Resource quality determines the evolution of resistance and its genetic basis

Katherine E. Roberts¹*, Sean Meaden ¹ §, Stephen Sharpe, Suzanne Kay, Toby Doyle, Drew Wilson, Lewis J. Bartlett⁴, Steve Paterson³, Mike Boots¹,²

¹. Biosciences, University of Exeter, Penryn Campus, UK. TR10 9FE
². Integrative Biology, University of California, Berkeley, USA. CA 94720
³. Institute of Integrative Biology, University of Liverpool, UK. L69 7ZB
⁴. Department of Biology, Emory University, Atlanta, GA, 30322, USA

* These authors contributed equally to this work
* Corresponding author: k.roberts@exeter.ac.uk

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/MEC.15621

This article is protected by copyright. All rights reserved
Abstract

Parasites impose strong selection on their hosts, but the level of any evolved resistance may be constrained by the availability of resources. However, studies identifying the genomic basis of such resource-mediated selection are rare, particularly in non-model organisms. Here, we investigated the role of nutrition in the evolution of resistance to a DNA virus (PiGV), and any associated trade-offs in a lepidopteran pest species (Plodia interpunctella). Through selection experiments and whole genome re-sequencing we identify genetic markers of resistance that vary between the nutritional environments during selection. We do not find consistent evolution of resistance in the presence of virus but rather see substantial variation among replicate populations. Resistance in a low nutrition environment is negatively correlated with growth rate, consistent with an established trade-off between immunity and development, but this relationship is highly context dependent. Whole genome resequencing of the host shows that resistance mechanisms are likely to be highly polygenic and although the underlying genetic architecture may differ between high and low nutrition environments, similar mechanisms are commonly used. As a whole, our results emphasise the importance of the resource environment on influencing the evolution of resistance.

Introduction

Parasites and pathogens impose strong selection on their hosts resulting in the evolution of a range of defence mechanisms. For example, invertebrates possess an effective innate immune system that is capable of fighting infections from a wide variety of pathogens (Kingsolver, Huang, & Hardy, 2013; Sackton et al., 2007; Viljakainen, 2015). Hosts can also use a range of pathways to prevent or tolerate infection, including behavioural or physiological changes (Curtis, de Barra, & Aunger, 2011; Lefevre, Williams & de Roode, 2011; Råberg, Graham, & Read, 2009). The host strategy most likely to evolve, or be maintained, will ultimately depend upon the resources available to the host, as resistance mechanisms can be both costly to initiate and subsequently maintain (Cotter, Simpson, Raubenheimer, & Wilson, 2011; Knutie, Wilkinson, Wu, Ortega, & Rohr, 2017; Kraaijeveld & Godfray, 1997; Lochmiller & Deerenberg, 2000). Such resource...
availability can vary due to both temporal and spatial differences e.g. seasonality, population density and patchiness of accessibility, and in terms of both the quantity and quality of required resources. A core component of resource availability is nutrition, which is likely to significantly influence the evolution of resistance to parasites due to the energetic demands involved in both its development and maintenance. Greater resistance is predicted to evolve under higher resource environments for two reasons. Firstly, reduced competition for resources should allow organisms to invest more in resistance mechanisms. Secondly, higher resources can lead to greater population density and therefore greater transmission events and chance of infection, resulting in stronger selection for resistance (d. C. Lopez-Pascua & Buckling, 2008; Gómez, Bennie, Gaston, & Buckling, 2015). Resistance mechanisms may also, in principle, be specific to host nutritional status, where a resource threshold is required for a resistance mechanism to be induced and functionally useful.

Resistance mechanisms typically come at a price: either through the activation of induced defence mechanisms (Graham, Allen, & Read, 2005; Moret & Schmid-Hempel, 2001; Sadd & Siva-Jothy, 2006) or through the maintenance of a constitutively expressed defence when parasites are absent (Bartlett, Wilfert, & Boots, 2018; M Boots & Begon, 1993; Fuxa & Richter, 1992; Kraaijeveld & Godfray, 1997; Kurt A McKeen, Yourth, Lazzaro, & Clark, 2008). A classic experimental example of such occurred when Drosophila evolved resistance to parasitoid attack, but the cost of resistance was only apparent when food was in limited supply, with an apparent trade-off being between resistance and larval competitiveness for limited food resources (Kraaijeveld & Godfray, 1997). Costs such as these may lead to the stable maintenance of polymorphism within populations (Antonovics & Thrall 1994; Juneja & Lazzaro 2009, Bowers, Boots & Begon 1994, Boots and Haraguchi 1999). However, there are examples of population level resistance in the wild. For example, Cydia pomonella appeared to evolve high levels of resistance to a baculovirus after its application as a bio-control agent (Asser-Kaiser et al., 2007).

The use of controlled laboratory experiments in order to understanding how variation in resistance mechanisms evolve, or are maintained, is therefore critical for predicting the evolution of resistance in more variable wild populations. To date, many studies have relied on traditional model systems such as Drosophila and critically are carried out on a single, often very high
nutritional quality diet to identify the genetic basis of variance in traits of interest (Jha et al., 2015; Michalak, Kang, Sarup, Schou, & Loeschcke, 2017; Shahrestani et al., 2017; Turner & Miller, 2012).

