31 when $T_B = 1.5$ days (Figure 2, B, D, and F), and 0.03 when $T_B = 1$ day (not shown).

Positively skewed variation in host susceptibility (T_S) tends to increase the probability of footrot establishment (Figure 2, A and B). The influence is strongest for populations containing a large proportion of highly susceptible individuals (low α). For $\alpha = 0.25$, the risk of footrot persistence for ≥3 years increases from 0.86 for homogeneous populations to 0.96 when $T_B = 12$ days (Figure 2A) and substantially more, *i.e.*, from 0.31 to 0.95 when $T_B = 1.5$
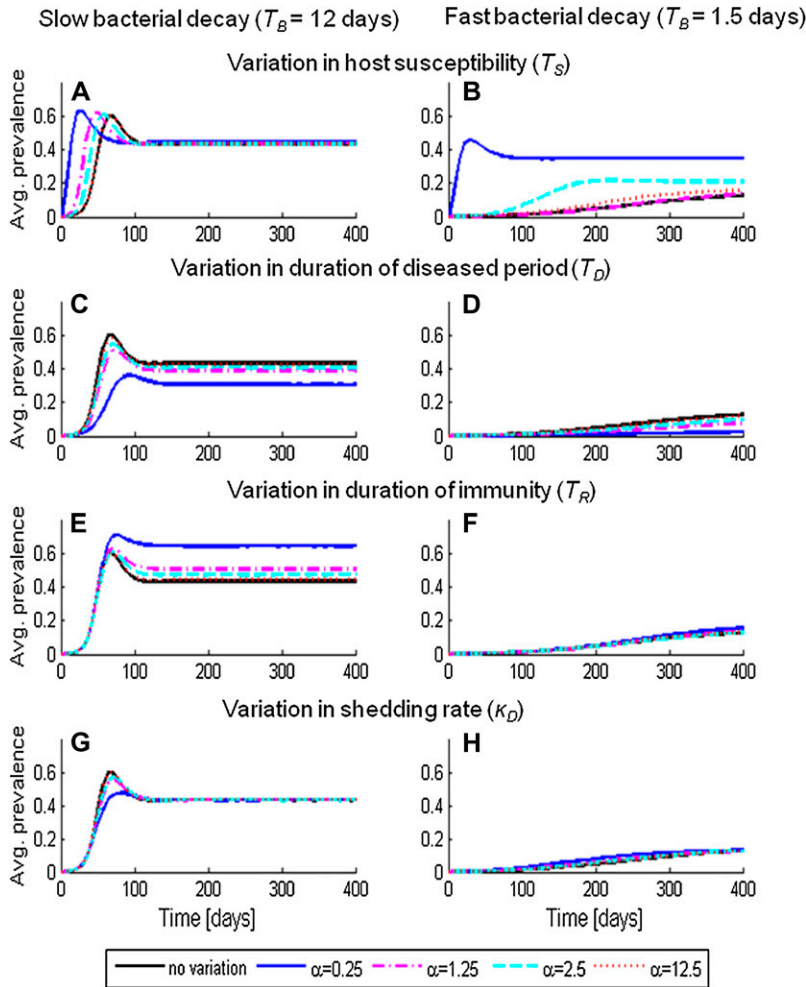
Figure 3 (A–H) Impact of variation in epidemiological traits on the average footrot prevalence over time. The average prevalence at time *t* was determined from all replicates avoiding extinction at time *t*. Please see File S1 for the impact of variation in all epidemiological traits for which variation was introduced.

days (Figure 2B). As α increases, the difference between heterogeneous and homogenous populations for the probability of disease establishment decreases, and for $\alpha = 12.5$ the difference is negligible (Figure 2, A and B).

