The genomic basis of evolved virus resistance is dependent on environmental resources

Katherine Roberts[1¶], Sean Meaden [1 ¶], Stephen Sharpe, Suzanne Kay, Toby Doyle, Drew Wilson, Lewis J. Bartlett[4], Steve Paterson[3], Mike Boots[1,2*]

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

Abstract

Parasites impose strong selection on their hosts, but the level of resistance evolved 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 associated trade-offs, in a lepidopteran pest species (*Plodia interpunctella*). Through selection experiments and whole genome sequencing we identify putative mechanisms of resistance that depend on the nutritional environment during selection. We find that the evolution of resistance is specific to diet, with adaptation to a low nutrition diet constraining resistance when challenged with the pathogen on a high nutrition diet. Resistance in a low nutrition environment is negatively correlated with growth rate, consistent with an established trade-off between immunity and development. Whole genome resequencing of the host shows that resistance mechanisms are highly polygenic and suggests evidence for trade-offs at the genetic level. Critically when populations evolve in high resource conditions, resistance is linked to metabolic and immune pathways, however it is more closely associated with cytoskeleton organisation when selected under low nutrition. Our results emphasise the importance of resources on 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 range of pathogens

(Kingsolver *et al*. 2013, Sackton *et al.* 2007, Viljakainen *et al.* 2015). Hosts can also use a range of pathways to prevent or tolerate infection, including behavioural or physiological changes (Raberg *et al.* 2009; Curtis *et al.* 2011, Lefevre et al. 2012). The host strategy most likely to evolve, or be maintained, will ultimately depend upon the resources available to the host, as resistance mechanisms are both costly to initiate and maintain (Cotter, Simpson, Raubenheimer, & Wilson, 2011; Knutie, Wilkinson, Wu, Ortega, & Rohr, 2017; Lochmiller & Deerenberg, 2000). Such resource availability can vary due to both temporal and spatial differences e.g. seasonality, population density and patchiness of resource availability, and in terms of the quantity and quality of required resources. A core component of resource availability is nutrition, which is likely to greatly impact the evolution of resistance to parasites. For example, 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 (Lopez-Pascua & Buckling 2008; Gómez *et al.* 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.

Such resistance mechanisms typically come at a price: either through the activation of induced defence mechanisms (Graham, Allen, & Read, 2005; Moret & Schmid-Hempel, 2000; Sadd & Siva-Jothy, 2006) or through the maintenance of a constitutively expressed defence when parasites are absent (Boots & Begon, 1993; Fuxa & Richter, 1992; Kraaijeveld &

Godfray, 1997; McKean, Yourth, Lazzaro, & Clark, 2008). Such costs may lead to the stable maintenance of polymorphism within a population (Antonovics & Thrall 1994; Juneja & Lazzaro 2009, Bowers, Boots & Begon 1994, Boots and Haraguchi 1999). Therefore, understanding how such mechanisms evolve, or are maintained, in controlled laboratory conditions is important for predicting the evolution of resistance in more variable wild populations. Here we examined the role of nutrition in the evolution of resistance to a DNA virus in its insect host in response to natural oral infection using an experimental evolution approach. We use the Indian Meal Moth, *Plodia interpunctella* and its naturally occurring granulosis baculovirus (PiGV) as a model system, where we have previously demonstrated that there is a resource-dependent cost to the evolution of resistance (Boots, 2011). Both the level of resistance attained and the associated costs depended on the selection environment, suggesting that different resistance mechanisms may be forced to evolve in different environments (Boots, 2011). We therefore evolve populations for multiple generations on two different nutritional environments, either in the presence or absence of the viral pathogen. We then tested the strength of the populations' resistances to the viral pathogen, and the larval development across the nutritional levels in order to quantify any potential trade-offs. Finally, we used whole-genome resequencing of populations to perform a genome scan for candidate resistance genes. Resequencing experiments are a powerful tool to identify the genetic basis of observed traits, including those involved in host-parasite interactions (Eoche-Bosy et al., 2017; Martins et al., 2014). To date many studies rely on model systems such as *Drosophila* and critically use only 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). By studying multiple nutritional levels in a non-model organism we aim to tease apart the contribution of diet to resistance and any relevant trade-offs in a 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 granulosis baculovirus PiGV. In order to establish genetically diverse and homogenous selection lines, we initially established a large outbred population of *Plodia interpunctella* by outcrossing existing laboratory strains with a number of populations received from the USDA. The initial set up of selection was based on the methods of (Boots, 2011). Briefly, for the virus selection lines PiGV was mixed into the food medium in which moths both feed on and reproduce within. 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 (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 (see supplementary for full methods for resources). 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, 4 control (no virus) and 4 virus populations were established, on each of the two food resources. Sixty, 3-day post eclosion moths of mixed sex were placed in a 500ml Nalgene pot using an excess of each food mixture (200g). These 16 populations (4 x Control-Low food, 4 x Virus-Low food, 4 x Control-High food and 4 x Virus-High food) constituted one block of the experiment. This set up was repeated for 5 replicate blocks to give 20 separate populations of each of the potential selection regimes. All populations were maintained in incubators at 27°C, 16 Light :8 Dark cycle, and pots were rotated around the incubator in order to control of any effects of incubator position. The day of first eclosion of each pot was noted, and three days post this first eclosion moths were moved onto the next generation as a way controlling for the effect of food on developmental time and ensuring the median day of eclosion was always used to generate the parents of the next generation (Boots, 2011). The populations were maintained for 12 generations in this manner after which they were assayed for their viral resistance and life history traits.

Phenotypic assays

After 12 generations all populations were relaxed from their selection regime. Populations were split onto two different food types; a high quality 0% MC food (common garden environment i.e. both high and low nutrition treatment populations were reared and assayed on a common diet with no addition of MC) or the food type that they had been selected on for the course of the experiment (10% MC or 55% MC). None of the food for this "relaxed" generation contained any virus but the population set up was otherwise the same

as for the selection regime. From these populations, second-generation (to avoid maternal effects), third-instar larvae were either bioassayed with a viral solution, to look at virus resistance, or allowed to develop individually for life history measures. Both assays were carried out in on individuals housed in a segmented 25 well petri dish with an excess of relaxed generation food. The infection assay followed the protocol of Boots & Roberts (2012) where third instar larvae were removed from each population and starved for two hours before being orally dosed with a freshly prepared virus solution diluted with distilled water, 0.1% Coomassie Brilliant Blue R dye (ingestion is indicated by the presence of blue dye in the gut) and 2% sucrose (to encourage feeding). For this experiment, each of the relaxed populations was dosed at 5 different virus concentrations, highest dose of $2.5x10^{-4}$ % virus solution to dye solution, with four further 1:10 dilutions. A control solution of the blue dye, sucrose solution was also used as a control for dosing protocol. Approximately 25 larvae were dosed at each dose of virus, from each population. Larvae were kept 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 were collected 25 larvae were again individually placed into the 25 well Petri dishes containing high quality resource and allowed to develop in standard incubators 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 weighted and recorded.