Here we examined the role of nutrition in the evolution of resistance to a DNA virus in an insect host's response to a naturally occurring oral infection. Using an experimental evolution approach, we use the Indian Meal Moth, Plodia interpunctella and its naturally occurring granulovirus (PiGV) as a model. We have previously demonstrated that there is a phenotypic, resource-dependent cost, in age and size of maturity associated with the evolution of resistance in this system (Boots, 2011). The level of resistance attained and the associated costs appeared to depend on the selection environment, this may suggest that different resistance mechanisms could be forced to evolve in different environments (Boots, 2011). To investigate if this was indeed the case, and if the genetic mechanisms of resistance could be driving the differing levels of resistance and their associated costs, we evolved populations for multiple generations on two differing resource qualities, either in the presence or absence of a viral pathogen. We then tested the strength of those populations’ resistances to the same viral pathogen, and their larval development across these nutritional levels in order to quantify any potential trade-offs between investment in resistance and investment in growth. Finally, we used whole-genome resequencing of the populations to perform a genome scan for candidate resistance genes using a pool-seq approach (Eoche-Bosy et al., 2017; Martins et al., 2014). We predict that in a resource rich environment a population could be able to invest in energetically costly resistance mechanisms and compensate for any associated cost, which may not be possible in a nutrient poor environment. Any evolved resistance in these two environments may show a different underlying genetic mechanism. By studying nutritional constraints on resistance using this insect model we aim to tease apart the contribution of parasite exposure and resource availability on the evolution of resistance, and assess any relevant trade-offs in this broader context.

Materials and Methods

Selection Experiment
Replicate selection lines of the Indian Meal Moth, *Plodia interpunctella* were set up at two different nutritional quantities, both in the presence and absence of a natural pathogen, the granulovirus PiGV. In order to establish genetically diverse and homogenous selection lines, we initially established a large outbred population of *P. interpunctella* by outcrossing existing laboratory strains with a number of populations received from the USDA. Briefly, moths were donated from USDA research units (Florida and Byron) and full crossings were carried out with the two existing laboratory populations. First the males of one population were crossed with females of a second population and then *vice versa*, so every combination of male and female pairs was mated. Once all potential male / female pairs were mated their resulting offspring were then crossed, before their offspring were mixed to create a new “Kernow” population, this population was allowed to establish in the lab for five generations before experimental set up (see supplementary methods for full description and schematic of establishment). The initial set up of selection experiment was based on the methods of Boots (2011). For the virus selection lines, PiGV was mixed into the food medium in which moths both feed on, and reproduce within. The viral inoculum for the selection experiment was an LD20 determined by a population level dose response assay carried about using a replication of the final experimental set up but using a range of viral concentrations. The lethal dose was then established using the number of individuals surviving to adulthood. The virus used for the course of the whole experiment was extracted from a large population of laboratory moths that were infected to bulk produce the virus (laboratory strain originally the viral isolate from (Boots & Begon, 1993)). This was purified by ultra-centrifugation in sucrose gradients, before being aliquot out and frozen so that a fresh aliquot could be used for establishment of each generation and avoid any potential viral degradation. This approach was taken as there may be variation in infectivity between solutions. Larvae become orally infected, which is the natural route of infection, through ingesting the infective viral particles whilst they feed. There is therefore a strong selection pressure on all larvae across instar stages.

The resource-level quality of the moth’s food is precisely controlled by the addition of methyl cellulose (an indigestible bulking agent) to the medium (Boots & Begon, 1994). The resource levels to establish our selection lines were determined based on the methods of Boots (2011).
The basic food consisted of a cereal base (50% Ready Brek ©, 30% wheat bran, and 20% ground rice), brewer's yeast, honey, and glycerol. To produce the two selection line food levels 10% of the mix was replaced with methyl cellulose (MC) to give the high-quality resource level, and 55% food mix was replaced with MC for the low-quality diet.

Initially four replicates of each treatment were established; virus unexposed populations (i.e. no virus present in the food) consisted of four populations on low quality food (Control - Low Food), and four on high quality food (Control - High food). This was repeated for populations to be evolved in the presence of virus; four on low quality food (Virus - Low Food), four on high quality food (Virus - High food). In order to maintain population levels within each selection line, 60 three-day post eclosion moths, of mixed sex, were placed in a 500ml Nalgene pot using an excess of each food mixture (200g). This does not account for any potential difference in fecundity between lines; therefore for all selection treatments the same number of control populations was established. These 16 populations (4x Control-Low food, 4x Virus-Low food, 4x Control-High food and 4x Virus-High food) constituted one block of the experiment. This set up was repeated for five replicate blocks to give 20 separate populations of each of the potential selection regimes (See supplementary for full schematic of selection line set up). All populations were maintained in incubators at 27°C, 16 Light: 8 Dark cycle, and pots rotated on their shelf daily, and around the incubator every generation in order to control for any effects of incubator position.

To prevent selection on earlier eclosion the day of first eclosion for each population was recorded and three days after this all eclosed moths were removed. The following days’ newly eclosed moths were then used to establish the next generation (~60 adults). The populations were maintained for 14 generations in this manner after which they were assayed for their viral resistance and life history traits (development rates).

Phenotypic assays
After the 14 generations all populations were “relaxed” from their selection regime. From each of the populations ~120 adults were collected and split onto two different food types (~60 adults on each); a “common garden” (0% MC, the high quality food used for maintaining laboratory Plodia populations) or the food type that they had been selected on for the course of the
experiment (10% MC or 55% MC “selection diet”) (Boots 2011). This common garden approach allows for investigation of the genetic bases of a trait by controlling for any phenotypic plasticity (de Villemereuil et al. 2016). In order to control for any maternal effects from this common garden generation, second-generation third instar larvae were collected for both the viral bioassay and the phenotypic - life history assay. Third instar larvae are assigned based on head capsule width. None of the food for this “relaxed generation” contained any virus but the population set up was otherwise the same. Therefore, for all experimentally evolved populations, there was a replicate population established and subsequently assayed on a common garden environment with no addition of MC. (Full experimental set up please see supplementary figure).

All assays are carried out on individuals housed in a segmented 25-well petri dish with an excess of food. The infection assay followed the protocol of Boots & Roberts (2012) where third instar larvae were removed at random from each population and starved for two hours before being orally inoculated with a freshly prepared virus suspension, diluted with distilled water, 0.1% Coomassie Brilliant Blue R dye and 2% sucrose (to encourage feeding). Ingestion is indicated by the presence of blue dye in at least half of the gut, at which point larvae are removed from the viral suspension. For this experiment, each of the of the relaxed populations were inoculated at five different viral concentrations, with the highest inoculum at 2.5x10^{-4} % virus suspension to dye solution, with four further 1:10 dilutions. This initial inoculum was determined to be one 1:100 dilution of an LD100 of 3rd instars in a preliminary dose response test. A control suspension of the blue dye, sucrose solution was also used as a control for dosing protocol.

