Single innate immune effectors control ecological microbiome bacteria across evolutionary timescales

M.A. Hanson[1,2*], L. Grollmus[1], and B. Lemaitre[1*]

1 Global Health Institute, School of Life Science, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.
2 Disease Ecology and Evolution, Biosciences, University of Exeter, Penryn, United Kingdom.

Corresponding authors: M.A. Hanson (M.Hanson@exeter.ac.uk), B.L. Lemaitre (bruno.lemaitre@epfl.ch)

Abstract

Antimicrobial peptides are host-encoded antibiotics that combat invading microbes and help shape the microbiome in plants and animals. However, little is known about how the host antimicrobial peptide repertoire is evolutionarily adapted to its microbiome. Here we characterize the function and evolution of the *Diptericin* antimicrobial peptide family of Diptera. Using mutations affecting the two *Diptericins* (*Dpt*) of *D. melanogaster*, we reveal the specific role of *DptA* against the pathogen *Providencia rettgeri* and *DptB* against the gut mutualist *Acetobacter*. Strikingly, presence of *DptA*- or *DptB*-like genes across Diptera correlates with the presence of *Providencia* and *Acetobacter* in their environment. Moreover, presence of *DptA*- and *DpB*-like genes is sufficient to predict host resistance against infection by these bacteria across a broad range of fly species. Our study explains the evolutionary logic behind the bursts of rapid evolution of an antimicrobial peptide family, and reveals how the host immune repertoire adapts itself to changing microbial environments.

Introduction

Animals live in the presence of a complex network of viruses and microbes known as the microbiome. The relationship between host and microbe can vary from mutualist to pathogen, which is often context-dependent [1]. To ensure presence of beneficial microbes and prevent infection by pathogens, animals produce many innate immune effectors as a front-line defence. Chief amongst these effectors are antimicrobial peptides (AMPs): small, cationic, host-encoded antibiotics that combat invading microbes in plants and animals [2–5]. While many studies have shown important roles for AMPs in regulating the microbiome [6], presently, we cannot define why animals have the particular repertoire of AMPs their genome encodes.

Innate immunity has been characterized extensively in *Drosophila* fruit flies [7,8]. Antimicrobial peptide responses are particularly well-characterized in this insect [3,9,10], including recent dissections of these effectors through systematic gene deletion [11–16]. In *Drosophila*, AMP genes are transcriptionally regulated by the Toll and Imd NF-κB signalling

pathways [8]. Recent work has shown individual effector genes can play prominent roles in defence against specific pathogens [12,14,15,17–20]. Consistent with this, population genetics studies have highlighted genetic variants in AMPs correlated with susceptibility against specific pathogens. A landmark study in *Drosophila* found that a serine/arginine polymorphism at residue 69 in one of the two fruit fly Diptericins, "S69R" of DptA (Fig. 1A), is associated with increased susceptibility to *Providencia rettgeri* bacterial infection [21]. Mutant studies later showed that flies lacking both *Diptericin* genes ("Dpt^{SK1}" flies lacking *DptA* and *DptB*) are as susceptible to *P. rettgeri* infection as Imd pathway mutants, while flies collectively lacking five other AMP families nevertheless resist infection as wild-type [11]. Like these studies in *Drosophila*, a recent study of Persian domestic cats uncovered a striking correlation between a G49E polymorphism in the AMP Calprotectin, and susceptibility to severe "ringworm" fungal skin disease [22]. These examples highlight the importance of single AMPs in combatting specific infections, and similar AMP variation is common across animals [23–26]. However, while *P. rettgeri* is an opportunistic pathogen found in wild flies, and ringworm fungus is common in certain cat breeds, it is unclear if these AMPs are actually evolving to selection imposed by these microbes. Given many recent studies highlighting the role of AMPs beyond infection [3,27], other fitness trade-offs could also explain AMP evolution.

It is now clear that antimicrobial peptides shape the microbiome [6], but defining if or how the host immune repertoire itself is shaped by the microbiome has been challenging. Here we characterize the function and evolution of the *Diptericin* gene family across dipteran lineages, revealing that these AMPs were selected to control ecologically-relevant microbes.