Positively skewed variation in the duration of the diseased period (T_D) and in shedding rates (κ_D) tends to decrease the risk of footrot establishment in the population. As for variation in susceptibility (T_S), the influence is substantially stronger for low α and becomes negligible for high values of α (Figure 2, C–F). However, in contrast to what is observed for host variation in susceptibility, variation in host infectivity (T_D or κ_D) can substantially affect the risk of disease establishment in all environments. For example, for $T_B = 12$ days, the probability of disease persistence for ≥3 years decreases from 0.86 for homogeneous populations to 0.46 and 0.45, for variation in T_D and κ_D with $\alpha = 0.25$, respectively (Figure 2, C and E). For $T_B = 1.5$ days, the risk of footrot persistence in populations with the same degree of variation in T_D and κ_D, respectively, decreases from 0.31 to 0.08 (Figure 2, D and F) and becomes virtually zero for $T_B = 1$ day (*i.e.*, disease persisted in <5 of 10,000 replicates; results not shown). There is little difference between the probabilities of disease establishment associated with host

variation in T_D or κ_D (Figure 2, C–F). Variation in other epidemiological parameters (T_L, T_C, κ_C, and T_R; Table 1) was, however, found to have little effect on the risk of disease establishment, regardless of the shape of variation and of the simulated environmental conditions.

Impact of variation on disease prevalence over time

Figure 3 shows the predicted influence of host variation on footrot prevalence over time in populations avoiding stochastic disease extinction. Average prevalence profiles, and the impact of host variation on them, depend strongly on bacterial survival rates. If bacteria survive in the soil for a long time, average footrot prevalence rapidly increases toward its peak within a few weeks after the introduction of the infection into the population, after which it gradually declines toward an endemic equilibrium (Figure 3, A, C, E, and G). In contrast, when bacteria decay fast, prevalence in populations in which footrot could establish generally increases gradually toward an endemic equilibrium with a lower average proportion of animals infected than for slow bacterial decay (Figure 3B shows the trend toward the equilibrium).

There is little difference in the prevalence patterns of homogeneous and heterogeneous populations when underlying

distributions are symmetric and dispersion is moderate (*e.g.*, $\alpha = 12.5$ in Figure 3). However, skewed distribution with great dispersion for host susceptibility (T_S) and the duration of immunity (T_R) tend to increase disease prevalence (Figure 3, A, B, E, and F), whereas similarly skewed distributions for the duration of the diseased period (T_D) and for the corresponding shedding rate (κ_D) tend to decrease its prevalence (Figure 3, C, D, G, and H). The degree of influence depends strongly on the environmental conditions affecting pathogen survival. For example, if bacteria decay slowly (*e.g.*, $T_B = 12$ days, Figure 3A), variation in host susceptibility may affect disease progression in the population at the early stages, but has little effect on the peak prevalence or on the prevalence at the endemic equilibrium. In contrast, when bacterial decay is fast (*e.g.*, $T_B = 1.5$ days, Figure 3B), variation in host susceptibility affects prevalence levels at any stage of the epidemics, with greater dispersion generally leading to a more rapid increase in prevalence, higher prevalence levels, and faster convergence toward the equilibrium.

Variations in the duration of the diseased period (T_D) and shedding rate (κ_D), as well as in the duration of immunity (T_R), affect predicted disease prevalence only in environmental conditions favoring bacterial survival and when dispersion and skewness are sufficiently strong (Figure 3, C–H). The presence of a large proportion of individuals with fast recovery (low T_D) decreases prevalence levels compared to those observed for homogeneous populations with the same mean recovery rate at any stage of the epidemics (Figure 3C), whereas a large proportion of individuals with little or no immunity (low T_R) affect prevalence levels only after prevalence has already peaked (Figure 3E) and variation in shedding rates (κ_D) affects only peak prevalence (Figure 3G). Variation in the duration of the latent stage (T_L) or the asymptotic carrier stage (T_C), as well as in the corresponding shedding rates (κ_C), has little impact on the simulated prevalence patterns (results not shown).

