

The Effect of Developmental Pleiotropy on the Evolution of Insect Immune Genes

Alissa M. Williams[†], Thi Minh Ngo[†], Veronica E. Figueroa, and Ann T. Tate  *

Department of Biological Sciences and Evolutionary Studies Initiative, Vanderbilt University, Nashville, Tennessee

*Corresponding author: E-mail: a.tate@vanderbilt.edu.

[†]These authors contributed equally to this work.

Accepted: 04 March 2023

Abstract

The pressure to survive ever-changing pathogen exposure explains the frequent observation that immune genes are among the fastest evolving in the genomes of many taxa, but an intriguing proportion of immune genes also appear to be under purifying selection. Though variance in evolutionary signatures of immune genes is often attributed to differences in gene-specific interactions with microbes, this explanation neglects the possibility that immune genes participate in other biological processes that could pleiotropically constrain adaptive selection. In this study, we analyzed available transcriptomic and genomic data from *Drosophila melanogaster* and related species to test the hypothesis that there is substantial pleiotropic overlap in the developmental and immunological functions of genes involved in immune signaling and that pleiotropy would be associated with stronger signatures of evolutionary constraint. Our results suggest that pleiotropic immune genes do evolve more slowly than those having no known developmental functions and that signatures of constraint are particularly strong for pleiotropic immune genes that are broadly expressed across life stages. These results support the general yet untested hypothesis that pleiotropy can constrain immune system evolution, raising new fundamental questions about the benefits of maintaining pleiotropy in systems that need to rapidly adapt to changing pathogen pressures.

Key words: molecular evolution, Toll pathway, insect immunity, evolutionary constraint, adaptive evolution.

Significance

Pleiotropy, where one gene affects multiple discrete traits, presents an interesting puzzle for evolutionary biologists because mutations that are adaptive for one trait could antagonize the function of another. We hypothesized that pleiotropy in genes shared by immune and developmental signaling pathways could constrain rapid adaptation of immune systems. Our results suggest that pleiotropy can constrain immune system evolution in the fruit fly *Drosophila melanogaster*, raising new fundamental questions about the benefits of maintaining pleiotropy in systems that need to rapidly adapt to changing pathogen pressures.

Introduction

Over evolutionary time, organisms have developed defense mechanisms against microbial pathogens and parasites which counter-adapt, in turn, to maintain successful infection strategies. Host immune systems put selective pressure on microbes to evade host recognition, repel antimicrobial

effectors, and even manipulate immune signaling components to dampen host defenses (Schmid-Hempel 2008; Heil 2016). Hosts that cannot circumvent these mechanisms could suffer massive fitness costs from infection. As a result, pressure from pathogens and parasites represents a major driving force in molecular evolution (Paterson et al. 2010).

© The Author(s) 2023. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.

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 reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

How should we expect selection to act on immune system genes? Host adaptation to microbial pressure should drive positive, directional selection or, in the face of co-evolutionary negative frequency dependence, balancing selection that maintains polymorphism in populations (Casals et al. 2011; Sackton 2019). Studies in species as diverse as humans (Mukherjee et al. 2009; Casals et al. 2011), non-human mammals (Seabury et al. 2010; Areal et al. 2011), and insects (Sackton et al. 2007; Obbard et al. 2009; Rottschaefer et al. 2015) have found evidence for both positive and balancing selection in immune system recognition and effector genes (Unckless et al. 2016). For example, Obbard et al. (2009) found that *Drosophila melanogaster* immune genes, as a class, have higher rates of adaptive substitution than location-matched non-immune genes. However, these trends were driven by a few particularly rapidly evolving genes associated with a subset of immune signaling pathways, while purifying selection was surprisingly prevalent on immune genes in other pathways. If parasites frequently target or evade signaling components, why would not those targets show rapid adaptation?

The answer may depend on a crucial but underappreciated quality of immune systems. Genetic pleiotropy arises when a single gene product contributes to multiple discrete phenotypic traits, and many components of immune pathways appear to be pleiotropic. Since the discovery of the Toll pathway, for example, numerous studies (and indeed Nobel prizes) have recognized its conserved dual role in development and innate immune system signaling (Lemaitre et al. 1997; DiAngelo et al. 2009; Anthony et al. 2018) and proposed that this could impose constraints on immune system evolution (Obbard et al. 2009; Tan et al. 2021). More broadly, a recent study estimated that approximately 17% of human genes affect multiple discrete phenotypic traits, and functional enrichment analysis of this pleiotropic gene set revealed immune system functions to be among the most over-represented processes (Sivakumaran et al. 2011). When a pleiotropic mutation affects uncorrelated traits, opposing forces of selection on each trait can reduce the efficacy of selection and resist the fixation of adaptive substitutions (Fraïsse et al. 2018). Thus, the adaptive evolution of pleiotropic immune genes may be constrained by the deleterious effects of substitutions on other traits.

Pleiotropy between development and immunity is particularly intriguing because a developmental program must be carried out faithfully for an organism to progress through its life cycle, resulting in purifying selection on genes involved in embryonic and early life development. Indeed, developmental pleiotropy (defined by the number of genetic interactions [Stark et al. 2006]) has been shown in *D. melanogaster* to constrain positive selection in early-expressed genes due to a higher number of functional

interactions in those genes that render mutations deleterious (Artieri et al. 2009). We hypothesize that developmental pleiotropy could constrain immune gene evolution, particularly for genes involved in the most complex stages of development (Tian et al. 2013), leading to an underrepresentation of signatures of positive selection on immune genes relative to theoretical expectations.

Insects can serve as particularly valuable models for studying the evolutionary consequences of developmental and immunological pleiotropy due to their discrete life stages, a wealth of genomic resources, and availability of studies on immune gene function (i5K Consortium 2013; Palmer and Jiggins 2015; Viljakainen 2015). The canonical components of an insect innate immune response include microbial recognition, signal transduction to initiate cellular and humoral responses, and production of effector molecules for pathogen clearance (Lemaitre and Hoffmann 2007). Many genes and signaling pathways previously identified as core participants in these processes are also broadly conserved among species (Waterhouse et al. 2007), including two of the best studied pathways, Toll and Imd, which coordinate expression of antimicrobial peptides and other pathogen-clearing effectors (Ferrandon et al. 2007; Tanji et al. 2007). While the Toll pathway is the most recognized example of developmental and immunological pleiotropy in insect immune systems, previous work has highlighted potential pleiotropy within other pathways (Tate and Graham 2015). For example, the same components of the melanization pathway responsible for tanning the insect cuticle after each larval molt are also used for melanizing parasitoid eggs and neutralizing pathogenic fungi, leading to allocation issues when an insect needs to accomplish both at once (McNeil et al. 2010; Parker et al. 2017). Thus, pleiotropy is likely to interfere with the deployment of immune responses if a host needs to use a gene product for both development and immunity in the same life stage. Even if these functions are segregated into different life stages, however, could pleiotropy still constrain immune system evolution?

We predict that immune genes that have a pleiotropic developmental function will be more likely to experience evolutionary constraint, as defined by slower rates of evolution and a lower frequency of positive selection, than immune genes that have no known developmental function. Further, we predict that pleiotropic genes that are crucial to multiple developmental stages will be the most constrained, relative to genes involved in more specific and less conserved developmental processes. To investigate these predictions, we combine transcriptional and functional genomics data from fruit flies (*Drosophila* spp.) to characterize the overall and immune pathway-specific degree of pleiotropy among immune and developmental genes. We then analyze the rates of evolution in immune genes using genomic data from 12 sequenced *Drosophila*

Table 1

The Extent of Pleiotropy as Defined With Different Annotation Methods

| | Definition | Pleiotropic | Immune Non-pleiotropic | Dev Non-pleiotropic |
|---|---|-------------------------------|------------------------|------------------------|
| 1 | Immune = all Immune GO + previous citations. + DE (808) Dev = all Dev GO | 354 (43.8%) $d_N/d_S = 0.062$ | 454 $d_N/d_S = 0.085$ | 2992 $d_N/d_S = 0.063$ |
| 2 | Immune = Immune Response GO + previous citations + DE (753) Dev = all Dev GO | 299 (39.7%) $d_N/d_S = 0.063$ | 454 $d_N/d_S = 0.085$ | 3047 $d_N/d_S = 0.063$ |
| 3 | Immune = Immune Response GO + previous citations + DE (753) | 52 (6.9%) $d_N/d_S = 0.051$ | 701 $d_N/d_S = 0.077$ | 210 $d_N/d_S = 0.056$ |

NOTE.—GO, Gene Ontology annotation terms; DE, differentially expressed via transcriptional analyses; Dev, developmental; Previous citations, genes or gene lists manually or computationally identified as having immune system functions in *Drosophila*. d_N/d_S , median d_N/d_S value of that class.

species; we also evaluate the 6 species in the melanogaster group separately. Empirical support for our predictions would raise the question of why evolution would maintain pleiotropy between development and immunity given the potential for conflict and constraint. On the other hand, if pleiotropic immune genes are not more constrained than non-pleiotropic ones, this study could inspire future investigations into compensatory evolution and the role of network architecture in minimizing evolutionary conflict.

Results

Extent of Developmental Pleiotropy in Immune Genes

To determine the prevalence of developmental pleiotropy among immune genes, we started by curating separate lists of immune and developmental genes in *D. melanogaster*. Previous studies have employed various methods to curate gene lists, ranging from using only Gene Ontology (GO) annotations (Fraisse et al. 2018) to compiling experimentally confirmed and/or computationally predicted immune gene orthologs (Early et al. 2017). Taking these different approaches into account, we employed several sources to assemble a comprehensive suite of genes that participate in immunity (table 1 and Materials and Methods). In total, we assembled a list of 808 immune genes, of which 551 genes have known canonical roles in immunity and 107 genes play a role in immune system development, as annotated by GO and previous studies (Early et al. 2017). The degree of overlap between different immune gene list sources can be found in supplementary figure S1, Supplementary Material online. The list of developmental genes contains 3,346 genes, of which 262 genes are annotated specifically as “embryonic development” genes and 508 as “post-embryonic development.” Some embryonic development genes also participate in post-embryonic development (overlap visualized in supplementary fig. S2, Supplementary Material online).