164

165 **DNA extraction and Sequencing**

166 Our methods for studying the genetic basis of PiGV resistance was to use a 'pool-seq'
167 approach where individual larvae from a population are pooled and the subsequent
168 extracted DNA is sequenced to generate estimates of allele frequencies within a population.
169 This approach has been developed and validated in a number of papers (Kofler, Langmüller,
170 Nouhaud, Otte, & Schlötterer, 2016; C Schlötterer et al., 2015; Christian Schlötterer, Tobler,
171 Kofler, & Nolte, 2014) and is an efficient way of comparing large numbers of populations.
172 Genomic DNA from each population was extracted using a Blood and Tissue DNA extraction
173 kit (Qiagen, UK). First, 50 larvae from each population were fully homogenised in ATL lysis
174 buffer, and after Proteinase-K digestion the max volume for the column was taken through
175 for the rest of the extraction protocol. (25mg tissue was the max for the column and 180uL
176 of lysate equated to 25mg of original tissue). In parallel, DNA was extracted from 8 individual
177 larvae in order to generate a high confidence SNP dataset, using the QIAGEN Genomic-tip
178 20/G standard protocol (Qiagen, UK). All samples were sequenced at the University of
179 Liverpool from Illumina TruSeq Nano libraries with 350bp inserts using 125bp paired-end
180 reads on an Illumina HiSeq2500 platform. Reads were quality filtered to remove adapter
181 sequences, reads shorter than 10bp and reads with a minimum window score of 20 using
182 cutadapt (version 1.2.1) (Martin, 2011) and Sickle (version 1.2 (Joshi & Fass, 2011)). Reads
183 were mapped to the *Plodia interpunctella* reference genome (described here and available
184 from LepBase.org) using Bowtie2. GATK's HaplotypeCaller program was used to generate
185 high confidence SNP markers from sequences obtained from individual larvae. Allele
186 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 using Samtools mpileup. Sequence data has been deposited in the 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≥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

To test the role of diet and exposure to PiGV resistance we used a linear mixed effect model. We used the proportion of surviving larvae at the median assay dose as the response term as this dosage exhibited the largest variance (0.042, Fig. 1A). All selection lines were assayed in both the common garden diet (with no MC replacement) or their respective selection diet (10% or 55% replacement). Selection treatment (PiGV exposure vs. control), assay diet (common garden vs. home) and selection diet (high vs. low) and interactions among these variables were fitted as fixed effects with block (population start date) included as a random effect and a binomial error structure applied. Checks of model residuals showed that the data conform to model assumptions. ANOVA was used to determine p-values following model simplification using AIC . Post-hoc comparisons were made using Tukey's all-pair comparison with p-values adjusted using the Holm-Bonferroni method. For the developmental data, a similarly structured model was applied, in this case including resistance as a fixed effect and mean growth rate (per population) as the response term.

Pool-seq genome wide association test

Association tests were run by iterating a binomial 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. 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 ~450,000 and 250,000 SNPs (common garden diet and low nutrition respectively). P-values were corrected for a false discovery rate using the Benjamini and Hochberg correction and plotted across the length of the genome to identify regions associated with PiGV resistance. Genetic structure was assessed with the program Baypass using subsets of the data (49686 markers per group, 20 groups total) to assess FMD statistics (distance between covariance matrices, (Förstner & 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 associated SNPs were extracted and linked to the *Plodia interpunctella* predicted gene set. In this case significance was defined as $P < 0.0001$ after false discovery rate correction (Benjamini-Hochberg method) in order to reduce spurious matches. Orthologous genes between *P. interpunctella* and *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) using Fisher's exact test. A second approach was to search the BLAST database directly for the best hits to the *P. interpunctella* genes of

255 interested. The resulting best hits were extracted and used to search the UniProt database
256 for gene ontology terms and functional characterizations.
257
258 **Results**
259
260 *Evolution of resistance to PiGV is diet dependent*
261
262 By comparing virus (PiGV) exposed and unexposed controls selection treatments across
263 multiple diets, we can test a number of specific hypotheses regarding the role of nutrition on
264 resistance evolution. For example, comparing exposed and unexposed controls on their
265 local diet (high or low nutrition) tests whether resistance evolved during the selection
266 experiment. Testing these same populations on a common garden diet tests whether any
267 evolved resistance mechanisms work across environments. A comparison of unexposed
268 controls on a common garden diet allows us to test an effect of diet itself in shaping
269 resistance. Finally, comparing exposed populations selected on a high or low nutrition diet
270 and assayed on a common garden diet tests whether there is a difference in effectiveness
271 between diet-specific resistance mechanisms.
272
273 In order to answer these questions, we first assessed resistance to PiGV across all diets and a
274 gradient of doses in order to identify the dosage that maximized variation (Fig. 1A, stock
275 dilution of 2.5e-06). At this dosage we found a strong three-way interaction between the
276 selection diet, assay diet and the selection treatment (exposure to the virus vs. control) on
277 resistance (GLMM, $\chi^2_9 = 16.03$, $p < 0.001$, Fig. 1B). To understand the drivers of this effect we

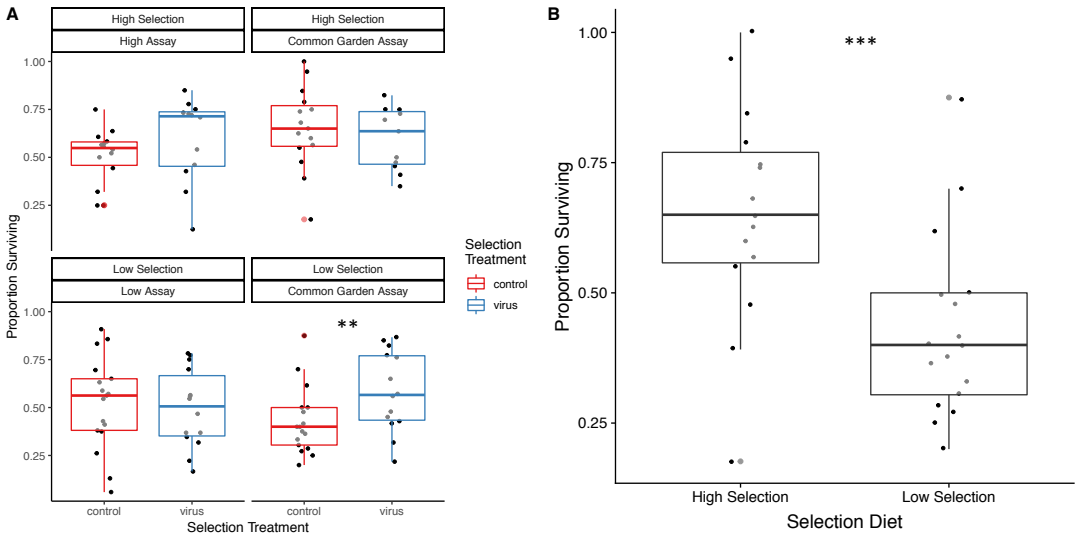
278 used post-hoc testing to compare survivorship among the contrasts that test the hypotheses
279 outlined above (Table 1).
280