Impact of variation on the overall disease severity in the population

The influence of host variation on the predicted overall footrot severity, described by the total number of incidences occurring during the simulated 3-year time period, was found to depend similarly on the degree of dispersion and on environmental conditions as was observed for disease risk and prevalence profiles. In particular, when bacteria decay fast (T_B low), a large proportion of highly susceptible individuals poses the greatest risk for severe outbreaks, whereas when bacterial decay is slow (T_B high), a large proportion of individuals not developing immunity causes the most severe outbreaks. For example, for $T_B = 1.5$ days, the average number of incidences increased from 4387 for homogeneous populations to 13,776 in populations with large variation in T_S ($\alpha = 0.25$). The same variation in T_R produced on average 5361 incidences. In contrast, for $T_B = 12$ days, the average number of incidences for homogeneous populations and populations with the same large variation

in T_S were similar (*i.e.*, 16,670 and 17,564, respectively), whereas populations with large variation T_R experienced a substantially higher average number of incidences (*i.e.*, 24,465 incidences).

Impact of variation on observable host infection characteristics

Table 2 shows the predicted impact of host heterogeneity, when expressed in different environments, on the frequencies and number of days that animals are infected per year. As would be expected, average number of infected days and frequency of infections were generally higher in environments with high contamination risk (high values of T_B). With some exceptions (outlined below), standard deviations in these characteristics were also higher for higher values of T_B, implying that genetic variation is generally more strongly expressed in more infectious environments (Table 2). Also, as would be expected, individual variation in observable infection characteristics generally increases with increasing degree of heterogeneity in the epidemiological parameters. Note that individual variation in observable infection characteristics also occurs for homogeneous population due to stochasticity.

Introducing heterogeneity shows similar footprints on individual infection characteristics as was observed for the risk of disease establishment and disease severity on a population level: positively skewed host variation in susceptibility and immunity (T_S, T_R) has a detrimental effect on observable infection characteristics, whereas positively skewed variation in the diseased period (T_D) has the opposite effect, and the effect size generally increases with increasing heterogeneity and skewness. Also, heterogeneity in T_S mainly affects averages and variances in the observable infection characteristics when environmental contamination is low, whereas heterogeneity in T_D and T_R shows their impact on infection characteristics in all simulated environments (Table 2). Variation in other epidemiological parameters, in particular shedding rates (κ_D and κ_C), was found to have no significant effect on average or variation in host infection characteristics (results not shown). The results in Table 2 also illustrate that variation in a particular epidemiological trait may affect various infection characteristics simultaneously. For example, increasing the proportion of individuals with a short expected duration of the diseased period (*i.e.*, decreasing α for T_D in Table 2) tends to decrease not only the average number of days per year that individuals are infected, but also, given that infection is a stochastic event, the average frequency of infections per year.

Discussion

The important role of host genetic heterogeneity on the spread of infectious diseases has long been recognized, but to our knowledge this study is the first to examine systematically how the type and shape of host variation influence disease risk and severity. Our model results suggest that the

Table 2 Impact of the source and shape of host variation on individual infection characteristics for different environmental conditions represented by T_B

Distribution parameters (α, θ)	Average (SD) no. days infected during year 3[a]		Average (SD) no. times infected during year 3[a]	
	T_B = 12 days	T_B = 1.5 days	T_B = 12 days	T_B = 1.5 days
Benchmark model (no variation)				
	157 (42.1)	56 (34.7)	5.43 (1.39)	1.87 (1.02)
Variation in susceptibility (T_S)				
(0.25, 80,000)	161 (47.7)	131 (51.1)	5.39 (1.48)	4.27 (2.02)
(1.25, 16,000)	158 (44.7)	93 (51.4)	5.43 (1.42)	3.18 (1.68)
(2.50, 8,000)	158 (43.9)	80 (45.1)	5.43 (1.40)	2.73 (1.43)
(12.5, 1,600)	158 (43.2)	62 (37.1)	5.43 (1.40)	2.08 (1.11)
Variation in duration of the diseased period (T_D)				
(0.25, 80)	112 (78.1)	35 (49.1)	5.08 (1.90)	0.73 (0.76)
(1.25, 16)	142 (53.6)	48 (42.9)	5.46 (1.77)	1.51 (0.98)
(2.50, 8)	149 (56.1)	52 (40.3)	5.49 (1.66)	1.69 (1.01)
(12.5, 1.6)	156 (46.3)	55 (36.3)	5.46 (1.46)	1.84 (1.03)
Variation in duration of immune period (T_R)				
(0.25, 120)	237 (57.8)	68 (40.7)	5.53 (1.80)	2.26 (1.24)
(1.25, 24)	184 (62.6)	59 (36.7)	5.87 (2.17)	1.97 (1.10)
(2.50, 12)	172 (61.4)	58 (35.9)	5.75 (1.72)	1.93 (1.07)
(12.5, 2.4)	161 (47.4)	56 (35.0)	5.52 (1.55)	1.87 (1.03)