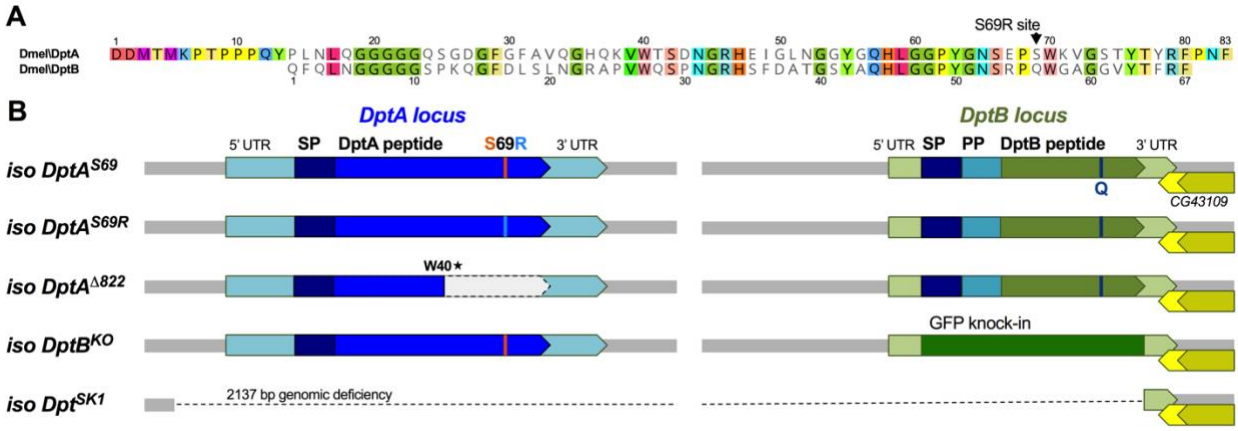


Figure 1: Diptericin loci of D. melanogaster. *A) Protein alignment of D. melanogaster DptA and DptB mature peptides. The $DptA^{S69R}$ site that regulates susceptibility to P. rettgeri is noted. These two Diptericin mature peptides are ~52% identical. B) The two Diptericin genes are located in tandem on Chromosome 2R (cytogeneic map 55F) with a span of only 1130bp between the end of the DptA gene region and the start of the DptB gene region. Amino acid residues at the S69R polymorphic site of DptA or its equivalent in DptB (Q) are annotated. The $DptA^{\Delta 822}$ premature stop site is noted (W40□). The Dpt^{SK1} deficiency removes 2137bp encompassing DptA and the coding sequence of DptB. DptB encodes a secreted propeptide (PP) not found in DptA. "SP" = Signal Peptide.*

Results

Diptericin B is specifically required for defence against Acetobacter bacteria

We previously showed that a strain of A*cetobacter* grows out of control in the gut of *Relish* mutant flies (*Rel*^{E20}) lacking the Imd pathway, or in flies carrying deletions removing 14 AMP genes (*ΔAMP14*) [28]. Here we identify this *Acetobacter* species as *A. sicerae strain BELCH* (Fig. 2supp1). *Acetobacter* species are common gut symbionts of *D. melanogaster*, suggesting this strain is likely a non-pathogenic component of the gut microbiome. Consistent with this, oral infection with *A. sicerae* does not cause mortality, even in *ΔAMP14* flies (Fig. 2supp2). However, pricking flies with a needle contaminated with *A. sicerae* kills *ΔAMP14* and *Rel*^{E20} flies [17,28], and this systemic infection also causes a severe abdominal bloating phenotype that precedes mortality (shown later). These observations suggest that AMPs regulated by the Imd pathway are required to control *A. sicerae*. We therefore used flies carrying overlapping sets of AMP mutations [11], including a *Diptericin* mutant panel with mutations affecting each of the two *Diptericins* (Fig. 1B), to narrow down which AMP(s) protect the fly against *A. sicerae* systemic infection.

Ultimately, deleting just *DptB* fully recapitulates the susceptibility of *ΔAMP14* flies. Both *Dpt*^{SK1} flies lacking both *Diptericin* genes and *DptB*^{KO} flies suffered 100% mortality after infection, with survival curves mirroring *ΔAMP14* and *Rel*^{E20} flies. *DptB*^{KO} flies also present abdominal bloating with the same frequency as *Rel*^{E20} and *ΔAMP14* flies (Fig. 2A-B). On the other hand, flies with the *DptA*^{S69R} polymorphism, lacking functional *DptA* (*DptA*$^{\Delta 822}$), or even *ΔAMP8* flies collectively lacking five other AMP gene families (*Drosocin, Attacin, Defensin, Metchnikowin, and Drosomycin*), resisted infection more comparable to wild-type. We further confirmed the specificity of *DptB* in defence against *A. sicerae* using RNA interference (RNAi), where ubiquitous knockdown *of DptB* causes mortality and bloating upon infection with *A. sicerae* (Fig. 2supp2B-C). Finally, *DptB* mutants display increased *A. sicerae* loads pre-empting mortality (Fig. 2C), suggesting a direct role for *DptB* in suppressing *A. sicerae* growth.