Approximately 25 larvae were inoculated at each dilution of virus, from each population. Larvae were again kept individually in 25-well petri dishes in incubators and the numbers of subsequently infected larvae were recorded as a binary response; on visual inspection infected larvae are clearly visible because of their opaque white colour due to a build-up of viral occlusion bodies. As PiGV is an obligate killer, there is no tolerance to infection. We therefore refer to resistance here as the proportion of individuals surviving following viral challenge.

At the same time as the larvae for the infection assay are collected 25 larvae are also individually placed into the 25 well Petri dishes containing high quality resource and allowed to develop in standard incubator conditions. The time to pupation was checked daily and the day that a brown
pupa was seen it was removed from its silk case and its weight recorded. A combination of the
time to pupation and the weight at pupation can then be used to calculate a mean growth rate
(calculated as weight at pupation divided by time to pupation).

**DNA extraction and Sequencing**

For studying the genetic basis of PiGV resistance we used a ‘pool-seq’ approach where individual
larvae from a population are pooled and the subsequent extracted DNA is sequenced to generate
estimates of allele frequencies within a population. This approach has been developed and
validated in a number of papers (Kofler, Langmüller, Nouhaud, Otte, & Schlötterer, 2016; C
Schlötterer, Kofler, Versace, Tobler, & Franssen, 2015; Christian Schlötterer, Tobler, Kofler, &
Nolte, 2014) and is an efficient way of comparing large numbers of populations. Genomic DNA
from each population was extracted using a Blood and Tissue DNA extraction kit (Qiagen, UK).
Fifty larvae from each population were fully homogenised in ATL lysis buffer, and after
Proteinase-K digestion the max volume for the column was taken through for the rest of the
extraction protocol. (25mg tissue was the max for the column and 180uL of lysate equated to
25mg of original tissue). In parallel, DNA was extracted from 8 individual larvae in order to
generate a high confidence SNP dataset, using the QIAGEN Genomic-tip 20/G standard protocol
(Qiagen, UK). All samples were sequenced at the University of Liverpool from Illumina TruSeq
Nano libraries with 350bp inserts using 125bp paired-end reads on an Illumina HiSeq2500
platform. Reads were quality filtered to remove adapter sequences, reads shorter than 10bp and
reads with a minimum window score of 20 using a combination of cutadapt (version 1.2.1)
(Martin, 2011) and Sickle (version 1.2 (Joshi & Fass, 2011). Reads were mapped to the *Plodia
interpunctella* reference genome (described here and available from LepBase.org) using Bowtie2.
GATK’s HaplotypeCaller program was used to generate high confidence SNP markers from
sequences obtained from the eight individual larvae. Allele frequency counts were filtered to
exclude SNPs with coverage greater than the median plus 3 standard deviations, in order to
exclude sequencing errors that could occur from mapping to collapsed repeats in the assembly.
This SNP dataset was used as a reference dataset to generate allele frequencies at each marker
per population, using the pool-seq data and *Samtools mpileup*. Sequence data has been
deposited in the European Nucleotide Archive (ENA) under accession number PRJEB27964.
Additional sequencing was undertaken to improve the scaffold lengths of the assembly using a proximity ligation method at Dovetail Genomics (Santa Cruz, CA, USA). This method creates chromatin cross links on input DNA, followed by proximity ligation to mark the physical proximity of sequences to each other (Putnam et al., 2016).

**De novo assembly and annotation of the *P. interpunctella* genome**

In order to reduce heterozygosity prior to genome assembly, a line of *Plodia* was generated by full-sib matings for 10 generations. DNA was extracted using Qiagen GenomeTip and used to make Illumina TruSeq PCR-free, paired-end libraries with insert sizes of c. 350bp, 450bp and 600bp and sequenced on an Illumina MiSeq platform to generate c. 18Gbp of 2x250bp reads and c. 40Gbp of 2x100bp reads on the Illumina HiSeq2000 platform. Nextera mate-pair libraries with 3Kbp and 10Kbp insert sizes were sequenced on the Illumina 2000 platform with c. 50m pairs of 100bp reads from each library. Illumina polyA-ScriptSeq RNA libraries were prepared from 15 individuals and sequenced on the Illumina 2000 platform with c. 45m pairs of 100bp reads from each library. Illumina MiSeq reads were trimmed to $Q\geq 30$ and adaptors removed using *Sickle* and *Perl* and assembled using *Newbler* (Roche GS-Assembler v2.6) with flags set for large genome and a heterozygote sample. Mate-pair reads were first mapped to these contigs using *Bowtie2* (Langmead & Salzberg, 2012) to remove duplicates and wrongly orientated reads, and scaffolded into contigs using *SSPACE* (Boetzer, Henkel, Jansen, Butler, & Pirovano, 2011). Gap filling was achieved using *GapFiller* for 2x 250bp and 2x 100bp paired-end reads and run for three iterations. RNAseq data were mapped to scaffolds within the assembled genome greater than 3Kbp using TopHat2 to identify transcribed regions and splice junctions. These, together with RNAseq data assembled using *Trinity*, were passed to the MAKER pipeline (Cantarel et al., 2008) to predict genes.