The source of host variation is represented by variation in different epidemiological parameters (*e.g.*, T_S, T_D, and T_R), whereas the shape is represented by different parameter values α of the gamma distribution. The scale parameter θ was chosen so that the population mean $\alpha\theta$ is the same as for homogeneous populations. Values and parameters are defined in Table 1. For other epidemiological parameters (κ_C, κ_D, T_L, and T_C) for which host variation was assumed (see Table 1), the predicted impact of variation on the predicted infection characteristics was found to be negligible.

[a] Averages were taken over all replicates with positive footrot prevalence.

impact of host heterogeneity on relevant epidemiological characteristics is largely controlled by the shape of underlying distributions describing individual variation. Whereas differences between homogeneous and heterogeneous populations in predicted epidemiological characteristics are generally subtle and often negligible when heterogeneity is described by symmetric distributions in underlying epidemiological traits, they can be large when underlying distributions are skewed and leptokurtic and dispersion is high. In particular, by representing host heterogeneity with right-skewed gamma distributions, our model results predict that the influence of a large proportion of individuals with low values in a particular trait (*e.g.*, low resistance, low shedding rate, short duration of diseased period, or short-lived immunity) on the disease epidemiology outweighs the influence of a small proportion of individuals with extremely high values for these traits. Thus, the presence of few high shedders does not necessarily increase disease risk if the majority of individuals have low pathogen shedding rates. Also, our results suggest that understanding the source of genetic variation, *i.e.*, identifying the epidemiological trait(s) for which variation occurs, is important when making predictions for epidemiological outcomes and interpreting epidemiological data. Distributions of the same shape for different epidemiological traits often resulted in different predictions for disease risk and severity. The impact of variation in a particular trait depends strongly on the role of this trait in the progression of the disease in the population. For example, both the duration of the infectious period (T_D and T_C) and bacterial shedding rates (κ_D and κ_C) may be considered as measures of host infectivity in field studies. However, whereas shedding rates are incorporated directly into epidemiological models, time periods are usually incorporated through their inverses, as these provide the required transition rates. Hence, even if variation in shedding rates and duration of the infectious period are represented by distributions of similar shape, their impact on the epidemiology can be quite different.

Our study differs from the majority of previous theoretical studies in that it focuses on an infectious disease that is transmitted through an environmental source instead of through direct contact with infectious individuals. This type of disease merits attention as it constitutes a substantial proportion of livestock diseases of major concern for animal and human health. Our results agree with those from studies on directly transmitted diseases in that host genetic heterogeneity can significantly affect the risk of disease emergence (*e.g.*, Hethcote and Van Ark 1987; Lloyd-Smith *et al.* 2005; Yates *et al.* 2006) and persistence (*e.g.*, Doeschl-Wilson *et al.* 2009) and the course of emerging epidemics (*e.g.*, May and Anderson 1988; Springbett *et al.* 2003; Nath *et al.* 2008). However, our results provide novel evidence that, for a disease transmitted through the environment, the influence of host genetic variation on disease emergence and severity can strongly depend on environmental conditions. These were found to affect more the degree than the direction of influence. For example, heterogeneity in host susceptibility was found to increase disease risk and severity in all modeled environments but its impact was much stronger when

environmental contamination was low (Figures 2, A and B, and 3, A and B). Also, except for populations with large heterogeneity in susceptibility, predicted variances in observable phenotypic infection characteristics were generally lower in these conditions (Table 2), indicating that genetic variation may not be fully observed in environments with low exposure. These results complement the analytical expressions derived by Bishop and Woolliams (2010), which predict that incomplete exposure to infection (*e.g.*, due to low disease prevalence in the population) leads to a downward bias in estimates of genetic variation.