Genes that appear in both the immune and developmental gene lists were labeled as “pleiotropic.” When considering immune genes as those identified by all methods including manually curated, GO-annotated, and differentially expressed genes, we found 354 immune genes

(43.8%) to be pleiotropic (table 1, row 1). When constraining the definition of the immune gene to those that directly contribute to an immune response while excluding genes participating in the development of the immune system, 299 (39.7%) genes are considered pleiotropic (table 1, row 2). Under the most conservative definition of development (only genes that directly participate in embryonic development or 7.8% (262/3,346) of all annotated developmental genes), 52 immune genes (6.9%) still meet the definition of pleiotropy (table 1, row 3). The full list of immune, developmental, and pleiotropic genes under different categorization methods is included in supplementary table S1, Supplementary Material online. Note that although we used several methods to compile a list of pleiotropic genes, the conclusions generated throughout this study are robust to different categorical definitions of immunity, development, and pleiotropy, as evidenced by median d_N/d_S values (table 2). Therefore, from this point on, for simplicity, we refer to our immune gene group as those defined using the sources from table 1, row 2, which comprises Immune Response GO-annotated genes, immune genes employed in previous large-scale studies, and a core set of genes differentially expressed in ten bacterial infections (Troha et al. 2018).

Comparison of Pleiotropic and Non-Pleiotropic Immune Gene Characteristics

Immune genes can be categorized into different classes, such as recognition, signaling, and effector, depending on their canonical function in an immune response. We were curious whether certain classes of immune genes are more likely to have a pleiotropic status than others. We divided immune genes into major categories, relying on both annotation from previous studies (Sackton et al. 2007; Early et al. 2017) and manual annotation based on gene description in FlyBase (supplementary table S2, Supplementary Material online). According to this classification system, the number of genes confirmed to each category includes 33 recognition genes, 123 signaling genes, and 27 effector genes (supplementary fig. S3, Supplementary Material online). As represented in figure 1A, the signaling immune class contains the highest

Table 2

Compiled Results for Statistical Values Across Primary and Downsampled Data Sets Using 12-Species and 6-Species Concatenated and Individual Gene Data

| | τ | d_N/d_S values | | | | α | ω_a | ω_{na} |
|-------------------------------|--------|------------------|-------------------|-----------|------------------|----------|------------|---------------|
| | | 12-species | 12-species Indiv. | 6-species | 6-species Indiv. | | | |
| | | Concat. | Genes | Concat. | Genes | | | |
| Full data sets | | | | | | | | |
| Non-pleiotropic immune | 0.731 | 0.098 | 0.085 | 0.1 | 0.077 | 0.647 | 0.16 | 0.091 |
| Pleiotropic | 0.67 | 0.077 | 0.063 | 0.089 | 0.06 | 0.774 | 0.178 | 0.052 |
| Non-pleiotropic developmental | 0.691 | 0.078 | 0.063 | 0.078 | 0.057 | 0.714 | 0.152 | 0.062 |
| Downsampled data sets | | | | | | | | |
| Non-pleiotropic immune | 0.732 | | 0.085 | | 0.078 | 0.635 | 0.161 | 0.092 |
| Pleiotropic | 0.67 | | 0.063 | | 0.061 | 0.778 | 0.182 | 0.052 |
| Non-pleiotropic developmental | 0.692 | | 0.064 | | 0.058 | 0.722 | 0.156 | 0.06 |

NOTE.—Values for each statistic are medians for each category.

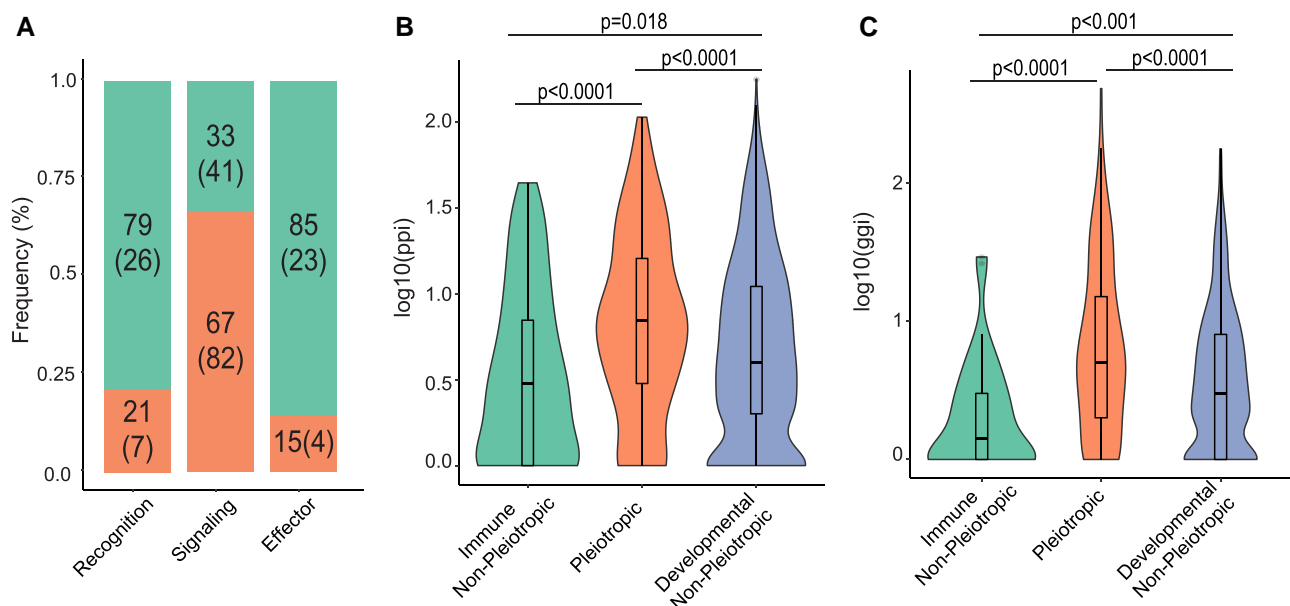


Fig. 1.—Overall characterization of pleiotropic and non-pleiotropic immune genes. Each immune gene was assigned a “gene class” (A) depending on their canonical function in an immune response. For each class, the percentage of pleiotropic (those with developmental roles; bottom bars) and non-pleiotropic genes (top bars) was determined (big number: proportion; number in parentheses: number of genes in that category). The number of known protein–protein interactions (ppi; B) and number of known gene–gene interactions (ggi; C) were also calculated for genes annotated as immune non-pleiotropic, pleiotropic for development and immunity, or developmental non-pleiotropic, represented on a log-scale and statistically analyzed using Kruskal–Wallis tests for overall significance followed by post hoc pairwise Dunn tests (Benjamini–Hochberg–adjusted *P* values on figure).

proportion of pleiotropic genes (66.67%, $n = 123$), and the different groups contain a significantly different proportion of pleiotropic genes overall ($\chi^2 = 37.94$, $P < 0.0001$). Moreover, using the PANTHER pathway database, we found that pleiotropic genes are, on average, associated with more pathways than non-pleiotropic ones (supplementary table S3, Supplementary Material online).

We also wanted to know whether our curated immune-developmental pleiotropic genes exhibit characteristics associated with alternative definitions of pleiotropy, such as a high number of associated protein–protein interactions and gene–gene interactions that reflect activity at the molecular level. When comparing pleiotropic and non-pleiotropic immune genes (fig. 1B and C), we do find that pleiotropic genes

have significantly more protein–protein interactions (Kruskal–Wallis w/Dunn post hoc test, $P_{\text{adj}} = 3.8e^{-05}$) and more gene–gene interactions (Kruskal–Wallis w/Dunn post hoc test, $P_{\text{adj}} = 6.3e^{-07}$). Moreover, pleiotropic genes are associated with more Biological Processes (Wilcoxon test, $P < 2e^{-16}$) and Molecular Functions (Wilcoxon test, $P < 2e^{-16}$) GO terms than non-pleiotropic genes (supplementary fig. S4, Supplementary Material online).

Expression Specificity Across Stages and Tissues Between Pleiotropic and Non-Pleiotropic Genes

To investigate the hypothesis that broadly expressed pleiotropic genes are under stronger evolutionary constraint than specific ones, we determined gene expression specificity across life stages and tissues for pleiotropic and non-pleiotropic immune genes using the τ specificity index ([Yanai et al. 2005], see Materials and Methods). A large τ value indicates specific expression while a small value indicates broad expression across stages or tissues. While we could not confidently determine whether any given gene plays only a developmental or immunological role or both at any given stage, genes involved in development at multiple life stages may present a temporal as well as evolutionary constraint on the immunological function of that gene.

We found that, in uninfected insects, there was a statistically detectable difference between pleiotropic immune genes (median $\tau = 0.670$) and non-pleiotropic immune genes (fig. 2A, median $\tau = 0.731$; Kruskal–Wallis w/Dunn test, $P_{\text{adj}} = 0.0009$), but not between pleiotropic immune genes and non-pleiotropic developmental genes (median $\tau = 0.691$; Kruskal–Wallis w/Dunn test, $P_{\text{adj}} = 0.07$). These results indicate broader expression across stages in the pleiotropic gene class relative to the non-pleiotropic gene classes (table 2). The unbalanced size of each gene category did not affect the results, as confirmed by downsampling through bootstrapping (supplementary methods, Supplementary Material online; supplementary fig. S5, Supplementary Material online).

