

## RESEARCH ARTICLE

# The genetic basis of variation in immune defense against *Lysinibacillus fusiformis* infection in *Drosophila melanogaster*

Brittney R. Smith, Kistie B. Patch, Anjali Gupta, Emma M. Knoles, Robert L. Unckless<sup>1</sup>\*

Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas, United States of America

\* [unckless@ku.edu](mailto:unckless@ku.edu)**OPEN ACCESS**

**Citation:** Smith BR, Patch KB, Gupta A, Knoles EM, Unckless RL (2023) The genetic basis of variation in immune defense against *Lysinibacillus fusiformis* infection in *Drosophila melanogaster*. PLoS Pathog 19(8): e1010934. <https://doi.org/10.1371/journal.ppat.1010934>

**Editor:** Alex Andrianopoulos, University of Melbourne, AUSTRALIA

**Received:** October 18, 2022

**Accepted:** June 29, 2023

**Published:** August 7, 2023

**Copyright:** © 2023 Smith et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The data that support the findings of this study are available in the online supporting material and at Figshare ([https://figshare.com/projects/The\\_genetic\\_basis\\_of\\_variation\\_in\\_immune\\_defense\\_against\\_Lysinibacillus\\_fusiformis\\_infection\\_in\\_Drosophila\\_melanogaster/163012](https://figshare.com/projects/The_genetic_basis_of_variation_in_immune_defense_against_Lysinibacillus_fusiformis_infection_in_Drosophila_melanogaster/163012)).

**Funding:** This work was funded by NIH grant R01-AI139154 to RLU. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

The genetic causes of phenotypic variation often differ depending on the population examined, particularly if the populations were founded by relatively small numbers of genotypes. Similarly, the genetic causes of phenotypic variation among similar traits (resistance to different xenobiotic compounds or pathogens) may also be completely different or only partially overlapping. Differences in genetic causes for variation in the same trait among populations suggests context dependence for how selection acts on those traits. Similarities in the genetic causes of variation for different traits, on the other hand, suggests pleiotropy which would also influence how natural selection shapes variation in a trait. We characterized immune defense against a natural *Drosophila* pathogen, the Gram-positive bacterium *Lysinibacillus fusiformis*, in three different populations and found almost no overlap in the genetic architecture of variation in survival post infection. However, when comparing our results to a similar experiment with the fungal pathogen, *B. bassiana*, we found a convincing shared QTL peak for both pathogens. This peak contains the *Bomanin* cluster of *Drosophila* immune effectors. Loss of function mutants and RNAi knockdown experiments confirms a role of some of these genes in immune defense against both pathogens. This suggests that natural selection may act on the entire cluster of *Bomanin* genes (and the linked region under the QTL) or specific peptides for specific pathogens.

## Author summary

Like most traits, the way individuals respond to infection vary among individuals within a population. Some of this variation is caused by genetic differences in the host. Over the past decade, two prominent resources were developed to assess genetic variation for complex traits of the fruit fly, *Drosophila melanogaster* and map the genetic variants responsible. We recently described a strain of *Lysinibacillus fusiformis* bacteria that was isolated from fruit flies and is moderately virulent when flies are infected via septic injury. We mapped genetic variation in resistance *L. fusiformis* using these mapping resources. We find that among the resources, different genetic changes were associated with immune defense. However, we also found that within a resource, the same region of the genome

**Competing interests:** The authors have declared that no competing interests exist.

was associated with resistance to both *L. fusiformis* and a fungal pathogen. These results suggest that different populations adapt differently to the same pathogens, but two distinct pathogens share similar causes of genetic variation within a single population.

## Introduction

Infection is ubiquitous, but the ways in which different organisms fight infection with different pathogens are likely as varied as the organisms themselves. Within species, there is often enormous genetic variation in the ability to survive infection and it is not clear whether the same host variants will provide wide ranging protection against several pathogens or whether specific host alleles provide narrow protection against one or a few pathogens [1–10]. If certain host alleles are indeed generally better at fighting infection, it follows that those alleles must either be increasing in frequency on their way to fixing in the population *or* that their fitness benefits are context dependent. Perhaps alleles conferring resistance are energetically more costly or cause dysbiosis of the native microbiota [11–13].

*Drosophila* and other invertebrates lack an adaptive immune system, which allows us to focus on how the innate immune system functions in infection and disease [14]. As a result, some of the initial insight about immune defense pathways were gained by studying *Drosophila* and such advances continue today [14–17]. Pioneering work in the 1980s provided insight into the Toll and Imd pathways which are largely conserved in vertebrates [18]. Recently, researchers have worked out more specific associations between immune effectors and infection outcomes [5,19,20]. In addition to the functional genetic work on the immune system, others have focused on the role of natural genetic variation in resistance to infection [1,21–23]. In some cases, those studies have inspired a deeper functional genetic dissection of the roles of specific immune genes in infection [24].

Insect innate immunity can be broadly grouped into cellular and humoral immunity. Cellular immunity involves the differentiation of blood cells that have several roles in directly attacking and/or immobilizing pathogens [25]. In contrast, the humoral immune response involves signaling cascades that generally lead to enormous induction of immune effectors [26]. In *Drosophila*, the canonical view is that the Toll pathway is activated in response to Gram-positive and fungal infection and stimulates expression of immune and antimicrobial peptide genes (*Bomanins*, *Daishos*, *Drosomycins*, etc.) [14,27,28]. This canonical view suggests that the Imd pathway is activated in response to Gram-negative infection and stimulates the expression of antimicrobial peptides (*Diptericins*, *Cecropins*, *Defensins*, *Attacins*, etc). Other pathways including JNK and JAK/STAT also play important roles and likely interact with Toll and Imd [29,30].

Dissecting the genetic bases of phenotypic variation in a phenotype such as immune defense yields two important results that are fundamentally different than candidate gene knockout approaches. First, because the associated approaches are unbiased, they often reveal genes previously unknown to be involved in immune pathways. Second, an understanding of the genetic causes of phenotypic variation informs both genetics and evolution. There are two commonly used *D. melanogaster* panels used for mapping genetic variation in phenotypes. The first is the *Drosophila* Genetic Reference Panel (DGRP) which was derived from a set of gravid females collected at a farmer's market in North Carolina, USA [31]. The offspring of those gravid females were inbred for several generations, then 205 lines were sequenced and made available to the public. This allows for a standard genome-wide association study (GWAS) approach to determining the genetic basis of phenotypic variation and has been used

widely for several different phenotypes [32–37], including response to infection [21,38,39]. The second panel is the *Drosophila* Synthetic Population Resource (DSPR) [40,41]. This set of lines (actually, two mostly independent sets) were derived from eight inbred founder lines that were allowed to interbreed for several generations, then were inbred. Founders were sequenced to high depth and the recombinant inbred lines (RILs) were genotyped at low depth to assign founder genotypes across the genome. This approach allows a quantitative trait locus (QTL) approach [42,43]. The DSPR lines have not been used for studies of genetic variation in infection, except for the fungal insect pathogen *Beauveria bassiana* [44], *Drosophila* Sigma Virus [24], and *Drosophila* C Virus [45]. Mapping genetic variation for resistance to infection (or any trait) requires that there be variants at high enough frequency in the mapping population to have statistical power to see an effect. These three mapping panels are derived from three pools of segregating alleles and whether the variants that influence susceptibility to infection are likely to be present in multiple panels depends on the frequency of those variants in natural populations, and in the case of the DSPR, the extent to which they are represented in the recombinant inbred lines. Therefore, a locus that shows significant association in one panel but little or no association in another panel may lack sufficient genetic variation at that locus in the second panel to find a significant association. This does not mean that the locus is unimportant for the trait in that population, just that the genetic variation is insufficient to show an association.

Our study investigates natural variation in defense against the Gram-positive *Lysinibacillus fusiformis*, which was isolated from *Drosophila* in the lab [46]. *L. fusiformis* is the first member of the *Bacillaceae* family to be used in a mapping study in *Drosophila*. It is an outlier for bacterial infection because it contains lysine, alanine, glutamic acid, and aspartic acid in its peptidoglycan instead of the standard Diaminopimelic acid (DAP) in most Gram-negative bacteria or l-lysine-(Lys) for Gram-positive bacteria [47]. It is a generalist, found associated with human respiratory illness [48], in symbiotic relationships with ticks [49], and in *Drosophila* stocks [46]. We measure survival 2- and 5- days after septic injury with *L. fusiformis* in the DGRP population and both DSPR populations (A and B). We also perform RNA-sequencing of flies 12-hours post infection and find that induction of gene expression after infection with the Gram-positive *L. fusiformis* is better correlated with induction of gene expression after Gram-negative microbes than with other Gram-positive microbes. This may be because the peptidoglycan, which is responsible in part for immune induction, of *L. fusiformis* is *not* similar to that of most Gram-positive microbes. Though the DGRP analysis reveals an excellent candidate gene for susceptibility to *L. fusiformis*, we find very little overlap in significantly associated SNPs with other immune-related studies. In contrast, the DSPR A population has a QTL peak that overlaps with a single peak in the *B. bassiana* study. While this peak contains several genes, we explore the possibility that variation in the same genes is responsible for infection differences for both *L. fusiformis* and *B. bassiana*. Our study highlights the importance of having relatively high frequency genetic variants to be able to draw conclusions about similarities among mapping studies, and the likelihood that haplotypes may have similar susceptibilities to very different pathogens.

## Results

### ***L. fusiformis* is moderately virulent and leads to induction of genes involved in immune defense**

We infected males from the standard *D. melanogaster* lab line, Canton S, with *L. fusiformis* at an optical density (OD<sub>600</sub>) of 3.977 and measured survival for seven days post infection (DPI). At this dose, 76% of flies (51 of 67) survived to 7 days, while all flies pricked with a sterile

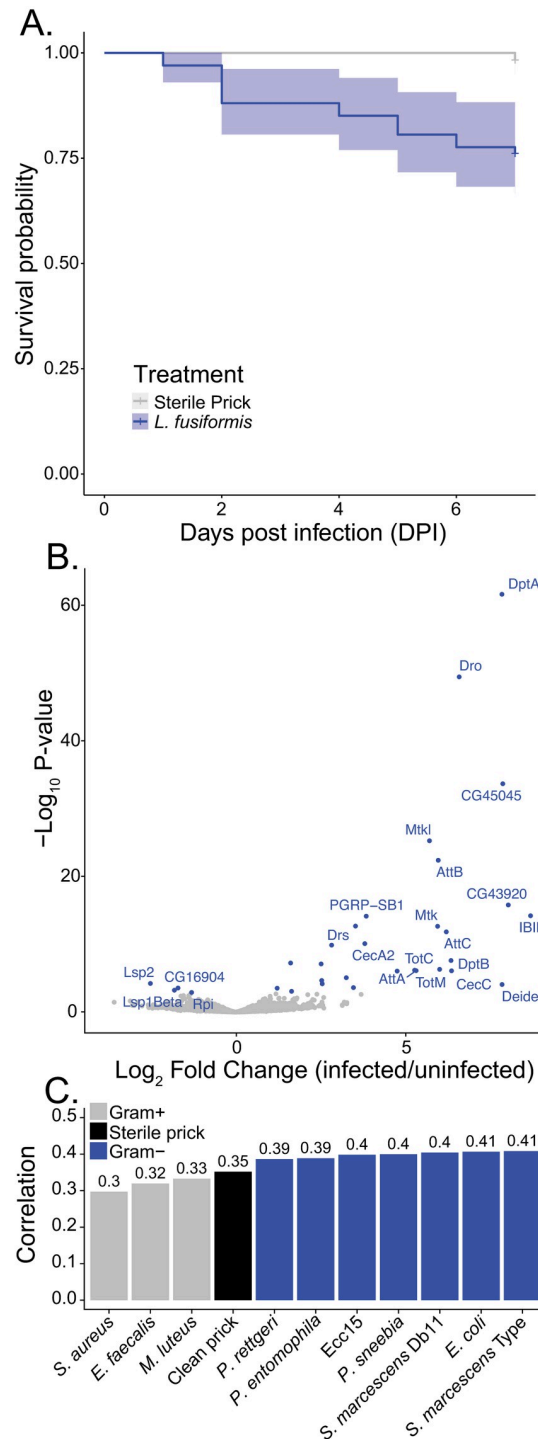
needle survived (Fig 1A). This represents a significant, but moderate reduction in survival (Cox Proportional Hazard test, hazard ratio = 16.1,  $P = 0.007$ ). Note that to achieve sensible hazard ratio, we called one sterile pricked fly dead at day 7. At this point, we could have elected to use a higher dose for subsequent infection experiments, but the  $OD_{600}$  around 4.0 led to sizeable variation in survival among different lines (both DGRP and DSPR had lines with survival rates close to zero), so we continued using this dosage throughout the remainder of the experiment.

We performed RNA-sequencing of pools of males 12 hours post-infection with *L. fusiformis* to examine changes in gene expression upon infection (Fig 1B). The most highly induced genes are expected based on other studies [50–52] and are mostly antimicrobial peptides in the Toll and Imd pathways. We used an adjusted P-value threshold of 0.1 for Gene Ontology enrichment analysis, which yielded 36 genes differentially expressed. Gene Ontology categories significantly enriched included several related to immune defense: response to bacteria, response to external biotic stress, defense response, immune response, defense response to Gram-positive bacterium, defense response to Gram-negative bacterium (see the entire list in S1 Table).