Previous studies often differ in their predictions for the effect of host variation owing largely to different definitions of host genetic heterogeneity (Springbett *et al.* 2003; Lloyd-Smith *et al.* 2005; Yates *et al.* 2006; Nath *et al.* 2008). The study closest to ours in terms of definition of host heterogeneity is that of Lloyd-Smith *et al.* (2005) who modeled host variation in infectiousness for diseases transmitted through direct contact also with continuous gamma distributions. They established that the probability of extinction increases with increasing dispersion and skewness due to the high proportion of individuals contributing very little to the spread of the disease at the early stages. Our model predicts the same trend for diseases transmitted through environmental sources. However, the models of Lloyd-Smith *et al.* (2005) also predict for populations avoiding extinction more severe outbreaks as dispersion and skewness increase. In contrast, our model predicts that variation in infectiousness, when represented by the duration of the infectious period or by different shedding rates, tends to decrease disease prevalence at all times. The discrepancy could be attributed to the different modes of disease transmission considered in both models. Compared to directly transmitted diseases, the influence of few highly infectious individuals on the spread of disease is diffused when transmission occurs through an environmental source, whose contamination level is the cumulative contribution of all infectious individuals.

Several studies on directly transmitted diseases have shown that the impact of genetic variation on the risk of disease emergence depends on to the source of genetic heterogeneity. For example, assuming a finite locus model for genetic variation, with one locus controlling the transmission coefficient (representing host susceptibility to infection) and another locus controlling the recovery rate, Nath *et al.* (2008) found that genetic variation in the transmission coefficient resulted in higher probabilities of disease emergence, whereas variation in the recovery rate had no such impact. Defining host variation also by few (*i.e.*, two) distinct homogeneous subgroups, Yates *et al.* (2006) have further shown that variation in host infectivity can greatly reduce the probability of disease emergence. Our model predicts similar trends if host susceptibility, recovery, or infectivity are controlled by multiple loci. However, in contrast to the results of Nath *et al.* (2008) and our results, Yates *et al.* (2006) predict that variation in susceptibility will not increase the risk of disease establishment, in particular when

bacterial decay is fast. A likely cause for this discrepancy is that Yates *et al.* (2006) simulated heterogeneous populations by assigning 100 times higher susceptibility to 10% of the individuals than to the rest of the population, thus approximating a negatively skewed distribution in susceptibility rather than the positively skewed distributions considered here. Combining the results of both studies would thus suggest that compared to homogeneous populations with the same average susceptibility, heterogeneous populations with a large proportion of highly susceptible individuals have a higher risk of disease emergence, but the emergence risk will not decrease if the majority of individuals are resistant.

For diseases where a significant fraction of the population become asymptomatic carriers of the infection, such as footrot in sheep (Depiazzi *et al.* 1998), a significant role in the persistence of the disease in the population has been attributed to these carriers (Anderson and May 1991). We found, however, that, for the specific set of parameters and distributions used, the presence of few individuals that become asymptomatic carriers over a considerably long time period (*i.e.*, several months) has no significant influence on the predicted epidemic outcomes.

Here we have assumed that individual variation occurs for one epidemiological parameter at a time. Although this approach allowed us to systematically examine the contributions of different sources of variation, in reality genetic variation is likely to occur simultaneously in a number of different epidemiological traits. We also carried out additional simulations assuming variation in more than one parameter at a time. The results are more difficult to interpret, but, in essence, symmetric variation with small dispersion in multiple parameters still led to subtle differences in model predictions between homogenous and heterogeneous populations, whereas high dispersion generally led to larger discrepancies between both types of populations.