contrast	estimate	SE	d.f.	z-ratio	p value
Exposed High vs. Control High on High assay	0.30	0.17	Inf	1.71	0.2612
Exposed Low vs. Control Low on Low assay	-0.14	0.16	Inf	-0.86	0.7822
Exposed High vs. Control High on C.G. assay	-0.35	0.18	Inf	-1.93	0.2154
Exposed Low vs. Control Low on C.G. assay	**0.57**	**0.16**	**Inf**	**3.51**	**0.0022**
Control High vs. Control Low on C.G. assay	**-1.02**	**0.16**	**Inf**	**-6.32**	**0.0000**
Exposed High vs. Exposed Low on C.G. assay	-0.09	0.18	Inf	-0.51	0.7822

P value adjustment: Holm-Bonferroni method for 6 tests.

281
282 We found the largest effect to be driven by diet itself, for example when comparing just
283 unexposed control populations we find that those selected on a high nutrition diet produce
284 larvae more likely to survive when facing viral challenge on a common garden diet (Table 1.;
285 Fig. 1). This is interesting as it suggests a trade-off between being able to survive in a
286 nutritionally limited environment and resistance to a parasite.
287
288 Secondly, we found that larvae from populations exposed to PiGV and selected on a low
289 nutrition diet showed greater survival than their counterpart controls, but only when
290 assayed on the common garden diet (Fig 1, panel A). This suggests that the method of
291 resistance being employed by populations evolved in the two environments is different
292 across the environments. It is interesting to note that we did not observe any difference
293 between larvae from exposed and unexposed populations who were selected on a low

294 nutrition environment, when we assayed them on nutrient limited food. Yet the clear
295 difference on the common garden diet suggests there is selection on resistance between
296 these populations. We also did not identify any differences between exposed and control
297 populations from the high nutrition treatment (Fig. 1, panel A). This suggests that the
298 differences seen in the low nutrition treatments may potentially be the result of a loss of a
299 costly resistance mechanism that is non-functional in a low nutrient environment.
300



301
302
303
304 Figure 1.
305 A) Resistance of larval populations when assayed on the diet they were selected on, high
306 nutrition or low nutrition (selected), or a common garden diet. B) Resistance of control
307 populations on their selected diets at the median dose of PiGV.
308

309 *Diet determines developmental trade-offs*

310

311 To assess the effects of diet and PiGV exposure on growth rate we assayed all populations
312 on the diet they were evolved on and a common garden diet, as in the resistance assays. We
313 found no significant interactions between assay diet, selection diet, and population
314 exposure, but did find significant independent effects of assay diet and population exposure
315 (i.e. no effect of selection diet) (Fig. 2). Assay diet had the largest effect on growth rate with
316 the fastest growth rates occurring on the common garden diet compared to selected diets
317 ($\chi^2 = 157.06$, df = 6, $p < 0.0001$). There was also a significant effect of PiGV exposure during
318 selection, with populations selected for virus resistance exhibiting quicker growth rates than
319 control populations ($\chi^2 = 5.05$, df = 6, $p < 0.025$), which is counter to the trade-off observed
320 previously in this system (Boots, 2011).
321
322 As we observed wide variation in growth rates within selection treatments we correlated the
323 mean growth rate of each population to resistance on either their selected or common
324 garden diet irrespective of selection treatment. In this case, we found a strong negative
325 correlation between resistance and growth rate for the low nutrition populations on their
326 selected diet ($\chi^2 = 7.35$, df = 6, $p = 0.0067$, Fig.2), but no directional correlation for the high
327 nutrition populations on their selected diet ($\chi^2 = 0.028$, df = 6, $p = 0.866$, Fig. 2). On the
328 common garden diet we found a significant interaction between measured resistance and
329 selection diet (high or low nutrition) as predictors of growth rate ($\chi^2 = 4.72$, df = 7, $p =$
330 0.0298, Fig. 2). This demonstrates that the nutritional environment larvae are selected on
331 leads to fundamentally different costs of resistance. In this experiment low nutrient
332 environments selected for a form of resistance that is traded off with growth rate, whereas
333 in the high nutrient environment selection for resistance came at no cost to growth rate.

