JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12977

Genotype-by-genotype interactions between an insect and its pathogen

A. I. HUDSON, A. E. FLEMING-DAVIES, D. J. PÁEZ & G. DWYER

Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA

Keywords:

baculovirus;
genotype × genotype interactions;
gypsy moth (*Lymantria dispar*);
host–pathogen specificity;
LdNPV.

Abstract

Genotype-by-genotype (G×G) interactions are an essential requirement for the coevolution of hosts and parasites, but have only been documented in a small number of animal model systems. G×G effects arise from interactions between host and pathogen genotypes, such that some pathogen strains are more infectious in certain hosts and some hosts are more susceptible to certain pathogen strains. We tested for G×G interactions in the gypsy moth (*Lymantria dispar*) and its baculovirus. We infected 21 full-sib families of gypsy moths with each of 16 isolates of baculovirus and measured the between-isolate correlations of infection rate across host families for all pairwise combinations of isolates. Mean infectiousness varied among isolates and disease susceptibility varied among host families. Between-isolate correlations of infection rate were generally less than one, indicating nonadditive effects of host and pathogen type consistent with G×G interactions. Our results support the presence of G×G effects in the gypsy moth–baculovirus interaction and provide empirical evidence that correlations in infection rates between field-collected isolates are consistent with values that mathematical models have previously shown to increase the likelihood of pathogen polymorphism.

Introduction

Genotype-by-genotype (G×G) interactions play an important role in the coevolution of hosts and parasites, through effects on epidemiologically important traits such as disease transmission and resistance (Lambrechts *et al.*, 2006). These interactions both affect and are affected by coevolution between host and pathogen (e.g. Scanlan *et al.*, 2011, 2015). Ecological theory suggests that G×G interactions may also affect life history traits such as host lifespan (Kirchner & Roy, 2000). Studies of the quantitative genetics of G×G interactions, however, are logistically challenging, because they require measurements of both host and pathogen, and so there are few examples in the literature on animal pathogens. As a consequence, G×G interactions in animal host–pathogen interactions have been quantified only for a small number of well-known systems such as *Anopheles* mosquitos and malaria (Lambrechts *et al.*, 2005), or *Daphnia magna* and its bacterial pathogens (Carius *et al.*, 2001).

Extensive work has also been carried out using gene-for-gene systems in plant–pathogen interactions, particularly in the model host plant *Arabidopsis thaliana* (Hammond-Kosack & Jones, 1997). In gene-for-gene systems, resistance occurs when a plant's resistance allele at a particular locus is expressed after the expression of a pathogen's virulence allele at a particular locus (Hammond-Kosack & Jones, 1997). Genetic polymorphism in pathogen virulence and host resistance can, however, result in gene-for-gene interactions that in principal produce pathogens able to infect a wide range of host genotypes, or even pathogens that are universally infective (Frank, 1994). Thus, gene-for-gene interactions allow for generalist pathogens as well as specialists, and so can allow a range of levels of host–pathogen specificity (Dybahl and Storfer 2003). Notably, G×G interactions may result from interactions between many genes and do not require specific gene-for-gene coevolution. This is important because multilocus gene-for-gene models allow for asymmetrical G×G

Correspondence: Asher I. Hudson, Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, USA.
Tel.: +530 752 3670; fax: +530 752 1537;
e-mail: aihudson@ucdavis.edu

interactions (Sasaki, 2000), such that pathogen strains may arise that are universally infectious.

In contrast to gene-for-gene models, matching-allele models require a higher degree of host–pathogen specificity. In these models, each virulence allele in a pathogen produces receptor proteins that match cell-surface proteins produced by a susceptibility allele in a host (Frank, 1994). In the strictest version of a matching-allele model, each strain of the pathogen is strictly specific to particular strains of the host, and so pathogens cannot be generalists (Agrawal & Lively, 2002).

The genetics of host–pathogen interactions are of course not limited to gene-for-gene and matching-allele models. In particular, the literature also includes inverse gene-for-gene models, in which infectivity is triggered by a host susceptibility allele (Fenton *et al.*, 2009), and inverse matching-allele models, in which resistance occurs when the proteins produced by a host allele match those produced by a pathogen allele (Nuismer 2006). Each of these different genetic architectures leads to different predictions about coevolutionary trajectories, as do models that combine different architectures (Agrawal and Lively 2002). A full understanding of the dynamics of host–pathogen coevolution thus requires a mechanistic understanding of the genetic basis of resistance and virulence in a range of host–pathogen systems. An important first step in achieving such an understanding is to quantify G×G interactions, which is the approach taken here.

Quantification of G×G effects is thus essential for understanding coevolution in host–pathogen systems. In particular, Heath & Nuismer (2014) used simulations to show that statistical evidence for G×G interactions almost certainly indicates that a host–pathogen interaction is driven by interactions between host and pathogen genetics. Several studies have therefore found that both pathogen infectiousness and host susceptibility are genetically based (Hill *et al.*, 1991; Schmid-Hempel & Koella, 1994; Ebert & Mangin, 1997; Zhan *et al.*, 2002), but evidence is limited, particularly in insect–pathogen systems such as the one that is the focus of our study.

Part of the problem is that measuring G×G effects in nonmodel organisms is usually logistically difficult, often because of a lack of knowledge of the underlying genes in the system, or because the interaction involves many genes. These obstacles are practically important because the traditional method of detecting G×G effects within a quantitative genetics approach is to implement a mixed-effects model (Falconer and Mackay 1996). Detecting weak G×G effects using these models is difficult because when both the pathogen and the host have many genotypes, very large sample sizes are needed to achieve the statistical power necessary to make inferences. Adding to this difficulty, the very coevolutionary cycles that are often the focus of interest can produce allele frequencies that make detection of statistical G×G using mixed-effects linear models unlikely (Heath & Nuismer, 2014).

In this study, we evaluated the potential for G×G interactions between the gypsy moth and its baculovirus, using two quantitative genetics methods. First, we used the standard approach of fitting linear mixed-effects models to data from laboratory experiments with full-sibling gypsy moth families and field-collected baculovirus isolates. Second, we used a novel method of estimating correlations in infection rates between pairs of virus isolates. The latter method was more successful at detecting weak G×G effects.