It appears that the induced antimicrobial peptides are more squarely induced by the Imd pathway instead of the Toll pathway, suggesting induction of gene expression is more like Gram-negative bacteria than Gram-positive bacteria. The *Lysinibacillus* genus are Gram-positive bacteria, but like others in the Bacillales order, the peptidoglycan is not l-lysine-(lys) type. Instead, most *Lysinibacillus* examined have the A4alpha (Lys-Asp) peptidoglycan [53,54]. Since the humoral response in *D. melanogaster* is driven in part by recognition of peptidoglycan type [55,56], the expected relative responses of the Toll and Imd responses to Bacillales is less clear cut than for most bacteria that fit more neatly into Gram-positive and Gram-negative categories based on peptidoglycan [14]. To determine whether the transcriptional response to *L. fusiformis* infection is indeed more “Gram-negative-like”, we calculated Pearson correlation coefficients for  $\log_2$  Fold Change (infected vs. uninfected) between our *L. fusiformis* data and 12-hour post infection gene expression for 11 bacterial infection treatments from the Troha *et al.* [50] data. Although all correlations were significant ( $P < 0.0001$ ), there was stronger correlation between *L. fusiformis* and Gram-negative bacteria (median = 0.4) than any of the Gram-positive bacteria (median = 0.32, Fig 1C). The same pattern is true when we only considered significantly differentially expressed genes (correlations for Gram-positive range from 0.798 to 0.832 and for Gram+ range from 0.910 to 0.949) or only significantly induced genes (correlations for Gram-positive range from 0.560 to 0.653 and for Gram+ range from 0.786 to 0.903), and these correlations are slightly stronger if we compare induction after *L. fusiformis* infection to the average induction after Gram-positive infection (Pearson correlation coefficient = 0.336) or Gram-negative infection (Pearson correlation coefficient = 0.433). Thus, *L. fusiformis* is a moderately virulent bacterial pathogen that induces a strong innate immune response that is more like the transcriptional response to Gram-negative than Gram-positive bacteria.

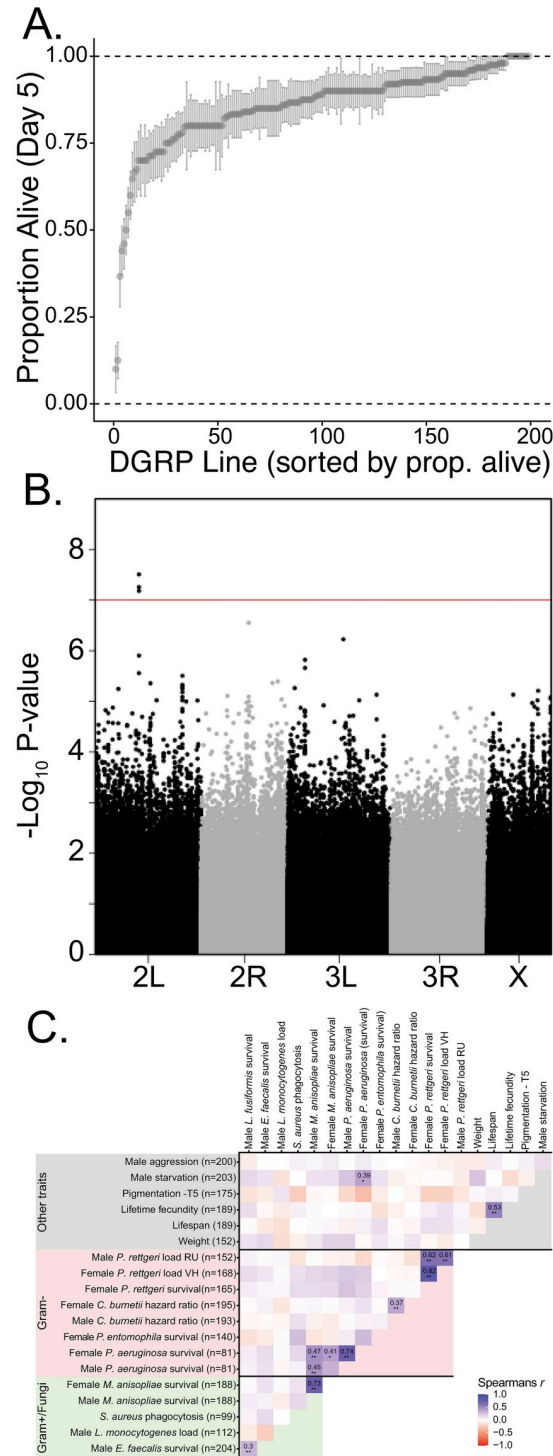
### Genetic variation for immune defense against *L. fusiformis* in the DGRP

We aimed to map the genetic causes of variation in resistance to infection in three (DGRP and two DSPR) populations. All three populations showed genetic variation for survival after infection with *L. fusiformis* (Figs 2A, 3A and 4A). Raw survival in the DGRP population ranged from 10% to 100% surviving 5 days post infection (DPI). We also examined survival 2 days post infection, which was much higher overall and showed less variation (S1 Fig). Like several other studies of immune defense in the DGRP, we found that line 321 was particularly



**Fig 1. *Lysinibacillus fusiformis* infection is moderately virulent.** A) Canton S male survival after infect with *L. fusiformis* ( $OD_{600} = 3.977$ , blue,  $n = 67$ ) or sterile prick (gray,  $n = 59$ ) with 95% confidence intervals. B) Volcano plot of gene expression in Canton S males 12 hours post infection with  $OD_{600} = 3.977$  *L. fusiformis* vs. uninfected (no sterile prick). Blue points are significantly differentially expressed (FDR-adjusted P-value < 0.05) and genes of interest are labeled. C) Pearson correlation coefficients for  $\log_2$  fold change between infections with *L. fusiformis* and microbes from Troha et al. 2018.

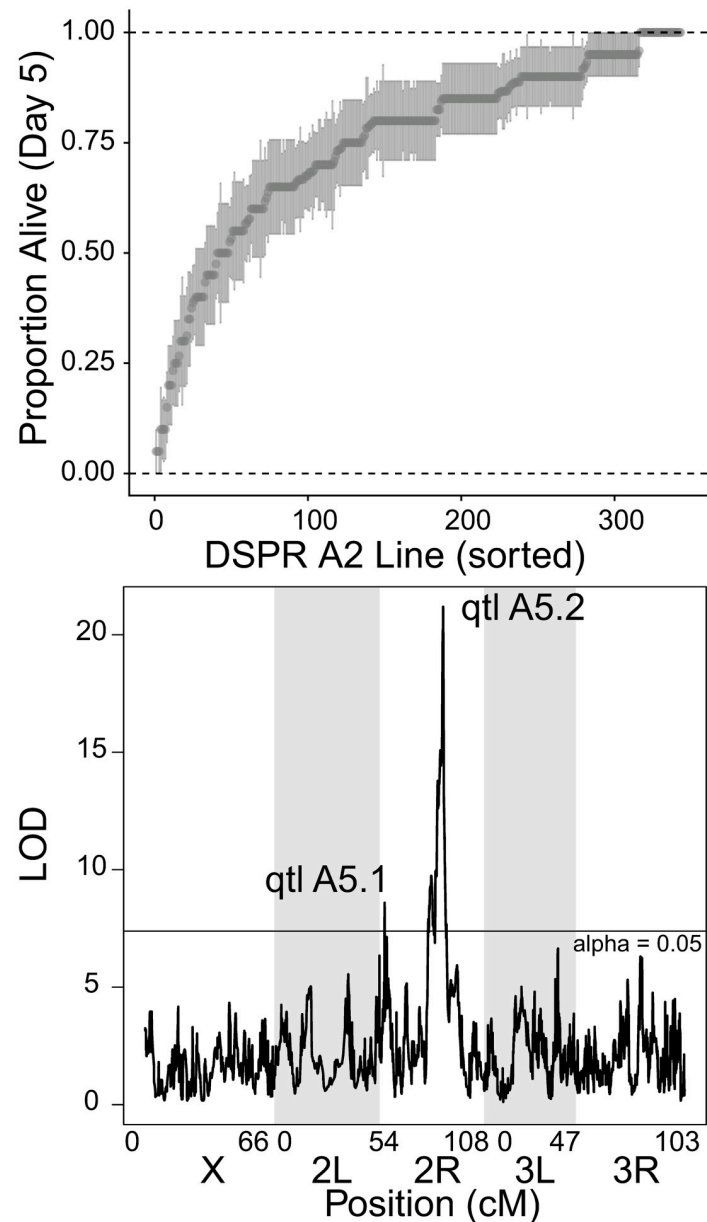
<https://doi.org/10.1371/journal.ppat.1010934.g001>



**Fig 2. Genetic variation in immune defense against *L. fusiformis* in the DGRP.** A) Sorted raw line mean proportion alive 5 days post infection with standard error of the proportion. B) Manhattan plot of genome wide associations with 5-day survival random effects. C) Correlation of line effects from several DGRP immune studies. Sample sizes in parentheses.

<https://doi.org/10.1371/journal.ppat.1010934.g002>



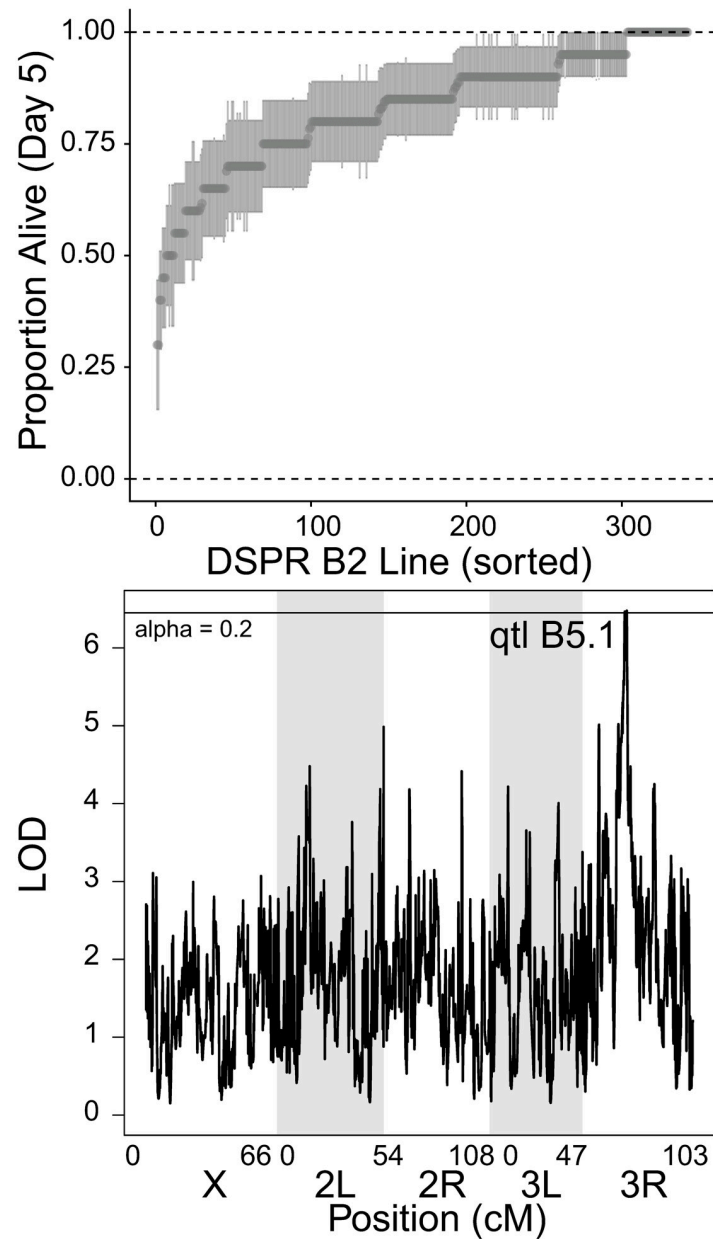


**Fig 3. Genetic architecture of immune defense against *L. fusiformis* in DSPR population A2.** A) Raw line means for proportion alive at day 5 post infection with standard error of the proportion sorted by mean proportion. B) QTL plot for line effects of survival post infection with *L. fusiformis*. Two QTLs are labeled, horizontal line shows a significance threshold with  $\alpha = 0.05$ .

<https://doi.org/10.1371/journal.ppat.1010934.g003>

susceptible to infection, so we calculated line effects with and without line 321 (S2 Table and S2 Fig). Survival after infection in *Drosophila* is an inherently noisy phenotype, so it is important to note how much of the variation in survival is due to genetic variation (the heritability). The heritability in the DGRP was 0.802 for survival 5 DPI (0.759 if line 321 is removed) and 0.924 for survival 2 DPI (0.890 if line 321 is removed).

For associations in the DGRP, we used the online DGRP mapping tool because it accounts for inversions, *Wolbachia* infection status and the relatedness of lines. We performed the



**Fig 4. Genetic architecture of immune defense against *L. fusiformis* in DSPR population B2.** A) Raw line means for proportion alive at day 5 post infection with standard error of the proportion sorted by mean proportion. B) QTL plot for line effects of survival post infection with *L. fusiformis*. Two QTLs are labeled, horizontal line shows a significance threshold with  $\alpha = 0.2$ .

<https://doi.org/10.1371/journal.ppat.1010934.g004>

mapping for 5-day survival and 2-day survival both with and without line 321). We chose an FDR-adjusted cutoff of  $p < 0.10$  as our significance threshold, and few SNPs were significant at that cutoff (Figs 2B and S3 and S3–S6 Tables). Below we discuss the different mapping approaches but put the most faith in mapping results of 5-day survival when we discarded line 321. S4 Fig show that when line 321 is included (S4A and S4C Fig), the QQ plots show a characteristic inflation of significance (lower p-values) over expectations. This is likely due to the relatively small sample size in the DGRP meaning that sites that are shared between line 321



and other lines by chance (and not genetic linkage) will show inflated significance. In contrast, the omitting line 321 and examining survival 5 days post infection yields a QQ plot that mostly fits expectations with a few outliers that are more significant than expectations.