We also found that the most stage-specific pleiotropic genes, determined by the top quartile in τ value, disproportionately exhibit maximal expression during the embryonic stage (43% among specific pleiotropic genes vs. 3.6% among specific non-pleiotropic immune genes) while the most specific non-pleiotropic immune genes exhibit a relatively even distribution of maximal expression across subsequent stages (fig. 2B, supplementary table S4, Supplementary Material online). At the tissue level, pleiotropic genes are also expressed more broadly than non-pleiotropic immune genes, and this trend is consistent throughout all life stages (fig. 2C). We found no significant differences in tissue expression specificity between developmental genes and pleiotropic genes except in the adult stage (fig. 2C), where developmental genes showed more specific patterns of expression.

Evolutionary Rates Among Different Gene Categories

To address whether pleiotropic genes are more evolutionarily constrained than non-pleiotropic genes, we calculated d_N/d_S values using codeml site model M0 in PAML v4.9j (Yang 2007), which assigns a single d_N/d_S value to an entire tree (see Materials and Methods). We ran this PAML model for concatenations of genes in 12 *Drosophila* species (*Drosophila ananassae*, *Drosophila erecta*, *Drosophila grimshawi*, *Drosophila mojavensis*, *Drosophila persimilis*, *Drosophila pseudoobscura*, *Drosophila sechellia*, *Drosophila simulans*, *Drosophila virilis*, *Drosophila willistoni*, and *Drosophila yakuba*; supplementary table S5, Supplementary Material online), where each concatenation represented one of three categories of genes: non-pleiotropic immune, pleiotropic, and non-pleiotropic developmental. Genes for each concatenation were defined using table 1, row 2, and after quality control, these concatenations contained 356, 231, and 2,067 genes, respectively. We also ran codeml site model M0 on each individual gene included in the concatenations; these model runs were successful for 348 non-pleiotropic immune genes, 227 pleiotropic genes, and 2,037 non-pleiotropic developmental genes (see Materials and Methods).

The model runs on the concatenated gene lists yielded d_N/d_S estimates of 0.098 for non-pleiotropic immune genes, 0.077 for pleiotropic genes, and 0.078 for non-pleiotropic developmental genes (table 2). Meanwhile, model runs on individual genes yielded median d_N/d_S estimates (fig. 3A) of 0.085, 0.063, and 0.063 respectively, and these three categories exhibited significantly different d_N/d_S distributions based on a Kruskal–Wallis test (fig. 3A, $\chi^2 = 66.53$, $P = 3.57e^{-15}$). Pairwise comparisons of distributions of individual gene d_N/d_S values were calculated using post hoc Dunn tests adjusted for multiple comparisons. The comparison between pleiotropic genes and developmental non-pleiotropic genes does not show a statistically significant difference ($P = 0.95$), but non-pleiotropic immune genes have a significantly higher median d_N/d_S value than both non-pleiotropic developmental genes ($P = 1.8e^{-15}$) and pleiotropic genes ($P = 4.2e^{-08}$). The differences in sample sizes between categories did not affect the results, as confirmed by downsampling through bootstrapping (supplementary methods, Supplementary Material online; supplementary fig. S6, Supplementary Material online). Within the pleiotropic gene set, we found that the most specifically stage-expressed genes (top τ quartile, e.g., fig. 2B) had significantly lower d_N/d_S ratios than the most broadly expressed pleiotropic genes (bottom τ quartile; $n = 41/\text{quartile}$, Wilcoxon test, $P = 0.023$, fig. 3B).

To account for possible saturation of d_S across the 12 species phylogeny and/or differences in selection across clades, we repeated the above PAML analyses for the 6 species in our data set that were part of the melanogaster

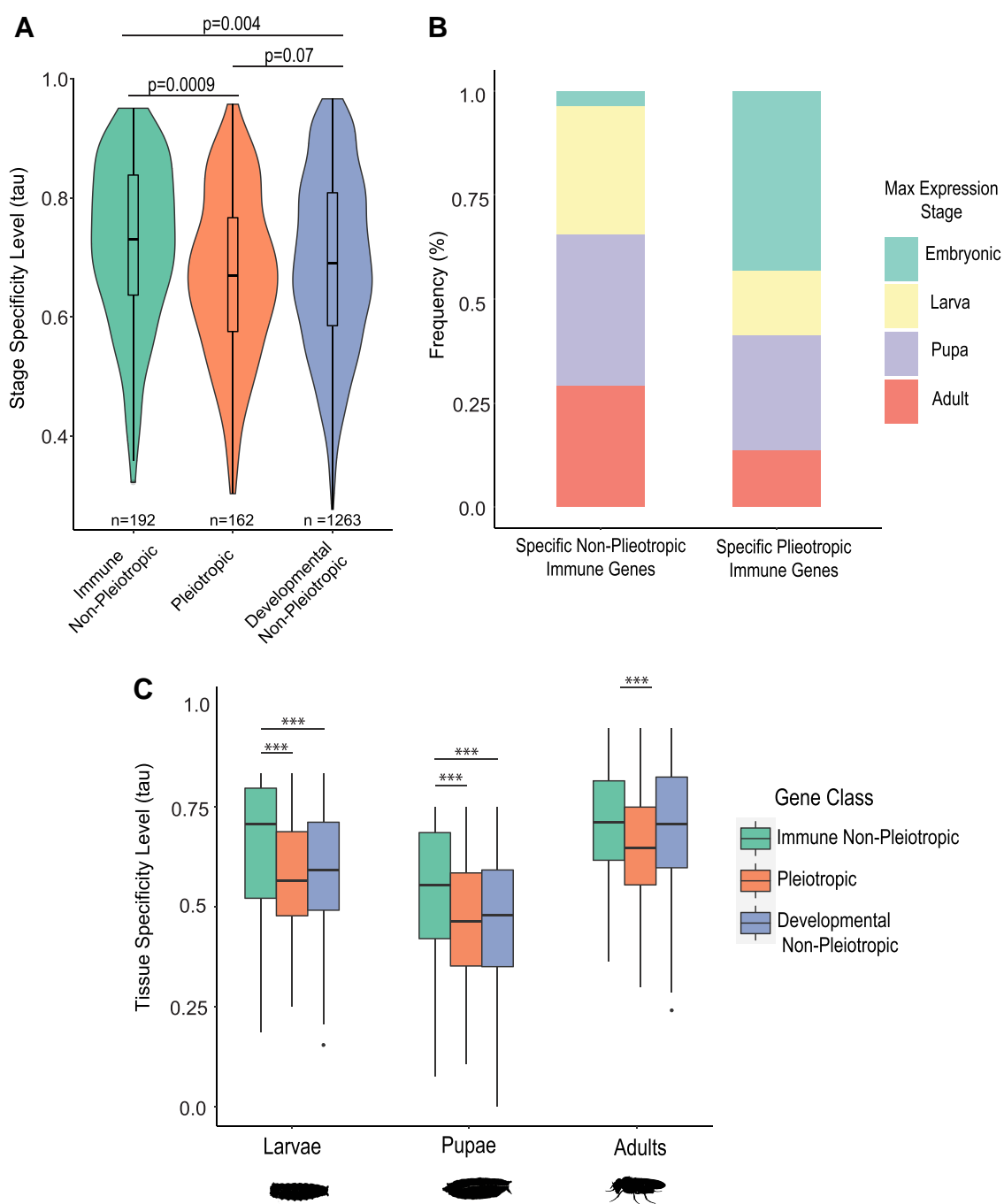


Fig. 2.—Comparison of relative life stage and tissue specificity of gene expression among immune, developmental, and pleiotropic genes. The stage specificity τ value, which varies from 0 (broadly expressed across all stages) to 1 (expressed in only one stage), was calculated for genes within each class (A). For the non-pleiotropic and pleiotropic immune gene group (B), the genes within the top 25th percentile of τ value were characterized as “specific genes,” and the stage with the highest expression for each gene was determined and tallied for the whole group. To compare tissue gene expression specificity between pleiotropic and non-pleiotropic genes within each life stage (C), the τ value (tissue specificity level) was calculated for each gene across tissues. Differences among groups were statistically analyzed using Kruskal–Wallis tests for overall significance followed by post hoc pairwise Dunn tests (Benjamini–Hochberg–adjusted P values on figure; *** indicates $P_{\text{adj}} < 0.001$).

group (*D. ananassae*, *D. erecta*, *D. melanogaster*, *D. sechellia*, *D. simulans*, and *D. yakuba*). Model runs on this 6-species data set yielded d_N/d_S estimates of 0.100 for

non-pleiotropic immune genes, 0.089 for pleiotropic genes, and 0.078 for non-pleiotropic developmental genes using the concatenated alignments (table 2). The individual