Several important implications arise from our results concerning both the analysis and interpretation of disease data and the evaluation of control strategies, in particular genetic selection for increasing disease resistance. The latter has long been considered a viable alternative to conventional disease control, but its benefits in terms of reducing disease risk and severity are poorly understood. Our results suggest that accurate predictions of selection response for disease risk and severity require a thorough understanding of the structure of genetic variation, *i.e.*, of the source of genetic variation and the shape of the corresponding distributions. Genetic control strategies in livestock currently focus on exploiting genetic variation in disease resistance as a whole to reduce disease risk and severity. Our results, however, indicate that variation in different epidemiological traits, such as infectiousness or duration of immunity, can influence very differently both disease risk and severity. The question therefore arises whether genetic variation in these alternative traits can be identified using currently available disease data and statistical methods. For example, a recent

simulation study (Lipschutz-Powell *et al.* 2010) demonstrated that classic quantitative genetics models applied to binary incidence data cannot identify genetic variation in infectivity, as an individual's infectivity is expressed in its group members rather than in the individual itself. Further efforts in more detailed recording of disease data and in the development of adequate genetic models would be required, before this important source of genetic variation can be exploited.

A novel insight emerging from our study is that the shape of genetic variation influences not only disease establishment and prevalence, but also observable infection characteristics of individuals that may constitute the phenotypes in genetic data analysis. Genetic analyses of disease data often assume an underlying normally distributed individual liability for disease prevalence (Robertson and Lerner 1949; Falconer and Mackay 1996). Although this assumption may hold for diseases for which many genes have small additive effects on epidemiological parameters, skewed leptokurtic distributions may be more appropriate for diseases where a number of genes have large effects and allele frequencies are unbalanced or in cases where the relationship between genetic effects and the epidemiological traits in consideration is not linear. Our results suggest that greater attention should be given to the assumptions concerning liability distributions in genetic analyses of disease data, as disease risk and prevalence depend strongly on the shape of the distributions. Wrong assumptions could not only lead to biased estimates of genetic parameters, but also produce wrong predictions for the selection response. Recent advances in genomic studies could provide the opportunity to accurately describe the genetic architecture that defines the shape of the distribution in epidemiological traits of interest.

Our model results further emphasize that the consequences of genetic selection on both the population mean and the variation should be taken into consideration when predicting disease risk and prevalence in future generations. For example, selecting against high susceptibility would not only increase the average disease resistance in future generations, but also decrease the proportion of highly susceptible individuals in the frequency distribution for susceptibility. On the basis of our predictions, the associated changes in the shape of the distribution would decrease the future risk of disease establishment and disease severity much more than would be anticipated from classical quantitative genetics theory alone, which does not account for changes in the distributions' shapes. Finally, the results of our study elucidate that the choice of the most appropriate selection strategy would depend on the parent and offspring environment. For example, whereas our results suggest that selection against high susceptibility would be the most efficient means to reduce disease risk and prevalence in environments with low contamination risk, selection for fast recovery would be advantageous if the contamination risk is high. Given thus the results from this study, among the next steps to be taken in the genetic control of infectious disease are a more accurate description of genetic heterogeneity among individuals and the development of statistical methods to account for different types and shapes of genetic variation in the analysis of disease data and in genetic–epidemiological models.

Acknowledgments

Literature Cited

Abbot, K. A., 2000   The epidemiology of intermediate footrot. Ph.D. Thesis, University of Sydney, Australia.

Abbot, K. A., and J. R. Egerton, 2003   Effect of climatic region on the clinical expression of footrot of lesser clinical severity (intermediate footrot) in sheep. Aust. Vet. J. 81: 756–762.

Anderson, R. M., and R. M. May, 1984   Spatial, temporal and genetic heterogeneities in host populations and the design of immunisation programmes. IMA J. Math. Appl. Med. Biol. 1: 233–266.

Anderson, R. M., and R. M. May, 1991   *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, Oxford.

Bansal, S., B. T. Grenfell, and L. A. Meyers, 2007   When individual behaviour matters: homogeneous and network models in epidemiology. J. R. Soc. Int. 4: 879–891.

Bishop, S. C., and J. A. Woolliams, 2010   On the genetic interpretation of disease data. PLoS ONE 5(1): e8940.

Bishop, S. C., M. De Jong, and D. Gray, 2003   Opportunities for incorporating genetic elements into the management of farm animal diseases: policy issues (background study paper No. 18). Food and Agriculture Organisation of the United Nations (FAO). http://dad.fao.org/en/refer/library/reports/bsp18e.pdf.