The FDR-adjusted significance threshold of 0.1 yielded only a handful of significant SNPs (S3–S6 Tables). Sixteen SNPs were significantly associated with survival 5 DPI when all lines were considered. Eleven intronic SNPs were in *eyf/CG9967*, *Pvr*, *CG44153*, *Trim9*, *CG15611* (4 SNPs), *trio*, *CG32698* (actually a 2bp deletion), and *Frq1*. One SNP was in the exon of the non-coding RNA *CR44310*, one SNP caused a silent change in *CG14971*, and 3 SNPs were intergenic. When we omitted line 321, only 3 SNPs were significantly associated with 5-day survival—all three were intronic SNPs in *Pvr*. Note that the SNPs significantly associated when all lines were included but not significantly associated when 321 was omitted were mostly (13 out of 14) still tested but had low enough minor allele frequencies that power was lost and P-values were above our significance threshold. Eight SNPs were significantly associated with 2-day survival when all lines were considered. This included 4 intronic SNPs in *Zasp52*, *trio*, *CG32364*, and *Plod* as well as 4 SNPs in the 3' UTR of *CG15564*. There were no significantly associated SNPs 2-days post infection when line 321 was omitted. The SNPs in *Pvr* significantly associated at 5-days post infection were nominally significant at 2-days post infection ( $\sim 3.0 \times 10^{-4}$  when all lines were considered and  $\sim 4.0 \times 10^{-5}$  when line 321 was omitted). With our conservative cutoff for significance, only one SNP (3L:1030175, intronic in the Rho guanine exchange factor, *trio*) was significantly associated at both 2-days and 5-days post infection. Similarly, only one SNP (2L:8232234, intronic in *Pvr*) was significantly associated at 5-days post infection whether line 321 was included or not.

We investigated genes connected to significantly associated SNPs in a few ways. First, we examined whether any of these genes were differentially expressed 12 hours after infection based on our RNA-seq data. None were significantly differentially expressed even if we consider the nominal P-values, so genes connected to significantly associated SNPs are not significantly induced or repressed upon infection at the 12-hour time point. Relying on transcriptional responses to infection assumes that genes must be induced to have a role in infection. This is clearly not always true but provides one line of corroborating evidence.

Next, we examined annotated functions of these genes. Two stand out. First, *Pvr* (*PDGF- and VEGF-receptor related*) is involved in the regulation of several innate immune functions including apoptosis, humoral immune response, JNK signaling and wound healing [57–61]. *Pvr* is a negative regulator of humoral immune response, so if these SNPs are associated with alternative splicing or otherwise influence *Pvr* expression, it would be a plausible connection. We performed RNAi knockdowns of *Pvr* using ubiquitous drivers (which produced no offspring), and hemocyte and fat body drivers. Though not significant, the knockdown genotypes survived better after infection with *L. fusiformis* than controls for the hemocyte driver and one of the fat body drivers (S5 Fig). We used a higher dose ( $OD_{600} \sim 10$ ) in these experiments after pilot experiments suggested *Pvr* knockdown enhanced survival. Given that *Pvr* is a negative regulator of humoral immunity, the increased survival makes sense, but we hesitate to draw any conclusions from this data since the differences were not significant. The other gene of interest is *CG44153*, which is predicted to enable cell-cell adhesion mediator activity and is active in the plasma membrane [62]. It stands out because it was also identified in two other DGRP immunity association studies: survival time after infection with *Pseudomonas aeruginosa* [21], and microbiota-related diet shift [63]. However, *CG44153* also was reported to harbor genetic variation associated with several other non-immune phenotypes, so it is possible that there is something about genetic variation at the locus that makes it prone to false positive associations.

### Little evidence for correlations among immune phenotypes in the DGRP

The DGRP lines have been available to the public for more than a decade, and in that time, there have been several studies attempting to map variation in immune defense against different pathogens. This allows us to begin to determine whether immune defense against different pathogens is correlated. Are lines that are better at fighting infection against one pathogen, generally better at fighting infection? If so, is this true universally or within categories such as Gram-positive bacteria? A general correlation in immune defense would suggest that at least some of the genetic basis of that variation is also shared for these different pathogens. There is some evidence for this in the overlap in associations in this study and those of Chapman *et al.* and Wang *et al.* [1,21]. However, it is also possible that resistance is much more specific for each pathogen. This is likely given some of the specificity of associations—particularly for *Providencia rettgeri* and *Diptericin* [5,20,38].

We calculated Spearman's rank correlations for 14 DGRP immune-related data sets plus an additional 6 non-immune phenotypes (Fig 2C). Though there are several positive correlations, few are significantly correlated after Benjamini & Hochberg correction. Correlations tend to be between the same microbe using different experimental design (different sex or survival vs. bacterial load) or between different microbes measured in the same laboratory (*E. faecalis* and *L. fusiformis* were both measured in our laboratory at the same basic time; *M. anisopliae* and *P. aeruginosa* were reported as part of the same experiment in Wang *et al.* [21]). However, there is an almost complete lack of significant correlation between immune phenotypes and any of the other phenotypes, the exception being a positive correlation between female survival after *P. aeruginosa* and male starvation survival. Given the general lack of correlation between phenotypes, we decided not to look at the overlap in significantly associated SNPs between immune phenotypes. Most of the phenotypes we chose utilized nearly all DGRP lines (170–200 of the 205 DGRP) lines. However, the number of pairwise observations for some of these correlations might be small enough that power to detect significant correlations is reduced. Differences in the lines used in different experiments also enhances the likelihood that traits measured in the same (or related) labs show correlation since those labs are likely to use a similar set of lines. Nevertheless, *Pseudomonas* infection phenotypes ( $n = 81$  for *P. aeruginosa*) were correlated with each other and with an enteric *Pseudomonas entomophila* infection. Our correlation results suggest that the DGRP immune phenotypes that have been reported do show some correlation, but it is difficult with the available data to determine if these correlations are due to phylogenetic relatedness of the microbes used or are an artifact of something about the experimental design (how phenotypes were defined, dose, overlap in lines utilized, etc.) or labs where the experiments were performed.

### Genetic variation for immune defense against *L. fusiformis* in the DSPR

The DSPR consists of two populations (A and B) that were derived from eight different founder genotypes. Each of those populations was further split into two parallel sets (*e.g.* A1 and A2) which were derived from the same founders, but different interbreeding populations. We infected 343 lines from population A2 with *L. fusiformis* ( $OD_{600} \sim 4.0$ , mean = 22.3 male flies per line, median = 20 male flies per line, standard deviation = 6.47 male flies per line). The overall survival rate 5 days post infection was 74.76% and the line mean raw survival was 75.16% (median = 80%, standard deviation = 21.11%, Figs 3A and S6A for Day 2). Heritability was 0.828 for survival 5 DPI and 0.928 2 DPI.

We used the same logistic regression model as described above to calculate line effects for each DSPR population A line (S7A Fig and S7 Table), and used the DSPRqtl package [64] to determine regions of the genome associated with variation in immune defense against *L.*

*fusiformis*. We determined a significance threshold ( $\alpha = 0.05$ ) by permuting the data for 1000 iterations. This yielded several QTL peaks with LOD scores above the threshold, but they concentrate in two regions (Fig 3B). The first peak (qtl A5.1) was centered at Chr2R:3,420,000 (LOD drop 2 range 3,360,000 to 3,470,000, note this is release 5 of the *D. melanogaster* genome, LOD = 8.59) and was associated with 10.96% of the variation in the phenotype. We extracted the effects of individual founder genotypes (S8A Fig) and found that founder A1 genotypes at this region had significantly higher survival than the other four genotypes represented. This includes all or part of 21 genes, none of which have obvious immune function, nor were any differentially expressed in our RNA-seq data after p-value correction. One gene, *Cyt-b5* showed an almost 50% reduction in expression upon infection, leading to a nominal p-value of 0.02. Importantly it is annotated to be involved in metabolism of xenobiotic compound, but the naïve expectation if it is breaking down microbial toxins would be induction, not suppression upon infection.

The second peak (qtl A5.2) had a higher LOD score and was also much broader. We focused on the main peak (highest LOD score) which was centered at Chr2R:14,720,000 (LOD drop range from 14,680,000 to 14,770,000, LOD = 21.19) and was associated with 24.89% of the variation in phenotype. This peak contains all or part of 19 genes including two genes involved in immune defense: *DptA* and *DptB*. The founder genotype from A5 is associated with significantly lower survival than the other represented genotypes (S8B Fig). If we expand to look at genes under the three adjacent (actually overlapping) peaks with LOD scores above 20, we also include the gene, *Jabba*, which functions in immune defense. One more QTL peak still overlapping these but with a LOD score of 14.7 (peak 20) contains the *Bomanin* gene cluster (eight small, secreted peptides involved in immune defense) as well as *Imd* (adaptor protein for the *Imd* pathway). Line A5 harbors two unique nonsynonymous SNPs in *Jabba* exon 6 as well as a loss of stop codon in *BomT2* (a *Bomanin*). Line A2 (not susceptible) contains a 1748bp insertion of a Jockey element in the intron of *BomS4* (the intron is 64bp in the other lines). The two *Diptericins* are tempting candidates, but neither have any variation unique to line A5 in either coding or regulatory regions. Furthermore, the DGRP lines segregate for null alleles of *DptA* and those lines are right in the middle of the line effects (51<sup>st</sup> and 53<sup>rd</sup> percentile). There is also a single line with a null mutation (premature stop) in *DptB* in the DGRP (line 223) and it showed 100% survival. In the DSPR, line A3 has a 37bp deletion that starts in exon 1 of *DptB* and ends in the intron, and that genotype has no apparent deficit in immune defense against *L. fusiformis*. We also investigated variation in other genes with known immune functions. Prophenoloxidase gene, *PPO1*, is involved in the melanization reaction. While line A5 (susceptible DSPR line) did have a single amino acid difference compared to other lines (a conservative arginine to lysine at the 420<sup>th</sup> residue), line A7 includes single base deletion leading to a frameshift and a premature stop codon only 44 amino acids into the otherwise 690 residue protein. We chose not to investigate *PPO1* further since prior work showed no effect of null PPO1 mutants except in response to infection with *Staphylococcus aureus* [65,66]. Finally, we examined the coding sequence of *Imd* and found only a single amino acid difference (A2 has a glycine to serine change at residue 103 compared to the others) among any of the four lines (A2, A5, A6 and A7) with representation at QTL A5.2.

Line effects and mapping in the DSPR A2 population were correlated when comparing Day 2 and Day 5 survival (S7A, S7B, S7F and S8 Figs and S8 Table, Pearson correlation coefficient for line effects = 0.908, Pearson correlation coefficient for LOD scores = 0.957). While the peaks overlapped significantly, the Day 2 survival did nothing to further filter to a smaller set of candidates. We therefore utilize Day 5 survival data for population A2.

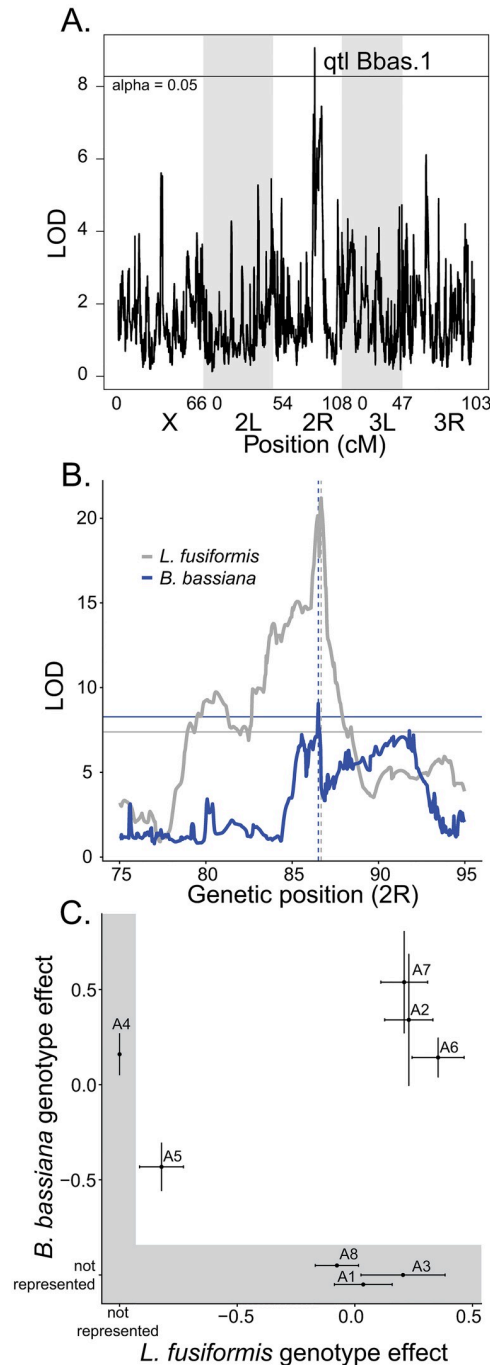
We infected 391 lines from the B2 DSPR panel with a mean of 18.81 (range 10–20, median = 20) individuals infected per line. The mean survival was 0.828 (median = 0.85, st.

dev. = 0.134). There was again ample variation in survival across lines (Figs 4A and S7C and S9 Table). Heritability was 0.674 for survival 5 DPI and 0.760 for survival 2 DPI. Thus, both heritability and the range of survival post infection are higher in the DSPR A panel (Figs 3A vs 4A), which may have resulted in reduced power for mapping in the B panel. The QTL scan did not result in any significant peaks with  $\alpha = 0.05$  (Fig 4B). In fact, we had to relax  $\alpha$  to 0.20 to get significant peaks. There were 3 overlapping peaks on Chr3R between 15,470,000 and 16,770,000—a region with 145 genes (qtl B5.1). This peak explained 7.35% of the variance. There are several genes plausibly associated with immune defense. First is *Stat92E*, the transcription factor of the JAK/STAT pathway (not significantly induced, but expression was up 25.7%). Second are a cluster of *Turandots*, immune/stress effectors regulated by JAK/STAT. Of the five *Turandots* under the peak, *TotA*, *TotC* and *TotX* were all significantly upregulated after infection and there was no expression data for *TotB* or *TotC*. The founder genotypes under these peaks show that lines B5 and B7 have significantly lower survival than the other lines (S9C Fig). Scanning through variation in these genes, we did not find any sites limited to B5 and B7 that were immediately good candidates for being causal. For panel B2, the correlation between LOD scores for Day 2 and Day 5 was lower and what we observed in population A (S9 and S10 Tables). In fact, there was an additional peak at Day 2 (qtl B2.1, S10 Fig). This peak is only significant with  $\alpha = 0.1$  and is on 3L at 2,880,000 (LOD score = 7.57, 2 LOD drop range: 2,850,000 to 3,100,000, 8.53% of variance explained). Founder line B3 genotypes had a significantly lower survival than others (S8D Fig). The peak includes 27 genes including *spz-5* (a paralog of the canonical *Toll* pathway *Spatzle* gene). These genes were not differentially expressed in the transcriptome data.