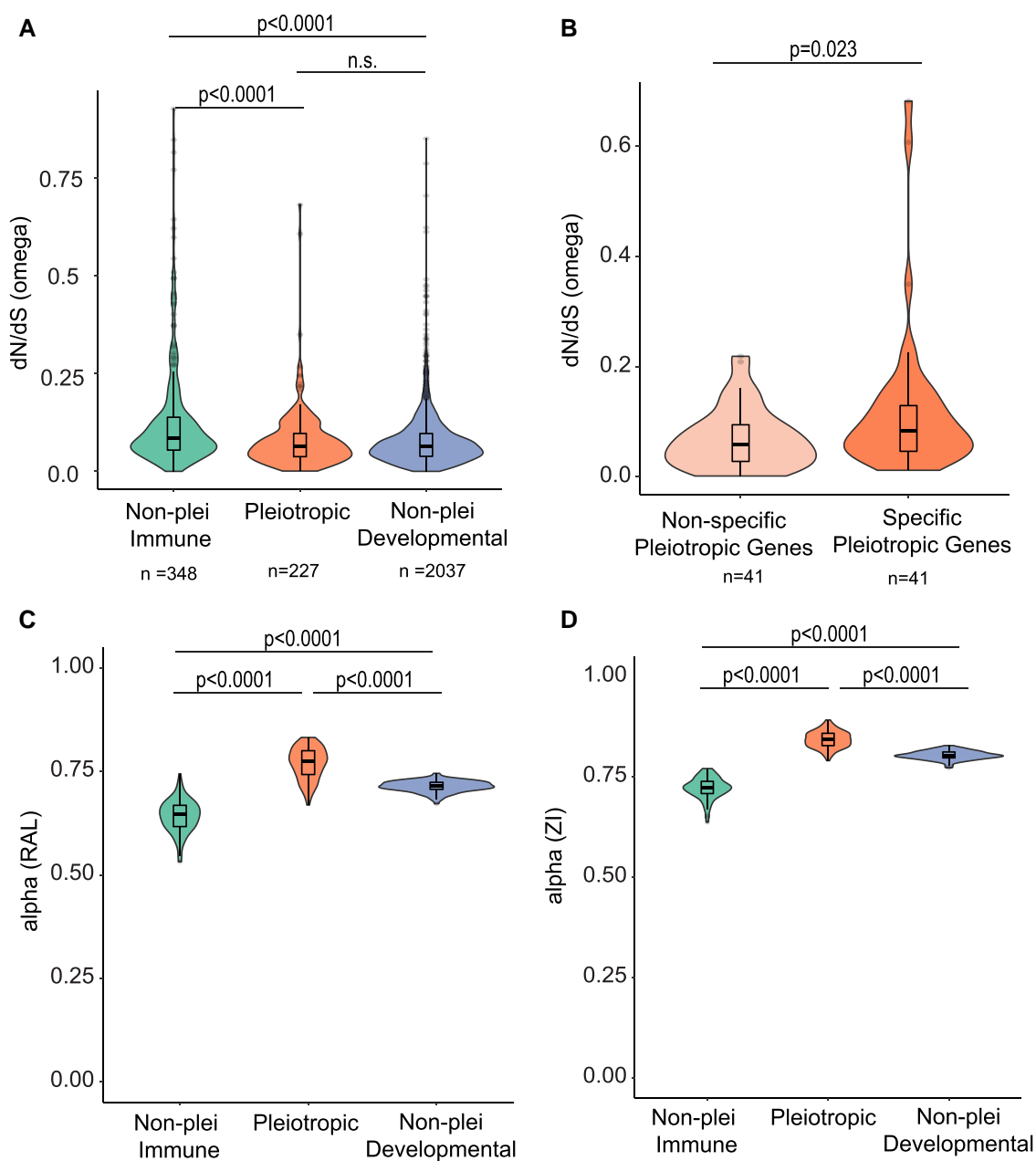


Fig. 3.—Associations between genetic pleiotropy, stage specificity, and evolutionary statistics. d_N/d_S values (A) were compared among non-pleiotropic immune genes, genes with pleiotropic roles in development and immunity, and developmental genes with no known pleiotropic role in immunity. d_N/d_S values were also compared between pleiotropic genes that scored within the top and bottom quartiles of stage-specific expression (B), where non-specific pleiotropic genes are broadly expressed across life stages ($\tau \leq 0.576$) while the top quartile is specifically or maximally expressed in fewer stages ($\tau \geq 0.767$). The alpha values of genes in each category from the Raleigh (C) and Zambia (D) populations both illustrate higher proportions of adaptive substitutions within pleiotropic genes. Differences among groups were statistically analyzed using a Kruskal–Wallis test (A, C, D) followed by post hoc Dunn tests (P values BH-adjusted) or a Wilcoxon test (B). P values reproduced on the figure; n.s. = not significant ($P_{adj} > 0.05$).

gene model runs for the 6-species data set yielded median d_N/d_S estimates of 0.077, 0.060, and 0.057, respectively (supplementary fig. S7A, Supplementary Material online, table 2), and these three categories exhibited significantly different d_N/d_S distributions based on a Kruskal–Wallis test (supplementary fig. S7A, Supplementary Material

online, $\chi^2 = 46.73$, $P = 7.12e^{-11}$). As with the 12-species data set, the comparison between pleiotropic genes and developmental non-pleiotropic genes does not show a statistically significant difference ($P = 0.25$), but non-pleiotropic immune genes once again have a significantly higher median d_N/d_S value than both non-pleiotropic

developmental genes ($P < 2.2e^{-16}$) and pleiotropic genes ($P < 2.2e^{-16}$). Downsampling via bootstrapping confirmed these results (supplementary methods, Supplementary Material online; supplementary fig. S7B, Supplementary Material online).

Evidence for Positive Selection Across Gene Categories

To determine whether there is evidence for positive selection in any of the three gene categories, we ran codeml site models M7 and M8 in PAML v4.9j (Yang 2007) on each concatenation (see Materials and Methods). Model M7 splits the codons in the alignment into ten groups, where each group contains 10% of the full alignment and has a d_N/d_S value constrained to be less than one. Model M8 splits the alignment into 11 groups, where the proportion of the alignment represented by each group varies; the first 10 groups in M8 have d_N/d_S values constrained to be less than 1, while group 11 can have a d_N/d_S value greater than 1 (representing positive selection in that group of codons). These two models are compared using a likelihood ratio test with two degrees of freedom to determine whether a model allowing for positive selection is a better fit for the data than a model that does not.

A likelihood ratio test between the two models provided significant evidence for positive selection in a fraction of sites within the concatenated alignments of each of the three categories ($P < 0.001$ for all). In the case of the non-pleiotropic immune gene concatenation, the proportion of sites in the eleventh category was 0.007 with an omega value of 5.37. The proportion of sites in the eleventh category for the pleiotropic gene concatenation was 0.015 with an omega value of 1.37. The non-pleiotropic developmental gene concatenation yielded a similar result as the pleiotropic one, with a proportion of 0.018 and omega value of 1.29. The three proportions calculated by model M8 were all statistically different from one another ($\chi^2 = 1034.6$, $P < 2.2e^{-16}$), and each pairwise comparison of proportions was statistically different even after Bonferroni correction ($P < 2.2e^{-16}$ for all three). We also ran models M7 and M8 on concatenated sequences from the six-species data set; likelihood ratio tests for all three categories were significant for this data set as well ($P < 0.001$ for all). Based on all these results, we used MultiDFE to explore positive selection in more depth.

Evidence of Adaptive Evolution Across Gene Categories

The PAML results indicated that non-pleiotropic immune genes had higher d_N/d_S values than either pleiotropic genes or non-pleiotropic developmental genes; the latter two categories were not statistically different from one another (fig. 3A). To help determine whether this difference in d_N/d_S values was driven by adaptive evolution and/or relaxed selection, we used MultiDFE to calculate the proportion of

substitutions that are adaptive (α), the rate of adaptive substitution (ω_a), and the rate of non-adaptive substitution (ω_{na}) for 100 bootstrap replicates of each of the three categories separately for the Raleigh population of *D. melanogaster* (RAL). We obtained site frequency spectra (SFS) from PopFlyData in the iMKT package (Murga-Moreno et al. 2019), and final values of α , ω_a , and ω_{na} were determined using a Jukes–Cantor correction (table 2).

We found that there were significant differences in α across categories (fig. 4A; $P < 2.2e^{-16}$). Median values of α for the non-pleiotropic immune genes, pleiotropic genes, and non-pleiotropic developmental genes were 0.647, 0.774, and 0.714, respectively. Post hoc Dunn tests revealed that there were significant differences in pairwise comparisons of α distributions even after Bonferroni correction ($P < 0.001$ in all cases). For both populations, the median α value was highest in the pleiotropic gene class, followed by the non-pleiotropic developmental gene class and then by the non-pleiotropic immune gene class.

There were also significant differences in ω_a across categories (fig. 4B; $P = 3.202e^{-13}$). Median values of ω_a for non-pleiotropic immune genes, pleiotropic genes, and non-pleiotropic developmental genes were 0.160, 0.178, and 0.152, respectively. Post hoc Dunn tests found that all pairwise comparisons of ω_a distributions were significant after Bonferroni correction ($P < 0.001$ in all cases), where ω_a was the highest for the pleiotropic category.

Additionally, there were significant differences in ω_{na} across categories in both categories (fig. 4C). Median values of ω_{na} for non-pleiotropic immune genes, pleiotropic genes, and non-pleiotropic developmental genes were 0.091, 0.052, and 0.062, respectively. Post hoc Dunn tests found that all pairwise comparisons of ω_{na} distributions were significant after Bonferroni correction ($P < 0.001$ in all cases), where ω_{na} was lowest in the pleiotropic category and highest for the non-pleiotropic immune category.

For all three values (α , ω_a , and ω_{na}), the unbalanced size of each gene category did not affect the results, as confirmed by re-running the MultiDFE analyses by summing the same number of SFS per category (supplementary methods, Supplementary Material online; supplementary fig. S8, Supplementary Material online).