**Phenotypic data analysis**

We first grouped the data according to selection treatment, selection diet, and assay diet, and tested for an interaction between group and viral dose on larval survival (Fig. 1A see supplementary for model fit to dose-response data). Treatments were grouped rather than using interaction terms, as there were insufficient degrees of freedom to model 3-way interactions.
We therefore focused on the median viral assay dilution, as this dosage exhibited the largest variance, and aids interpretability between treatments. To test the role of diet and exposure to PiGV resistance we used a linear mixed effect model, with the ratio of infected: uninfected larvae as the response term. All selection lines were assayed in both the common garden diet (0% MC) and their respective selection food (10% or 55% replacement). Selection treatment (PiGV exposure vs. unexposed control), diet the phenotypic assays were carried out on (common garden 0% vs. the diet the populations were selected on “selection food”) and selection diet (high quality vs. low quality) and interactions among these variables were fitted as fixed effects with block (population start date) and population ID included as random effects and a binomial error structure applied using the ‘lmer’ package (Bates, Mächler, Bolker, & Walker, 2015) in R (version 3.5.1) (R Development Core Team, 2005). An observation level random effect (OLRE) was added to the model to account for overdispersion (as suggested in (Harrison, 2014). Post-hoc comparisons were made by assessing the relevant contrast for each hypothesis using the ‘emmeans’ package (Lenth, 2019; Searle, Speed, & Milliken, 1980) with p-values adjusted using the Benjamini-Hochberg method. For the developmental data a linear mixed effect model was also used, including resistance (proportion uninfected at the median dose) as a fixed effect and mean growth rate (calculated as weight at pupation divided by time to pupation, for each population) as the response term. Again selection treatment, assay diet and selection diet (high quality vs. low quality “selection food”) and interactions among these were fitted as fixed effects, with block and population ID included as random effects. A Gaussian error structure was applied based on the distribution of growth rate values. All analysis was carried out in RStudio (version 3.5.1 (Team, 2016)).

**Pool-seq genome wide association test**

Association tests were run by iterating a GLM on each SNP marker using the alternative and reference allele count as the response variable and the proportion of surviving larvae as the explanatory variable. We scaled allele counts by the size of the pool used (n=50) to account for variation in coverage across SNP markers and used a quasibinomial error structure to account for overdispersion, which has been shown to substantially inflate p-values (Wiberg, Gaggiotti, Morrissey, & Ritchie, 2017). P-values were computed using stepwise ANOVA. Any SNPs where a
model failed to converge, or that resulted in a regression containing data points with a Cook’s distance greater than 1, were discarded. This resulted in a filtered dataset of 988137 and 956647 SNPs (common garden diet and low nutrition diet respectively). P-values were corrected for a false discovery rate (FDR) using the Benjamini and Hochberg correction and plotted across the length of the genome to identify regions associated with PiGV resistance. Underlying genetic structure can confound the results of association studies, we therefore quantified genetic structure was with the program Baypass (Mathieu Gautier, 2015). BayPass is a program designed to estimate covariance structure among population allele frequencies originating from a shared history of the populations being studied. Subsets of the data (49686 markers per group, 20 groups total) were used to assess FMD statistics (distance between covariance matrices, (Fürstner & Moonen, 2003).

Putative function analysis

Functional analysis of the candidate loci resulting from the association tests was conducted in two ways. Firstly, all genes containing significantly associated SNPs were extracted and linked to the \textit{P. interpunctella} predicted gene set. Here, significance was defined as P < 0.05 after false discovery rate correction (Benjamini-Hochberg method) in order to reduce spurious matches. Orthologous genes between \textit{P. interpunctella} and \textit{Drosophila melanogaster} were identified using InParanoid (version 4.1). The resulting UniProt codes from matched genes were used for gene set enrichment analysis using the AmiGO service (http://amigo.geneontology.org/amigo) that allows the Gene Ontology (GO) database to be searched for analogous genes that can then be sorted depending on their functional characteristics. A Fisher’s exact test is then employed using the full list of orthologs as a reference. A second approach selected whole scaffolds that showed clear peaks in the Manhattan plot and is used to generate functional predictions of proteins using the PANNZER2 program (Törönen, Medlar, & Holm, 2018). The most abundant gene ontology terms were then ranked.

Results

\textit{Evolution of resistance to PiGV is diet dependent}

This article is protected by copyright. All rights reserved
By comparing populations that were evolved with the presence of virus (PiGV exposed) to unexposed control populations across the diets, we can test a number of specific hypotheses regarding the role of nutrition on resistance evolution. Firstly, comparing exposed populations and unexposed controls, on the diet they were evolved on (high or low quality) tests whether resistance evolved during the selection experiment. Then testing these same populations on a common garden diet allows us to ask whether any resistance mechanisms that may have evolved are effective across environments, or if there is a difference in effectiveness between diet-specific resistance mechanisms.

To answer these questions we first assessed resistance to the virus across a gradient of PiGV dilutions for each selection treatment (excluding the control dose of zero, where we found no infection, as thus prevented model convergence) we found a significant interaction between each treatment group and viral dilution on infection rates (Quasibinomial GLM, $F_{1,554} = 2.16$, $p = 0.036$, see supplementary dose response figure of model fitted data). We wanted to identify the PiGV dilution that maximized variation (Fig. 1A highlighted box), which was a stock dilution of $2.5 \times 10^{-6}$ (Fig. 1B). At this dilution we found a significant three-way interaction between the selection diet, assay diet, and the viral selection treatment (exposure to the virus vs. unexposed controls) on resistance (GLMM, $\chi^2_9 = 7.27 \ p < 0.01$). To understand the potential drivers of this effect we used post-hoc testing to compare survivorship among the contrasts that test the hypotheses outlined above (Table 1).

We found the largest effect to be driven by diet itself. When comparing only the unexposed control populations, we find that those selected on a high nutrition diet produce larvae more likely to survive than those larvae selected on the low nutrition diet when facing viral challenge on the common garden diet (Table 1.). This is notable as it suggests a trade-off may exist between being able to survive in a nutritionally limited environment and being exposed to parasite pressure.

We found a borderline significant difference ($p = 0.08$) between larvae from populations selected with exposure to PiGV whilst on low nutrition diets, and their counterpart unexposed controls,
again when in the common garden (Fig. 1B). It is interesting that we did not observe any
difference between larvae from virus evolved and unexposed control populations who were
selected on a low nutrition environment, when assayed on the same low quality selection food.
We also did not find any significant differences between exposed and unexposed control
populations from the high quality treatment (Fig. 1B). These differences suggest that depending
on the nutritional environment there may be indirect selection for resistance. Our results appear
complex and heavily determined by diet regardless of viral exposure. These results may
tentatively suggest that populations selected on the nutrient limited diet may be under such
limitation that they have lost a resistance mechanism when it was no longer needed, and that
this constraint is not imposed on populations selected on a high nutrition diet.