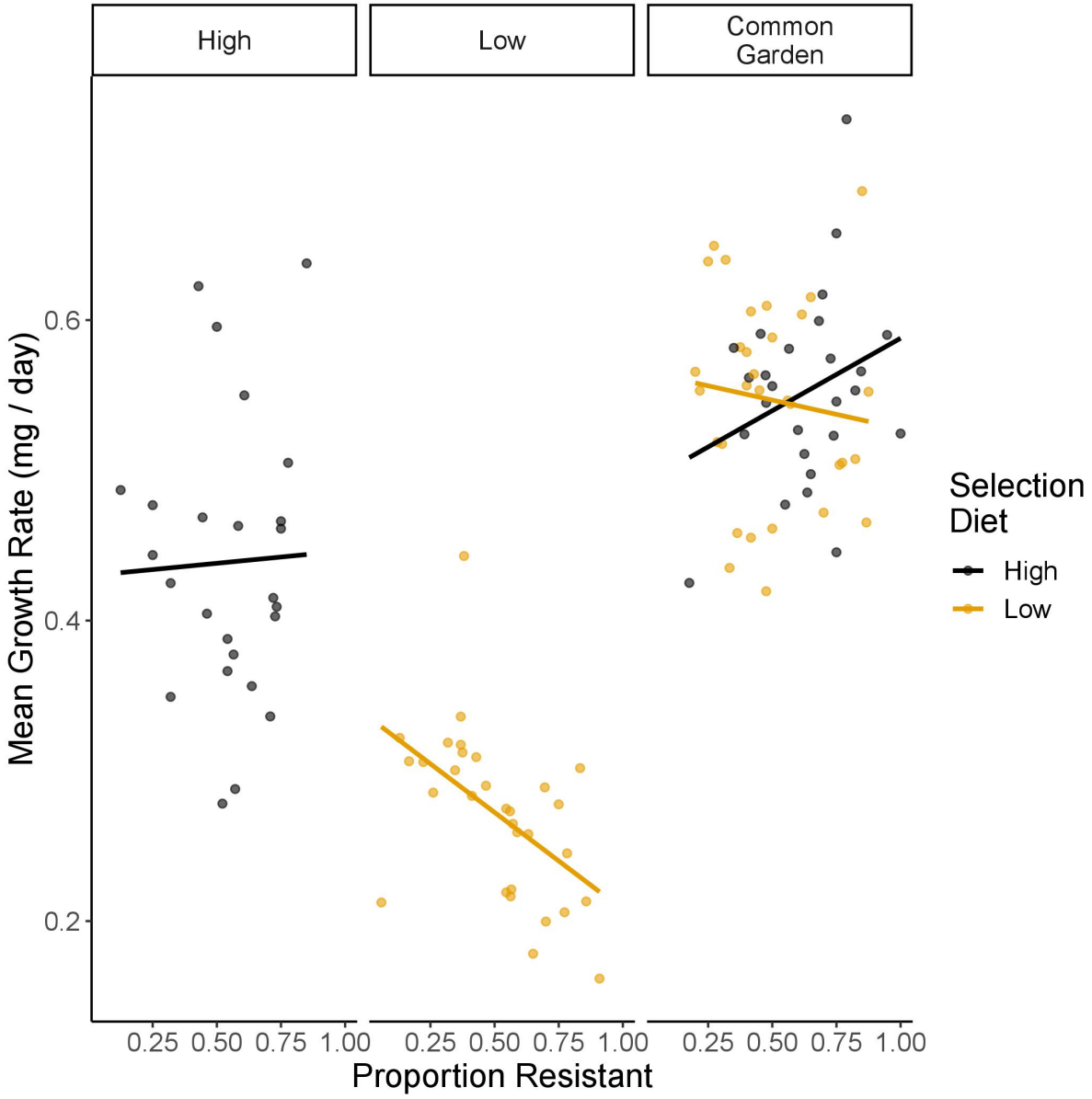
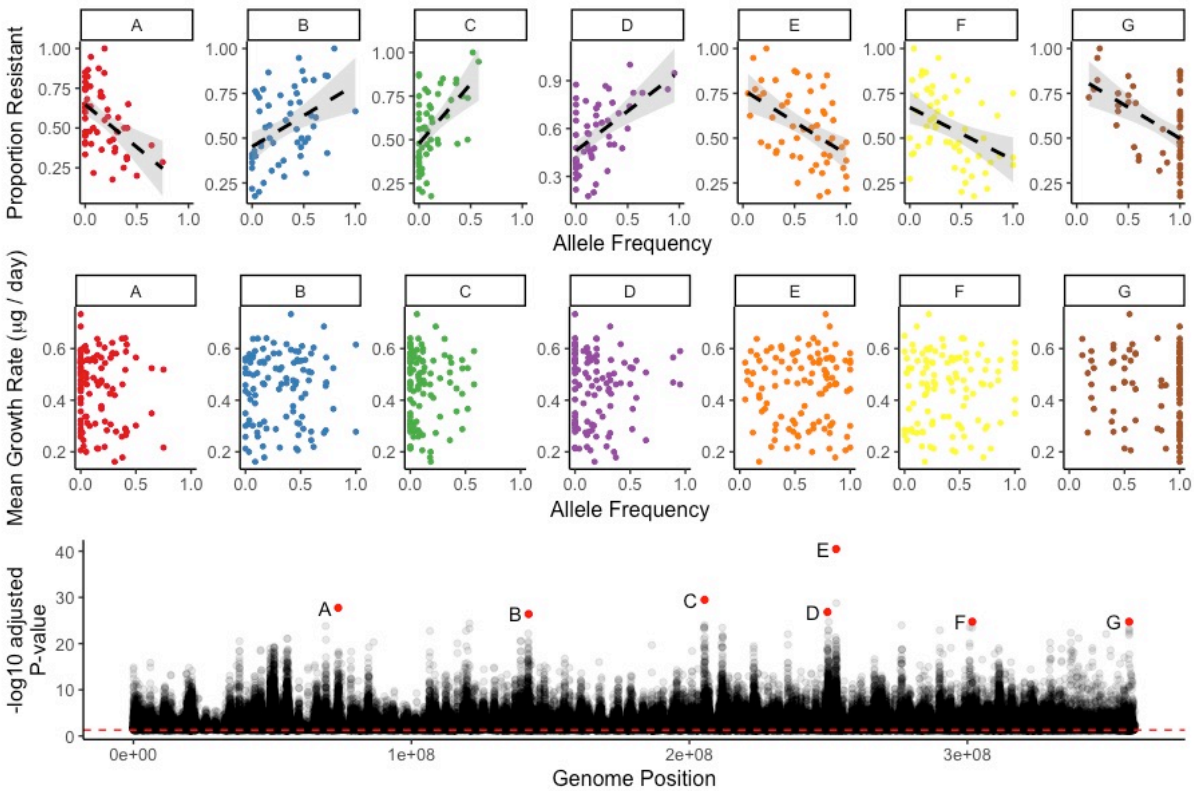


Figure 2.
Nutrient dependent trade-offs between mean population growth rate and resistance under all diets. Points represent raw data and lines are model predictions. Yellow points denote larval populations selected on a low-nutrition diet, whilst black denote populations selected on a high nutrition diet. Both resistance and growth rate are population level estimates based on either survivorship or development assays of 25 larvae per population.

Identifying the genomic basis of resistance

346 To identify the genomic basis of resistance to PiGV infection we employed a genome wide
347 scan in order to associate specific loci with resistance. This method tests the association
348 between allele frequency at each SNP present in our dataset and the resistance of each
349 population. We first assessed SNPs that predicted resistance on a common dietary
350 background i.e. the proportion of surviving larvae when assayed on the common garden
351 diet. We identified a number of candidate SNPs that were strongly correlated with
352 resistance (Fig. 3). Whilst these SNP markers were strongly correlated with PiGV resistance-
353 required by the association test used, they were not correlated with development.
354



355

356 Fig 3. Whole genome scan for SNPs associated with PiGV resistance, regardless of diet
357 (bottom panel). By definition of the methods used for the association test, the allele
358 frequency of these SNPs must be correlated with resistance (top panel). No correlations are
359 seen between allele frequency and growth rate (middle panel).
360

Whilst assessing the correlations of individual SNPs is useful for identifying strong effect loci, identifying such 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. Following gene set enrichment analysis, we found that the gene ontology term most overrepresented was calcium ion homeostasis (15-fold enrichment). Neuromuscular synaptic transmission and immune response pathways were also overrepresented and showed 8 and 7-fold enrichments respectively (Fig. 4).