Evidence of Positive Selection in Immune Signaling Pathways

The high overall frequency of pleiotropy among immune signaling genes (fig. 1A) prompted us to examine the distribution of d_N/d_S along the three major insect immune signaling pathways (fig. 5: Imd, Toll, and Jak/STAT) to further investigate whether there are certain components that tend to be pleiotropic or show discernable patterns of ω values. We also ran codeml site models M7 and M8 in PAML on these individual pathway components across the

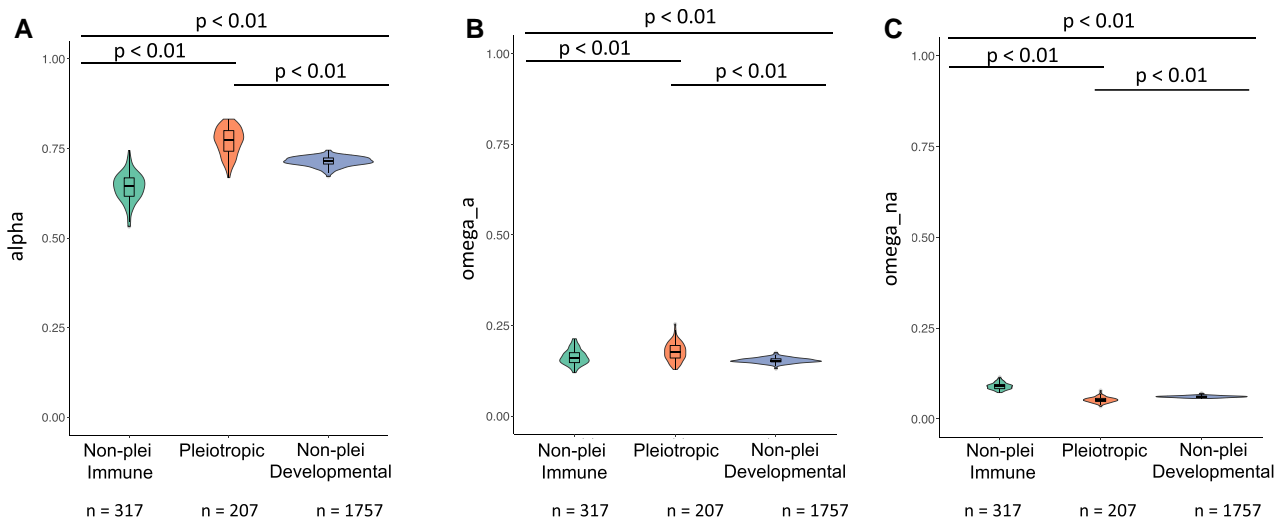


Fig. 4.—Distributions in the *D. melanogaster* Raleigh (RAL) population of (A) α values, (B) ω_a values, and (C) ω_{na} values. α , ω_a , and ω_{na} values were calculated using MultiDFE on 100 bootstrap replicates of summed site frequency spectra (SFS) for each gene category. Distributions were compared using a Kruskal–Wallis test followed by post hoc Dunn tests in R.

12-species data set to determine whether any harbored strong evidence of positive selection.

As illustrated in figure 5, extracellular signaling components tend to be non-pleiotropic, while intracellular signaling components are consistently pleiotropic. The exceptions are immune-specific adapters within the IMD signaling pathway (e.g., Tab2 and Kenny) that interact with pleiotropic proteins. There were no clear patterns with regard to overall d_N/d_S distribution along these pathways, as both the intracellular and extracellular compartments contain proteins with relatively low and high d_N/d_S values. While any estimate of positive selection for individual genes through comparison of model M7 and M8 outputs will be underpowered and thus overly conservative because of the small number of sites, our analysis did still identify several genes in these pathways that contain sites undergoing positive selection (fig. 5 gold stars). Most of these genes are extracellular (ModSP, Sphinx1/2, Upd3) or involved in pathogen recognition (e.g., PGRP-LC and PGRP-LA) and thus conform to the typical profile for immune genes experiencing rapid evolution. However, we also found that the pleiotropic intracellular caspase Dredd exhibited statistical evidence of positive selection (model 8: 5.7% of sites with average $\omega = 1.22$, $P = 0.0006$), providing a salient candidate for future studies of pleiotropy.

Discussion

Researchers have long recognized that some immune genes, such as those in the Toll pathway, play double duty in development (Lemaitre et al. 1996) and proposed that it might constrain immune system evolution (Obbard et al. 2009). Pleiotropy seems like it would be a liability

for a host, for multiple reasons—what if a gene product cannot be deployed to fight a parasite because it is already being fully allocated to development? Should not purifying selection on developmental genes constrain the rate of adaptation against parasite pressure, putting the host at a disadvantage during coevolution with rapidly evolving parasites? In this study, we investigated the relationship between immunity-development pleiotropy and signatures of molecular evolution in *D. melanogaster* immune genes. Our results provide clear quantitative evidence for the notion that pleiotropy between development and immunity is actually quite common (Tate and Graham 2015). Moreover, immune genes involved in development exhibit stronger signatures of evolutionary constraint than non-pleiotropic immune genes, particularly if they are broadly expressed across life stages, consistent with our hypothesis of evolutionary constraint.

In terms of d_N/d_S values, the highest median value was for the non-pleiotropic immune gene class, while the pleiotropic and non-pleiotropic developmental gene classes had more similar medians relative to one another. Interestingly, in most comparisons, these latter two classes had medians that did not statistically differ from one another (fig. 3A, supplementary figs. S6 and S7A, Supplementary Material online) in either the 12-species or 6-species data set. This observation suggests that genes with both immune and developmental functions are similar to developmental-only genes rather than immune-only genes (or an intermediate between the two groups) in terms of evolutionary constraint. We also found that among the three gene categories, pleiotropic immune genes had the highest α and ω_a values and the lowest ω_{na} values (fig. 4, supplementary fig. S8, Supplementary Material online), suggesting that

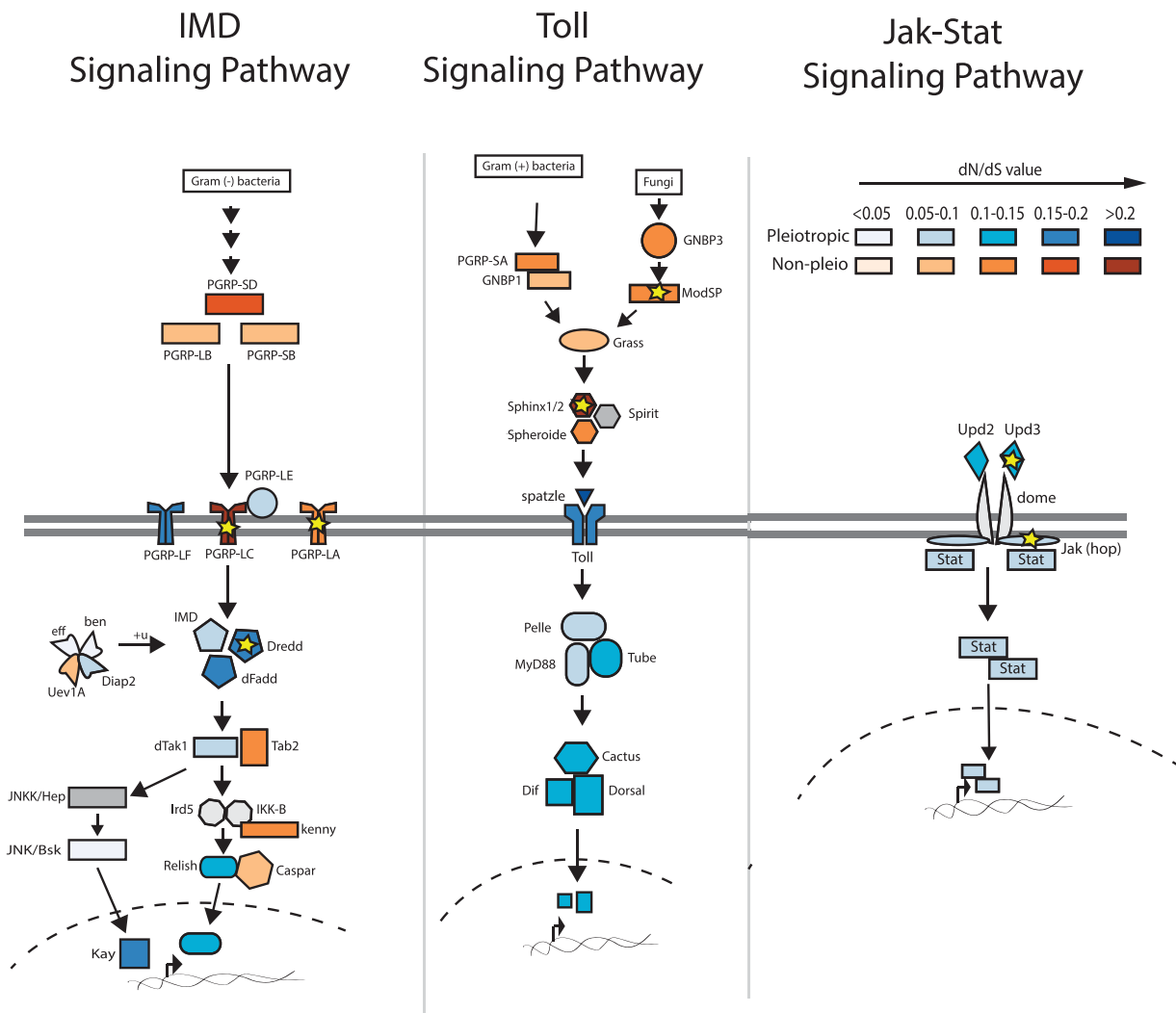


Fig. 5.—Examining the pleiotropy status and d_N/d_S levels for genes participating in major insect immune signaling pathways. The color indicates whether it has pleiotropic roles in development and immunity (blue) or functions exclusively in immunity (orange). Each color is shaded according to the d_N/d_S level of each gene, with the darker shade representing a higher ω value within the gene's respective pleiotropic or non-pleiotropic group. Pathway components reflect annotated genes from KEGG. Components for which no pleiotropy status available (e.g., JNKK and Spirit) are shown in gray. Yellow stars indicate genes that have a positively selected fraction of sites ($d_N/d_S > 1$) as determined by comparison of PAML models M7 and M8 outputs (see Materials and Methods).

increased d_N/d_S values in the non-pleiotropic immune category are at least partially due to an increase in relaxed selection relative to the pleiotropic category. A higher proportion of adaptive substitutions driven by both a higher rate of adaptive substitution and a lower rate of non-adaptive substitution in the pleiotropic category is consistent with stronger purifying selection in those genes compared to non-pleiotropic immune genes.