After revealing the critical importance of *DptB* in defence against *A. sicerae*, we tested if *DptB* has a broader role in the control of other *Acetobacter* species. To this end, we infected flies with a panel of diverse *Acetobacter* species including *A. aceti*, *A. indonesiensis, A. orientalis, A. tropicalis,* and *A. pomorum*. While these *Acetobacter* species displayed different levels of virulence, *DptB* appears to play a major role to promote survival and/or prevent bloating in all virulent *Acetobacter* species (Fig. 2supp3, Fig. 2supp4).

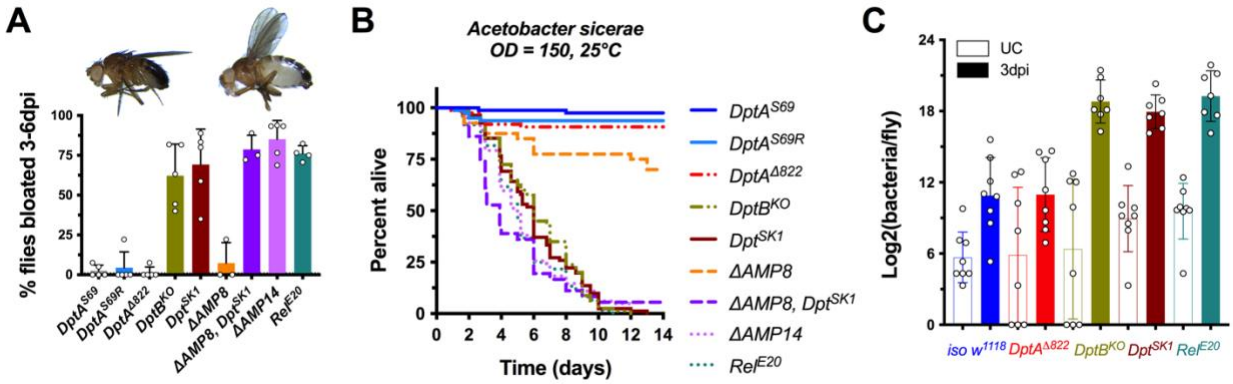


Figure 2: DptB is specifically required for defence against A. sicerae. *A) Flies lacking DptB bloat after A. sicerae infection. Each data point reflects the average from one replicate experiment (~20 flies). B) Survival after systemic infection shows DptB alone controls defence against A. sicerae. C) A. sicerae bacterial load increases prior to mortality. Each data point reflects the average of 5 pooled flies.*

Collectively, these results indicate that *DptB* is an AMP of specific importance in defence against a broad range of *Acetobacter* species, revealing another example of high specificity between an innate immune effector and a microbe relevant to host ecology. As *Acetobacter* are common in fermenting fruits, the major ecological niche of *Drosophila*, *DptB* might be especially important for flies to colonize this niche.

Diptericin A is specifically required to defend against P. rettgeri

The Gram-negative bacterum *P. rettgeri* has been found in the hemolymph of wild-caught flies [21]. This bacterium displays moderate pathogenicity to wild-type flies upon systemic infection, suggesting it is an opportunistic pathogen in *Drosophila*. Previous studies have shown that the Imd pathway, and specifically *Diptericins,* plays a major role in surviving *P. rettgeri* infection. A genome-wide association (GWAS) further highlighted a strong correlation between the residue at the serine/arginine $DptA^{S69R}$ polymorphism and resistance against this bacterium: flies encoding arginine were more susceptible than flies encoding serine at this site. However, it is unclear if *DptB* plays a role in controlling *P. rettgeri* infection, as previous studies did not assess *DptB* ([11,21] and Fig. 3supp1).