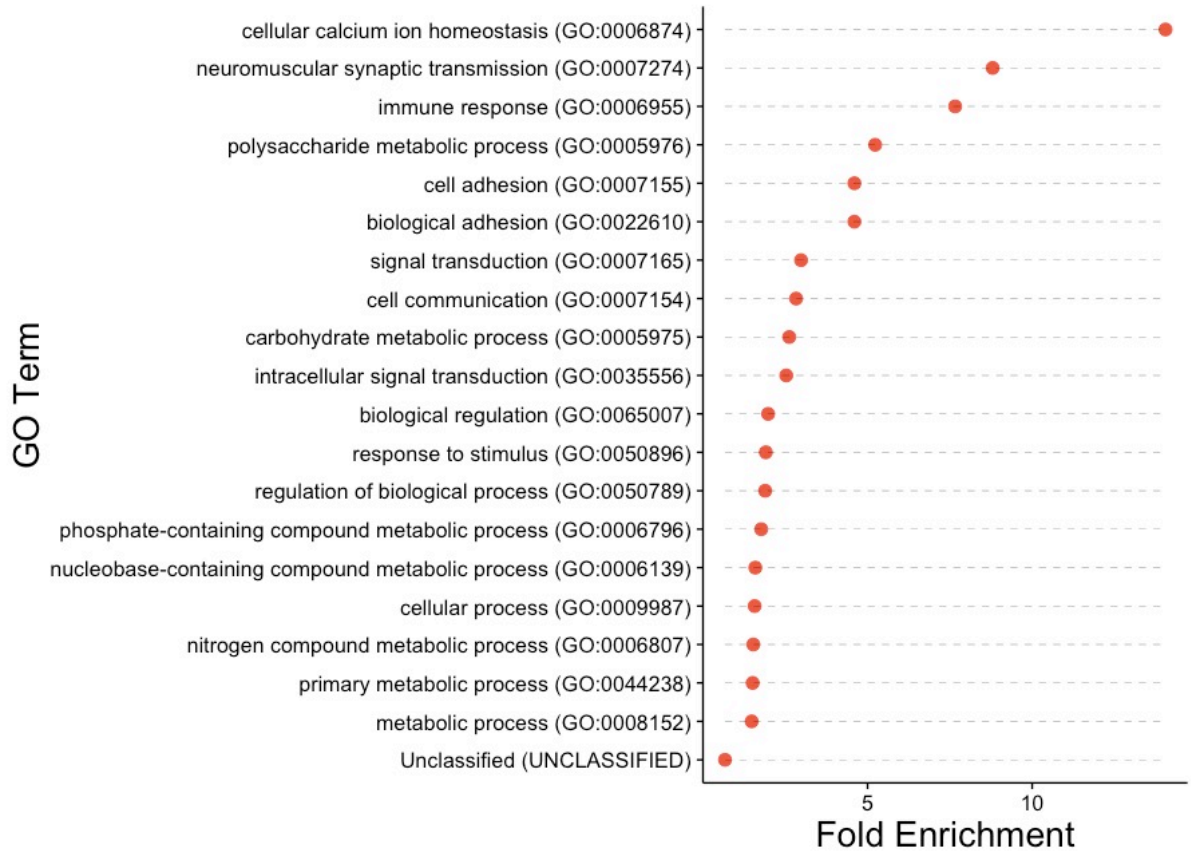


Figure 4.
Gene ontology terms significantly overrepresented in the SNPs associated with PiGV resistance in larvae assayed on a common garden diet.

Diet-specific resistance mechanisms

375 Our phenotypic data provided strong evidence that resistance evolves in an environmentally
376 dependent manner, suggesting distinct genomic routes to resistance under specific
377 nutritional conditions. This was particularly the case when we examined populations evolved
378 on a nutrient limited diet but assayed on a high resource, common garden diet. In this
379 situation, control populations exhibited much lower resistance than exposed populations. To
380 identify the mechanism that provided this difference in resistance we repeated the
381 association test for this subset of populations. Again, we found a large number of SNPs
382 strongly correlated with resistance, suggesting a polygenic trait. Enrichment analyses were
383 used to categorise these markers into biologically meaningful processes. In this case,
384 different pathways were over-represented compared to the previous analysis that compared
385 resistance on each population's respective selection diet. For populations selected on a low
386 nutrition diet but assayed on the common garden diet, cytoskeleton organization had a 5-
387 fold enrichment followed by signal transduction and cell communication (Fig. S1). As a
388 further demonstrative example, we investigated a scaffold that contained a high density of
389 SNPs associated with PiGV resistance under these conditions (Fig. 5). After running a blast
390 search on all genes on this scaffold we found a combination of developmental and metabolic
391 genes interspersed with genes linked to viral and innate immunity as well as apoptosis
392 (Table S2). Such linkage is suggestive of a pleiotropic effect or correlated selection, either of
393 which could lead to the trade-off between immunity and development we observe.
394
395
396
397
398

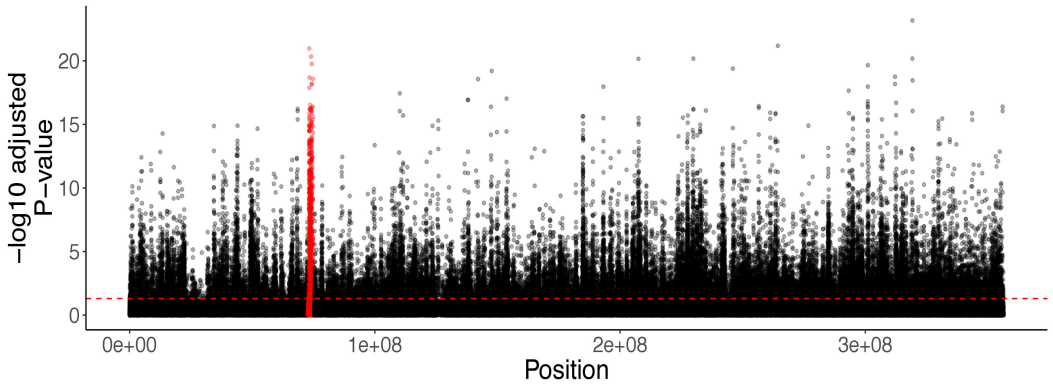


Figure 5. Manhattan plot of SNPs associated with resistance to PiGV on a low nutrition diet. Scaffold 23 is highlighted in red due to its high density of significantly associated SNPs. The putative functions of these genes are available in Table S2.