To further investigate overlap between the DGRP and DSPR populations, we looked for DGRP SNPs significantly associated with survival after *L. fusiformis* infection that were under the three focal QTL peaks (A5.1, A5.2 and B5.1). For the A populations, there were no SNPs that were significantly associated with survival after infection (adjusted  $P > 0.85$  where adjustment accounted for just those SNPs under the QTL peak). For the B population, there was a single SNP that, after multiple test correction was associated at  $p = 0.091$ , but with multiple corrections for the entire data set, the adjusted  $p$ -value was 1.0. We conclude that any causative genetic variation segregating in the DGRP in the regions of the three DSPR peaks was either absent or at too low frequency to result in a significant association.

## Defense against *L. fusiformis* shares a similar genetic architecture with defense against a fungal pathogen

We reanalyzed the data from Shahrestani *et al.* [44], where DSPR panel A1 was infected with *Beauveria bassiana* (a fungal pathogen) and survival proportion was subsequently mapped. We performed the same logistic regression (described above), which is different from their approach, and used other variables (all random effects) as they described. Our heritability estimate was 0.921 which is much higher than what they estimated (0.53). We cannot examine correlations among lines for immune defense against *B. bassiana* and *L. fusiformis* because the experiments were done in different lines (set A1 vs. A2). However, we can determine whether there is significant overlap in QTLs, which would suggest a shared architecture for immune defense. Like *L. fusiformis*, there was significant variation in the random effects of line for survival after *B. bassiana* infection (S7E Fig and S11 Table). In contrast to what Shahrestani *et al.* [44] reported, we did find one peak significantly associated with survival after *B. bassiana* infection at  $\alpha = 0.05$  (Fig 5A and S12 Table). The peak centered at chromosome 2R:14,660,000 (Fig 5B, qtl Bbas.1, 2 LOD drop range 14,590,000–14,700,000, LOD score = 9.06, 13.15% of variance explained). The A5 founder genotype had significantly lower survival than other lines (Fig 5C).



**Fig 5. A shared genetic architecture of immune defense between *L. fusiformis* and *B. bassiana*.** A) *B. bassiana* QTL analysis yields a single peak. B) The peaks on 2R for *B. bassiana* (blue, qtl Bbas.1) and *L. fusiformis* (gray, qtl A5.2) overlap. Horizontal lines represent significance threshold ( $\alpha = 0.05$ ), dashed vertical lines show the locations of the highest peaks. C) Founder genotype effects (plus standard error) for both experiments. Note that founder genotypes not present at that locus in one experiment are over a gray background.

<https://doi.org/10.1371/journal.ppat.1010934.g005>

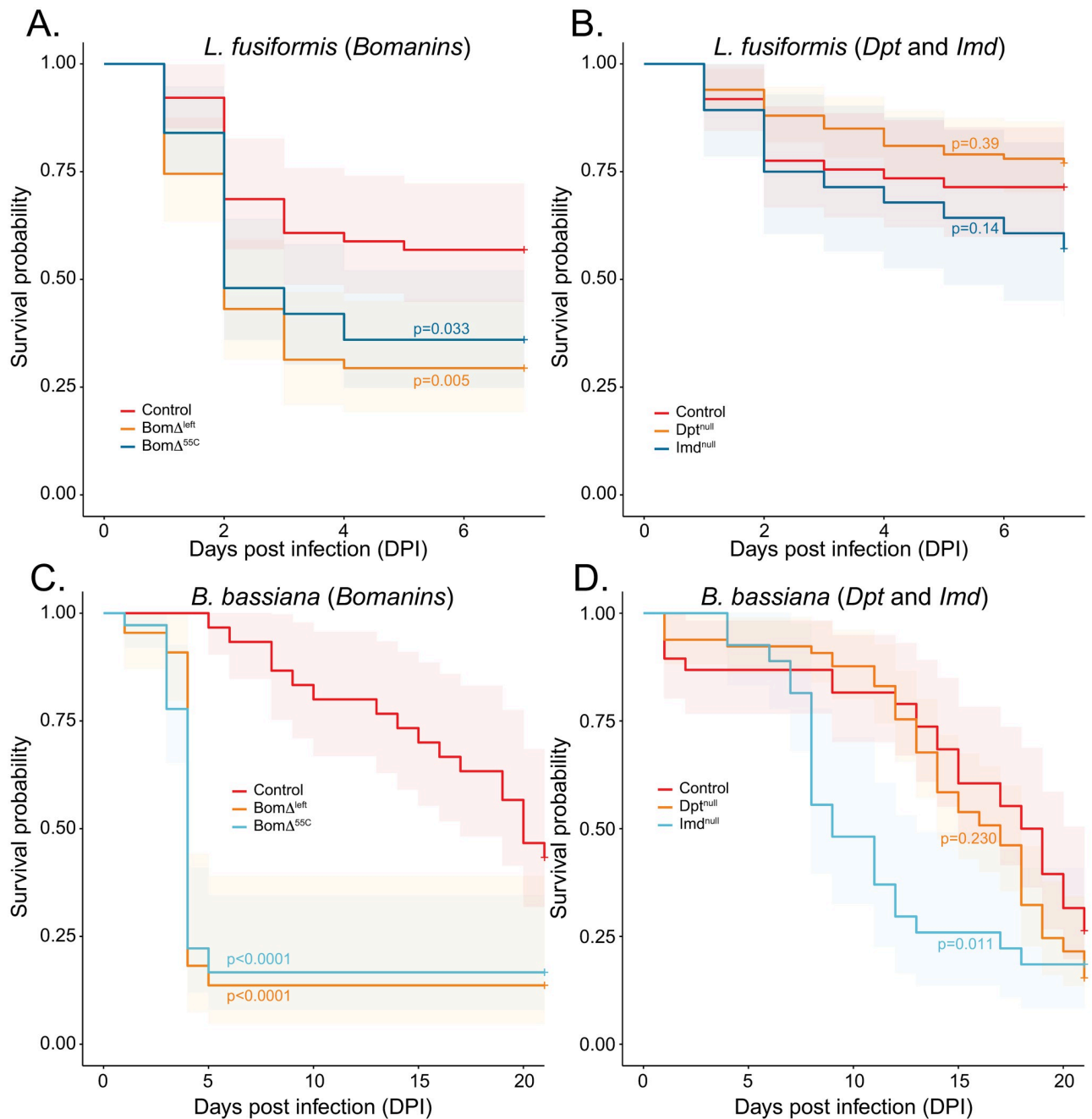


Thus, the *B. bassiana* peak is quite similar to the location of qtl A5.2 for *L. fusiformis*, even though the two experiments were performed in different subsets of the A population (Fig 5B). To test whether this overlap is more than expected by chance, we performed a Fisher's Exact Test with the fraction of physical positions with LOD score greater than the threshold score ( $\alpha = 0.05$ ). Day 5 survival for *L. fusiformis* had 289 of 11768 significant, while *B. bassiana* had a single position significant of 11768. That one position significant for *B. bassiana* was also significant for *L. fusiformis*, which is a significant excess over the expectation (expected overlap = 0.0246, Fisher Exact Test P-value < 0.0001). Furthermore, the founder genotype effects for both phenotypes clearly show reduced survival for the A5 genotype compared to all other genotypes present (this is true whether we match the peaks as in Fig 5C or pick the most significant peak for each genotype). The DSPR experiments lead us to conclude that different panels (DSPR A2 and B2), which have different founder genotypes, have different genetic architectures for immune defense against *L. fusiformis*, but in two different mapping experiments performed in two different labs with very different pathogens, there appears to be a shared genetic architecture for resistance to infection (DSPR A2 *L. fusiformis* vs. DSPR A1 for *B. bassiana*).

### Infections of mutant and RNAi knockdown flies show the potential role of individual genes in immune defense against both *L. fusiformis* and *B. bassiana*

The number of genes under the shared QTL peak in the DSPR A population are too numerous (96 genes plus tRNAs and lncRNAs) to test individually. We therefore tested several reasonable candidate genes including 4 genes in the *Bomanin* cluster and *Jabba* by RNAi knockdown plus knockout lines for genes encoding all *Bomanins* a subset of *Bomanins*, *Imd*, and *Diptericin A*. The *Bom $\Delta$ <sup>55C</sup>* line contains an excision of all 10 *Bomanin* paralogs in the cluster on chromosome 2R, and the *Bom $\Delta$ <sup>left</sup>* contains an excision of the six "left-most" paralogs: *BomT1*, *BomBc1*, *BomS1*, *BomS4*, *BomBc2*, and *BomS2*. Infections with *L. fusiformis* confirm a role for *Bomanin* genes in survival after infection since both *Bom $\Delta$ <sup>55C</sup>* and *Bom $\Delta$ <sup>left</sup>* show significantly reduced survival compared to control (Fig 6A). In contrast, neither *Dpt* nor *Imd* null mutants had any effect on survival after *L. fusiformis* infection (Fig 6B). *B. bassiana* infections also revealed an important role for the *Bomanin* cluster as again both mutants showed severely reduced survival after infection (Fig 6C). Interestingly, *Imd* null mutants also showed reduced survival after infection with *B. bassiana* compared to controls, suggesting a role for *Imd* in defense against fungal pathogens (Fig 6D).

With evidence that the *Bomanin* cluster was generally important for infection with *L. fusiformis* and *B. bassiana*, we sought to determine whether specific *Bomanin* copies had specific protective benefits against infection using fat body RNAi of individual *Bomanins*. After infection with *L. fusiformis* or *B. bassiana*, both *BomS2* and *BomBC2* knockdown flies stand out as being more susceptible to infection than controls (Fig 7, S13–S17 Tables,  $P < 0.001$  in each comparison). Other members of the *Bomanin* cluster had little to no influence on survival after infection with either pathogen. Given the similarity in sequence of some of these paralogs, we tested whether RNAi knockdowns were specific to specific *Bomanins*. While the two bicipital forms (*BomBc1* and *BomBc2*) were specific, *BomS1* did show some reduced induction even when *BomS2* was targeted by RNAi (see methods). *Jabba* fat body knockdowns also had little impact on survival after infection. Note that the two AttP40 RFP RNAi knockdowns, used as additional controls, showed intermediate survival. Each still show significantly higher survival compared to *BomBc2* ( $P < 0.001$ ) for both *L. fusiformis* and *B. bassiana*. Taken together, it appears that specific *Bomanin* paralogs (*BomS2* and *BomBc2*) provides some defense against *L. fusiformis* and *B. bassiana* infection. However, the best approach would be confirmed CRISPR



**Fig 6.** CRISPR/Cas9 and classical mutants reveal roles of *Bomanin* cluster and *Imd* in immune defense against both *L. fusiformis* (A and B) and *B. bassiana* (C and D). In all cases, controls are in red, other and knockouts of genes of interest are in other tones.