We therefore infected our panel of *Diptericin* mutants with *P. rettgeri* to determine the contribution of *DptB* in defence against this microbe (Fig. 3A). We confirmed the $DptA^{S69R}$ allele reduces survival after *P. rettgeri* infection [21], here with a controlled genetic background ($p < .001$). The survival rate of $DptA^{\Delta 822}$ flies also paralleled mortality of Dpt^{SK1} flies lacking both *Diptericin* genes ($P = .867$). Initially, we found that $DptB^{KO}$ flies also showed higher susceptibility to *P. rettgeri* ($P = 1e\text{-}5$), correlated with higher bacterial load (Fig. 3B). However, we realized $DptB^{KO}$ flies have only ~57% the induction of the *DptA* gene compared to the $DptA^{S69}$ wild-type 7 hours post-infection (hpi) (Fig. 3C). $DptB^{KO}$ flies

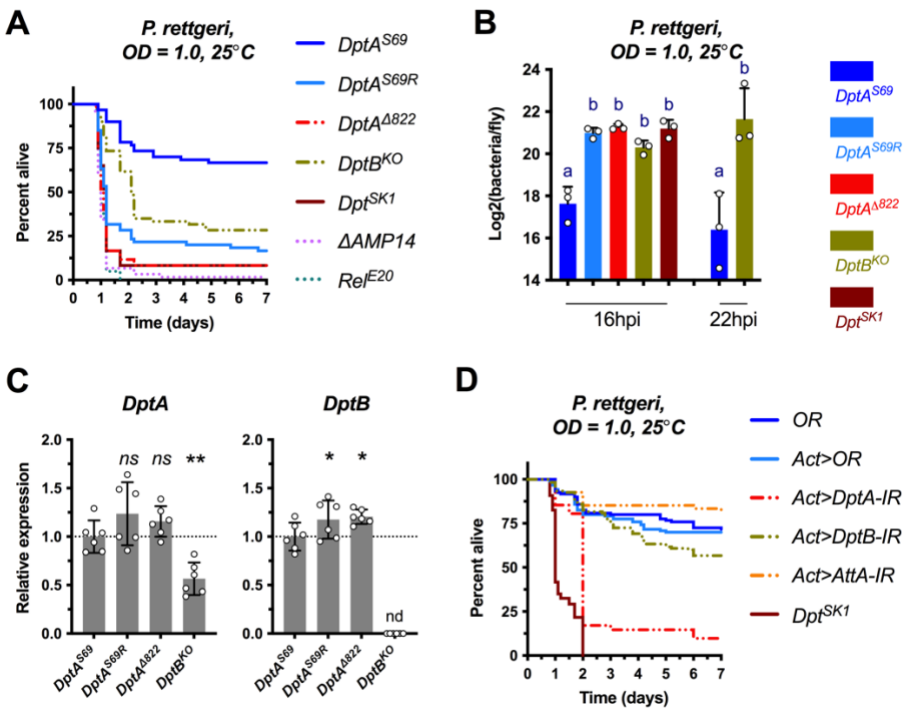


Figure 3: DptA is specifically required for defence against P. rettgeri. *A-B) Survival (A) and bacterial loads (B) of isogenic Diptericin mutants after infection with P. rettgeri.. Each point in (B) is the average bacterial load of 5 pooled flies. C) DptA expression in DptB^{KO} flies reaches only ~57% of DptA^{S69} wild-type in flies collected at 7hpi. D) Silencing DptA by RNAi causes a marked susceptibility to P. rettgeri while DptB silencing causes only a minor susceptibility.*

have an otherwise normal immune response (Fig. 3supp2), suggesting this lower induction of *DptA* is due to a cis-genetic background effect. Indeed, silencing *DptB* using RNAi did not significantly affect survival against *P. rettgeri* (*Act>OR vs. Act>DptB-IR*: *P* = .067; Fig. 3D, Fig. 3supp3). We therefore conclude that *DptB* does not have a major effect on resistance to *P. rettgeri*.

Taken together, use of our panel of *Diptericin* mutants shows that *DptA* plays a major role in defence against *P. rettgeri,* but not *A. sicerae.* Conversely, *DptB* plays a major role against *A. sicerae*, but not *P. rettgeri*. Thus, each of these two *Diptericin* genes are highly specific effectors explaining most of the Imd-mediated defence of *D. melanogaster* against either bacterium.