PiGV resistance is a polygenic trait

The large number of SNPs associated with resistance in both the common garden and low nutrition diets suggests that resistance to PiGV infection is likely to be a highly polygenic trait. 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. When we further assessed the proximity of SNPs, using improved scaffold lengths (increased from 0.5 to 5Mbp on average) from proximity-ligation sequencing we found many independent peaks of selection along the larger scaffolds (Fig. S2). This suggests that the genomic location of SNPs is not an artefact of the 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

420 than a signature of selection. Whilst our populations were all derived from a single ancestral
421 population, that had been out crossed repeatedly prior to the selection experiment, it is
422 possible that the observed differences in allele frequency were the result of an underlying
423 population structure. To rule this out we used a Bayesian approach to identity population
424 structure naively on independent subsets of the SNP data. We identified very weak
425 population structure suggesting that our results are unlikely to be spurious and found no
426 clustering of populations that would be indicative underlying population structure (Fig. S3).
427 We also found high reproducibility of results independent of which subset was used,
428 suggesting our methods were robust (FMD always < 0.6, see Gautier 2015).
429
430 **Discussion**
431
432 We used an evolve and re-sequence experiment (Schlötterer, Kofler, Versace, Tobler, &
433 Franssen, 2015) to identify the resource dependence of the genetic basis of resistance to an
434 insect DNA virus in an insect model system. We demonstrate that the evolution of resistance
435 is diet specific, with populations selected on a nutrient limited diet utilising a different form
436 of immunity. Surprisingly, we found no increase in resistance to PiGV exposure in
437 populations selected on a high resource diet. In contrast, we found significantly higher
438 survival in exposed populations selected on a low resource diet relative to controls.
439 However, this effect was only apparent when such populations were assayed on a common
440 garden diet, not when assayed on their low nutrition diet. Taken together, these results
441 suggest that the resistance mechanism employed on low food is likely more complex than
442 our bioassay can detect. For example, selection is acting on all life-stages during the

443 selection experiment, but our assay by necessity only tests resistance at the 3rd instar stage.
444 This difference in resistance suggests that exposed populations selected on a low nutrient
445 diet have evolved a low nutrition-specific resistance mechanism that is ineffective when
446 reared and assayed on the common garden diet.
447
448 We also observed a trade-off between resistance and growth rate in nutrient limited
449 populations. These results suggest that low nutrition environments may broadly limit the
450 evolution of effective immune mechanisms. Whilst this nutrition-specific trade-off has been
451 demonstrated before in two separate selection experiments (Boots & Begon, 1993; Boots,
452 2011) and shown to be genetic by comparison amongst inbred lines (Bartlett, Wilfert, &
453 Boots, 2018), ours is the first study to use whole-genome resequencing to identify the
454 genetic basis of such resistance. This demonstrates the genomic basis of one of the most
455 well characterized genotypic trade-offs that shapes resistance to pathogens.
456
457 Our genome scan for resistance of all populations, regardless of selection diet, identified a
458 number of biological processes involved in resistance to PiGV. A well-established
459 phenomenon in baculovirus infections is apoptosis following cellular infection (reviewed in
460 Schultz & Friesen 2009; Rohrmann 2013) and is thought to be a key immune response given
461 that many baculoviruses encode anti-apoptosis genes (Clem, Hardwick, & Miller, 1996),
462 including PiGV (Harrison, Rowley, & Funk, 2016). Although we did not identify apoptosis
463 genes directly, the most overrepresented biological process was calcium ion homeostasis,
464 which has previously been linked with baculovirus induced apoptosis (Xiu, Peng, & Hong,
465 2005). Other overrepresented biological processes include immune response and many

466 metabolic and regulatory processes. This is understandable given that after the rapid global
467 shutdown of mRNA and protein production following baculovirus infection, energy
468 metabolism escapes this shutdown (Nguyen, Nielsen, & Reid, 2013). As such, our findings are
469 broadly in line with existing mechanisms described in high nutrition diet.
470
471 Interestingly, when we picked out a number of candidate SNPs for further investigation (Fig.
472 3), we identified some potential targets for selection. Notably, SNP A (Fig. 3) encodes a
473 mediator of RNA polymerase II transcription. This is of interest as all baculoviruses studied to
474 date carry a protein that negatively regulates RNA polymerase II and is therefore likely to be
475 a crucial component of baculovirus infectivity (Nguyen et al., 2013). We also identified a
476 sodium and chloride dependent GABA transporter and a glycine receptor subunit, but it is
477 difficult to link these to any specific baculovirus infection pathways. Finally, 3 of the
478 candidates we selected were hypothetical proteins or in non-coding regions. This is
479 potentially a result of working with a non-model organism for which genomic information is
480 limited and may warrant further investigation.
481
482 As we identified a different phenotypic response in those populations selected on a low
483 nutrition diet, we repeated the genome scan on just these populations to identify the
484 putative resistance mechanism. In this instance, the most overrepresented biological
485 process was cytoskeleton organization. This is interesting as baculoviruses are thought to
486 manipulate the actin cytoskeleton during nucleocapsid transport and this is vital for
487 successful infection and replication (Volkman, 2007). Baculoviruses use a cytoskeletal
488 component (actin filaments) to both reach the nucleus and for transmission from the

nucleus following nucleocapsid production (Marek, Merten, Galibert, Vlak, & van Oers, 2011; Ohkawa, Volkman, & Welch, 2010). As such, there is likely to be strong selection on resisting such manipulations. Other overrepresented biological processes included signal transduction, cell communication and cellular process. Furthermore, when we selected a single scaffold (scaffold 23) that was highly correlated with resistance in a low nutrition environment, we found it includes genes for innate immunity, intracellular virus transport, apoptosis and development (GO terms, see table S1). The close genetic linkage of such genes goes some way to explain the trade-off between resistance and development we observed. Finally, the disparity we see between resistance on a high nutrition diet vs. a nutrient limited diet suggests that under nutrient limitation, insects are forced to invest in intracellular resistance mechanisms that are traded off against growth rate. This is reinforced by our development data, where we find a much weaker trade-off when the same nutrient limited populations are assayed on a high-quality diet.

The number of SNPs correlated with observed resistance suggests that resistance is highly polygenic. To verify this, we used Hi-C scaffolding to improve scaffold lengths and still found many independent regions correlated with resistance. Taken together, this suggests that that resistance on both diets is a highly polygenic trait. A number of complex traits have previously been found to be highly polygenic in insects (Jha et al., 2015; Kang, Aggarwal, Rashkovetsky, Korol, & Michalak, 2016), suggesting that the mechanism of resistance in our system may also be complex, rather than a small number of typical immune genes as is typical in RNA virus immunity (Magwire et al., 2012). A recent study reports that a complex genetic architecture of many interacting genes can lead to genetic redundancy, with many

competing beneficial alleles thus allowing rapid evolutionary responses. As such, it may be adaptive to rely on the complexity of polygenic traits (Barghi et al., 2018).