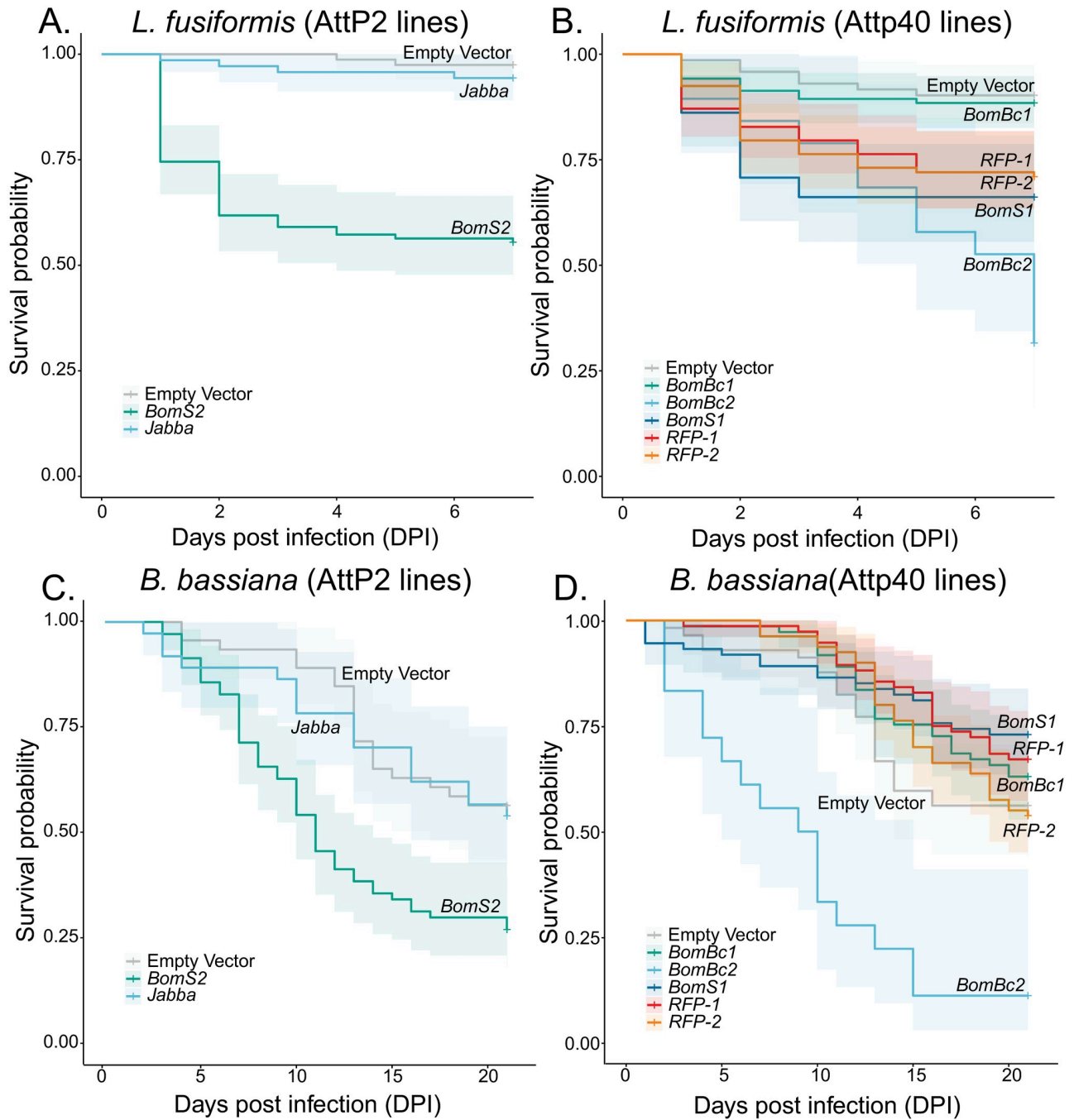
<https://doi.org/10.1371/journal.ppat.1010934.g006>

knockouts of each specific *Bomanin* paralog, or even better, recreation of the specific mutations observed in the DSPR lines.

## Discussion

Populations face infection challenge from innumerable different pathogens. The genetic factors that influence susceptibility to infection may vary from population to population for a





**Fig 7. Fat body RNAi knockdowns reveal roles of specific *Bomanin* genes in immune defense against both *L. fusiformis* (A and B) and *B. bassiana* (C and D).** In all cases, empty vector controls are in gray, other controls are in reddish tones and knockdowns of genes of interest are in blue/green tones.

<https://doi.org/10.1371/journal.ppat.1010934.g007>

specific pathogen, and from pathogen to pathogen within a specific host population. We examined genetic variation for susceptibility to the Gram-positive bacterium, *L. fusiformis* infection in three *Drosophila* mapping populations and find no evidence that the same genetic factors contribute to phenotypic variation in susceptibility across the three populations. We also

compared the genetic causes of phenotypic variation for susceptibility to *L. fusiformis* and the fungal pathogen, *B. bassiana* [44]. Surprisingly, even though these experiments were conducted by different researchers in different labs with different subpopulations, there was striking concordance of a single quantitative trait locus (QTL) for both pathogens. Although that QTL peak was broad, it contains a cluster of immune peptides, some of which are induced upon infection, and show increased susceptibility to both infections upon RNAi knockdown.

We hoped to determine the extent to which different genotypes are correlated in how they defend against infection. Are some genotypes universally more resistant to infection? Is correlation stronger when pathogens are more closely related? Do correlations suggest tradeoffs among immune phenotypes (pathogen X vs. pathogen Y) or between immune phenotypes and other life history traits? The DGRP lines are widely used for measuring phenotypic variation and attempting to map the genetic causes of that variation. As such, there are several studies on both immune and life history traits to examine genotypic correlations. Although we did find evidence of correlation, it was difficult to tease apart whether these relationships were truly biological or were artifacts—stronger correlations tended to be between phenotypes measured in the same lab or institution (different pathogens in the same lab, the same phenotype for each sex). In fact, there were no significant negative correlations, which would be expected with strong tradeoffs. The only exception to the general lack of signal (positive or negative) was the positive correlation between male starvation stress resistance and female *P. aeruginosa* survival. Although this might suggest that some genotypes are better at responding to general stress (starvation or infection) than others, it is curious that the correlation is only significant for one infection type and not even in males infected with *P. aeruginosa*. We therefore hesitate to draw conclusions from this correlation.

In contrast to the DGRP experiments, we find a striking shared QTL peak comparing infection with both *L. fusiformis* and *B. bassiana* in the DSPR lines. We are unable to measure the correlation between genotypes since the two experiments were performed in two different DSPR subpopulations. Note that both subpopulations were initiated from the same set of eight founder genotypes, so the genetic variation present in the two subpopulations is similar. The shared peak for two very different pathogens could mean several things. First, there are many genes under that peak and even many paralogs of the genes encoding *Bomanin* peptides. Therefore, it is possible that the genetic basis of increased susceptibility to *L. fusiformis* in the A5 background is different from the genetic basis of increased susceptibility to *B. bassiana* in the A5 background. If the phenotypic variation for the two pathogens was caused by different, and unique genetic changes, we imagine that natural selection might act on these linked variants as a single unit. Second, there may be a shared genetic basis for resistance to infection. Perhaps a single *Bomanin* (or small subset of *Bomanins*) are responsible for most of the response to infection, but in a non pathogen-specific way. At this point, we know very little about how specific *Bomanins* act in the immune system (Table 1), so whether individual *Bomanin* paralogs act generally or broadly is not clear. If single *Bomanins* or even clusters of *Bomanins* act generally, times of high pathogen pressure should select for resistant (either due to coding variation or higher expression) genotypes. Such selection would be relaxed or reversed (if *Bomanins* are costly) in the absence of infection.

Our approach enabled us to also determine whether the same genetic factors influence immune defense in different populations. We acknowledge that none of the three populations are ideal populations: DGRP are inbred lines that survived the lab environment after capture at a farmer's market, DSPR are derived from several lines from around the world and made into a "synthetic" population. Nonetheless, there was no evidence that any of the three populations shared genetic variation in any particular region that influenced survival after *L. fusiformis* infection. This may be because pathogen pressure varied in the populations from which

Table 1. Summary of *Bomanin* gene immune phenotypes.

<i>Bomanin</i>	Old name	FlySick Induction*	<i>L. fusiformis</i> induced	Under QTL peak	<i>L. fusiformis</i>	<i>B. bassiana</i>	<i>E. faecalis</i> [1]	<i>C. glabrata</i> [2]	<i>A. fumigatus</i> mycotoxin [3]
<i>BomT1</i>	CG43202	G+, Pe	2.89**	Yes			No		
<i>BomBc1</i>	IM23	G+, Ec	0.70	Yes	No	No	Yes		
<i>BomS1</i>	IM1	G+, Ec	1.83**	Yes	Yes	No	No		
<i>BomS4</i>	CG18107		-0.02	Yes			No		Yes
<i>BomBc2</i>	CG15067		-0.38	Yes	Yes	Yes			
<i>BomS2</i>	IM2	G+	0.45	Yes	Yes	Yes			
<i>BomS3</i>	IM3	G+	0.23	Yes				Yes	
<i>BomT2</i>	CG16836		0.94	Yes				No	
<i>BomS5</i>	CG15065	G+	1.13**	Yes				Yes	
<i>BomS6</i>	CG15068		-0.05	Yes				Moderate	Yes
<i>BomT3</i>	CG5778		-1.19**	No (Chr3R)					
<i>BomBc3</i>	CG5791	Ml	-0.04	No (Chr3R)					

\*FlySick induction indicates which of the ten microbes resulted in at least 10-fold higher expression compared to uninfected individuals. “G+” means all Gram-positive bacteria (*M. luteus*—Ml, *E. faecalis*, *S. aureus*) induced at least 10-fold. Two Gram-negative bacteria out of 7 also led to a 10-fold increase in expression: *Pseudomonas entomophila* (Pe) and *Escherichia coli* (Ec). “*L. fusiformis* induced” shows the fold induction and adjusted significance (\* <0.05, \*\* <0.01). “Under QTL peak” refers to whether the *bomanin* gene occurs under the *L. fusiformis* QTL peak. For each of the additional columns, “Yes” or “No” refers to whether genetic experiments showed a significant influence of that specific gene on defense against each pathogen.

<https://doi.org/10.1371/journal.ppat.1010934.t001>

these lines were derived. It may also be that immune alleles tend to turn over rapidly as populations coevolve with pathogens meaning that the alleles best suited to fight infection in one population are missing from another population. Finally, the small founder populations in the DSPR may preclude anything but high frequency variants from being detected in multiple experiments.

Several recent studies have attempted to characterize the role of individual *Bomanin* genes, the entire cluster of 10 genes, or subsets of those genes (summarized in Table 1) [1–3,27,50]. This began with Clemmons *et al.* [27] who characterized the gene family and showed that deletion of the entire cluster rendered flies highly susceptible to infection with *E. faecalis*, *C. glabrata* and *F. oxysporum*, but not the Gram-negative *E. cloacae*. The authors also showed the first evidence that there might be some specificity of activity since deleting the six left paralogs of the *Bomanin* cluster resulted in high susceptibility to *E. faecalis* and *F. oxysporum*, but not *C. glabrata*. To investigate further, Lindsay *et al.* [2] infected flies with four *Bomanin* transgenes on a *Bomanin* cluster partial deletion with *C. glabrata*. Two transgenes (*BomS3* and *BomS5*) restored survival to wildtype levels, one transgene (*BomS6*) showed intermediate restoration of survival, and one (*BomT2*) showed no restoration of wildtype survival. These experiments show that *BomS3* and *BomS5* are not just necessary, but sufficient to protect against infection with *C. glabrata* (at least among the *Bomanins*). As expected, the expression of these genes after infection correlates with their ability to restore wildtype survival. When expressed with the *BomS3* promoter, *BomS6* was not susceptible to *C. glabrata* infection, but *BomT2* showed no difference in susceptibility. Xu *et al.* [3] showed that *BomBc1*, *BomS3* and *BomS4* were significantly induced upon *A. fumigatus* infection and that, as a whole, the *Bomanin* cluster enhances both resistance to, and tolerance of, infection with *A. fumigatus*. *BomBc1*, *BomS3* and *BomS6* increased tolerance to the fungal toxin, restrictocin, whereas *BomS6* and *BomS1* increased tolerance to another toxin, verruculogen.

Here, we showed that natural variation in immune defense against both the Gram-positive *L. fusiformis* and the fungal *B. bassiana* map to a QTL that contains the *Bomanin* cluster and

that knockouts of the entire cluster or the left six copies and knockdown of specific *Bomanin* copies result in enhanced susceptibility (note that we do not differentiate between tolerance and resistance). Thus, our study adds to the growing evidence for the importance of the *Bomanin* cluster in immune defense, but also shows that naturally occurring genetic variation is associated with variation in immune defense against two different pathogens. Unfortunately, most of these studies utilize a subset of the *Bomanins* and use different techniques (CRISPR/Cas9 knockouts, transgenic lines, RNAi), so a systematic understanding of the specificity of individual *Bomanins* or the subclasses (tailed, short and bicipital) remains elusive, but would be necessary to truly understand paralog specificity.

This study highlights the context-dependence of the causes of genetic variation in immune defense. Within the same species, different populations harbor different genetic variants that result in variation in immune defense. Surprisingly, however, within a population, two very different pathogens share a QTL for susceptibility to infection, with a single founder genotype being significantly more susceptible than others. Although our approach does not have the resolution to determine whether the causal genetic variants are the mutation for both *L. fusiformis* and *B. bassiana* or are just genetically linked, given the high effective rate of recombination in *D. melanogaster* overall and that this region experiences exceptionally high recombination [67,68], there is a reasonable chance that the same mutation underlies both immune phenotypes. This linkage may allow natural selection to act more broadly to improve infection outcomes when susceptibility to such disparate pathogens is governed by a single locus.