Our results show evidence of G×G interactions: some baculovirus isolates were more infectious in particular gypsy moth full-sib families, and the infection risk for a given host with a given virus isolate was not entirely predicted by that isolate's mean infectiousness across families or by the host family's overall susceptibility. We also found evidence of positive correlations in infectivity between virus isolates, providing evidence for levels of host–pathogen specificity in the gypsy moth–baculovirus system that appear to be consistent with gene-for-gene models.

Materials and methods

Study organisms

As in many insect baculoviruses, the gypsy moth virus is transmitted when larvae accidentally consume infectious particles or 'occlusion bodies' while feeding on the foliage of host plants (Cory & Myers, 2003). Infected larvae typically die 6–21 days later, releasing large numbers of occlusion bodies onto the foliage, completing the cycle of transmission in nature. Because infection occurs through virus ingestion, larvae can be easily fed specific doses of the virus in controlled experiments, as described below.

Gypsy moth rearing protocol and virus exposure

To ensure host variation, we used 21 full-sibling gypsy moth families that were produced by mating unrelated adult males and females, following standard mating protocols described in Páez *et al.*, 2015. The adults were raised from egg masses collected in Roscommon, Michigan, whereas virus isolates were collected from six sites throughout Michigan in 2000 (Fleming-Davies *et al.*, 2015). The 16 virus isolates used here are genetically distinct and exhibit distinct infection patterns in the wild (Kennedy, 2012). All egg masses were first surface-sterilized in a 4% formalin solution, which eliminates external contamination (Dwyer & Elkinton, 1995). Larvae were then raised in the laboratory on approximately 100 mL of artificial diet in plastic cups under constant density conditions, at a constant 25 °C,

a 14-h : 10-h light to dark cycle and 60% humidity, following standard procedures (Dwyer & Elkinton, 1995).

We then exposed larvae to the virus while controlling for both larval developmental stage and the identity of the virus isolate. All larvae were exposed to the virus while in the 4th larval stage or 'instar', to minimize differences in susceptibility associated with developmental stage (Hoover *et al.*, 2002). To ensure developmental synchrony, we used insects that had reached the fourth instar within a 72-h period.

Because our aim was to detect G×G interactions between insect families and virus isolates, each larva received the same dose of virus, so that infection rates would be determined by G×G interactions and not by differences in dose. Larvae were starved for 24 h before virus exposure and then provided with 3 mm^3 cubes of artificial diet contaminated with a water solution of 3 μL containing 900 occlusion bodies of a particular virus isolate. To ensure that all individuals received the intended dose, larvae that failed to consume the entire cube after 24 h were discarded. Larvae from each full-sib family were allocated to 16 groups, each corresponding to a different virus isolate.

We exposed groups of 12 individuals from each gypsy moth family to each of the 16 virus isolates. After accounting for individuals that did not eat the entire dose, the mean number of insects per host family and per virus isolate was 7.60 ± 3.54. Every exposure combination (isolate × family) was represented, but two host–pathogen combinations included only a single individual host. Sample sizes for each pairwise combination of host and isolate are listed in Table S1. Following exposure, larvae were reared separately on virus-free diet until death or pupation.

Infected larvae usually display a distinctive 'melted' phenotype as their cuticle dissolves, due to the action of viral enzymes, and so virus infection was first determined visually (Cory & Myers, 2003). Whenever the cause of death was uncertain, we necropsied insects under a light microscope at 400×, as occlusion bodies are easily visible at this magnification. In the few cases in which necropsies were inconclusive, we discarded the individual larvae from our data. Controls consisted of larvae that were fed cubes of artificial diet with 3 μL of water without virus. Of 12 control larvae selected haphazardly from different families, none died of baculovirus infection and 11 survived to pupation.

Statistical methods

We first tested for G×G interactions by using a generalized random-effects model with host fate (survived or died) as the binomial response variable and with the host's full-sib family, the pathogen isolate and the interaction between full-sib family and pathogen isolate as random effects. G×G interactions can only occur if the variation in the fraction of infected insects is explained by variation across both insect families and pathogen isolates. We therefore also used this model to estimate the extent to which the variation in the fraction of infected insects was due to variation across host families and pathogen isolates.

In the generalized random-effects model, host mortality was assumed to follow a binomial error distribution with a logit link function. Since the residual variance associated with logit probabilities of binary data cannot be observed, we followed standard practice in fixing the residual variance of the logit probability at 1 (Hadfield, 2010), and we then estimated the total variance in the fraction infected as $V_T = V_F + V_I + 1 + \pi^2/3$, where V_F and V_I are the variances due to host family and pathogen isolate, respectively, 1 is the fixed residual variation and $\pi^2/3$ is the variance associated with a logistic distribution. When appropriate we calculated broad-sense heritabilities of host susceptibility as $H = 2 \times V_F/V_T$ (Falconer and Mackay 1996).

In addition to testing for G×G interactions using random-effects modelling, we also examined the between-isolate correlation of infection. A correlation of 1 between isolates would indicate that the fraction of infected insects in a given host family exposed to a given pathogen isolate can be predicted simply by knowing, for insects in the same family, the fraction that became infected when exposed to a different pathogen isolate. Such cases would exclude the possibility of a G×G interaction, because infection rates within a host family would be effectively the same across all pathogen isolates. Correlations < 1 would in contrast indicate G×G interactions, because they could not be predicted by host genotype alone, instead depending on the pathogen isolate as well. In addition, a negative correlation would indicate that if a particular host family is relatively resistant to a given pathogen isolate, then that family would instead be relatively susceptible to a second pathogen isolate. A negative correlation would therefore suggest ecological specialization of the pathogen to particular host genotypes, and so resistance to a given pathogen isolate would have a cost, in the form of an increased susceptibility to other isolates.

To calculate correlations in infection rates, we used Markov chain Monte Carlo methods and the MCMCglmm package in R (Hadfield, 2010). We treated rates of host survival when exposed to different pathogen isolates as different host traits and then used a bivariate mixed-effects model to test for correlations between these traits. Thus, in our mixed models, host survival for a given pair of pathogens consists of a pair of values, which vary as a function of a global fixed intercept and a random host family effect. In these bivariate models, the random effects are described by a 2 × 2 matrix with the variances in survival probability

across families on the diagonal elements and the covariances on the off-diagonal elements. We then calculated correlations as the ratio between the covariance and the square root of the product of the family variances. We estimated correlation values and 95% credible intervals from the posterior distribution of the estimates.