Our systematic curation of transcriptional data, GO terms, and functional evidence from *D. melanogaster* revealed that about 40–44% of immune genes are pleiotropic with development. This estimate aligns with a phenotypic screening study in mammals that more

generally classified approximately 65% of screened alleles as pleiotropic across a range of phenotypes (De Angelis et al. 2015) [18–21]. Upon analyzing the different immune gene classes for their prevalence of pleiotropy (fig. 1A), we found that immune signaling genes are most likely to participate in developmental functions. This is expected since a signaling pathway is capable of activating the transcription of multiple genes, as opposed to, for example, effector genes which likely only interact with microbial pathogens or have specific immune functions. Further, genes annotated as pleiotropic through our classification method also exhibited significantly higher values of molecular parameters associated with pleiotropy (Alvarez-Ponce et al.

2017), as they have more protein–protein and gene–gene interactions (fig. 1B and C) and are expressed more broadly across life stages and tissues (fig. 2B and C). Although these interactions may not directly reflect immune or developmental activities, it suggests that the pleiotropic genes might participate in different processes by interacting with more molecular partners. The broader expression of pleiotropic genes across stages compared to non-pleiotropic genes suggests that one or both of the immune and developmental functions are required throughout ontogeny. Finally, among the most specifically expressed immune genes (fig. 2B), pleiotropic genes were disproportionately expressed in embryos and pupae—key developmental stages—while the maximum expression of non-pleiotropic genes was more evenly distributed among post-embryonic life stages. This may reflect decoupling of immunological regulation across life stages, which could allow the different life stages to independently optimize immune responses over evolutionary time as they are exposed to different parasites and ecological conditions (Fellous and Lazzaro 2011; Critchlow et al. 2019; Rolff et al. 2019). In the future, it would be interesting to clarify the extent to which pleiotropic genes exhibit temporal segregation of developmental processes and immune roles in different life stages, as opposed to simultaneous participation in both functions in one or more stages.

Our results suggest a significant association between pleiotropy status and the rate of molecular evolution in immune system genes. Other studies that have considered the general relationship between signatures of molecular evolution and molecular pleiotropy have reached contrasting conclusions. In some cases, pleiotropy, as defined by connectivity in protein–protein or gene co-expression networks, is negatively correlated with molecular evolution rates (Alvarez-Ponce et al. 2017; Masalia et al. 2017) as we observe in our study. Meanwhile, others have detected very minimal or no correlation (Hahn et al. 2004; Fraïsse et al. 2018). The variance in these results could be attributed to differences in study organisms, different experimental contexts, and the inherent differences in the various definitions of pleiotropy. For example, our definition of pleiotropy focused on two primary traits rather than considering the entire constellation of traits that might push estimates of pleiotropy in immune systems even higher. The two traits we chose, however, cover the extreme ends of evolutionary rate predictions, as development is thought to be one of the most conserved processes (Artieri et al. 2009), while immunity is consistently identified as one of the most rapidly evolving systems across studied taxa (Obbard et al. 2006; Areal et al. 2011).

We found that α values, which represent that the proportion of substitutions drive by positive selection, were significantly higher in pleiotropic genes than in the other two categories, driven by both higher rates of adaptive

substitution and lower rates of non-adaptive substitution. These results reflect key conclusions from a recent study demonstrating that virus-interacting proteins that participate in diverse cellular processes, which are otherwise more evolutionarily constrained, also showed higher rates of adaptation relative to those that are not known to interact with viruses (Enard et al. 2016). We speculate that when mutations occur in pleiotropic proteins that have antagonistic effects on immunity or development, compensatory substitutions could arise to resolve this conflict. For example, a previous study suggested that the presence of a non-synonymous mutation greatly increases the chance of finding other substitutions nearby, possibly reflecting the correlated evolution of codons within a protein module (Callahan et al. 2011). Because our analyses are not domain specific, we cannot parse signatures of selection on regions within a pleiotropic gene that might provide specific immune or developmental functions or that could be closely associated with compensatory mutations. Although such analysis would require very specific knowledge of the effect of each mutation on immune and development phenotypes, future analyses could focus on a subset of genes with well-defined protein domain structures and protein–protein interaction data to refine the functional and evolutionary significance of pleiotropic activity. Our analysis suggests that Dredd and Jak (fig. 5, gold stars) would be good candidates for such an analysis, while previous studies have identified evidence of positive selection in Dnr1 (Han et al. 2013) and other signaling proteins in *D. melanogaster* (Jiggins and Kim 2007) and related species (Begun and Whitley 2000) that would also provide powerful options for connecting selection at specific sites to the function of pleiotropic proteins.

Across immune pathways, intracellular components are disproportionately pleiotropic compared to extracellular components (fig. 5). Interestingly, however, we observed that many pleiotropic intracellular signaling components associate with non-pleiotropic adapters or interact with proteins that exhibit higher rates of adaptation, which could provide a way to modify pleiotropic protein function in specific immunological contexts to relieve antagonism (Kinsler et al. 2020). This analysis raises new questions for future investigation: how can a signaling pathway balance its role in multiple biological processes? What are the key players and their characteristics that affect how a pathway is used across several contexts or life stages?

Overall, our study serves as the first one to systematically quantify the degree of pleiotropy in a specific biological context and investigate correlations between pleiotropy and rates of molecular evolution in immune systems. These results lay the groundwork for future work to tease apart the mechanistic framework of these pleiotropic patterns to understand how genetic architecture shapes the mode and tempo of immune system evolution and their influence on immune phenotypes.

Materials and Methods

Immune and Developmental Gene List Curation

We curated a comprehensive list of genes representing immunity by combining several resources, starting with a manually curated list from previous immune studies (Lemaitre and Hoffmann 2007; Early et al. 2017), which include most experimentally validated “canonical” immune genes. Separately, we appended GO-annotated genes under the term “immune system process” (GO:0002376) to the list. We further sub-divided genes under this GO term into either “Immune Response” or “Immune Development” genes to differentiate between genes that play direct roles in mounting an immune response and genes contributing to the development and maturation of the immune system. Finally, we added to our list a core set of immune genes from (Troha et al. 2018), which comprises 252 genes that show differential expression across infection with ten different bacterial species of variable virulence.

For each immune gene, we also assigned an immune gene class—recognition, signaling, or effector—based on the gene’s known function in the immune system. If a gene has not been assigned a class in previous studies, we manually assign it a class based on the gene description from FlyBase. For a detailed description of each gene class definition, see [supplementary methods, Supplementary Material online](#).

Separately, we created a list of GO-annotated developmental genes by querying the term “Developmental Process” (GO:0032502), while separately annotating genes belonging to the child term “embryonic morphogenesis” (GO:0048698). All GO annotation queries were conducted through FlyBase (Thurmond et al. 2019). A full list of genes in each group is included in [supplementary table S1, Supplementary Material online](#), and visualization of the degree of overlap between different resources is in [supplementary figure S1, Supplementary Material online](#).

Pleiotropy Categorization

Pleiotropy refers to the phenomenon where a single gene influences multiple traits. However, the definition of “trait” can be ambiguous across different biological contexts, and thus, pleiotropy can manifest at different levels and be detected by various methods (Paaby and Rockman 2013; Tyler et al. 2016). At the molecular level, pleiotropy can refer to the multiple biochemical roles that a gene can have and is frequently measured as the number of physical interacting partners (Hahn et al. 2004). At the developmental or phenotypic level, pleiotropy can involve genes affecting distinct phenotypes or biological processes, as measured by the number of stage or tissues in which such genes are expressed (Artieri et al. 2009). Lastly, under an evolutionary perspective, pleiotropy can refer to the separate

components of fitness that a gene might modulate, a well-known example being the antagonistic pleiotropy model for the evolution of aging (Williams 1957). Though many interpretations of pleiotropy exist, in this study, we are specifically concerned about pleiotropic genes at the phenotypic level. In particular, we focused on genes annotated to play roles in both immune and developmental processes. As such, if a gene is annotated as functioning in both immunity and development from the lists curated from the method described above, it was considered pleiotropic. A full list of pleiotropic genes is included in [supplementary table S1, Supplementary Material online](#).

For comparison purposes, we also calculated molecular metrics of pleiotropy for each gene in the genome regardless of annotated function in immunity or development. These measurements include expression stage specificity (described below), number of associated Biological Processes GO terms, number of associated Molecular Functions GO terms, number of protein–protein interactions, and number of gene–gene interactions. All raw data files were obtained through the FlyBase [ftp server](#), and the latest version of each file was downloaded (March 2020, [supplementary methods, Supplementary Material online](#)).

Categorization of Stage and Tissue Specificity

Genes with functions limited to specific tissues or life stages (and particularly later life stages) may have less pervasive effects on organismal fitness (Cutter and Ward 2005; Artieri et al. 2009), possibly buffering evolutionary constraint from pleiotropy. To calculate expression specificity, we applied the following equation (Yanai et al. 2005) to expression level data of all *D. melanogaster* genes in all stages (embryo, larva, pupa, and adult) and tissues ([supplementary methods, Supplementary Material online](#)):

$$\tau = \frac{\sum_{j=1}^n 1 - \log(A_j)/\log(A_{\max})}{n - 1}.$$

In this equation, n is the number of stages or tissues. A_j is the expression level at stage/tissue j , and A_{\max} is the maximum expression level of stages/tissues. Lower tau (τ) values signify specific expression in a certain stage/tissue, while a higher one indicates broad expression across all stages/tissues (Fraïsse et al. 2018). Tau values for all of the genes used in the analysis are provided in [supplementary table S6, Supplementary Material online](#). In addition to plotting all tau values, we also used bootstrapping to plot the same number of tau values per gene class (we used the lowest number of genes, which was 162 for the pleiotropic class) to account for variation in sample size ([supplementary methods, Supplementary Material online](#)).

Pathway Annotation

We used the PANTHER database to annotate our gene lists to pathway, if available. In short, all genes are compiled into a list of IDs, which is then used as a query in PANTHER (<http://pantherdb.org/>). We then downloaded the annotations and computed the total number of unique pathways associated with each gene group (pleiotropic vs. non-pleiotropic).