We have demonstrated that adaptation to a low nutrient diet can have profound effects on the underlying genetic architecture of virus resistance. We highlight a potential trade-off at the molecular level 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. Further work will be required to fully characterize these mechanisms and functional validation of mutants in genome edited insects may soon be possible. Our results have implications for understanding wild insect populations and more broadly the role of nutrition across environments on pathogen resistance.

Acknowledgements

SM and KR were supported by a NERC (UK) grant NE/J009784/1 awarded to MB and SP.

Data Availability

The data that support the findings of this study will be openly available in DataDryad prior to publication. Sequence data is available from the European Bioinformatics Institute (EBI), under accession number PRJEB27964.

References

Antonovics, J., & Thrall, P. H. (1994). The Cost of Resistance and the Maintenance of Genetic Polymorphism in Host-Pathogen Systems. *Proceedings of the Royal Society B: Biological*

Sciences, *257*(1349), 105–110. http://doi.org/10.1098/rspb.1994.0101

Barghi, N., Tobler, R., Nolte, V., Jaksic, A. M., Mallard, F., Otte, K., … Schloetterer, C. (2018). Polygenic adaptation fuels genetic redundancy in Drosophila. *BioRxiv*, 332122. http://doi.org/10.1101/332122

Bartlett, L. J., Wilfert, L., & Boots, M. (2018). A genotypic trade-off between constitutive resistance to viral infection and host growth rate. *Evolution*, *72*(12), 2749–2757. http://doi.org/10.1111/evo.13623

Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D., & Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*, *27*(4), 578–579. http://doi.org/10.1093/bioinformatics/btq683

Boots, M. (2011). The Evolution of Resistance to a Parasite Is Determined by Resources. *The American Naturalist*, *178*(2), 214–220. http://doi.org/10.1086/660833

Boots, M., & Begon, M. (1993). Trade-offs with resistance to a granulosis virus in the Indian meal moth, examined by a laboratory evolution experiment. *Functional Ecology*.

Boots, M., & Begon, M. (1993). Trade-Offs with Resistance to a Granulosis Virus in the Indian Meal Moth, Examined by a Laboratory Evolution Experiment. *Functional Ecology*, *7*(5), 528. http://doi.org/10.2307/2390128

Boots, M., & Begon, M. (1994). Resource limitation and the lethal and sublethal effects of a viral pathogen in the Indian meal moth, Plodia interpunctella. *Ecological Entomology*.

Boots, M., & Haraguchi, Y. (1999). The Evolution of Costly Resistance in Host-Parasite Systems. *The American Naturalist*, *153*(4), 359–370. http://doi.org/10.1086/303181

Boots, M., & Roberts, K. E. (2012). Maternal effects in disease resistance: poor maternal environment increases offspring resistance to an insect virus. *Proceedings of the Royal Society of London B: Biological Sciences*, *279*(1744).

Bowers, R., Boots, M., the, M. B.-P. of, & 1994, undefined. (n.d.). Life-history trade-offs and the evolution of pathogen resistance: competition between host strains. *Royalsocietypublishing.Org*.

Cantarel, B. L., Korf, I., Robb, S. M. C., Parra, G., Ross, E., Moore, B., … Yandell, M. (2008). MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, *18*(1), 188–96. http://doi.org/10.1101/gr.6743907

Clem, R. J., Hardwick, J. M., & Miller, L. K. (1996). Anti-apoptotic genes of baculoviruses. *Cell Death and Differentiation*, *3*(1), 9–16.

Cotter, S. C., Simpson, S. J., Raubenheimer, D., & Wilson, K. (2011). Macronutrient balance mediates trade-offs between immune function and life history traits. *Functional Ecology*, *25*(1), 186–198. http://doi.org/10.1111/j.1365-2435.2010.01766.x

Curtis, V., de Barra, M., & Aunger, R. (2011). Disgust as an adaptive system for disease avoidance behaviour. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1563), 389–401. http://doi.org/10.1098/rstb.2010.0117

d. C. LOPEZ-PASCUA, L., & BUCKLING, A. (2008). Increasing productivity accelerates host–parasite coevolution. *Journal of Evolutionary Biology*, *21*(3), 853–860. http://doi.org/10.1111/j.1420-9101.2008.01501.x

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 adaptation to plant resistance in the potato cyst nematode *Globodera pallida*. *Molecular Ecology*, *26*(18), 4700–4711.

582 http://doi.org/10.1111/mec.14240
583 Förstner, W., & Moonen, B. (2003). A Metric for Covariance Matrices. In *Geodesy-The*
584 *Challenge of the 3rd Millennium* (pp. 299–309). Berlin, Heidelberg: Springer Berlin
585 Heidelberg. http://doi.org/10.1007/978-3-662-05296-9_31
586 Fuxa, J. R., & Richter, A. R. (1992). Virulence and multigeneration passage of a nuclear
587 polyhedrosis virus selected for an increased rate of vertical transmission. *Biological*
588 *Control*, *2*(3), 171–175. http://doi.org/10.1016/1049-9644(92)90055-I
589 Gautier, M. (2015). Genome-Wide Scan for Adaptive Divergence and Association with
590 Population-Specific Covariates. *Genetics*, *201*(4), 1555–1579.
591 http://doi.org/10.1534/genetics.115.181453
592 Gómez, P., Bennie, J., Gaston, K. J., & Buckling, A. (2015). The Impact of Resource Availability
593 on Bacterial Resistance to Phages in Soil. *PLOS ONE*, *10*(4), e0123752.
594 http://doi.org/10.1371/journal.pone.0123752
595 Graham, A. L., Allen, J. E., & Read, A. F. (2005). Evolutionary Causes and Consequences of
596 Immunopathology. *Annual Review of Ecology, Evolution, and Systematics*, *36*(1), 373–
597 397. http://doi.org/10.1146/annurev.ecolsys.36.102003.152622
598 Harrison, R., Rowley, D., & Funk, C. (2016). The Complete Genome Sequence of Plodia
599 Interpunctella Granulovirus: Evidence for Horizontal Gene Transfer and Discovery of an
600 Unusual Inhibitor-of-. *PLoS One*.
601 Jha, A. R., Miles, C. M., Lippert, N. R., Brown, C. D., White, K. P., & Kreitman, M. (2015).
602 Whole-Genome Resequencing of Experimental Populations Reveals Polygenic Basis of
603 Egg-Size Variation in *Drosophila melanogaster*. *Molecular Biology and Evolution*, *32*(10),
604 2616–2632. http://doi.org/10.1093/molbev/msv136
605 Joshi, N., & Fass, J. (2011). Sickle: A sliding-window, adaptive, quality-based trimming tool
606 for FastQ files.
607 Juneja, P., & Lazzaro, B. P. (2009). Population genetics of insect immune responses. In *Insect*
608 *Infection and Immunity* (pp. 206–224). Oxford University Press.
609 http://doi.org/10.1093/acprof:oso/9780199551354.003.0013
610 Kang, L., Aggarwal, D. D., Rashkovetsky, E., Korol, A. B., & Michalak, P. (2016). Rapid genomic
611 changes in Drosophila melanogaster adapting to desiccation stress in an experimental
612 evolution system. *BMC Genomics*, *17*(1), 233. http://doi.org/10.1186/s12864-016-2556-
613 y
614 Knutie, S. A., Wilkinson, C. L., Wu, Q. C., Ortega, C. N., & Rohr, J. R. (2017). Host resistance
615 and tolerance of parasitic gut worms depend on resource availability. *Oecologia*,
616 *183*(4), 1031–1040. http://doi.org/10.1007/s00442-017-3822-7
617 Kofler, R., Langmüller, A. M., Nouhaud, P., Otte, K. A., & Schlötterer, C. (2016). Suitability of
618 Different Mapping Algorithms for Genome-wide Polymorphism Scans with Pool-Seq
619 Data. *G3 (Bethesda, Md.)*, *6*(11), 3507–3515. http://doi.org/10.1534/g3.116.034488
620 Kraaijeveld, A. R., & Godfray, H. C. (1997). Trade-off between parasitoid resistance and larval
621 competitive ability in Drosophila melanogaster. *Nature*, *389*(6648), 278–80.
622 http://doi.org/10.1038/38483
623 Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature*
624 *Methods*, *9*(4), 357–359. http://doi.org/10.1038/nmeth.1923
625 Lochmiller, R. L., & Deerenberg, C. (2000). Trade-offs in evolutionary immunology: just what
626 is the cost of immunity? *Oikos*, *88*(1), 87–98. http://doi.org/10.1034/j.1600-