In the absence of G×G effects, true between-isolate correlations would be equal to one, and the effects of isolate and family would be additive. Observed between-isolate correlations, however, are not expected to be 1 even if the true correlation is 1, due to the errors inherent in a binomial process in a limited sample. Likewise, even if the true correlation were −1, we would not expect the observed between-isolate correlations to be exactly −1.

Accordingly, to determine the probability that the observed correlations were < 1 and > −1, we compared the observed distribution of between-isolate correlations to two different null distributions of correlations, based on simulations that assumed either no G×G effects (correlation = 1) or strong G×G effects (correlation = −1). To create the first null distribution, for each pair of isolates, we assumed that isolate and family effects were additive, and we used the observed data from isolate 1 to predict the expected probability of infection per family for isolate 2. We then drew 10 000 random samples from a binomial distribution, using the expected probability of infection under these assumptions and the actual experimental sample sizes per family for isolate 2. Next, we computed the between-isolate correlation using each of those samples for isolate 2 and the observed data for isolate 1. This gave 10 000 bootstrapped correlations. We then repeated this procedure starting with the data for isolate 2 and predicting the infection outcomes for isolate 1 under the assumption of no G×G effects, and additive effects of isolate and family. We thus calculated bootstrapped correlations for all combinations of isolates, and we pooled the resulting correlations to create a null distribution of the expected between-isolate correlations for all isolates under the assumption of no G×G effects (correlation = 1).

To create the second null distribution, which assumes strong G×G effects (correlation = −1), we first used the observed data from isolate 1 to predict the expected probability of infection per family. We then normalized the linear model coefficients for each model by subtracting the mean, and we predicted the probability of infection for each family for isolate 2, by using the normalized coefficients in an inverse normal distribution. A host family that was more susceptible than average to isolate 1 was thus assumed to be more resistant than average to isolate 2 by an equal amount, consistent with a perfect negative correlation between isolates 1 and 2. We then followed the same steps as for the first null distribution to create a distribution of correlations under the assumption of strong G×G effects (correlation = −1).

We then compared our observed distribution of between-isolate correlations to the two null distributions to determine whether the observed correlations were consistent with a between-isolate correlation of one, the expectation in the absence of G×G effects, and conversely, whether they were consistent with a between-isolate correlation of negative one, the expectation in the presence of strong G×G effects. Note that data from each virus isolate were used multiple times in the pairwise analysis, and so the pairwise correlations are not independent and were therefore not used in statistical tests.

Results

Our results support G×G effects in this host–pathogen interaction. First, our linear random-effects model of infection risk suggests that there is a weak but statistically significant interaction between host family and pathogen isolate. Second, correlations in infection rates between different isolates showed that host and pathogen effects on infection are not merely additive, as would be expected in the absence of G×G. We summarize the results of these analyses in the following paragraphs.

In total, we exposed 2556 insects in 21 full-sib families to the 16 virus isolates, which resulted in an overall 48% infection rate. The fraction infected varied significantly among host families (Fig. 1a, ranging from 0.129 ± 0.0814 to 0.677 ± 0.122) and among virus isolates (Fig. 1b, ranging from 0.169 ± 0.0560 to 0.716 ± 0.0687). Of the variation in the fraction of infected insects, 9.1% was due to variation across full-sib families, with a broad-sense heritability equal to 0.18 (CI = 0.08, 0.35). This estimate is close to an earlier estimate based on a different natural population of gypsy moths (Páez *et al.*, 2015). Similarly, we found that 13% (CI = 6, 26%) of the variance in the fraction of infected insects was due to variation across pathogen isolates. Variation in the fraction of infected individuals is thus explained both by genetic effects arising from host susceptibility and by differences in infectiousness among pathogen isolates.

Visual inspection of the raw infection data suggests that there are at least small G×G effects in this host–pathogen interaction (Fig. 2). For example, the most infectious isolate (the bottom row in Fig. 2) exhibited high infectiousness for most host families but failed to infect many individuals from some otherwise highly susceptible families (e.g. the 8th column from the right in Fig. 2). No isolate had a constant infection rate across all families (Fig. 3), and most isolates varied considerably across host families. Visual inspection further reveals a strong effect of host family within some but not all isolates (Fig. 4).

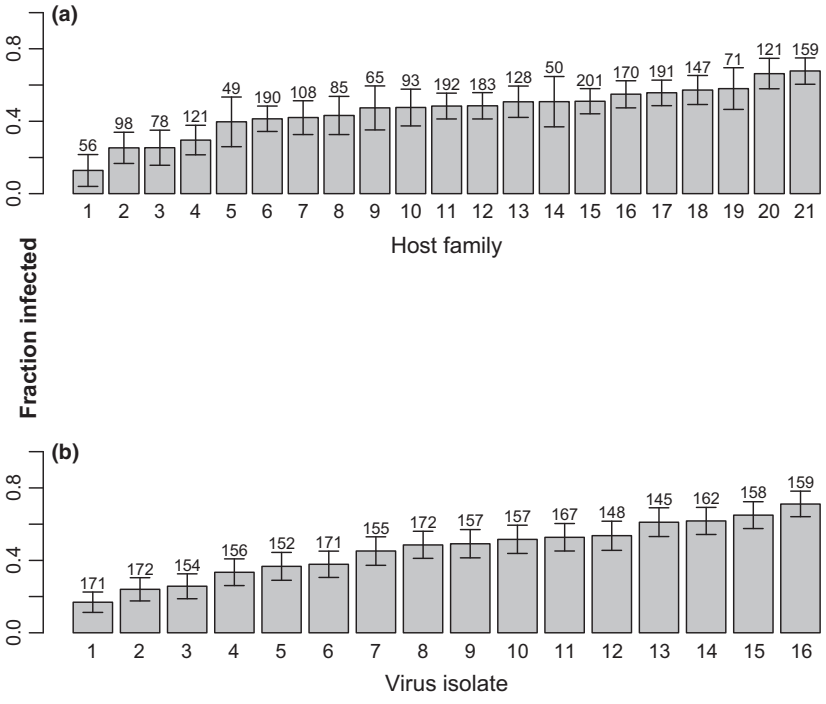

Fig. 1 Fraction infected varies by host family (a) and virus isolate (b). Bars in both panels are means ± 2SE fraction infected, and numbers above each bar give the sample sizes. Fraction infected per full-sib host family (a) is averaged across virus isolates for each family, and families are ordered from the least susceptible to the most susceptible. Fraction infected per virus isolate (b) is averaged across host families for each isolate, and isolates are ordered from least infectious to most infectious.