## Methods

### *Drosophila* strains and husbandry

*Drosophila* Synthetic Population Resource (DSPR) lines from populations A1 and B1 are maintained by Stuart Macdonald at the University of Kansas [40]. The DSPR population lines were derived from two different sets of founder inbred lines. Each population was interbred among lines for several generations, then inbred to create the recombinant inbred lines used in these experiments. *Drosophila* Genetic Reference Panel (DGRP) lines were obtained from Brian Lazzaro at Cornell University and from the Bloomington *Drosophila* Stock Center (Bloomington, IN). Canton S is a standard laboratory *Drosophila* line. All *Drosophila* stocks were maintained on standard molasses media at 23 degrees Celsius on a cycle of 12 hours light to 12 hours dark. We refer to both sets of lines by their numeric assignment (for the DGRP we refer to line 321, not RAL\_321 or line\_321).

### Microbes and microbiology

*Lysinibacillus fusiformis* (strain Juneja) is a natural pathogen of *Drosophila* obtained from Brian Lazzaro at Cornell University [46]. Bacteria were grown from single colonies in LB liquid media overnight to stationary phase, then concentrated to an optical density (600nm) of about 4.0. *Beauveria bassiana* (strain GHA) was obtained from Parvin Shahrestani (California State University at Fullerton). Spores were grown in malt extract agar, then resuspended with a 10ml syringe, filtered through a 0.22  $\mu\text{m}$  Miracell filter and resuspended in 20% glycerol at 4.3e08 spores per mL as measured on a hemocytometer.

### General infection protocol

Flies were tipped to new media and cleared after 2–4 days of egg laying to control larval density. Ten to fourteen days later, emerging adults were moved to new media vials. Three- to seven-day old males were infected after light CO<sub>2</sub> anesthesia by pricking in the thorax with a

needle dipped in bacteria or fungal suspension or LB broth (or 20% glycerol) as a sterile control. The number surviving 2- and 5-days post infection was noted. For Canton S infections and RNAi knockdowns, we noted survival daily. Infections were performed over several weeks with 3–4 days of infection each week. A single researcher performed all infections ( $n = 22174$ ) for the GWAS and QTL experiments, and a different researcher performed all infections for the RNAi experiments. These experiments were done in batches of subsets of lines since it would not be possible to infect all lines in one batch. Each subset of lines was chosen randomly and batch effects were included in statistical models.

### Infection dynamics in Canton S

The effect of *L. fusiformis* infection on Canton S survival was assessed using a Cox Proportional Hazard test implemented in the R packages *survival* and *survminer* [69,70] with treatment (infected vs. sterile prick) as the independent variable.

### RNA-sequencing of infection in Canton-S males

We used Canton S flies to compare expression 12 hours after infection with *L. fusiformis* to those uninfected (no sterile prick). Each treatment contained 3 replicates of 3 flies per replicate. We flash froze flies in liquid nitrogen, then extracted RNA using a Zymo Quick-RNA Microprep kit (R1050, Zymo Research, Irvine, CA). We then constructed RNA libraries using the Lexogen QuantSeq 3' mRNA-Seq Library Prep kit (Lexogen GmbH, Vienna, Austria). We sequenced the samples on an Illumina MiSeq Micro SR100 run. Raw Reads were quality filtered using the fastx toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/); RRID:SCR\_005534), then mapped with STAR version 2.7.10b [71,72] to the *Drosophila melanogaster* (r6.18) genome, processed with samtools version 1.9 [73,74] and analyzed with DESeq2 [75] using a simple infection vs. no prick design.

### Genetic variation for immune defense in the DGRP

We infected male DGRP flies with  $OD_{600} \sim 4.0$  *L. fusiformis*. Each vial consisted of eight to ten infected flies with one to five vials per line for the DGRP. This corresponds to an average of 36.0 flies (3–4 vials per line) infected per line in 199 DGRP lines.

We performed the DGRP genome wide associations with and without line 321. We used both approaches because line 321 showed unusually low survival at 5 days post infection (DPI) and others [1,21] noted and excluded 321 in their analysis of resistance to *Metarhizium anisopliae* and *Enterococcus faecalis* infections. However, line 309 actually had lower survival (0% at 5 DPI) and we hesitate to start dropping all lines with low survival since variation is the point of this type of mapping. We measured survival both 2- and 5-days post infection and performed the mapping experiment at both time points. We fit a mixed effects logistic regression model using the *lmer* function in *lme4* [76] and line effects were extracted using the *ranef* function. These values were used with the DGRP website [31] to perform association mapping. Thus, there were four total mapping experiments for the DGRP: 2-and 5-days post infection and with and without line 321. We calculated heritabilities using line variance and total variance from the *VarrCorr* function in *lme4*.

We next asked whether line effects correlated between DGRP infection experiments (and a few other traits). We retrieved line effects for both male and female 24 hour bacterial load and female survival post *Providencia rettgeri* infection [5,38], male and female *Coxiella burnetii* hazard ratio [77], female *Pseudomonas entomophila* enteric survival [78], male and female *Pseudomonas aeruginosa* survival [21], male and female *Metarhizium anisopliae* survival [21], *Staphylococcus aureus* phagocyte mobilization [79], *Listeria monocytogenes* bacterial load [80],



*Enterococcus faecalis* survival [1], male aggression [81], male starvation resistance [31], pigmentation (tergite 5) [82], lifespan and fecundity [83], and weight [35]. We calculated Spearman's rank correlations for these 20 phenotypes and grouped by non-immune, Gram-positive/fungal and Gram-negative pathogens. In these analyses, we used the negative of the load values which means that all immune phenotypes have higher values indicating better immune defense and lower values meaning worse immune defense.

### Genetic variation for immune defense in the DSPR

The DSPR experiment was similar to that for the DGRP, but with more lines and fewer individuals infected per line. We used from one to four vials (usually 2–3 vials), consisting of 8 to 10 individuals, per line in the DSPR population. This corresponds to an average of 22.3 flies infected per line in 343 lines from DSPR population A2 and 18.8 flies infected per line in 391 lines from DSPR population B2. For each population, we used a logistic regression for proportion alive at day 2 and day 5 with line and infection date as random effects.

These random effects of line were used directly in the DSPRqtl package in R [40,41]. We used the DSPRscan function with the model effect~1. To determine a significance threshold, we performed 1000 permutations using the same model. Peaks were determined using the LOD drop 2 method in the DSPRPeaks function and nearby and overlapping peaks were curated by hand.

We also retrieved raw data for survival after *B. bassiana* infection [44] and performed a logistic regression analysis with random effects of line, replicate nested in line, group and group nested in set (as in their analysis). This differs from the authors' approach who used arc-sin-square root transformed survival. All other analyses were performed as above.

### Mutant and RNAi knockdown experiments to examine the role of individual genes in immune defense

We followed up our QTL mapping by examining several genes that occur putatively associated with immune defense. These included null mutants (*Imd*, *Diptericin A*, *BomΔ<sup>55C</sup>* and *BomΔ<sup>left</sup>*) and their respective controls. We obtained the TRIP RNAi lines [84] from the Bloomington Drosophila Stock Center (Bloomington, Indiana, USA) and crossed to the C564 (RRID: BDSC\_6982) fat body-specific driver [85]. In the case of *Pvr*, we used two additional drivers. We also crossed the driver to an empty vector and RFP RNAi (when available) as negative controls. All lines are described in S18 Table. Infections were performed as described above for *L. fusiformis*, but *B. bassiana* were monitored for 21 days instead of 7 days, with flies transferred to new food every 3–4 days. In addition, we mock infected lines with either Luria-Bertani broth or 20% glycerol as a sterile wound control. Very few of these flies died from the wound. We assessed significant differences in survival using the *coxph* function from the *survival* package in R [69,86] with genotype and date as variables and the empty vector as the reference. Since the RNAi lines are on different backgrounds (AttP2 and AttP40), we performed separate analyses based on background. Since these *Bomanin* genes are paralogous and have some degree of sequence similarity, we performed qPCR to examine the degree of specificity for each knockdown. To do so, we first determined the RNAi target sequence for each focal *Bomanin* gene, then found the paralog with the sequence closest to that target sequence. We then asked whether both target gene and closest paralog (based on target sequence) were differentially expressed 8 hours post infection or no-infection control and between knockdown and RNAi control (BDSC\_36303 or BDSC\_36304). qPCR primers are listed in S19 Table. S10A Fig shows that the two short form *Bomanins* are nonspecific. *BomS1* shows reduced induction compared to control whether *BomS1* or *BomS2* was targeted. However, *BomS2* only shows reduced

induction compared to control when *BomS2* was targeted. For the bicipital form *Bomanins* (S10B Fig), both were specific—only showing decreased induction when the focal gene was targeted. We assessed significance with a simple *t* test of the difference in relative expression in infected vs. uninfected conditions:  $(RP49_{\text{infected-target}}) - (RP49_{\text{uninfected-target}})$ .

## Supporting information

**S1 Table. Gene Ontology enrichment for expression 12 hours after *L. fusiformis* infection in *D. melanogaster* Canton S males.**

(XLSX)

**S2 Table. Values for DGRP random effects from the logistic regression model. All lines included or line 321 removed and at 5- or 2-days post infection.**

(XLSX)

**S3 Table. SNPs significantly associated with survival 5 days post infection in the DGRP with all lines.**

(XLSX)

**S4 Table S. NPs significantly associated with survival 5 days post infection in the DGRP when line 321 is omitted.**

(XLSX)

**S5 Table. SNPs significantly associated with survival 2 days post infection in the DGRP with all lines.**

(XLSX)

**S6 Table. SNPs significantly associated with survival 2 days post infection in the DGRP when line 321 is omitted.**

(XLSX)

**S7 Table. Values for DSPR random effects from the logistic regression model for survival 5 days post infection with *L. fusiformis* in population A.**

(XLSX)

**S8 Table. Values for DSPR random effects from the logistic regression model for survival 2 days post infection with *L. fusiformis* in population A.**

(XLSX)

**S9 Table. Values for DSPR random effects from the logistic regression model for survival 5 days post infection with *L. fusiformis* in population B.**

(XLSX)

**S10 Table. Values for DSPR random effects from the logistic regression model for survival 2 days post infection with *L. fusiformis* in population B.**

(XLSX)

**S11 Table. Values for DSPR random effects from the logistic regression model for survival post infection with *B. bassiana* in population A.**

(XLSX)

**S12 Table. LOD scores for all genomic positions for all 5 mapping experiments (populations A2 and B2 for 2- and 5- days post infection for *L. fusiformis* and A1 for *B. bassiana*) in the DSPR.**

(XLSX)



**S13 Table. Raw data for RNAi infection with date dead and status (0 if alive at the end of the experiment).**

(XLSX)

**S14 Table. Genotype and block effects for survival analysis after infection with *L. fusiformis* for RNAi knockdown lines on the AttP2 background.**

(XLSX)

**S15 Table. Genotype and block effects for survival analysis after infection with *L. fusiformis* for RNAi knockdown lines on the AttP40 background.**

(XLSX)

**S16 Table. Genotype and block effects for survival analysis after infection with *B. bassiana* for RNAi knockdown lines on the AttP2 background.**

(XLSX)

**S17 Table. Genotype and block effects for survival analysis after infection with *B. bassiana* for RNAi knockdown lines on the AttP40 background.**

(XLSX)

**S18 Table. *Drosophila* Lines used for validation experiments.**

(XLSX)

**S19 Table. Primers used for qPCR experiments.**

(XLSX)

**S1 Fig. DGRP Day 2 raw survival with standard error of the proportion.** Survival measured as the proportion alive at Day 2 divided by the total number infected per line.

(TIF)

**S2 Fig. DGRP Random effects sorted by effect size from mixed effects logistic regression model.** Data is presented with all lines present or excluding line 321 (which had low survival) and at 5- or 2-days post infection.

(TIF)

**S3 Fig. Manhattan Plots for DGRP Genome Wide Association Studies at 5- or 2-days post infection and with or without line 321 included.**

(TIF)

**S4 Fig. QQ Plots for DGRP Mapping Experiments at 5- and 2-days post infection and with all lines or with line 321 excluded.**

(TIF)

**S5 Fig. Knockdown of *Pvr* in fat body and hemocytes does not significantly impact survival after infection with *L. fusiformis* (OD<sub>600</sub> ~10).** Three different drivers (r4 –fat body specific, He–hemocyte specific, and c564 –expressed in fat body, gut, and several male tissues) were crossed to a TRIP *Pvr* RNAi line (37520) or empty vector control (36304). Survival was monitored for 10 days post infection.

(TIF)

**S6 Fig. DSPR Day 2 Raw Survival plus standard error of the Proportion.** Raw survival is measured as proportion alive at 2 days post infection. A) Panel A, B) Panel B.

(TIF)

**S7 Fig. DSPR Line effects for (A-D) *L. fusiformis* survival day 2 or day 5 post infection in populations A2 and B2, E) *B. bassiana* and F) correlation between Day 5 and Day 2 line**

random effects in population A2 (gray) and B2 (blue).

(TIF)

**S8 Fig. DSPR Scan plot for Day 2 survival random effect in DSPR population A2.**

(TIF)

**S9 Fig. Founder genotype effects for significant QTL peaks in DSPR populations.** (A and B) Population A2 Day 5 peaks 1 and 2, C) Population B2 Day 5 peak 1, D) Population B2 Day 2 peak 1, E) Population A1 *B. bassiana* survival peak 1.

(TIF)

**S10 Fig. DSPR Scan plot for Day 2 survival random effect in DSPR population B2.**

(TIF)

**S11 Fig. Expression of *Bomanin* genes after RNAi knockdown to determine specificity of knockdown.** The Y axis depicts the log<sub>2</sub> fold induction 8 hours post infection compared to an uninfected (no sterile prick) individual. Relevant comparisons are presented with black bars and p values are from T tests. RNAi knockdown refers to the genes targeted by RNAi constructs and gene expression refers to the primers used to Measure gene expression.