Conington, J., M. H. M. Speijers, A. F. Carson, S. Johnston, and S. Hanrahan, 2010   Foot health in sheep—prevalence of hoof lesions in UK and Irish sheep [paper no. 340 (Abstract)]. Proceedings of the British Society of Animal Science, Belfast.

Daetwyler, H. D., B. Villanueva, and J. A. Woolliams, 2008   Accuracy of predicting the genetic risk of disease using a genome-wide approach. PLoS ONE 3(10): e3395.

Depiazzi, L. J., W. D. Roberts, C. D. Hawkins, M. A. Palmer, N. C. Pitman *et al.*, 1998   Severity and persistence of footrot in Merino sheep experimentally infected with a protease thermostable strain of *Dichelobacter nodosus* at five sites. Aust. Vet. J. 76: 32–38.

Doeschl-Wilson, A. B., R. Sawalha, S. Gubbins, and B. Villanueva, 2009   Implications of conflicting associations of the prion protein (PrP) gene with scrapie susceptibility and fitness on the persistence of scrapie. PLoS ONE 4: e7970.

Dushoff, J., and S. Levin, 1995   The effects of population heterogeneity on disease invasion. Math. Biosci. 128: 25–40.

Dwyer, G., J. S. Elkinton, and J. P. Buoncaccorsi, 1997   Host heterogeneity in susceptibility and disease dynamics: tests of a mathematical model. Am. Nat. 150(6): 685–707.

Dye, C., and G. Hasibeder, 1996   Population dynamics of mosquito-borne disease: effects of flies which bite some people more than others. Trans. R. Soc. Trop. Med. Hyg. 80: 69–77.

Egerton, J. R., 2000   Footrot and other foot conditions, pp. 243–249 in *Diseases of Sheep*, Ed. 3, edited by Martin, I. D. and I. D. Aitken. Blackwell Science, London.

Emery, D. L., D. J. Stewart, and B. L. Clark, 1984   The comparative susceptibility of five breeds of sheep to footrot. Aust. Vet. J. 61: 85–88.

Falconer, D. S., and T. F. C. Mackay, 1996   *Introduction to Quantitative Genetics*, Ed. 4. Longman, Harlow, UK.

Ferguson, N. M., and G. P. Garnett, 2000   More realistic models of sexually transmitted disease transmission dynamics: sexual partnership networks, pair models, and moment closure. Sex. Transm. Dis. 27: 600.

Gillespie, D. T., 1977   Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. 81: 2340–2361.

Hayes, B. J., and M. E. Goddard, 2001   The distribution of the effects of genes affecting quantitative traits in livestock. Genet. Sel. Evol. 33: 209–229.

Hethcote, H. W., 1978   An immunization model for a heterogeneous population. Theor. Popul. Biol 14: 338–349.

Hethcote, H. W., and W. Van Ark, 1987   Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation, and immunization programs. Math. Biosci. 84: 85–118.

Hirschhorn, J. N., and M. J. Daly, 2005   Genome-wide association studies for common diseases and complex traits. Nat. Rev. Genet. 6: 95–108.

Keeling, M. J., 1999   The effects of local spatial structure on epidemiological invasions. Biol. Sci. 226(1421): 859–867.

Keeling, M. J., and K. T. D. Eames, 2005   Networks and epidemic models. J. R. Soc. Int. 2: 295–307.

Levin, S. A., and R. Durrett, 1996   From individuals to epidemics. Philos. Trans. R. Soc. B 351: 1615–1621.

Lipschutz-Powell, D., J. A. Woolliams, P. Bijma, and A. B. Doeschl-Wilson, 2010   The importance of associative effects in the control of infectious disease through selection. Proceedings of the 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany, article 0491, p. 142.

Lohmueller, K. E., C. L. Pearce, M. Pike, E. S. Lander, and J. N. Hirschhorn, 2003   Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat. Genet. 33: 177.

Lloyd, A. L., J. Zhang, and A. M. Root, 2007   Stochasticity and heterogeneity in host-vector models. J. R. Soc. Int. 4: 851–863.

Lloyd-Smith, J. O., S. J. Schreiber, P. E. Kopp, and W. M. Getz, 2005   Superspreading and the effect of individual variation on disease emergence. Nature 438: 355–359.