0706.2000.880110.x

Magwire, M. M., Fabian, D. K., Schweyen, H., Cao, C., Longdon, B., Bayer, F., & Jiggins, F. M. (2012). Genome-Wide Association Studies Reveal a Simple Genetic Basis of Resistance to Naturally Coevolving Viruses in Drosophila melanogaster. *PLoS Genetics*, *8*(11), e1003057. http://doi.org/10.1371/journal.pgen.1003057

Marek, M., Merten, O.-W., Galibert, L., Vlak, J. M., & van Oers, M. M. (2011). Baculovirus VP80 protein and the F-actin cytoskeleton interact and connect the viral replication factory with the nuclear periphery. *Journal of Virology*, *85*(11), 5350–62. http://doi.org/10.1128/JVI.00035-11

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10. http://doi.org/10.14806/ej.17.1.200

Martins, N. E., Faria, V. G., Nolte, V., Schlötterer, C., Teixeira, L., Sucena, É., & Magalhães, S. (2014). Host adaptation to viruses relies on few genes with different cross-resistance properties. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(16), 5938–43. http://doi.org/10.1073/pnas.1400378111

McKean, K. A., Yourth, C. P., Lazzaro, B. P., & Clark, A. G. (2008). The evolutionary costs of immunological maintenance and deployment. *BMC Evolutionary Biology*, *8*(1), 76. http://doi.org/10.1186/1471-2148-8-76

Michalak, P., Kang, L., Sarup, P. M., Schou, M. F., & Loeschcke, V. (2017). Nucleotide diversity inflation as a genome-wide response to experimental lifespan extension in Drosophila melanogaster. *BMC Genomics*, *18*(1), 84. http://doi.org/10.1186/s12864-017-3485-0

Moret, Y., & Schmid-Hempel, P. (2000). Survival for immunity: the price of immune system activation for bumblebee workers. *Science (New York, N.Y.)*, *290*(5494), 1166–8.

Nguyen, Q., Nielsen, L., & Reid, S. (2013). Genome Scale Transcriptomics of Baculovirus-Insect Interactions. *Viruses*, *5*(11), 2721–2747. http://doi.org/10.3390/v5112721

Ohkawa, T., Volkman, L. E., & Welch, M. D. (2010). Actin-based motility drives baculovirus transit to the nucleus and cell surface. *The Journal of Cell Biology*, *190*(2).

Putnam, N. H., O'Connell, B. L., Stites, J. C., Rice, B. J., Blanchette, M., Calef, R., … Green, R. E. (2016). Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Research*, *26*(3), 342–50. http://doi.org/10.1101/gr.193474.115

Raberg, L., Graham, A. L., & Read, A. F. (2009). Decomposing health: tolerance and resistance to parasites in animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1513), 37–49. http://doi.org/10.1098/rstb.2008.0184

Rohrmann, G. (2013). Baculovirus infection: The cell cycle and apoptosis.

Sadd, B. M., & Siva-Jothy, M. T. (2006). Self-harm caused by an insect's innate immunity. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1600), 2571–2574. http://doi.org/10.1098/rspb.2006.3574

Schlötterer, C., Kofler, R., Versace, E., Tobler, R., & Franssen, S. U. (2015). Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Heredity*, *114*(5), 431–440. http://doi.org/10.1038/hdy.2014.86

Schlötterer, C., Tobler, R., Kofler, R., & Nolte, V. (2014). Sequencing pools of individuals — mining genome-wide polymorphism data without big funding. *Nature Reviews Genetics*, *15*(11), 749–763. http://doi.org/10.1038/nrg3803

Schultz, K. L. W., & Friesen, P. D. (2009). Baculovirus DNA replication-specific expression

672 factors trigger apoptosis and shutoff of host protein synthesis during infection. *Journal*
673 *of Virology*, *83*(21), 11123–32. http://doi.org/10.1128/JVI.01199-09
674 Shahrestani, P., Burke, M. K., Birse, R., Kezos, J. N., Ocorr, K., Mueller, L. D., … Bodmer, R.
675 (2017). Experimental Evolution and Heart Function in Drosophila. *Physiological and*
676 *Biochemical Zoology : PBZ*, *90*(2), 281–293. http://doi.org/10.1086/689288
677 Turner, T. L., & Miller, P. M. (2012). Investigating natural variation in Drosophila courtship
678 song by the evolve and resequence approach. *Genetics*, *191*(2), 633–42.
679 http://doi.org/10.1534/genetics.112.139337
680 Volkman, L. E. (2007). Baculovirus infectivity and the actin cytoskeleton. *Current Drug*
681 *Targets*, *8*(10), 1075–83.
682 Xiu, M., Peng, J., & Hong, H. (2005). Mitochondrial response and calcium ion change in
683 apoptotic insect cells induced by SfaMNPV. *Chinese Science Bulletin*, *50*(12), 1191–
684 1198. http://doi.org/10.1007/BF03183692
685
686
687
688