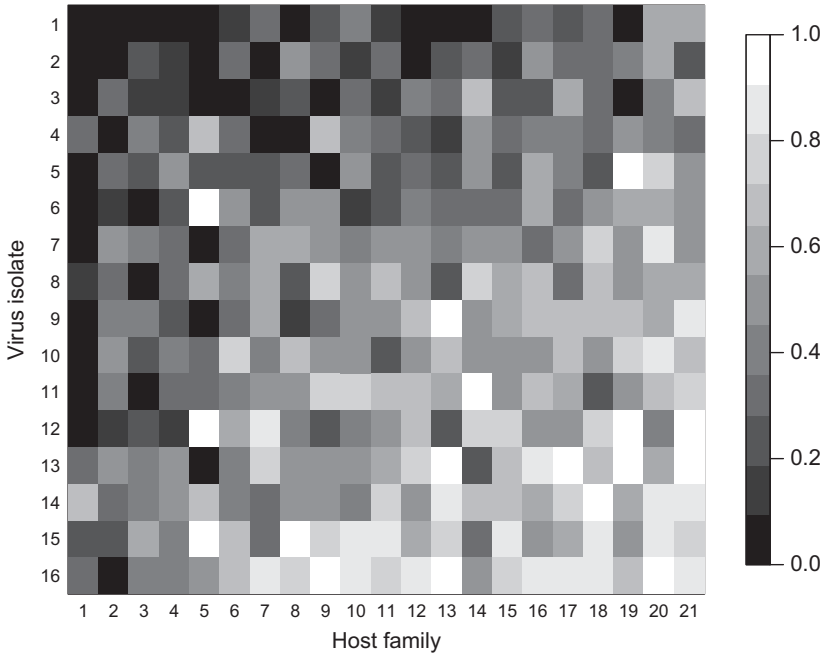

Fig. 2 Heat map of the fraction infected for each combination of host family and virus isolate. The greater the fraction infected, the brighter the square. The virus isolates are ranked from lowest fraction infected on the top to highest fraction infected on the bottom and the host families are ranked from least susceptible on the left to most susceptible on the right.

Correlations of infection rates between pairs of isolates further support the presence of G×G effects. Specifically, correlation in the fraction infected between pairs of isolates varied from −0.17 to 0.80 (Table 1, posterior medians from MCMC chains of pairwise bivariate linear mixed-effects model). Of the six pairs of isolates for which the median correlation was negative, in all cases the 95% HPD overlapped zero, indicating no statistically significant difference from zero (Table 1). Of the remaining 114 pairwise comparisons for which the median correlation was positive, 13 pairs of isolates had 95% HPD intervals that were entirely positive, whereas the remaining comparisons had intervals that included zero, and thus

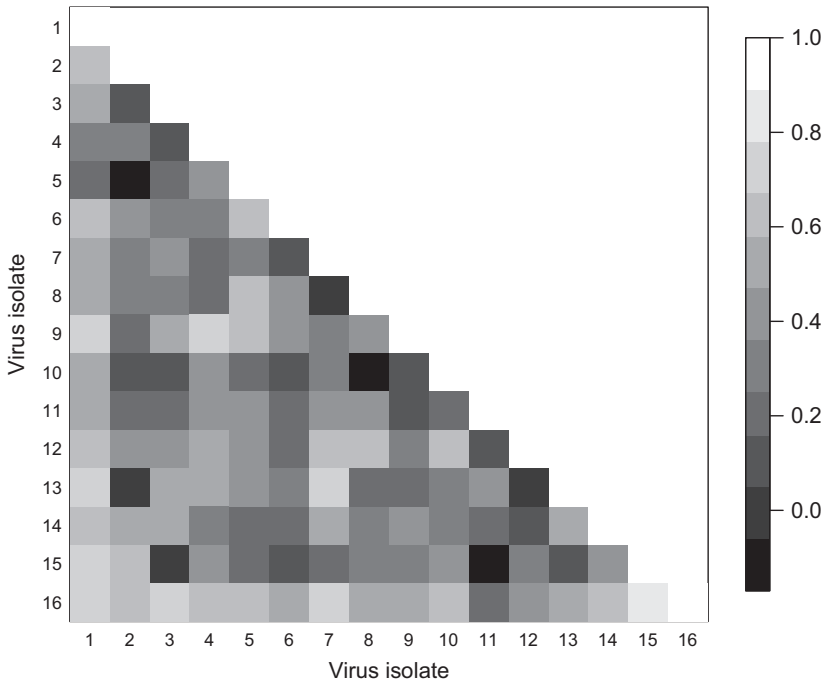

Fig. 3 A heat map showing the correlations between pairs of isolates. The isolates are ordered by fraction infected, with the most infectious isolates on the bottom and on the right. Brighter shades indicate higher between-isolate correlations, with the lowest observed correlation (−0.17) in black and the highest observed correlation (0.80) in white.

cannot be concluded to be different from zero (Table 1).

Overall, correlations between pairs of virus strains appeared to be normally distributed, with a mean of 0.331 and a standard deviation of 0.201 (Fig. S1). This result suggests that the pairwise correlations between isolates tend to be positive, but not close to 1. Comparison with the first bootstrapped null distribution, which assumes that between-isolate correlations are equal to 1, confirms this result, in that most of the observed correlations fall outside of 95% of the simulated values of a null distribution that assumes no G×G effects (Fig. 5). This suggests that most between-isolate correlations are significantly < 1, further supporting the presence of G×G interactions. However, comparison with the second bootstrapped distribution, which instead assumes that between-isolate correlations are equal to -1, suggests that our results are also inconsistent with strong G×G effects, because most of the observed correlations also fall outside of 95% of the simulated values when strong G×G effects are assumed (Fig. 5). Our results thus support the presence of weak G×G interactions in this system.