Diet determines developmental trade-offs
To assess the effect of diet and PiGV exposure on growth rate we assayed all populations on the
diet they were evolved on (selection food) and a common garden diet (0% M.C.), as in the
resistance assays. Using a mixed effect model with mean growth rate per population as the
response term, we found no significant interactions between assay diet, selection diet, and virus
exposure. However we did find significant effects of the assay diet (the populations selection
diet, versus common garden) and viral exposure of the population (i.e. there was no effect of
selection diet) (Fig. 2). Unsurprisingly, the diet the assay was carried out on had the largest effect
on growth rate; with the fastest growth rates occurring on the common garden diet compared to
the selection diets ($\chi^2 = 157.06$, df = 6, $p < 0.0001$, mean growth rates; common garden 0% MC
0.533-0.572 mg/day$^{-1}$; high quality 0.421-0.458 mg/day$^{-1}$ and low quality 0.266-0.276 mg/day$^{-1}$.
See supplementary table for full growth rates). There was also a significant effect of viral
exposure during selection, with populations facing direction selection for viral resistance
exhibiting quicker growth rates than unexposed control populations ($\chi^2 = 5.05$, df = 6, $p < 0.025$),
which is counter to the trade-offs observed previously in this system (Boots, 2011).

In order to identify any underlying trade-offs we aimed to test whether there was a correlation
between growth rate and survivorship, and if so, if this was consistent among treatment groups.
We correlated the mean growth rate of each population to their mean resistance level. When

This article is protected by copyright. All rights reserved
examining the survival of each viral treatment on the diet they were selected on, we find no correlation between survival and growth rate for either of the high food groups (virus exposed: GLM, $F_{1,11} = 0.0004, p = 0.98$; unexposed controls: GLM, $F_{1,12} = 0.03, p = 0.87$). However, in the low food populations we found a significant correlation in the virus exposed populations (GLM, $F_{1,13} = 18.6, p < 0.01$) but not in the unexposed controls (GLM, $F_{1,16} = 2.65, p = 0.12$). When we analyzed the populations when they were assayed on the common garden diet, we found no significant correlations in any of the selection treatments, i.e. there are no trade-offs between resistance and growth rate, (see supplementary table of treatment level correlations for all populations assayed on the common garden). This demonstrates that the nutritional environment larvae are selected on leads to fundamentally different costs of resistance. In this experiment low nutrient environments selected for a form of resistance that is traded off with growth rate, whereas in the high nutrient environment selection for resistance appeared to come at no cost to growth rate.

**Identifying the genomic basis of resistance**

To identify the genomic basis of resistance to PiGV infection we employed a genome wide scan in order to associate specific loci with resistance. This method tests the association between allele frequency at each SNP present in our dataset and the resistance of each population. We exploited the variation in phenotypic resistance we found by using both the exposed and unexposed populations in the GWAS. We first assessed SNPs that predicted resistance on a common dietary background i.e. the proportion of surviving larvae when assayed on the common garden diet. We identified a number of candidate SNPs that were strongly correlated with resistance (Fig. 3).

Whilst assessing the function of individual SNPs is useful for identifying strong effect loci, a large number of SNPs across many scaffolds is suggestive of a polygenic trait and therefore enrichment analyses may be more appropriate for functional inference. Enrichment analysis of these SNPs found no significant biological processes, molecular function or cellular component. However, a single protein class (oxygenase) was highly enriched (12.16 fold, adjusted p-value = 0.021). This is of note as a recent finding associated a down regulation of cytochrome P450 with baculovirus
infection, suggesting a key role of oxygenase enzymes in baculovirus resistance (Shrestha et al., 2019). A full list of all significantly associated genes is provided (Supplementary table S1). As a secondary approach we selected 4 scaffolds from the Manhattan plot (Fig. 3) that showed a high density of SNPs associated with resistance. For these scaffolds we carried out functional prediction of all genes present and ranked the most abundant gene ontology terms (Fig. 4). Notably, integral component of the membrane was the most common gene description suggesting a role for disrupting either viral entry into the cell or maturation of virus particles within the cell (Rohrmann, 2019).

**Diet-specific resistance mechanisms**

Our phenotypic data provide evidence for resistance evolving in an environmentally dependent manner, suggesting distinct genomic routes to resistance under specific nutritional conditions. When the SNP markers that were significantly associated with resistance for each treatment were compared (i.e. virus exposed on low versus high food, control unexposed on low versus high food etc.) there appeared to be different patterns in the SNPS associated, with little overlap between the different treatment groups. This suggests little correlated selection for resistance across the different diets (See supplementary Figure: Venn Diagram of shared SNPs). This was particularly the case when we examined populations evolved on a low nutrition diet but assayed on a higher resource, common garden diet. Here, unexposed control populations exhibited much lower resistance than exposed populations. To identify the mechanism that provided this difference in resistance we repeated the association test for this subset of populations using their survival on the low nutrition diet as the response term. Again, we found multiple SNPs strongly correlated with resistance, suggesting a polygenic trait (Fig. 5). However, not enough SNPs were identified to perform an enrichment analyses. Instead, we selected 4 scaffolds that showed a high density of SNPs associated with resistance. As previously we collapsed functional annotations into GO terms and tallied these terms. Interestingly we again found that integral membrane component was by far the most abundant GO term, despite different scaffolds being identified with the association test. Taken together these results suggest some partially shared resistance mechanism, but the exact genetic background likely dictates the magnitude of both the effect and the associated trade-off.
**PiGV resistance is a polygenic trait**

The large number of SNPs associated with resistance in both the common garden (high quality) and low nutrition diets suggests that resistance to PiGV infection is likely to be a highly polygenic trait (See supplementary Figure, Venn Diagram). However it is possible that the SNPs we associated with disease resistance are located in similar genomic regions, only appearing disparate due to the length of the scaffolds in our draft assembly. Therefore, we further assessed the proximity of SNPs, using improved scaffold lengths (increased from 0.5 to 5Mbp on average) from proximity-ligation sequencing, we then found many independent peaks of selection along the larger scaffolds (Supplementary Figure). Suggesting that the genomic location of SNPs is not an artefact of many small scaffolds that make up the assembly and that PiGV resistance is indeed a highly polygenic trait.