The Diptericin family shows multiple bursts of rapid evolution across Diptera

The high specificity of the two *Drosophila Diptericins* in defence against different ecologically-relevant microbes motivated us to ask if host ecology might explain *Diptericin* evolution. To address this question, we first reviewed the evolutionary history of *Diptericins* within *Drosophila* and across Diptera using newly-available genomic resources.

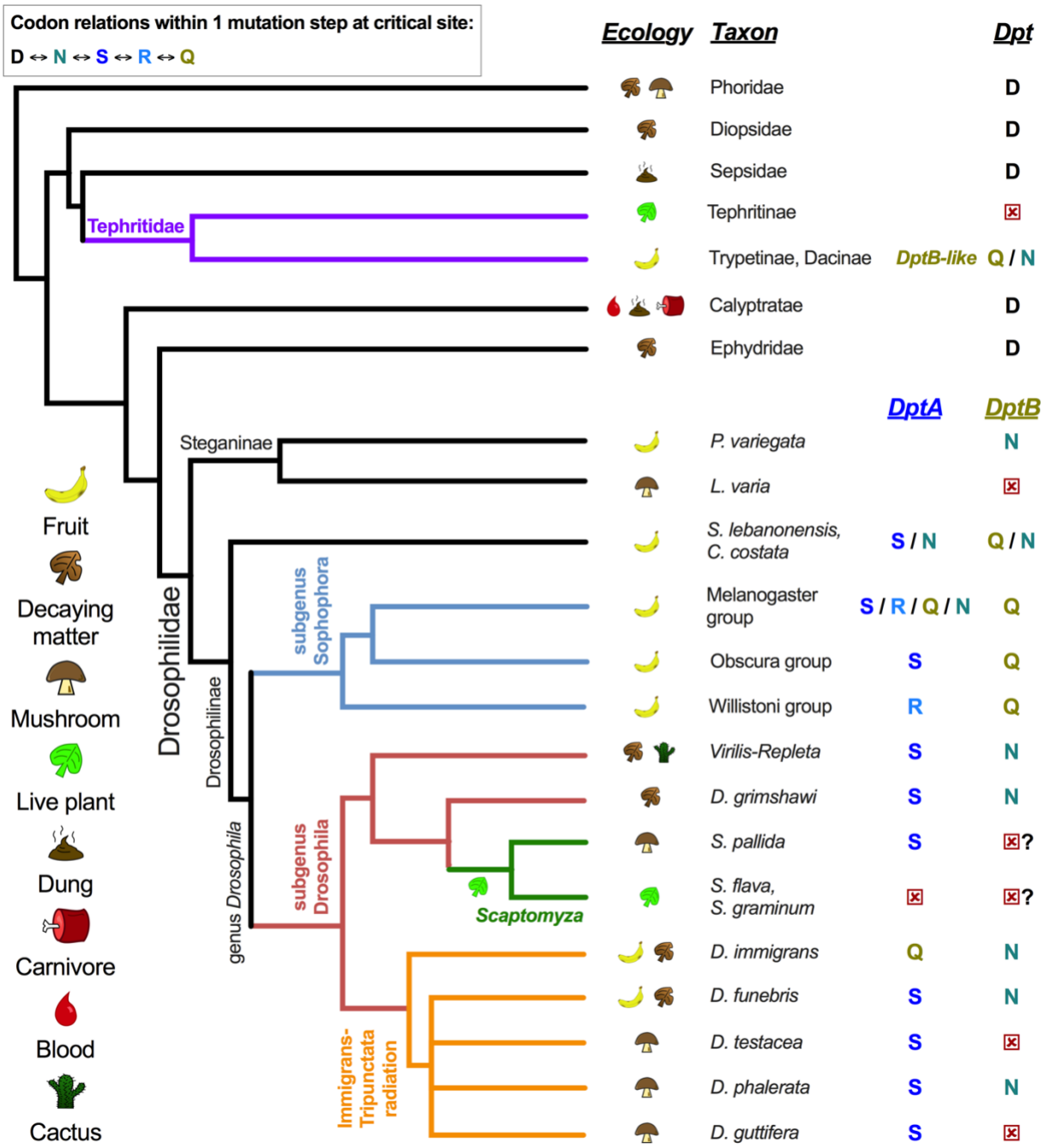


Figure 4: Diptericin evolution correlates with host ecology and presence of Acetobacter or Providencia species. *The presence of Diptericin(s), including the emergence of DptA-like and DptB-like genes, was screened in diverse dipteran species. The residue aligned with the DptA^{S69R} or DptB^{Q56N} polymorphism site is shown for each gene. Figure 4supp2 confirms that serine evolved by convergent evolution at this site in DptA at least twice, and explains how relatedness of codons for these five amino acids constrains their evolution (summary in top left). Fruit-feeding tephritids convergently evolved a DptB-like gene (Fig. 4supp1) including a parallel Q/N polymorphism. Within Drosophilidae (bottom part) three mushroom-feeding species have pseudogenized their DptB genes. In both Drosophilidae (specifically Scaptomyza) and Tephritidae, divergence from fruit-feeding to plant-feeding is also correlated with pseudogenization patterns in DptB-like genes. Plant-feeding Scaptomyza have also pseudogenized at least one of their DptA genes. Microbiome studies suggest absence of Providencia and Acetobacter in the host ecology is correlated with DptA and DptB pseudogenization respectively (Fig. 4supp3). Red [x] = pseudogenized, ? = likely pseudogenized given circumstantial evidence (see Fig. 4supp4,5).*

We began by tracing the evolutionary origins of the two *Drosophila Diptericin* genes. *Diptericins* are found in many dipteran species [29,30], and perhaps even beetles [31], indicating an ancient origin of this antibacterial peptide that predates much of the diversification of the "true flies". The *Drosophila* gene encoding the extant DptB-like protein structure was originally derived in the Drosophilidae ancestor through rapid evolution ([30,32] and Fig. 4supp1 for protein comparisons). Later, a duplication of *DptB* gave rise to the *DptA* locus in the Drosophilinae ancestor about ~50ma (date per [33]), which