Discussion

Our results show that correlations in infectivity between some gypsy baculovirus isolates are almost certainly positive but < 1, consistent with a gene-for-gene model, but inconsistent with a matching-allele model (Sasaki, 2000; Flores *et al.*, 2011). As emphasized by Heath & Nuismer (2014), data from cross-infection experiments are not, on their own, sufficient to allow us to distinguish between the gene-for-gene and matching-allele models. We nevertheless note that multilocus gene-for-gene models of pathogens often exhibit universal infectivity with small differences in host range (Frank, 1994; Sasaki, 2000), whereas in matching-allele models, pathogens often have high specificity, with some pathogen genotypes completely failing to infect some host genotypes (Frank, 1994). Our data therefore appear to provide more support for gene-for-gene models than for matching-allele models, but further research is clearly needed.

Our work also has several implications for the maintenance of variability in the gypsy moth–baculovirus interaction. First, previous work on the gypsy moth–baculovirus interaction showed that trade-offs between pathogen fitness components can allow pathogen coexistence (Fleming-Davies *et al.*, 2015), whereas trade-offs between host fitness components likely maintain host variability (Páez *et al.*, 2015). As part of the former work, however, the authors also showed that, although perfect correlations between pathogen infectiousness to different hosts are not inconsistent with pathogen coexistence, coexistence is more likely if correlations are < 1 (Fleming-Davies *et al.*, 2015). Our results thus provide important support for the hypothesis that a combination of trade-offs and imperfect correlations in infectiousness promote coexistence of baculovirus strains. In addition, the mechanism for two-pathogen coexistence in this population model (Fleming-Davies *et al.*, 2015)

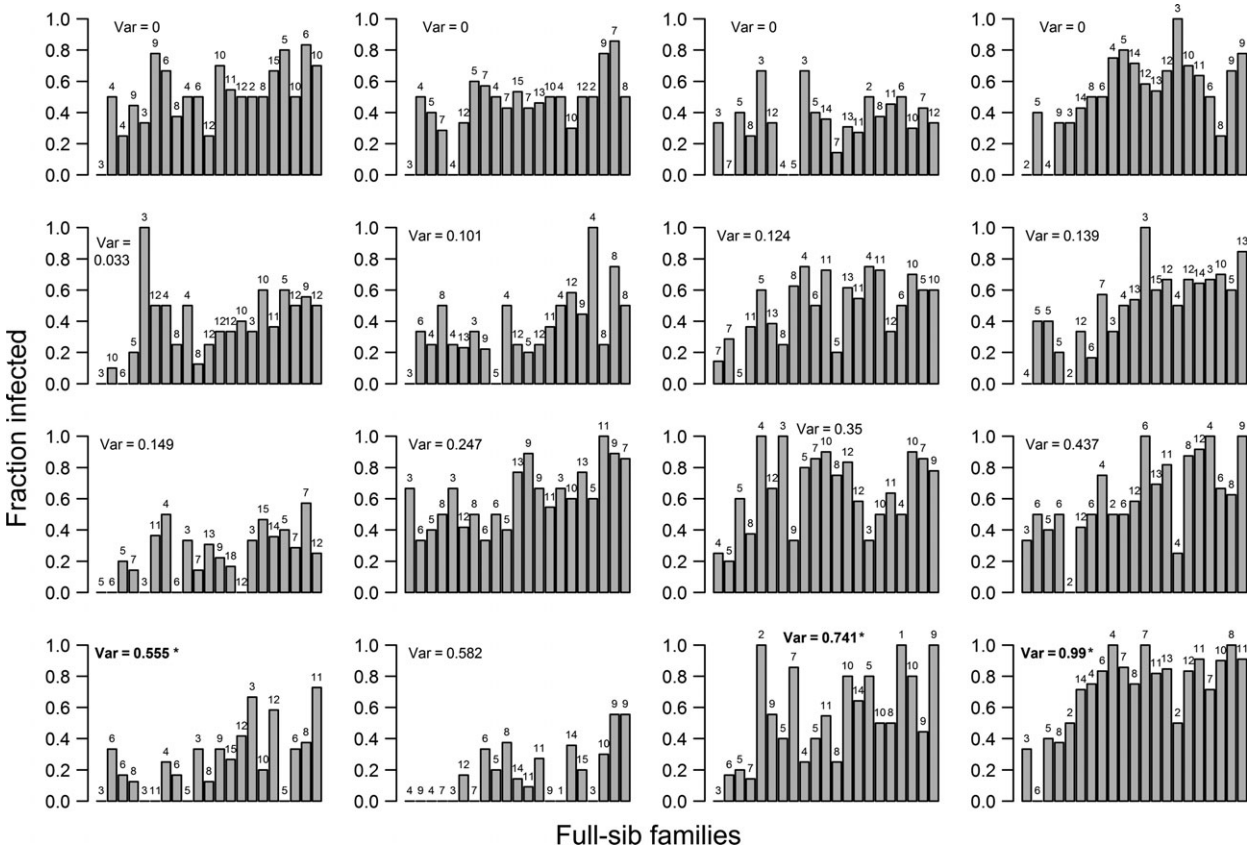

Fig. 4 Variation by host family within virus isolates. Each plot is a different isolate, with plots ordered left to right and top to bottom with increasing mean fraction infected across host families. The variance due to host families is given for each isolate (calculated from fitting a random-effects linear model to the data for that isolate). Bold font and an asterisk denote statistically significant among-family variation in infection rate for that isolate, as inferred from model comparisons using AIC. Host families are ordered by their mean susceptibility averaged across isolates, and the order of host families is thus the same for each subplot. Numbers above bars give sample sizes for that family and isolate.

would apply to any system with density-dependent quantitative variation in host resistance, and thus the result that imperfect correlations in infectiousness might promote pathogen polymorphism is also likely to apply to other systems.