Matthews, L. J., C. Low, D. L. Gally, M. C. Pearce, D. J. Mellor, *et al.*, 2006   Heterogeneous shedding of Escherichia coli O157 in cattle and its implications for control. Proc. Natl Acad. Sci. USA 103: 547–552.

May, R. M., and R. M. Anderson, 1988   The transmission dynamics of human immunodeficiency virus (HIV). Philos. Trans. R. Soc. B 321: 565–607.

Nath, M., J. A. Woolliams, and S. C. Bishop, 2008   Assessment of the dynamics of microparasite infections in genetically homogeneous and heterogeneous populations using a stochastic epidemic model. J. Anim. Sci. 86: 1747–1757.

Nicholas, F. W., 2005   Animal breeding and disease. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360(1459): 1529–1536.

Nieuwhof, G. J., J. Conington, L. Bűnger, W. Haresign, and S. C. Bishop, 2008   Genetic and phenotypic aspects of footrot lesion scores in sheep of different breeds and ages. Animal 2: 1289–1296.

Nieuwhof, G. J., J. Conington, and S. C. Bishop, 2009   A genetic epidemiological model to describe resistance to an endemic bacterial disease in livestock: application to footrot in sheep. Genet. Sel. Evol. 41: 1–12.

Raadsma, H. W., J. R. Egerton, D. Wood, D. Kristo, and F. W. Nicholas, 1994   Disease resistance in Merino sheep III. Genetic variation in resistance to footrot following challenge and subsequent vaccination with an homologous rDNA pilus vaccine under both induced and natural conditions. J. Anim. Breed. Genet. 111: 367–390.

Raadsma, H. W., G. A. Attard, F. W. Nicholas, and J. R. Egerton, 1995   Disease resistance in Merino sheep IV. Genetic variation in immunological responsiveness to fimbrial *Dichelobacter nodosus* antigens, and its relationship with resistance to footrot. J. Anim. Breed. Genet. 112: 349–372.

Renshaw, E., 1991   *Modelling Biological Populations in Space and Time*. Cambridge University Press, Cambridge, UK.

Robertson, A., and I. M. Lerner, 1949   The heritability of all-or-none traits: viability of poultry. Genetics 34: 395–411.

Segal, S., and A. V. S. Hill, 2003   Genetic susceptibility to infectious disease. Trends Microbiol. 11(9): 445–448.

Shen, Z., F. Ning, W. Zhou, X. He, C. Lin, *et al.*, 2004   Superspreading *SARS* events, Beijing, 2003. Emerg. Infect. Dis. 10: 256–260.

Sinclair, A. N., 1957   Studies on contagious footrot of sheep. Aust. Vet. J. 33: 202–206.

Springbett, A. J., K. Mackenzie, J. A. Woolliams, and S. C. Bishop, 2003   The contribution of genetic diversity to the spread of infectious diseases in livestock populations. Genetics 165: 1465–1474.

Yates, A., R. Antia, and R. R. Regoes, 2006   How do pathogen evolution and host heterogeneity interact in disease emergence? Proc. R. Soc. B 273: 3075–3083.

Woolhouse, M. E., C. Dye, J. F. Etard, T. Smith, J. D. Charlwood, *et al.*, 1997   Heterogeneities in the transmission of infectious agents: implications for the design of control programs. Proc. Natl. Acad. Sci. USA 94: 338–342.

Woolhouse, M. E., J. F. Etard, K. Dietz, P. D. Ndhlovu, and S. K. Chandiwana, 1998   Heterogeneities in *schistome* transmission dynamics and control. Parasitology 117: 475–482.

Communicating editor: D. W. Threadgill

GENETICS

Implications of Host Genetic Variation on the Risk and Prevalence of Infectious Diseases Transmitted Through the Environment

Andrea B. Doeschl-Wilson, R. Davidson, J. Conington, T. Roughsedge, M. R. Hutchings, and B. Villanueva

File S1

Supporting Data

File S1 is available for download as an Excel file at http://www.genetics.org/cgi/content/full/genetics.110.125625/DC1.