also evolved rapidly after the duplication (Fig. 4supp1,2 and [32]). Given these repeated bursts of rapid evolution, the distinct roles of *DptA* and *DptB* against *P. rettgeri* and *A. sicerae* are not necessarily surprising; indeed, the mature *D. melanogaster* Diptericins only show 52% similarity (Fig. 1A). Furthermore, we realized the *DptA* S69 residue of *D. melanogaster* is also present in the subgenus Drosophila due to convergent evolution: different codons are used by the subgenus Sophophora (e.g. AGC) and subgenus Drosophila (e.g. TCA) to produce DptA proteins with the S69 residue (Fig. 4supp2). Thus, this residue was derived at least twice, providing further evidence that adaptive evolution selects for serine at this site (complementing previous analyses [21,32]). Moreover, there is a high level of variation at this site across species: in addition to the serine and arginine alleles found in *D. melanogaster*, this polymorphic site can also encode either glutamine (Q) or asparagine (N) in *DptA* genes of other *Drosophila* species. Interestingly, the same Q or N residues are also seen at the aligned residue of DptB across *Drosophila* species (Q56N in *DptB*, Fig. 4). These four residues (S, R, Q, N) found in *Drosophila* Diptericins are derived from an ancestral aspartic acid residue (D) found in most other dipterans (Fig. 4, Fig. 4supp1,2).

Collectively, this analysis suggests that the extant *DptB* gene first evolved in the drosophilid ancestor, while *DptA* emerged from a duplication of *DptB* followed by rapid diversification. The $DptA^{S69}$ residue, which is key to resist *P. rettgeri* infection, was also derived in at least two independent lineages, and this site is highly polymorphic across genes and species. These repeated bursts of rapid evolution suggest these *Diptericin* genes have evolved in response to selective pressures affecting the drosophilid ancestor.

Diptericin evolution correlates with microbe presence in host ecology

Drosophilidae is a highly diverse lineage that includes species with various ecologies feeding on fruits, mushrooms, or even live plants. Moreover, there have been a number of studies characterizing the microbiomes of drosophilid species. This places us in a unique position to pair the microbial ecologies of flies with patterns in the evolution of their *Diptericins*, which have microbe-specific importance.