Alternatively, variation in the gypsy moth–baculovirus system may be due to local adaptation, in which host and pathogen types are structured geographically and adapt to each other. This mechanism is widely acknowledged to maintain host and pathogen diversity in many systems (Hammond-Kosack & Jones, 1997; Kawecki & Ebert, 2004). In the gypsy moth–baculovirus system, however, previous studies have found no evidence of geographic structure in host or pathogen (Elderd *et al.*, 2008; Fleming-Davies *et al.*, 2015), likely due to high migration rates of hosts. In addition, we found correlations between pathogen isolates that suggest gene-for-gene interactions, and local adaptation is unlikely in simulations using gene-for-gene models that include migration, unless there are also repeated colonizations and extinctions (Lively, 1999). Our results thus do not support the presence of local adaptation in the gypsy moth–baculovirus system, but further research into the population structure of this system could yield a more definitive conclusion.

Our data are also relevant to a third potential mechanism that could maintain genetic diversity in this system, namely frequency-dependent selection. The host–pathogen specificity resulting from G×G interactions could allow rare host or pathogen genotypes to have a selective advantage, leading to frequency-dependent selection (Hamilton, 1980; Carius *et al.*, 2001). This effect is especially likely to be important when no single family of host is most resistant to all isolates of the pathogen and when no single isolate of the pathogen is most infectious in all families of host, although this condition is not necessary for frequency-dependent selection (Sasaki, 2000; Carius *et al.*, 2001; Fenton *et al.*, 2009). Although we do not have direct evidence of frequency-dependent selection, our results show that

Table 1 Correlations of fraction infected between isolates. Values reported are posterior medians. Correlations in bold have 95% CIs entirely positive.

1.0																1	
0.8	1															2	
0.63	0.38	1														3	
0.5	0.1	0.5	1													4	
0.38	0.27	0.14	0	1												5	
0.19	−0.16	0.17	0.41	0.13	1											6	
0.63	0.4	0.34	0.3	0.59	0.19	1										7	
0.51	0.29	0.37	0.2	0.31	0.11	0.14	1									8	
0.55	0.35	**0.3**	0.19	0.65	0.46	−0.07	0.41	1								9	
0.69	0.19	0.5	0.75	0.63	0.42	0.27	0.26	−0.06	1							10	
0.5	0.12	0.16	0.36	0.18	0.15	0.14	0.42	0.44	0.1	1						11	
0.58	0.19	0.17	0.39	0.4	0.38	0.18	**0.62**	**0.64**	0.3	0.6	1					12	
0.64	0.41	**0.36**	0.49	0.5	0.41	0.36	0.75	**0.2**	0.17	0.3	0.39	1				13	
0.71	0.02	0.49	0.47	0.37	0.18	0.14	0.51	0.3	0.38	0.34	0.17	0.14	1			14	
0.63	0.58	0.57	0.03	0.42	0.2	0.14	0.19	0.36	0.32	0.39	−0.17	0.29	0.09	1		15	
0.7	0.69	0.64	0.7	0.64	**0.55**	0.51	0.69	**0.54**	0.49	0.62	**0.16**	**0.35**	**0.49**	**0.62**	1	16	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Virus isolate	

there is no single most resistant host family or single most infectious pathogen isolate. This lack of universal infectivity suggests that rare genotypes may have an advantage, thus increasing the likelihood that negative frequency-dependent selection is present in this system. In addition, the prevalence of different pathogen genotypes might affect the dose of a particular isolate consumed by the host, as more abundant genotypes would be expected to be present at higher densities on leaves. Because in our experiments, we used only one dose, further work is needed to determine the effect of dose on G×G interactions in this system.

Finally, our data suggest that there is a strong potential for coevolution to occur in this system, and models have shown that coevolution can help maintain variation in both host and pathogen (Sasaki & Godfray, 1999; Tellier and Brown 2007; Penman *et al.*, 2013; Boots *et al.*, 2014). Consistent with previous studies (Páez *et al.*, 2015), we found that gypsy moth variability is heritable in the broad sense, a necessary requirement for coevolutionary processes to maintain that variation. Moreover, host–pathogen and host–parasite models that include G×G interactions can display high levels of variation in susceptibility and resistance alleles (Gillespie & Turelli, 1989; Frank, 1994; Sasaki, 2000; Agrawal & Lively, 2002). Our expectation is therefore that the G×G interactions that we have documented in the gypsy moth–baculovirus system likely enhance the effects of coevolution in maintaining variation in the gypsy moth and its baculovirus. Understanding such effects, however, requires further theoretical and experimental work.

Intriguingly, only very small G×G effects were detected using a random-effects model, yet these G×G effects apparently lead to differences in between-isolate correlations, which in turn can have strong effects on the population dynamics in this system (Fleming-Davies *et al.*, 2015). Despite their importance to coevolutionary systems, empirical demonstrations of G×G effects are challenging; theoretical work has shown that even in systems experiencing strong coevolution, G×G effects may be difficult to directly measure (Heath & Nuismer, 2014). There is thus a great need for alternative analytical methods of detecting G×G interactions in host–pathogen interactions (Flores *et al.*, 2011). Our results suggest that calculating between-group correlations may be a particularly sensitive method for detecting small G×G effects, which may nevertheless have large ecological and evolutionary consequences. Our method is therefore worth considering alongside traditional random-effects models in future studies. This conclusion further suggests that an important next step in understanding host–pathogen coevolution is to develop robust predictions for between-group correlations under different coevolutionary scenarios such as matching-allele models and gene-for-gene models.