Compiling Sequences for PAML Analyses

Genes included in our analyses were chosen using the [table 1](#), row 2, inclusion criteria for non-pleiotropic immune (454 genes), pleiotropic (299 genes), and non-pleiotropic developmental (3,047 genes) lists. We used the FlyBase gene IDs to download coding sequences (CDSs) using the FlyBase Sequence Downloader tool (FB2021_05, released October 15, 2021) for *D. melanogaster* (Thurmond et al. 2019). We then obtained a list of orthologs from FlyBase for all 12 sequenced *Drosophila* species. Using custom scripts (https://github.com/alissawilliams/pleiotropy_Drosophila/tree/main/scripts), we parsed out FlyBase sequence IDs for 11 other *Drosophila* species (*D. ananassae*, *D. erecta*, *D. grimshawi*, *D. mojavensis*, *D. persimilis*, *D. pseudoobscura*, *D. sechellia*, *D. simulans*, *D. virilis*, *D. willistoni*, and *D. yakuba*) for the genes of interest using the *D. melanogaster* IDs. We used the Sequence Downloader tool from an archived version of FlyBase (FB2017_05, released October 25, 2017) to download CDSs for each gene of interest for each of the other 11 species.

We used another set of custom scripts to compile one sequence file for each gene of interest within each pleiotropy category. These scripts added one CDS per species to each file; in cases where more than one CDS was obtained for a single gene ID, the first CDS in the file of downloaded sequences was used. In cases of paralogy (i.e., where one species had multiple gene identifiers within a single orthogroup), the species with gene duplicates were excluded from the sequence file. After this step, 400, 294, and 2,549 sequence files contained at least two sequences for the non-pleiotropic immune, pleiotropic, and non-pleiotropic developmental groups, respectively.

Next, sequence files containing at least two sequences were aligned in codon space with the “*insi*” option in MAFFT v7.310 (Katoh and Standley 2013) using a custom script (https://github.com/dbsloan/perl_modules). Successful alignment occurred for 356 non-pleiotropic immune genes, 231 pleiotropic genes, and 2,067 non-pleiotropic developmental genes. These alignment files were trimmed in codon space using Gblocks v0.91b (Castresana 2000) with parameters $-t=c$ and $-b5=h$. These trimmed files were used in downstream PAML analyses.

Calculating Gene-wide d_N/d_S Values Using PAML

The trimmed sequence files were individually run through codeml site model M0 in PAML v4.9j (Yang 2007) to obtain d_N/d_S values for each gene. The codeml command was run using “*seqtype = 1*,” “*CodonFreq = 2*,” “*model = 0*,” “*NSsites = 0*,” and “*cleandata = 0*.” Constraint trees for each gene were built by starting with the known species tree for the 12 *Drosophila* species on FlyBase and eliminating any species not present in the particular sequence file. The site model M0 runs were successful for 348 of the 356 non-pleiotropic immune genes, 227 of the 231 pleiotropic genes, and 2,037 of the 2,067 non-pleiotropic developmental genes. d_N/d_S values across the three gene categories were compared using a Kruskal–Wallis test followed by post hoc Dunn tests in R (R Core Team 2012). We also used downsampling to account for different sample sizes in the different gene categories ([supplementary methods, Supplementary Material online](#)).

Detection of Positive Selection Using PAML Site Models

To detect positive selection in genes of the three categories, we used codeml site models M7 and M8 in PAML. The trimmed files for each category were concatenated into single alignments and run through codeml with parameters “*seqtype = 1*,” “*CodonFreq = 2*,” “*model = 0*,” “*NSsites = 78*,” and “*cleandata = 0*.” A constraint tree for the 12 *Drosophila* species was built based on the phylogeny provided on FlyBase (Thurmond et al. 2019). Within each class of genes, models M7 and M8 were compared using likelihood ratio tests ($df = 2$). Site model M0 (“*model = 0*,” “*NSsites = 0*”) was also run for each of the three concatenated gene sets using the same parameters as described in the previous section.

In addition to the concatenated sequences, we ran codeml site models M7 and M8 on individual pleiotropic and non-pleiotropic immune genes from the three KEGG-annotated immune signaling pathways ([fig. 5](#)).

PAML Analyses on the Melanogaster Group

In addition to using the “12-species data set” described above, we also conducted PAML tests on the melanogaster group (using 6 representatives: *D. ananassae*, *D. erecta*, *D. melanogaster*, *D. sechellia*, *D. simulans*, and *D. yakuba*) to account for possible dS saturation and/or differences in selection across clades (6-species data set). We used the set of 400 non-pleiotropic immune, 294 pleiotropic, and 2,549 non-pleiotropic developmental genes described above (those that had at least 2 sequences out of the 12 original species after filtering) to identify genes for which there were at least 2 sequences out of the 6 melanogaster group species. After this initial filtering, there were 385 non-pleiotropic immune, 291 pleiotropic, and 2,520

non-pleiotropic developmental genes represented. Of these, 362, 257, and 2,239, respectively, successfully aligned using the “*einsi*” option in MAFFT v7.310 as described above. We trimmed each of these individual alignments using the GBlocks parameters detailed above and ran them through PAML codeml site model M0 again, of which 360, 257, and 2,236 were successful, respectively. We also concatenated the trimmed, aligned files into a single alignment for each class of genes and ran these concatenations through PAML codeml site models M0, M7, and M8 as we did before. Finally, we also conducted downsampling via bootstrapping for this data set to account for differences in sample size (supplementary methods, Supplementary Material online).

Calculation of α , ω_a , and ω_{na} Using MultiDFE

To calculate the proportion of substitutions driven by positive selection (α), the rate of adaptive substitutions (ω_a), and the rate of non-adaptive substitutions (ω_{na}), we used PopFly data from the Raleigh (RAL) population of *D. melanogaster* (Hervas et al. 2017) in the iMKT package in R (Murga-Moreno et al. 2019) as input to the software package MultiDFE (<https://github.com/kousathanas/MultiDFE>). The MultiDFE input was in the form of SFS. The PopFly data was obtained from the file dsimDmelSites.tab provided by Jesús Murga-Moreno (Murga-Moreno et al. 2019). Of the 356 non-pleiotropic immune genes, 231 pleiotropic genes, and 2,067 non-pleiotropic developmental genes included in the concatenated alignments, the dsimDmelSites.tab contained 317, 207, and 1,757, respectively. We modified the code in the iMKT Jupyter notebook (<https://nbviewer.org/github/jmurga/iMKTData/blob/master/notebooks/dmelProteins.ipynb>, accessed June 1, 2022) to obtain raw counts of variants for each gene in each population. We then used bootstrapping to create 100 samples for each gene class in each population by summing variant counts as well as π , p_0 , d_i , d_0 , m_i , and m_0 from the iMKT PopFlyData table, where π = the number of non-synonymous polymorphisms, p_0 = the number of synonymous polymorphisms, d_i = the number of non-synonymous divergences, d_0 = the number of synonymous divergences, m_i = the total number of putatively selected sites, and m_0 = the total number of putatively neutral (Murga-Moreno et al. 2019). Divergence was measured by comparing the *D. melanogaster* population to *D. simulans*. We calculated the 0th column of each SFS (i.e., the number of sites with no observed variants) using the equations $m_i - \pi$ and $m_0 - p_0$ for non-synonymous and synonymous sites, respectively. Scripts used for this process are provided at https://github.com/alissawilliams/pleiotropy_Drosophila/tree/main/scripts.

We ran MultiDFE with the recommended parameters “-conpop 0,” “-sfsfold,” “1 -selmode 4,” “-nspikes 0,”

and “-ranrep 1” (Kousathanas and Keightley 2013) for each bootstrapped SFS file (<https://github.com/kousathanas/MultiDFE>, downloaded April 14, 2022). We then extracted the average fixation probability (fix_prob) for each bootstrap replicate from its respective.sfs.MAXL.out output file. Following Galtier (2016) (eqs. 15 and 16), fix_prob is equivalent to ω_{na} . We also calculated α and ω_a by plugging fix_prob from MultiDFE and the summed d_i and d_0 from PopFlyData into equations (10) and (11) from (Kousathanas and Keightley 2013). Values of d_i and d_0 were corrected using the Jukes–Cantor correction function provided on the MultiDFE GitHub page (<https://github.com/kousathanas/MultiDFE>, accessed April 14, 2022). Distributions of α , ω_a , and ω_{na} values were compared using a Kruskal–Wallis test followed by post hoc Dunn tests in R (R Core Team 2012) in cases where the Kruskal–Wallis test produced a significant result. To account for differences in sample size, we summed the same number of genes per class and re-ran MultiDFE (supplementary methods, Supplementary Material online).

Statistical Analysis

All statistical analyses were conducted in R (4.1.0). We used Shapiro tests to assess distribution normality in data sets. For comparison between multiple groups, we conducted Kruskal–Wallis tests followed by pairwise Dunn tests (in which all possible sets of two categories were compared) with Benjamini–Hochberg correction in cases where there was a significant difference between groups.

Supplementary material

Supplementary data are available at *Genome Biology and Evolution* online (<http://www.gbe.oxfordjournals.org/>).

Acknowledgments

We thank Nora Schulz, Stephanie Birnbaum, and other members of the Tate lab for comments and discussion. We also thank Seth Bordenstein and two anonymous reviewers for their helpful comments on an earlier version of this manuscript. Additionally, we thank Athanasios Kousathanas for his help with compilation and use of the MultiDFE software and Jesús Murga-Moreno for his help with the iMKT package and data set. This work was supported by the National Institute of General Medical Sciences at the National Institutes of Health (grant number R35GM138007 to A.T.T.).

Data Availability

Gene classifications, scripts, untrimmed and trimmed alignments, PAML output, and MultiDFE input and output are provided at https://github.com/alissawilliams/pleiotropy_Drosophila. Additional data are provided in the

supplementary file 1, Supplementary Material online, and supplementary file 2, Supplementary Material online.