We next carried out a systematic review of Diptera microbiome literature (Fig. 4supp3). *Acetobacter* bacteria, which are controlled by *DptB*, are regularly found across

species feeding on rotting fruits in *Drosophila* microbiome studies [34–37]. However, *Acetobacter sp.* appear to be largely absent from rotting mushrooms [38], and are not readily found in mushroom-feeding flies themselves [38,39]. Meanwhile, *Providencia* bacteria related to *P. rettgeri*, which is controlled by *DptA*, are common in species feeding on both rotting fruits and mushrooms. Strikingly, we observed that three drosophilid species with mushroom feeding ecology, *D. testacea*, *D. guttifera*, and *Leucophenga varia*, have independently pseudogenized their *DptB* genes (Fig. 4 and [30]). Thus, three independent *DptB* pseudogenization events have occurred in mushroom-feeding flies, an ecology that does not commonly expose them to *Acetobacter*.

We also noticed another *Drosophila* sublineage whose ecology lacks *Acetobacter*: *Scaptomyza* (Fig. 4 green branch with leaf). Specifically, the *Scaptomyza* ancestor derived a leaf-mining ecology, parasitizing live plants [40]. This plant-feeding ecology is still found in some extant species (e.g. *Scaptomyza flava, Scaptomyza graminum*), while others later shifted to ecologies including mushroom-feeding (*Scaptomyza pallida*). Of note, the *S. flava* microbiome shows little prevalence of either *Acetobacter* or *Providencia* [41]. Using available genomes, we checked if *Scaptomyza* species had pseudogenized any of their copies of *DptA* (two genes, *DptA1* and *DptA2*) or *DptB* (one gene). Notably, we found a premature stop codon in *DptA1* in both live plant-feeding species *S. flava* and *S. graminum*, but not the mushroom-feeding *S. pallida* (Fig. 4supp4). We also analysed the promoter regions of these *DptA* genes for presence of Relish NF-κB transcription factor binding sites ("Rel-κB" sites, motif from [42]), and confirm Rel-κB sites in the *S. pallida DptA1* promoter, suggesting this gene is immune-induced. Thus, *Scaptomyza DptA* genes show pseudogenization specifically in the plant-feeding species that lack *Providencia* in their present-day ecology. However, *DptA1* appears functional in *S. pallida*, a mushroom-feeding *Scaptomyza* species likely exposed to *Providencia* through its diverged ecology.

We next screened for *DptB* genes of *Scaptomyza*. We found *DptB* genes in all three species with no obvious loss-of-function mutations in coding sequence. However, *DptB* genes in all three *Scaptomyza* species lack κB sites in their promoter regions. Whether due to plant-feeding or mushroom-feeding, none of these *Scaptomyza* species have an ecology associated with *Acetobacter*. This lack of κB sites suggests *DptB* may have been pseudogenized at the expression level in this lineage. Of note, we confirmed that the κB site motif we used is conserved in *Scaptomyza* and close relatives, as it is present in the promoters of *Diptericins* in related species (Fig. 4supp4) and upstream of an *Attacin* AMP across *Scaptomyza* and relatives (Fig. 4supp5). Taken together, patterns in *Scaptomyza DptA* and *DptB* genes suggest these genes are pseudogenized *(DptA1)*, or in the process of pseudogenization *(DptB)*. Importantly, premature stops or κB site losses occur in genes correlated with presence or absence of *Providencia* or *Acetobacter* in their ecology. Overall,

our survey of *Diptericins* within Drosophilidae reveals a striking congruence between the loss of *DptA* or *DptB* and the presence of *Acetobacter* or *Providencia* in the host ecology across multiple lineages.

Finally, there is also evidence of convergent evolution towards *DptB*-like sequence in another lineage of "fruit flies": Tephritidae [30,32]; see Fig. 4supp1,2 for alignment and paraphyly of tephritid *Diptericins* clustering with drosophilid *DptB*. This family of Diptera is distantly related to Drosophilidae having a last common ancestor ~111ma. Many tephritid lineages feed on fruits like *Drosophila* (e.g. Trypetinae, Dacinae), but one lineage parasitizes live plants like *Scaptomyza* (Tephritinae: Fig. 4 purple branches). In light of the present study, the tephritid species that feed on fruit are the ones that have convergently evolved a *DptB*-like gene, and *Acetobacter* are commonly found in their microbiomes [43–45]. This convergent evolution of *DptB*-like sequence even includes a parallel Q/N trans-species polymorphism at the critical Diptericin residue, mirroring the DptB Q56N polymorphism found in Drosophilidae (Fig. 4supp1,2), compared to the ancestral D residue. Similar to *Scaptomyza* species, plant-parasitizing tephritids lack both *Acetobacter* and *Providencia* in their microbiomes [30], and have lost their *Diptericin* genes (Fig. 4 and [30]). Convergent evolution of *DptB*-like genes in Tephritidae parallels the derivation of *DptB*-like sequence in Drosophilidae, associated with fruit-feeding ecology where *Acetobacter* is a dominant member of the microbiome.