**Larval populations show little underlying population structure**

Genome-wide association studies can lead to spurious correlations as a result of underlying population structure where associations are a result of shared demographic history rather than a signature of selection. Whilst our populations were all derived from a single ancestral population that had been out crossed repeatedly prior to the selection experiment (Supplementary Figure: Establishing populations), it is possible that the observed differences in allele frequency were the result of an underlying population structure. To rule this out we used a Bayesian approach to identify population structure naively on independent subsets of the SNP data. We identified very weak population structure suggesting that our results are unlikely to be spurious, we found no clustering of populations that would be indicative of underlying population structure (Supplementary Figure: population clusters). Crucially, when we clustered the populations we found that these populations did not group by resistance, and therefore our results are unlikely to be driven by underlying population structure. We also found high reproducibility of results independent of which subset was used, suggesting our methods were robust (FMD always < 0.6, see Gautier (2015)).

**Discussion**
In order to investigate if the genetic basis of resistance was resource dependent in an insect–DNA virus system, we employed an evolve and re-sequence experiment (Schlötterer, Kofler, Versace, Tobler, & Franssen, 2015). We found that the evolution of resistance is resource specific, with populations selected on a nutrient limited diet appearing to utilise a different mode of resistance. Surprisingly, we found no response to direct selection for resistance, with no clear increase in resistance to PiGV in populations selected on a high resource diet. In contrast, we found greater resistance levels in control populations that were selected on a high resource than a low resource diet, suggestive of a trade-off between growth and resistance that is only apparent when resources are limited. There was also high variation across populations in the resistance levels that evolved, with the effect of nutritional environment being greater than that of selection with exposure to the virus.

The finding that diet had more of an effect than viral exposure in the evolution of resistance is striking, especially given that we used a strong viral selection pressure, and an outbred-genetically diverse starting population. This suggests that resistance may be being indirectly selected. We find control populations selected on a high nutrition diet more resistant than those on low, when assayed on the common garden diet. This suggests that through indirect mechanisms resistance is evolving as a correlated response with another trait. It is possible that the resistance mechanism employed on low food may be more complex than our bioassay can detect. For example selection acts on all life-stages during the selection experiment, but our assays by necessity, only test resistance at the 3rd instar stage (Boots & Begon, 1993). Even with this caveat, it appears populations selected on a low nutrient diet may have indirectly evolved a low nutrition-specific resistance mechanism that is ineffective when reared and assayed on the common garden. This raises questions over the mechanisms that drive population level changes in resistance, why would resistance appear lower when expressed on a high resource environment than in a lower one. A potential hypothesize is that it is driven by a stress response from a sudden change in environment. Given that the control populations selected on low food are more resistant when assayed on their selection food than the common garden suggests they have adapted to this food. Such a reduction in nutritional quality or availability has been shown to cause increased resistance in other systems. For example, a study in tree frogs found that a
low resource diet favored host resistance to a gut worm, contrary to predictions. Authors suggested this may be due to resource competition between the host and parasite, or an imbalance in specific components of the diet (Knutie et al., 2017). Similarly, a study of nutrition and immunity in the greater wax moth (G. mellonella) found a significant negative correlation between body mass and the strength of an immune response in larvae on high-energy food, with longer developmental time being associated with a stronger immune response (Krams et al., 2015). Whilst infected fruit flies showed increased resistance to one bacterial pathogen but decreased resistance to another, under self imposed dietary restriction, suggesting resistance also depends upon the pathogen to which hosts are exposed (Ayres & Schneider, 2009). Such a decrease in feeding rate in response to infection has been termed “illness-mediated anorexia” and suggests individuals may voluntarily change resource investment as an adaptive response to parasite challenge (reviewed in Hite, Pfenning, & Cressler, 2020).

The marked variation in the results of our phenotypic assays led us to analyze the genome scan data for resistance across all populations, regardless of the selection food they were evolved in, using the population level resistance as our response variable. This identified two likely factors that may be involved in determining any resistance to PiGV. The first was that a single class of protein was overrepresented for the genes that were significantly associated with resistance on the common garden diet. This was an oxygenase and is of interest as many oxidoreductases have been previously linked to insect resistance (Chen et al., 2014; Nguyen, Nielsen, & Reid, 2013; Shrestha et al., 2019; Sun et al., 2012). Oxidative stress induced by tissue damage during viral infection appears to be a ubiquitous pathology, implicating antioxidant enzymes in modulating immune function and resistance to viruses including influenza A infection in humans (reviewed in Camini, da Silva Caetano, Almeida, & de Brito Magalhães, 2017 and Liu et al., 2017). However, further work is required to elucidate its exact function in our study system. Secondly, in regions identified in both the common garden and the low nutrient food assay, the most common Gene Ontology term was an integral membrane component. This suggests that there is strong selection for either the blocking of baculovirus binding to the cell, or its localization at the cell membrane prior to cell lysis. Interestingly, although this was the most common GO term across diets, the scaffolds identified as being associated, differed. Suggesting that there may be multiple
routes to resistance and that nutrition may shape the pathway taken. Multiple resistance mechanisms have been found in another Lepidopteran - Baculovirus system following the use of the CpGV virus as a bio-control agent for *Cydia pomonella* (Asser-Kaiser et al., 2007; Cory, 2017; Gebhardt, Eberle, Radtke, & Jehle, 2014). Wild moth populations showed high levels of resistance - originally thought to be due to replication being prevented by a repeat in the viral pe38 gene (Gebhardt et al., 2014; Graillot et al., 2016; Graillot, Blachere-López, Besse, Siegwart, & López-Ferber, 2019; Lacey, Thomson, Vincent, & Arthurs, 2008; Lange & Jehle, 2003; Annette J. Sauer et al., 2017). Resistance was however found to be driven by either dominant, monogenic, and sex chromosome (Z) -linked inheritance of viral (CpGV) resistance, or alternatively an autosomal linked mechanism, or finally resistance being carried by Z-chromosomal and autosomal inheritance traits, that provide resistance against a set of viral isolates (Graillot et al., 2019; A. J. Sauer, Schulze-Bopp, Fritsch, Undorf-Spahn, & Jehle, 2017; Annette J. Sauer et al., 2017). As a whole this demonstrates that resistance against a baculovirus can be complex and include multiple mechanisms and inheritance pathways.