Our results thus show that G×G interactions are present in this system, which may help maintain variability

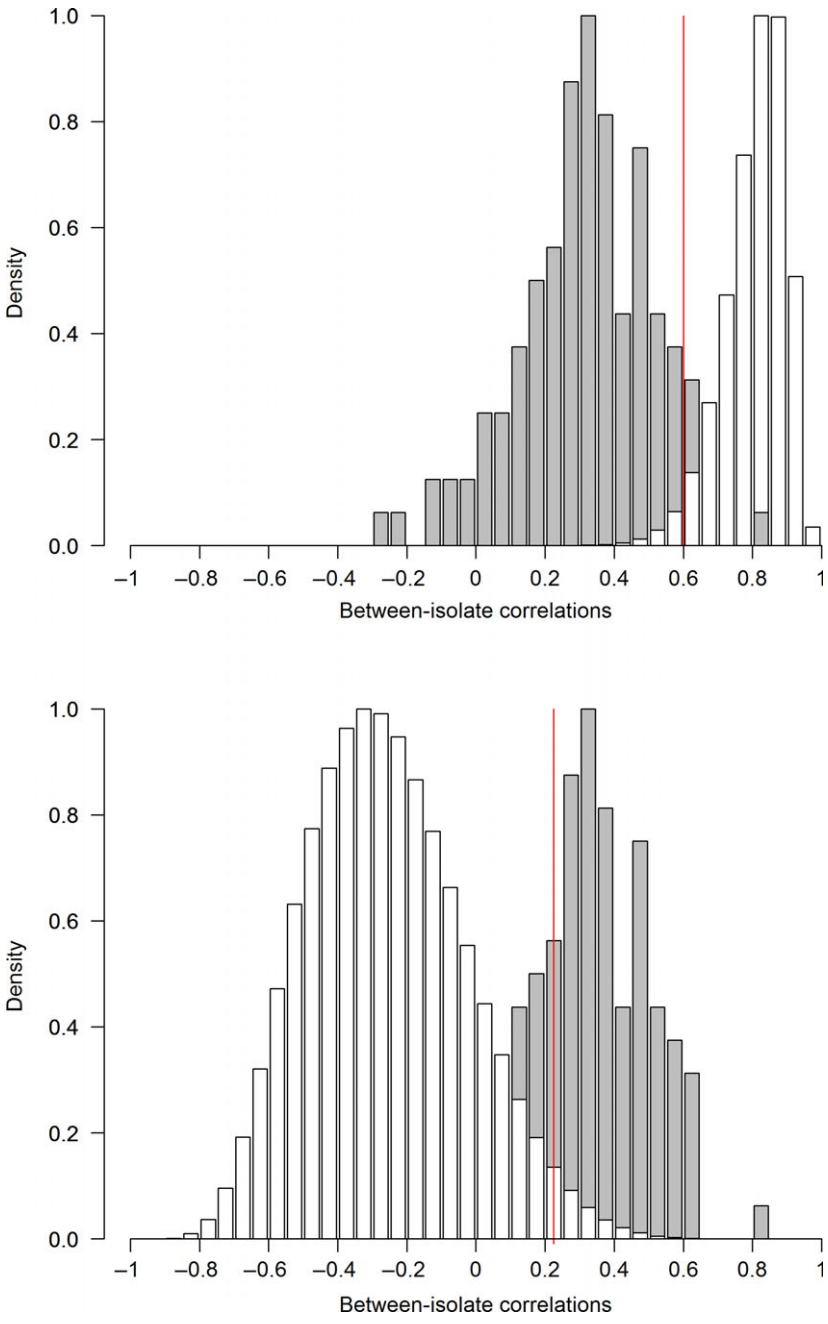

Fig. 5 Comparison of observed between-isolate correlations to two bootstrapped null distributions. In both graphs, the shaded bars show observed correlations, and the clear bars show the null distribution. In the upper graph, the null is based on the assumption that the correlation between isolates is +1, and the red line gives the value above which lie 95% of the simulated correlations. In the lower graph, the null is based on the assumption that the correlation between isolates is −1, and the red line gives the value below which lie 95% of the simulated correlations. For clarity, all distributions are normalized to a maximum density of 1.

in hosts and pathogens. In addition, our work provides an important counter-example to much of the invertebrate host–pathogen literature (e.g. Hufbauer & Via, 1999; Lively, 1989; Schönrogge *et al.*, 2006; but see Altizer, 2001; Prugnolle *et al.* 2006), which emphasizes local coadaptation driven by negative correlations in infectiousness between pathogen isolates, as in matching-allele models (Hamilton *et al.*, 1990; Frank, 1993; Agrawal & Lively, 2002). Although such negative correlations are typically invoked to explain pathogen coexistence, the positive correlations that we observed are not inconsistent with coexistence (Fleming-Davies *et al.*, 2015). Our work thus shows the importance of studying a wide range of systems to fully understand the diversity of coevolutionary effects that can be produced by variation in host–pathogen specificity.

References

Agrawal, A. & Lively, C.M. 2002. Infection genetics: gene-for-gene versus matching-alleles models and all points in between. *Evol. Ecol. Res.* **4**: 79–90.

Altizer, S.M. 2001. Migratory behaviour and host–parasite coevolution in natural populations of monarch butterflies infected with a protozoan parasite. *Evol. Ecol. Res.* **3**: 567–581.

Boots, M., White, A., Best, A. & Bowers, R. 2014. How specificity and epidemiology drive the coevolution of static trait diversity in hosts and parasites. *Evolution* **68**: 1594–1606.

Carius, H.J., Little, T.J. & Ebert, D. 2001. Genetic variation in a host-parasite association: potential for coevolution and frequency-dependent selection. *Evolution* **55**: 1136–1145.

Cory, J.S. & Myers, J.H. 2003. The ecology and evolution of insect baculoviruses. *Annu. Rev. Ecol. Evol. Syst.* **34**: 239–272.

Dwyer, G. & Elkinton, J.S. 1995. Host dispersal and the spatial spread of insect pathogens. *Ecology* **76**: 1262–1275.

Dybdahl, M.F. & Storfer, A. 2003. Parasite local adaptation: Red Queen versus Suicide King. *Trends Ecol. Evol.* **18**: 523–530.

Ebert, D. & Mangin, K.L. 1997. The influence of host demography on the evolution of virulence of a microsporidian gut parasite. *Evolution* **51**: 1828–1837.

Elderd, B.D., Dushoff, J. & Dwyer, G. 2008. Host-pathogen interactions, insect outbreaks, and natural selection for disease resistance. *Am. Nat.* **172**: 829–842.

Falconer, D.S. & Mackay, T.F.C. 1996. *Introduction to Quantitative Genetics*. Pearson Education, New York, NY.

Fenton, A., Antonovics, J. & Brockhurst, M.A. 2009. Inverse-gene-for-gene infection genetics and coevolutionary dynamics. *Am. Nat.* **174**: E230–E242.

Fleming-Davies, A.E., Dukic, V., Andreasen, V. & Dwyer, G. 2015. Effects of host heterogeneity on pathogen diversity and evolution. *Ecol. Lett.* **18**: 1252–1261.