Literature Cited

- Alvarez-Ponce D, Feyertag F, Chakraborty S. 2017. Position matters: network centrality considerably impacts rates of protein evolution in the human protein–protein interaction network. *Genome Biol Evol.* 9:1742–1756.
- Anthony N, Foldi I, Hidalgo A. 2018. Toll and Toll-like receptor signaling in development. *Development* 145:dev156018.
- Areal H, Abrantes J, Esteves PJ. 2011. Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC Evol Biol.* 11:368.
- Artieri CG, Haerty W, Singh RS. 2009. Ontogeny and phylogeny: molecular signatures of selection, constraint, and temporal pleiotropy in the development of *Drosophila*. *BMC Biol.* 7:42.
- Begun DJ, Whitley P. 2000. Adaptive evolution of relish, a *Drosophila* NF- κ B/ κ B protein. *Genetics* 154:1231–1238.
- Callahan B, Neher RA, Bachtrog D, Andolfatto P, Shraiman BI. 2011. Correlated evolution of nearby residues in *Drosophilid* proteins. *PLoS Genet.* 7:e1001315.
- Casals F, et al. 2011. Genetic adaptation of the antibacterial human innate immunity network. *BMC Evol Biol.* 11:202.
- Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* 17:540–552.
- Critchlow JT, Norris A, Tate AT. 2019. The legacy of larval infection on immunological dynamics over metamorphosis. *Philos Trans R Soc B Biol Sci.* 374:20190066.
- Cutter AD, Ward S. 2005. Sexual and temporal dynamics of molecular evolution in *C. elegans* development. *Mol Biol Evol.* 22:178–188.
- De Angelis MH, et al. 2015. Analysis of mammalian gene function through broad-based phenotypic screens across a consortium of mouse clinics. *Nat Genet.* 47:969–978.
- DiAngelo JR, Bland ML, Bambina S, Cherry S, Birnbaum MJ. 2009. The immune response attenuates growth and nutrient storage in *Drosophila* by reducing insulin signaling. *Proc Natl Acad Sci USA.* 106:20853–20858.
- Early AM, et al. 2017. Survey of global genetic diversity within the *Drosophila* immune system. *Genetics* 205:353–366.
- Enard D, Cai L, Gwennap C, Petrov DA. 2016. Viruses are a dominant driver of protein adaptation in mammals. *Elife* 5:e12469.
- Fellous S, Lazzaro BP. 2011. Potential for evolutionary coupling and decoupling of larval and adult immune gene expression. *Mol Ecol.* 20:1558–1567.
- Ferrandon D, Imler J-L, Hetru C, Hoffmann JA. 2007. The *Drosophila* systemic immune response: sensing and signalling during bacterial and fungal infections. *Nat Rev Immunol.* 7:862–874.
- Fraïsse C, Puixeu Sala G, Vicoso B. 2018. Pleiotropy modulates the efficacy of selection in *Drosophila melanogaster*. *Mol Biol Evol.* 36:500–515.
- Galtier N. 2016. Adaptive protein evolution in animals and the effective population size hypothesis. *PLoS Genet.* 12:e1005774.
- Hahn MW, Conant GC, Wagner A. 2004. Molecular evolution in large genetic networks: does connectivity equal constraint? *J Mol Evol.* 58:203–211.
- Han M, et al. 2013. Evolutionary rate patterns of genes involved in the *Drosophila* Toll and Imd signaling pathway. *BMC Evol Biol.* 13:245.
- Heil M. 2016. Host manipulation by parasites: cases, patterns, and remaining doubts. *Front Ecol Evol.* 4:1–15.
- Hervas S, Sanz E, Casillas S, Pool JE, Barbadilla A. 2017. Popfly: the *Drosophila* population genomics browser. *Bioinformatics* 33:2779–2780.
- i5K Consortium. 2013. The i5K initiative: advancing arthropod genomics for knowledge, human health, agriculture, and the environment. *J Hered.* 104:595–600.
- Jiggins FM, Kim KW. 2007. A screen for immunity genes evolving under positive selection in *Drosophila*. *J Evol Biol.* 20:965–970.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30:772–780.
- Kinsler G, Geiler-Samerotte K, Petrov DA. 2020. Fitness variation across subtle environmental perturbations reveals local modularity and global pleiotropy of adaptation. *Elife* 9:e61271.
- Kousathanas A, Keightley PD. 2013. A comparison of models to infer the distribution of fitness effects of new mutations. *Genetics* 193:1197–1208.
- Lemaitre B, Hoffmann J. 2007. The host defense of *Drosophila melanogaster*. *Annu Rev Immunol.* 25:697–743.
- Lemaitre B, Nicolas E, Michaut L, Reichhart J-M, Hoffmann JA. 1996. The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell* 86:973–983.
- Lemaitre B, Reichhart J-M, Hoffmann JA. 1997. *Drosophila* host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms. *Proc Natl Acad Sci USA.* 94:14614–14619.
- Masalia RR, Bewick AJ, Burke JM. 2017. Connectivity in gene coexpression networks negatively correlates with rates of molecular evolution in flowering plants. *PLoS One* 12:e0182289.
- McNeil J, Cox-Foster D, Slavicek J, Hoover K. 2010. Contributions of immune responses to developmental resistance in *Lymantria dispar* challenged with baculovirus. *J Insect Physiol.* 56:1167–1177.
- Mukherjee S, Sarkar-Roy N, Wagener DK, Majumder PP. 2009. Signatures of natural selection are not uniform across genes of innate immune system, but purifying selection is the dominant signature. *Proc Natl Acad Sci.* 106:7073–7078.
- Murga-Moreno J, Coronado-Zamora M, Hervas S, Casillas S, Barbadilla A. 2019. iMKT: the integrative McDonald and Kreitman test. *Nucleic Acids Res.* 47:W283–W288.
- Obbard DJ, Jiggins FM, Halligan DL, Little TJ. 2006. Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Curr Biol.* 16:580–585.
- Obbard DJ, Welch JJ, Kim K-W, Jiggins FM. 2009. Quantifying adaptive evolution in the *Drosophila* immune system. *PLoS Genet.* 5:e1000698.
- Paaby AB, Rockman MV. 2013. The many faces of pleiotropy. *Trends Genet.* 29:66–73.
- Palmer WJ, Jiggins FM. 2015. Comparative genomics reveals the origins and diversity of arthropod immune systems. *Mol Biol Evol.* 32:2111–2129.
- Parker BJ, Barribeau SM, Laughton AM, Griffin LH, Gerardo NM. 2017. Life-history strategy determines constraints on immune function. *J Anim Ecol.* 86:473–483.
- Paterson S, et al. 2010. Antagonistic coevolution accelerates molecular evolution. *Nature* 464:275–278.
- R Core Team. 2012. R: a language and environment for statistical computing. Vienna, Austria. <http://www.R-project.org/>
- Rolf J, Johnston PR, Reynolds S. 2019. Complete metamorphosis of insects. *Philos Trans R Soc B Biol Sci.* 374:20190063.
- Rottschaefer SM, et al. 2015. Population genetics of *Anopheles coluzzii* immune pathways and genes. *G3: Genes|Genomes|Genetics* 5:329–339.
- Sackton TB, et al. 2007. Dynamic evolution of the innate immune system in *Drosophila*. *Nat Genet.* 39:1461–1468.
- Sackton TB. 2019. Comparative genomics and transcriptomics of host–pathogen interactions in insects: evolutionary insights and future directions. *Curr Opin Insect Sci.* 31:106–113.
- Schmid-Hempel P. 2008. Parasite immune evasion: a momentous molecular war. *Trends Ecol Evol.* 23:318–326.
- Seabury CM, et al. 2010. Diversity and evolution of 11 innate immune genes in *Bos taurus taurus* and *Bos taurus indicus* cattle. *Proc Natl Acad Sci USA.* 107:151–156.

- Sivakumaran S, et al. 2011. Abundant pleiotropy in human complex diseases and traits. *Am J Hum Genet.* 89:607–618.
- Stark C, et al. 2006. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* 34:D535–D539.
- Tan W-H, et al. 2021. Population genomics reveals variable patterns of immune gene evolution in monarch butterflies (*Danaus plexippus*). *Mol Ecol.* 30:4381–4391.
- Tanji T, Hu X, Weber AN, Ip YT. 2007. Toll and IMD pathways synergistically activate an innate immune response in *Drosophila melanogaster*. *Mol Cell Biol.* 27:4578–4588.
- Tate AT, Graham AL. 2015. Dynamic patterns of parasitism and immunity across host development influence optimal strategies of resource allocation. *Am Nat.* 186:495–512.
- Thurmond J, et al. 2019. Flybase 2.0: the next generation. *Nucleic Acids Res.* 47:D759–D765.
- Tian X, Strassmann JE, Queller DC. 2013. Dictyostelium development shows a novel pattern of evolutionary conservation. *Mol Biol Evol.* 30:977–984.
- Troha K, Im JH, Revah J, Lazzaro BP, Buchon N. 2018. Comparative transcriptomics reveals CrebA as a novel regulator of infection tolerance in *D. melanogaster*. *PLoS Pathog.* 14:e1006847.
- Tyler AL, Crawford DC, Pendergrass SA. 2016. The detection and characterization of pleiotropy: discovery, progress, and promise. *Brief Bioinformatics* 17:13–22.
- Unckless Robert L, Howick Virginia M, Lazzaro Brian P. 2016. Convergent balancing selection on an antimicrobial peptide in *Drosophila*. *Curr Biol.* 26:257–262.
- Viljakainen L. 2015. Evolutionary genetics of insect innate immunity. *Brief Funct Genomics.* 14:407–412.
- Waterhouse RM, et al. 2007. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science* 316:1738–1743.
- Williams GC. 1957. Pleiotropy, natural selection, and the evolution of senescence. *Evolution* 11:398–411.
- Yanai I, et al. 2005. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* 21:650–659.
- Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 24:1586–1591.

Associate editor: Josefa Gonzalez