We find that resistance is polygenic from the number of SNPs correlated with observed resistance. To verify this, we used Hi-C scaffolding to improve scaffold lengths and still found many independent regions correlated with resistance. This suggests that resistance on both diets is a highly polygenic trait, as previously seen in other insect traits (Jha et al., 2015; Kang, Aggarwal, Rashkovetsky, Korol, & Michalak, 2016). We propose that the mechanism of resistance in our system may also be complex, rather than a small number of immune-linked genes, as is characteristic in RNA virus immunity (Magwire et al., 2012). A study of inbred *Drosophila* lines found significant genetic variation in resistance that was not restricted to the canonical immune genes, but was in genes affecting many different aspects of host physiology (Wang, Lu, St Leger, & St. Leger, 2017). Lines appeared to have maintained resistance through coordination of both morphological and physiological routes that function in different ways in different lines, and has parallels to the polygenic route we find. It has been suggested that an intricate genetic architecture of many interacting genes can lead to genetic redundancy, with many competing beneficial alleles, in turn allowing for rapid evolutionary responses (Barghi et al., 2018). The complexity of polygenic traits may explain why the variation in resistance we see,
both between and within selection lines, is maintained. Polygenic resistance that includes non-immune genes and pathways would theoretically also be much more beneficial if it reduces the rate of parasite co-evolution (Lazzaro, Sackton, & Clark, 2006). An increase in immunocompetence or resistance that trades-off with other fitness components, such as development rate, as investigated here, may in fact act as a mechanism for the maintenance of genetic variation in immune responses (Kraaijeveld & Godfray, 1997; Lazzaro et al., 2006; K. A. McKean & Nunney, 2001; Kurt A McKean et al., 2008). Given that the resistance mechanisms seem likely to be highly polygenic in this system may, in part, explain why such a complex picture emerges from the phenotypic assays, with the expression of trade offs being dependent on assay environment.

Association tests like those we used in this study, rely on consistent divergence of allele frequencies between treatments (Wiberg et al., 2017). Therefore if the same functional outcomes were reached by alternative genetic architecture then association tests would not identify these alleles. Evolutionary responses can also be highly stochastic and this is further complicated by the high degree of polygenicity identified in our genome scans. It is also important to note that whilst experimental selection and re-sequence studies are powerful frameworks to study evolution, recent studies have highlighted several important points to consider in interpreting results. Studies such as ours that find polygenic responses with multiple SNPs, could be driven by the fact that the SNPs are located in genomic regions that coincide with segregating inversions in the genome (Franssen, Nolte, Tobler, & Schlötterer, 2015). SNP number could also be elevated as a result of hitchhiking, where selection on low-frequency haplotypes also causes a strong selection signal on any linked neutral SNPs (Barghi & Schlötterer, 2019).

Whilst we used a large, genetically diverse founding population and each generation was established using the offspring of ~60 adults, as with any selection study there is the potential for genetic drift to influence results (Kennedy & Dwyer, 2018). We were unable to control or measure mating rates, sex ratios or final reproductive output- all of which could lead to drift. Although underlying population structure could differ between food treatments through indirect effects of the selection treatment on moth reproduction and subsequent offspring size or number, we found little evidence of this is in our structure analysis. Feeding rate may differ
across the treatments. It is also known that cannibalism occurs in larvae of this species and we could not monitor cannibalism rates which may differ between the diets and also have potential to alter population structures (Boots, Childs, Reuman, & Mealor, 2009; Briggs, Sait, Begon, Thompson, & Godfray, 2000).

We observed a trade-off between resistance and growth rate in populations evolved under nutrient limited conditions, suggesting low nutrition may augment the evolution of specific immune response mechanisms. Resource dependency of this trade-off has been demonstrated in this system before (Boots & Begon, 1993; Boots, 2011), with costs to resistant being more apparent on poor quality food. However the picture from our current study is more complicated, in addition to selecting for resistance through exposure to the virus (direct selection) there appears to be correlated responses to selection of diet (indirect selection). Previous results had suggested there may be a genetic basis for the trade-off in this system (Bartlett et al., 2018). To further investigate this potential *Plodia* populations were directly selected for their development rate, however a trade-off between development and viral resistance was not consistently expressed. This emphasized the importance of defining the trait under selection and taking a holistic view of selection pressures when interpreting results (Bartlett, Visher, Haro, Roberts, & Boots 2020). For example in this current study, assay diet had a significant effect on growth rate. This could be due to the common garden environment being of higher quality than either of the selection diets; had it been more differentiated than the selection foods a different manifestation of any trade off may have been seen.