Flores, C.O., Meyer, J.R., Valverde, S., Farr, L. & Weitz, J.S. 2011. Statistical structure of host–phage interactions. *Proc. Natl. Acad. Sci. USA* **108**: E288–E297.

Frank, S.A. 1993. Specificity versus detectable polymorphism in host-parasite genetics. *Proc. Biol. Sci.* **254**: 191–197.

Frank, S.A. 1994. Recognition and polymorphism in host-parasite genetics. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **346**: 283–293.

Gillespie, J.H. & Turelli, M. 1989. Genotype–environment interactions and the maintenance of polygenic variation. *Genetics* **121**: 129–138.

Hadfield, J.D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* **33**: 1–22.

Hamilton, W.D. 1980. Sex versus mon-sex versus parasite. *Oikos* **35**: 282–290.

Hamilton, W.D., Axelrod, R. & Tanese, R. 1990. Sexual reproduction as an adaptation to resist parasites. *Proc. Natl. Acad. Sci. USA* **87**: 3566–3573.

Hammond-Kosack, K. & Jones, J. 1997. Plant disease resistance genes. *Annu. Rev. Plant Biol.* **48**: 575–607.

Heath, K.D. & Nuismer, S.L. 2014. Connecting functional and statistical definitions of genotype by genotype interactions in coevolutionary studies. *Front. Genet.* **5**: 77. 1–7.

Hill, A.V.S., Allsopp, C.E.M., Kwiatkowski, D., Anstey, N.M., Twumasi, P., Rowe, P.A. *et al.* 1991. Common West African HLA antigens are associated with protection from severe malaria. *Nature* **352**: 595–600.

Hoover, K., Grove, M.J. & Su, S. 2002. Systemic component to intrastadial developmental resistance in *Lymantria dispar* to its baculovirus. *Biol. Control* **25**: 92–98.

Hufbauer, R.A. & Via, S. 1999. Evolution of an aphid-parasitoid interaction: variation in resistance to parasitism among aphid populations specialized on different plants. *Evolution* **53**: 1435–1445.

Kawecki, T.J. & Ebert, D. 2004. Conceptual issues in local adaptation. *Ecol. Lett.* **7**: 1225–1241.

Kennedy, D.A. 2012. Assessing the potential for evolution in the insect baculovirus *Lymantria dispar* nucleopolyhedrovirus. Ph.D. thesis, University of Chicago.

Kirchner, J.W. & Roy, B.A. 2000. Evolutionary implications of host–pathogen specificity: the fitness consequences of host life history traits. *Evol. Ecol.* **14**: 665–692.

Lambrechts, L., Halbert, J., Durand, P., Gouagna, L. & Koella, J. 2005. Host genotype by parasite genotype interactions underlying the resistance of anopheline mosquitoes to Plasmodium falciparum. *Malar. J.* **4**: 3.

Lambrechts, L., Fellous, S. & Koella, J. 2006. Coevolutionary interactions between host and parasite genotypes. *Trends Parasitol.* **22**: 12–16.

Lively, C.M. 1989. Adaptation by a parasitic trematode to local populations of its snail host. *Evolution* **43**: 1663–1671.

Lively, C.M. 1999. Migration, virulence, and the geographic mosaic of adaptations by parasites. *Am. Nat.* **153**: S34–S47.

Nuismer, S.L. 2006. Parasite local coadaptation in a geographic mosaic. *Evolution* **60**: 24–30.

Páez, D., Fleming-Davies, A. & Dwyer, G. 2015. Effects of pathogen exposure on life-history variation in the gypsy moth (*Lymantria dispar*). *J. Evol. Biol.* **28**: 1828–1839.

Penman, B.S., Ashby, B., Buckee, C.O. & Gupta, S. 2013. Pathogen selection drives nonoverlapping associations between HLA loci. *Proc. Natl. Acad. Sci. USA* **110**: 19645–19650.

Prugnolle, F., de Meeûs, T., Pointier, J.P., Durand, P., Rognon, A. & Theron, A. 2006. Geographical variations in infectivity and susceptibility in the host-parasite system *Schistosoma mansoni/Biomphalaria glabrata*: no evidence for local adaptation. *Parasitology* **133**: 313–319.

Sasaki, A. 2000. Host-parasite coevolution in a multilocus gene-for-gene system. *Proc. Biol. Sci.* **267**: 2183–2188.

Sasaki, A. & Godfray, H.C.J. 1999. A model for the coevolution of resistance and virulence in coupled host-parasitoid interactions. *Proc. Biol. Sci.* **266**: 455–463.

Scanlan, P., Hall, A.R., Lopez-Pascua, L.D. & Buckling, A. 2011. Genetic basis of infectivity evolution in a bacteriophage. *Mol. Ecol.* **20**: 981–989.

Scanlan, P., Hall, A.R., Blackshields, G., Friman, V.-P., Davis, M.R., Goldberg, J.B. *et al.* 2015. Coevolution with bacteriophages drives genome-wide host evolution and constrains the acquisition of abiotic-beneficial mutations. *Mol. Biol. Evol.* **32**: 1425–1435.

Schmid-Hempel, P. & Koella, J.C. 1994. Variability and its implications for host-parasite interactions. *Parasitol. Today* **10**: 98–102.

Schönrogge, K., Gardner, M.G., Elmes, G.W., Napper, E.K.V., Simcox, D.J., Wardlaw, J.C. *et al.* 2006. Host propagation permits extreme local adaptation in a social parasite of ants. *Ecol. Lett.* **9**: 1032–1040.

Tellier, A. & Brown, J.K. 2007. Stability of genetic polymorphism in host-parasite interactions. *Proc Biol Sci.* **274**: 809–817.

Zhan, J., Mundt, C.C., Hoffer, M.E. & McDonald, B.A. 2002. Local adaptation and effect of host genotype on the rate of pathogen evolution: an experimental test in a plant pathosystem. *J. Evol. Biol.* **15**: 634–647.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1 Histogram of the genetic correlations of infection risk for each pair of virus isolates.
Table S1 Sample sizes for each combination of host family and virus isolate.

Received 23 March 2016; revised 31 August 2016; accepted 2 September 2016