(TIF)

## Acknowledgments

We thank Brian Lazzaro for providing the *L. fusiformis* strain, Stuart Macdonald for the DSPR lines and advice on mapping, Steve Wasserman for the *Bomanin* knockout lines, and Unckless lab members for helpful comments and advice. The Bloomington Drosophila Stock Center provided several DGRP and RNAi lines. RNA sequencing was performed in the KU Genome Sequencing Center, supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number P30-GM145499.

## Author Contributions

**Conceptualization:** Brittny R. Smith, Robert L. Unckless.

**Data curation:** Robert L. Unckless.

**Formal analysis:** Robert L. Unckless.

**Funding acquisition:** Robert L. Unckless.

**Investigation:** Brittny R. Smith, Kistie B. Patch, Anjali Gupta, Emma M. Knoles, Robert L. Unckless.

**Methodology:** Brittny R. Smith, Robert L. Unckless.

**Project administration:** Robert L. Unckless.

**Resources:** Kistie B. Patch.

**Supervision:** Robert L. Unckless.

**Validation:** Brittny R. Smith, Kistie B. Patch, Robert L. Unckless.

**Writing – original draft:** Robert L. Unckless.

**Writing – review & editing:** Robert L. Unckless.

## References

1. Chapman JR, Dowell MA, Chan R, Unckless RL. The Genetic Basis of Natural Variation in *Drosophila melanogaster* Immune Defense against *Enterococcus faecalis*. *Genes (Basel)*. 2020; 11(2). Epub 32098395. <https://doi.org/10.3390/genes11020234> PMID: 32098395
2. Lindsay SA, Lin SJH, Wasserman SA. Short-Form Bomanins Mediate Humoral Immunity in *Drosophila*. *J Innate Immun*. 2018; 10(4):306–14. Epub 2018/06/20. <https://doi.org/10.1159/000489831> PMID: 29920489
3. Xu R, Lou Y, Tidu A, Bulet P, Heinekamp T, Martin F, et al. The Toll pathway mediates *Drosophila* resilience to *Aspergillus* mycotoxins through specific Bomanins. *bioRxiv*. 2022:2022.08.18.504437. <https://doi.org/10.15252/embr.202256036> PMID: 36322050
4. Magwire MM, Fabian DK, Schweyen H, Cao C, Longdon B, Bayer F, et al. Genome-wide association studies reveal a simple genetic basis of resistance to naturally coevolving viruses in *Drosophila melanogaster*. *PLoS Genet*. 2012; 8(11):e1003057. Epub 2012/11/15. <https://doi.org/10.1371/journal.pgen.1003057> PMID: 23166512
5. Unckless RL, Rottschaefer SM, Lazzaro BP. The complex contributions of genetics and nutrition to immunity in *Drosophila melanogaster*. *PLoS Genet*. 2015; 11(3):e1005030. Epub 2015/03/13. <https://doi.org/10.1371/journal.pgen.1005030> PMID: 25764027
6. Baker C, Antonovics J. Evolutionary determinants of genetic variation in susceptibility to infectious diseases in humans. *PLoS One*. 2012; 7(1):e29089. Epub 2012/01/05. <https://doi.org/10.1371/journal.pone.0029089> PMID: 22242158
7. Chapman SJ, Hill AV. Human genetic susceptibility to infectious disease. *Nat Rev Genet*. 2012; 13(3):175–88. Epub 2012/02/07. <https://doi.org/10.1038/nrg3114> PMID: 22310894.
8. Doeschl-Wilson AB, Davidson R, Conington J, Roughsedge T, Hutchings MR, Villanueva B. Implications of host genetic variation on the risk and prevalence of infectious diseases transmitted through the environment. *Genetics*. 2011; 188(3):683–93. Epub 2011/04/28. <https://doi.org/10.1534/genetics.110.125625> PMID: 21527777
9. Covid-Host Genetics Initiative. A first update on mapping the human genetic architecture of COVID-19. *Nature*. 2022; 608(7921):E1–E10. Epub 2022/08/03. <https://doi.org/10.1038/s41586-022-04826-7> PMID: 35922517
10. Covid-Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. *Nature*. 2021; 600(7889):472–7. Epub 2021/07/08. <https://doi.org/10.1038/s41586-021-03767-x> PMID: 34237774
11. McMillan LE, Adamo SA. Friend or foe? Effects of host immune activation on the gut microbiome in the caterpillar *Manduca sexta*. *J Exp Biol*. 2020; 223(Pt 19). Epub 2020/10/14. <https://doi.org/10.1242/jeb.226662> PMID: 33046577.
12. Metcalf CJE, Koskella B. Protective microbiomes can limit the evolution of host pathogen defense. *Evol Lett*. 2019; 3(5):534–43. Epub 2019/10/23. <https://doi.org/10.1002/evl3.140> PMID: 31636945
13. Vijendravarma RK, Narasimha S, Chakrabarti S, Babin A, Kolly S, Lemaitre B, et al. Gut physiology mediates a trade-off between adaptation to malnutrition and susceptibility to food-borne pathogens. *Ecol Lett*. 2015; 18(10):1078–86. Epub 2015/08/08. <https://doi.org/10.1111/ele.12490> PMID: 26249109.
14. Buchon N, Silverman N, Cherry S. Immunity in *Drosophila melanogaster*—from microbial recognition to whole-organism physiology. *Nat Rev Immunol*. 2014; 14(12):796–810. Epub 2014/11/26. <https://doi.org/10.1038/nri3763> PMID: 25421701
15. Ferrandon D, Imler JL, Hoffmann JA. Sensing infection in *Drosophila*: Toll and beyond. *Semin Immunol*. 2004; 16(1):43–53. <https://doi.org/10.1016/j.smim.2003.10.008> PMID: 14751763.
16. Lemaitre B. The road to Toll. *Nat Rev Immunol*. 2004; 4(7):521–7. <https://doi.org/10.1038/nri1390> PMID: 15229471.
17. Boman HG, Nilsson I, Rasmuson B. Inducible antibacterial defence system in *Drosophila*. *Nature*. 1972; 237(5352):232–5. <https://doi.org/10.1038/237232a0> PMID: 4625204.
18. Coscia M, Giacomelli S, Oreste U. Toll-like receptors: an overview from invertebrates to vertebrates. *Invertebrate Survival Journal*. 2011; 8(2):210–26.
19. Carboni AL, Hanson MA, Lindsay SA, Wasserman SA, Lemaitre B. Cecropins contribute to *Drosophila* host defense against a subset of fungal and Gram-negative bacterial infection. *Genetics*. 2022; 220(1). <https://doi.org/10.1093/genetics/iyab188> PMID: 34791204
20. Hanson MA, Dostalova A, Ceroni C, Poidevin M, Kondo S, Lemaitre B. Synergy and remarkable specificity of antimicrobial peptides in vivo using a systematic knockout approach. *Elife*. 2019; 8. Epub 2019/02/26. <https://doi.org/10.7554/eLife.44341> PMID: 30803481
21. Wang JB, Lu HL, St Leger RJ. The genetic basis for variation in resistance to infection in the *Drosophila melanogaster* genetic reference panel. *PLoS Pathog*. 2017; 13(3):e1006260. Epub 2017/03/04. <https://doi.org/10.1371/journal.ppat.1006260> PMID: 28257468

22. Sterken MG, van Sluijs L, Wang YA, Ritmahan W, Gultom ML, Riksen JAG, et al. Punctuated Loci on Chromosome IV Determine Natural Variation in Orsay Virus Susceptibility of *Caenorhabditis elegans* Strains Bristol N2 and Hawaiian CB4856. *J Virol.* 2021; 95(12). Epub 20210524. <https://doi.org/10.1128/JVI.02430-20> PMID: 33827942
23. Perlmutter JI, Chapman JR, Wilkinson MC, Nevarez-Saenz I, Unckless RL. A single amino acid polymorphism in natural Metchnikowin alleles of *Drosophila* results in systemic immunity and life history tradeoffs. *bioRxiv.* 2023. Epub 2023/01/31. <https://doi.org/10.1101/2023.01.16.524277> PMID: 36712113
24. Cogni R, Cao C, Day JP, Bridson C, Jiggins FM. The genetic architecture of resistance to virus infection in *Drosophila*. *Mol Ecol.* 2016; 25(20):5228–41. Epub 20160826. <https://doi.org/10.1111/mec.13769> PMID: 27460507
25. Strand MR. The insect cellular immune response. *Insect science.* 2008; 15(1):1–14.
26. Mylonakis E, Podsiadlowski L, Muhammed M, Vilcinskas A. Diversity, evolution and medical applications of insect antimicrobial peptides. *Philos Trans R Soc Lond B Biol Sci.* 2016; 371(1695). <https://doi.org/10.1098/rstb.2015.0290> PMID: 27160593
27. Clemmons AW, Lindsay SA, Wasserman SA. An effector Peptide family required for *Drosophila* toll-mediated immunity. *PLoS Pathog.* 2015; 11(4):e1004876. Epub 2015/04/29. <https://doi.org/10.1371/journal.ppat.1004876> PMID: 25915418
28. Cohen LB, Lindsay SA, Xu Y, Lin SJH, Wasserman SA. The Daisho Peptides Mediate *Drosophila* Defense Against a Subset of Filamentous Fungi. *Front Immunol.* 2020; 11:9. Epub 2020/02/11. <https://doi.org/10.3389/fimmu.2020.00009> PMID: 32038657
29. Shen R, Zheng K, Zhou Y, Chi X, Pan H, Wu C, et al. A dRASSF-STRIPAK-Imd-JAK/STAT axis controls antiviral immune response in *Drosophila*. *Cell Rep.* 2022; 40(4):111143. <https://doi.org/10.1016/j.celrep.2022.111143> PMID: 35905720.
30. Davies SA, Overend G, Sebastian S, Cundall M, Cabrero P, Dow JA, et al. Immune and stress response 'cross-talk' in the *Drosophila* Malpighian tubule. *J Insect Physiol.* 2012; 58(4):488–97. Epub 20120128. <https://doi.org/10.1016/j.jinsphys.2012.01.008> PMID: 22306292.
31. Mackay TF, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, et al. The *Drosophila melanogaster* Genetic Reference Panel. *Nature.* 2012; 482(7384):173–8. Epub 2012/02/10. <https://doi.org/10.1038/nature10811> PMID: 22318601
32. Rodrigues LR, Zwoinska MK, Wiberg RAW, Snook RR. The genetic basis and adult reproductive consequences of developmental thermal plasticity. *J Anim Ecol.* 2022; 91(6):1119–34. Epub 2022/01/22. <https://doi.org/10.1111/1365-2656.13664> PMID: 35060127
33. Akhund-Zade J, Lall S, Gajda E, Yoon D, Ayroles JF, de Bivort BL. Genetic basis of offspring number-body weight tradeoff in *Drosophila melanogaster*. *G3 (Bethesda).* 2021. Epub 2021/04/20. <https://doi.org/10.1093/g3journal/jkab129> PMID: 33871609
34. Zhang S, Pointer B, Kelleher ES. Rapid evolution of piRNA-mediated silencing of an invading transposable element was driven by abundant de novo mutations. *Genome Res.* 2020; 30(4):566–75. Epub 2020/04/03. <https://doi.org/10.1101/gr.251546.119> PMID: 32238416
35. Unckless RL, Rottschaefer SM, Lazzaro BP. A genome-wide association study for nutritional indices in *Drosophila*. *G3 (Bethesda).* 2015; 5(3):417–25. Epub 2015/01/15. <https://doi.org/10.1534/g3.114.016477> PMID: 25583649
36. Palu RAS, Ong E, Stevens K, Chung S, Owings KG, Goodman AG, et al. Natural Genetic Variation Screen in *Drosophila* Identifies Wnt Signaling, Mitochondrial Metabolism, and Redox Homeostasis Genes as Modifiers of Apoptosis. *G3 (Bethesda).* 2019; 9(12):3995–4005. Epub 2019/10/02. <https://doi.org/10.1534/g3.119.400722> PMID: 31570502
37. Carbone MA, Yamamoto A, Huang W, Lyman RA, Meadors TB, Yamamoto R, et al. Genetic architecture of natural variation in visual senescence in *Drosophila*. *Proc Natl Acad Sci U S A.* 2016; 113(43):E6620–E9. Epub 2016/10/30. <https://doi.org/10.1073/pnas.1613833113> PMID: 27791033
38. Howick VM, Lazzaro BP. The genetic architecture of defence as resistance to and tolerance of bacterial infection in *Drosophila melanogaster*. *Mol Ecol.* 2017; 26(6):1533–46. Epub 2017/01/19. <https://doi.org/10.1111/mec.14017> PMID: 28099780.
39. Palmer WH, Obbard DJ. Variation and Evolution in the Glutamine-Rich Repeat Region of *Drosophila* Argonaute-2. *G3 (Bethesda).* 2016; 6(8):2563–72. Epub 2016/06/19. <https://doi.org/10.1534/g3.116.031880> PMID: 27317784
40. King EG, Macdonald SJ, Long AD. Properties and power of the *Drosophila* Synthetic Population Resource for the routine dissection of complex traits. *Genetics.* 2012. <https://doi.org/10.1534/genetics.112.138537> PMID: 22505626