As far as we know this is the first study to use whole-genome resequencing to investigate the genetic basis of variation in the evolution of resistance, demonstrating the genomic basis of a well characterized trade-off between growth rate and resistance to pathogens. We have demonstrated that selection on a low nutrient diet can have significant effects on the underlying genetic architecture of virus resistance. We highlight a potential trade-off between the forms of resistance acquired and describe putative resistance mechanisms that vary by diet. Our pool-seq approach has allowed a high level of replication at the population level and provided insights into the genetic nature of resistance. As with all such gene association studies, particularly in non-
traditional model organisms, we can only allude to, and correlate our results with, potential genes and processes. While experimental evolution may be an effective method to identify traits under selection, it relies on large treatment effects, consistent selection across replicate populations and may be limited by high polygenicity in traits when combined with genomic association tests. Therefore, further work will be required to fully characterize these mechanisms and functional validation of mutants in genome-edited insects may soon be possible and be of great experimental value. Our results have implications for understanding wild insect populations and more broadly the role of nutrition across environments on pathogen resistance.

Acknowledgements
SM and KER were supported by a NERC (UK) grant NE/J009784/1 awarded to MB and SP. MB was also supported by National Institutes of Health (R01-109501 and R01-GM122061-03); the content of this study is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank S.E.F.Evison for beneficial comments on the manuscript, and the constructive reviewer comments, as well as the USDA for kind donation of Plodia stocks.

Data Availability
All the experimental data to support the findings of this study including all virus assay and development data is available at DataDryad. https://doi.org/10.5061/dryad.k98sf7m4g. The complete sequencing data in CRAM format is available from the European Bioinformatics Institute (EBI), under accession number PRJEB27964.

References


Society B: Biological Sciences, 279(1744), 4009–4014. doi:10.1098/rspb.2012.1073


Eoche-Bosy, D., Gautier, M., Esquibet, M., Legeai, F., Bretaudeau, A., Bouchez, O., ... Montarry, J. (2017). Genome scans on experimentally evolved populations reveal candidate regions for

This article is protected by copyright. All rights reserved


*G3 (Bethesda, Md.),* 6(11), 3507–3515. doi:10.1534/g3.116.034488


Naturally Coevolving Viruses in Drosophila melanogaster. *PLoS Genetics, 8*(11), e1003057. doi:10.1371/journal.pgen.1003057


This article is protected by copyright. All rights reserved


### Table 1

The estimated marginal means (EMM) of the contrasts between the survival of the different selected populations to the viral assay. Viral exposed or control – unexposed populations, from both diets (High or Low quality) are assayed on both their selection diet (High or Low quality selection food) and the common garden diet (C.G. 0% MC). The estimate refers to difference between estimated means between the populations as stated in the contrast column. Degrees of freedom (d.f.) are set to infinite when the model is tested against a normal distribution with a z-test.

<table>
<thead>
<tr>
<th>Contrasts</th>
<th>Estimate</th>
<th>Std Error</th>
<th>D.F.</th>
<th>z- ratio</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control High vs. Exposed High on High assay</td>
<td>0.36</td>
<td>0.32</td>
<td>infinite</td>
<td>1.12</td>
<td>&gt;0.90</td>
</tr>
<tr>
<td>Control Low vs. Exposed Low on Low assay</td>
<td>-0.15</td>
<td>0.30</td>
<td>infinite</td>
<td>-0.49</td>
<td>&gt;0.90</td>
</tr>
<tr>
<td>Control High vs. Exposed High on C.G. assay</td>
<td>-0.37</td>
<td>0.33</td>
<td>infinite</td>
<td>-1.11</td>
<td>&gt;0.90</td>
</tr>
<tr>
<td>Control Low vs. Exposed Low on C.G. assay</td>
<td>0.65</td>
<td>0.30</td>
<td>infinite</td>
<td>2.18</td>
<td>0.147</td>
</tr>
<tr>
<td>Control High vs. Control Low on C.G. assay</td>
<td>-1.08</td>
<td>0.29</td>
<td>infinite</td>
<td>-3.67</td>
<td>0.001</td>
</tr>
<tr>
<td>Exposed High vs. Exposed Low on C.G. assay</td>
<td>-0.06</td>
<td>0.33</td>
<td>infinite</td>
<td>-0.18</td>
<td>&gt;0.90</td>
</tr>
</tbody>
</table>

*P value adjustment: Holm – Bonferroni method for 6 tests*
Figure 1 A) Dose response of populations shown as proportion infected across the control inoculum (0 - Left hand side) and the five dilutions of PiGV virus (most concentrated –Right-hand side). Each dose is generated by a serial dilution of stock virus. See supplementary for model fits for dose response across treatments. Each population was assayed on the diet they were evolved on during the selection experiment “Selection Food”, high nutrition (left hand- dark grey) or low nutrition (right hand- light grey) and also on the common garden diet (C.G.) Highlighted box shows the dilution that showed highest levels of variation, expanded for clarity in (b). B) Expansion of highlighted box in (A) the proportion infected results at the dilution used for genomic analysis, 2.5 x 10^{-6} which showed the largest level of variation. Significant contrast from Table 1 is marked on figure by an asterisk. Each point represents the 25 larvae from each population.
Figure 2. Nutrient dependent trade-offs between mean population growth rate and resistance across diets. Each point represents the raw data of each population and lines are the model predictions. Triangle points denote the populations that were selected in the presence of PiGv (exposed populations) and circles denote control (unexposed populations). Yellow points denote larval populations selected on a low-nutrition section food, whilst black denote populations selected on a high nutrition diet. Both resistance and growth rate are population level estimates based on either survivorship or development assays of 25 larvae per population.
Figure 3. Whole genome scan for SNPs associated with PiGV resistance on a common garden diet (0% MC). Points denote position on the genome scaffold and the FDR adjusted p-values. Dashed line represents the level of significance (0.05). Shades show the individual scaffolds.
Figure 4. The most abundant Gene Ontology (GO) terms predicted from 4 scaffolds associated with viral (PiGV) resistance in larvae assayed on a common garden diet (0%MC). Functional predictions were made with PANNZER2 program.
Figure 5. Manhattan plot of SNPs associated with PiGV resistance on a low nutrition diet. Points denote position on the genome scaffold and the FDR adjusted p-values. Dashed line represents the level of significance (0.05). Shades show the individual scaffolds.