41. Long AD, Macdonald SJ, King EG. Dissecting complex traits using the *Drosophila* Synthetic Population Resource. *Trends Genet.* 2014; 30(11):488–95. Epub 2014/09/02. <https://doi.org/10.1016/j.tig.2014.07.009> PMID: 25175100
42. Everman ER, Cloud-Richardson KM, Macdonald SJ. Characterizing the genetic basis of copper toxicity in *Drosophila* reveals a complex pattern of allelic, regulatory, and behavioral variation. *Genetics.* 2021; 217(1):1–20. <https://doi.org/10.1093/genetics/iyaa020> PMID: 33683361
43. Najarro MA, Hackett JL, Smith BR, Highfill CA, King EG, Long AD, et al. Identifying Loci Contributing to Natural Variation in Xenobiotic Resistance in *Drosophila*. *PLoS Genet.* 2015; 11(11):e1005663. Epub 20151130. <https://doi.org/10.1371/journal.pgen.1005663> PMID: 26619284
44. Shahrestani P, King E, Ramezan R, Phillips M, Riddle M, Thornburg M, et al. The molecular architecture of *Drosophila melanogaster* defense against *Beauveria bassiana* explored through evolve and resequence and quantitative trait locus mapping. *G3 (Bethesda).* 2021; 11(12). Epub 2021/09/18. <https://doi.org/10.1093/g3journal/jkab324> PMID: 34534291
45. Cao C, Cogni R, Barbier V, Jiggins FM. Complex Coding and Regulatory Polymorphisms in a Restriction Factor Determine the Susceptibility of *Drosophila* to Viral Infection. *Genetics.* 2017; 206(4):2159–73. Epub 2017/06/21. <https://doi.org/10.1534/genetics.117.201970> PMID: 28630113
46. Smith BR, Unckless RL. Draft Genome Sequence of *Lysinibacillus fusiformis* Strain Juneja, a Laboratory-Derived Pathogen of *Drosophila melanogaster*. *Genome Announc.* 2018; 6(5). Epub 2018/02/14. <https://doi.org/10.1128/genomeA.01571-17> PMID: 29437109
47. Ahmed I, Yokota A, Yamazoe A, Fujiwara T. Proposal of *Lysinibacillus boronitolerans* gen. nov. sp. nov., and transfer of *Bacillus fusiformis* to *Lysinibacillus fusiformis* comb. nov. and *Bacillus sphaericus* to *Lysinibacillus sphaericus* comb. nov. *Int J Syst Evol Microbiol.* 2007; 57(Pt 5):1117–25. Epub 2007/05/03. <https://doi.org/10.1099/ijs.0.63867-0> PMID: 17473269.
48. Smith EC. Inoculation Experiments with *Bacillus fusiformis* isolated from Tropical Ulcer with Observations on the *Bacillus*. *J Hyg (Lond).* 1933; 33(1):95–102 3. Epub 1933/01/01. <https://doi.org/10.1017/s0022172400018428> PMID: 20475172
49. Ross BD, Hayes B, Radey MC, Lee X, Josek T, Bjork J, et al. *Ixodes scapularis* does not harbor a stable midgut microbiome. *ISME J.* 2018; 12(11):2596–607. Epub 2018/06/28. <https://doi.org/10.1038/s41396-018-0161-6> PMID: 29946195
50. Troha K, Im JH, Revah J, Lazzaro BP, Buchon N. Comparative transcriptomics reveals *CrebA* as a novel regulator of infection tolerance in *D. melanogaster*. *PLoS Pathog.* 2018; 14(2):e1006847. Epub 2018/02/03. <https://doi.org/10.1371/journal.ppat.1006847> PMID: 29394281
51. Ganesan S, Aggarwal K, Paquette N, Silverman N. NF-kappaB/Rel proteins and the humoral immune responses of *Drosophila melanogaster*. *Curr Top Microbiol Immunol.* 2011; 349:25–60. Epub 2010/09/21. [https://doi.org/10.1007/82\\_2010\\_107](https://doi.org/10.1007/82_2010_107) PMID: 20852987
52. Tzou P, Ohresser S, Ferrandon D, Capovilla M, Reichhart JM, Lemaitre B, et al. Tissue-specific inducible expression of antimicrobial peptide genes in *Drosophila* surface epithelia. *Immunity.* 2000; 13(5):737–48. Epub 2000/12/15. [https://doi.org/10.1016/s1074-7613\(00\)00072-8](https://doi.org/10.1016/s1074-7613(00)00072-8) PMID: 11114385.
53. Schleifer KH, Kandler O. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev.* 1972; 36(4):407–77. Epub 1972/12/01. <https://doi.org/10.1128/br.36.4.407-477.1972> PMID: 4568761
54. Miwa H, Ahmed I, Yokota A, Fujiwara T. *Lysinibacillus parviboronicipiens* sp. nov., a low-boron-containing bacterium isolated from soil. *Int J Syst Evol Microbiol.* 2009; 59(Pt 6):1427–32. Epub 2009/06/09. <https://doi.org/10.1099/ijs.0.65455-0> PMID: 19502328.
55. Kurata S. Peptidoglycan recognition proteins in *Drosophila* immunity. *Dev Comp Immunol.* 2014; 42(1):36–41. Epub 20130622. <https://doi.org/10.1016/j.dci.2013.06.006> PMID: 23796791
56. Michel T, Reichhart JM, Hoffmann JA, Royet J. *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. *Nature.* 2001; 414(6865):756–9. <https://doi.org/10.1038/414756a> PMID: 11742401.
57. Ishimaru S, Ueda R, Hinohara Y, Ohtani M, Hanafusa H. PVR plays a critical role via JNK activation in thorax closure during *Drosophila* metamorphosis. *EMBO J.* 2004; 23(20):3984–94. Epub 2004/10/01. <https://doi.org/10.1038/sj.emboj.7600417> PMID: 15457211
58. Wu Y, Brock AR, Wang Y, Fujitani K, Ueda R, Galko MJ. A blood-borne PDGF/VEGF-like ligand initiates wound-induced epidermal cell migration in *Drosophila* larvae. *Curr Biol.* 2009; 19(17):1473–7. Epub 2009/08/04. <https://doi.org/10.1016/j.cub.2009.07.019> PMID: 19646875
59. Bond D, Foley E. Autocrine platelet-derived growth factor-vascular endothelial growth factor receptor-related (Pvr) pathway activity controls intestinal stem cell proliferation in the adult *Drosophila* midgut. *J Biol Chem.* 2012; 287(33):27359–70. Epub 2012/06/23. <https://doi.org/10.1074/jbc.M112.378018> PMID: 22722927



60. Tsuzuki S, Matsumoto H, Furihata S, Ryuda M, Tanaka H, Sung EJ, et al. Switching between humoral and cellular immune responses in *Drosophila* is guided by the cytokine GBP. *Nat Commun*. 2014; 5:4628. Epub 2014/08/19. <https://doi.org/10.1038/ncomms5628> PMID: 25130174
61. Sopko R, Lin YB, Makhijani K, Alexander B, Perrimon N, Bruckner K. A systems-level interrogation identifies regulators of *Drosophila* blood cell number and survival. *PLoS Genet*. 2015; 11(3):e1005056. Epub 2015/03/10. <https://doi.org/10.1371/journal.pgen.1005056> PMID: 25749252
62. Larkin A, Marygold SJ, Antonazzo G, Attrill H, Dos Santos G, Garapati PV, et al. FlyBase: updates to the *Drosophila melanogaster* knowledge base. *Nucleic Acids Res*. 2021; 49(D1):D899–D907. Epub 2020/11/22. <https://doi.org/10.1093/nar/gkaa1026> PMID: 33219682
63. Jehrke L, Stewart FA, Droste A, Beller M. The impact of genome variation and diet on the metabolic phenotype and microbiome composition of *Drosophila melanogaster*. *Sci Rep*. 2018; 8(1):6215. Epub 2018/04/20. <https://doi.org/10.1038/s41598-018-24542-5> PMID: 29670218
64. King EG, Merkes CM, McNeil CL, Hooper SR, Sen S, Broman KW, et al. Genetic dissection of a model complex trait using the *Drosophila* Synthetic Population Resource. *Genome Res*. 2012; 22(8):1558–66. Epub 2012/04/10. <https://doi.org/10.1101/gr.134031.111> PMID: 22496517
65. Dudzic JP, Kondo S, Ueda R, Bergman CM, Lemaitre B. *Drosophila* innate immunity: regional and functional specialization of prophenoloxidases. *BMC Biol*. 2015; 13:81. Epub 2015/10/01. <https://doi.org/10.1186/s12915-015-0193-6> PMID: 26437768
66. Binggeli O, Neyen C, Poidevin M, Lemaitre B. Prophenoloxidase activation is required for survival to microbial infections in *Drosophila*. *PLoS Pathog*. 2014; 10(5):e1004067. Epub 24788090. <https://doi.org/10.1371/journal.ppat.1004067> PMID: 24788090
67. Fiston-Lavier AS, Singh ND, Lipatov M, Petrov DA. *Drosophila melanogaster* recombination rate calculator. *Gene*. 2010; 463(1–2):18–20. Epub 2010/05/11. <https://doi.org/10.1016/j.gene.2010.04.015> PMID: 20452408.
68. Singh ND, Arndt PF, Petrov DA. Genomic heterogeneity of background substitutional patterns in *Drosophila melanogaster*. *Genetics*. 2005; 169(2):709–22. Epub 2004/11/03. <https://doi.org/10.1534/genetics.104.032250> PMID: 15520267
69. Therneau T. A Package for Survival Analysis in R. 2022.
70. Kassambara A, Kosinski M, Biecek P. survminer: Drawing Survival Curves using 'ggplot2'. 2021.
71. Dobin A, Gingeras TR. Optimizing RNA-Seq Mapping with STAR. *Methods Mol Biol*. 2016; 1415:245–62. Epub 2016/04/27. [https://doi.org/10.1007/978-1-4939-3572-7\\_13](https://doi.org/10.1007/978-1-4939-3572-7_13) PMID: 27115637.
72. Dobin A, Gingeras TR. Mapping RNA-seq Reads with STAR. *Curr Protoc Bioinformatics*. 2015; 51:11 4 1–4 9. Epub 2015/09/04. <https://doi.org/10.1002/0471250953.bi1114s51> PMID: 26334920
73. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *Gigascience*. 2021; 10(2). Epub 2021/02/17. <https://doi.org/10.1093/gigascience/giab008> PMID: 33590861
74. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25(16):2078–9. Epub 2009/06/10. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
75. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15(12):550. Epub 2014/12/18. <https://doi.org/10.1186/s13059-014-0550-8> PMID: 25516281
76. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*. 2015; 67(1):1–48. <https://doi.org/10.18637/jss.v067.i01>
77. Guzman RM, Howard ZP, Liu Z, Oliveira RD, Massa AT, Omsland A, et al. Natural genetic variation in *Drosophila melanogaster* reveals genes associated with *Coxiella burnetii* infection. *Genetics*. 2021; 217(3). Epub 2021/04/01. <https://doi.org/10.1093/genetics/iyab005> PMID: 33789347
78. Bou Sleiman MS, Osman D, Massouras A, Hoffmann AA, Lemaitre B, Deplancke B. Genetic, molecular and physiological basis of variation in *Drosophila* gut immunocompetence. *Nat Commun*. 2015; 6:7829. Epub 2015/07/28. <https://doi.org/10.1038/ncomms8829> PMID: 26213329
79. Nazario Toole AE. Genome-wide association studies of phagocytosis and the cellular immune response in *Drosophila melanogaster* [Ph.D.]. Ann Arbor: University of Maryland, College Park; 2016.
80. Hotson AG, Schneider DS. *Drosophila melanogaster* Natural Variation Affects Growth Dynamics of Infecting *Listeria monocytogenes*. *G3 (Bethesda)*. 2015; 5(12):2593–600. Epub 2015/10/07. <https://doi.org/10.1534/g3.115.022558> PMID: 26438294
81. Shorter J, Couch C, Huang W, Carbone MA, Peiffer J, Anholt RR, et al. Genetic architecture of natural variation in *Drosophila melanogaster* aggressive behavior. *Proc Natl Acad Sci U S A*. 2015; 112(27):E3555–63. Epub 2015/06/24. <https://doi.org/10.1073/pnas.1510104112> PMID: 26100892

82. Dembeck LM, Huang W, Carbone MA, Mackay TF. Genetic basis of natural variation in body pigmentation in *Drosophila melanogaster*. *Fly (Austin)*. 2015; 9(2):75–81. Epub 2015/11/12. <https://doi.org/10.1080/19336934.2015.1102807> PMID: 26554300
83. Durham MF, Magwire MM, Stone EA, Leips J. Genome-wide analysis in *Drosophila* reveals age-specific effects of SNPs on fitness traits. *Nat Commun*. 2014; 5:4338. Epub 2014/07/09. <https://doi.org/10.1038/ncomms5338> PMID: 25000897.
84. Perkins LA, Holderbaum L, Tao R, Hu Y, Sopko R, McCall K, et al. The Transgenic RNAi Project at Harvard Medical School: Resources and Validation. *Genetics*. 2015; 201(3):843–52. Epub 2015/09/01. <https://doi.org/10.1534/genetics.115.180208> PMID: 26320097
85. Hrdlicka L, Gibson M, Kiger A, Micchelli C, Schober M, Schock F, et al. Analysis of twenty-four Gal4 lines in *Drosophila melanogaster*. *Genesis*. 2002; 34(1–2):51–7. Epub 2002/09/27. <https://doi.org/10.1002/gene.10125> PMID: 12324947.
86. Team RC. R: A language and environment for statistical computing.: R Foundation for Statistical Computing, Vienna, Austria.; 2022.