Collectively, our phylogenetic and ecological survey reveals multiple parallels between the host immune effector repertoire, ecology, and the associated microbiome. This suggests that these dipteran species have derived *DptA*- or *DptB*-like genes as their evolutionary solution to control important bacteria found in their microbiome. In contrast, specific *Diptericin* genes become superfluous if their hosts shift to ecologies lacking *Diptericin*-relevant microbes, leading to their pseudogenization. Thus, the host AMP repertoire can rapidly evolve to tailor itself for defence against newly-relevant microbiome members.

Variation in DptA or DptB predicts host resistance across species separated by 50 million years of evolution

Our study indicates that among the suite of immune genes involved in *Drosophila* host defence, the single AMP genes *DptA* and *DptB* are critically important against two environmentally relevant bacteria: *P. rettgeri* and *Acetobacter*. Moreover, our phylogeny-microbiome analysis reveals striking correlations in terms of gene emergence, retention, and loss. The critical polymorphic site also convergently evolved across, and segregates with, specific *Diptericin* lineages. If *DptA* and *DptB* really evolved to control *P. rettgeri* and *Acetobacter*, the outcomes of *P. rettgeri* and *A. sicerae* infection across species should be

readily predicted using just variation in these two *Diptericin* genes. We therefore chose 12 *Drosophila* species with variation in the polymorphic site in *DptA* and presence/absence of *DptB*, and infected them with *P. rettgeri* or *A. sicerae*. Of note, experiments in *D. melanogaster* suggest that $DptA^{S69R}$ affects defence against *P. rettgeri*, but how $DptA^{S69Q}$ or $DptA^{S69N}$ affects defence against this bacterium has never been tested. Similarly, the effect of $DptB^{Q56N}$ on defence is also untested, and so we have no *a priori* expectations about how these polymorphisms should affect resistance to either microbe. To analyze these experiments, we used a linear mixed-model approach, and we included *D. melanogaster* flies from our *Diptericin* mutant panel as experimental controls. This helped to calibrate our model for the expected effect size for variants of *DptA* or *DptB* within a single species or across species. We also conducted these experiments at 21°C to avoid heat stress to some species, which reduced *D. melanogaster* mortality compared to 25°C (Fig. 5supp1).

Summaries of fly species mortality are shown in Figure 5. As found in *D. melanogaster*, resistance to *P. rettgeri* is correlated with the presence of *DptA* having a serine-coding $DptA^{S69}$ allele across species. Indeed, $DptA^{S69R}$ found in either *D. melanogaster* or *D. willistoni* correlates with increased susceptibility to *P. rettgeri* (P = 2e-7). Moreover, *D. yakuba* with $DptA^{S69N}$ was also more susceptible, suggesting asparagine (N) is an immune-poor allele against *P. rettgeri* (P = 4e-3). Meanwhile, $DptA^{S69Q}$ flies (*D. suzukii, D. immigrans*) had, if anything, slightly improved survival against *P. rettgeri* compared to $DptA^{S69}$ flies (P = .04), suggesting glutamine (Q) is a competent defence allele against *P. rettgeri* in the protein backbone of DptA (Fig. 5A). Overall, our linear mixed-model suggests that ~71% of variation in susceptibility can be attributed to variation in *DptA* alone as a fixed effect (marginal R^2 = .713).

For infections with *A. sicerae*, we observed a similar correlation regarding the presence of *DptB* and resistance to *A. sicerae*. In particular, absence of *DptB* in the mushroom-feeding species *D. testacea* and *D. guttifera* correlates with increased susceptibility compared to their close relatives that feed on either mushrooms or fruits (*D. phalerata* and *D. funebris* respectively) (P = 3e-12). Of note, mushroom-feeding flies tested in our study generally displayed increased susceptibility to *A. sicerae* infection compared to flies with fruit-feeding ecologies (P = 3e-9). However, even within this more susceptible lineage, *DptB* loss still increases mortality to a similar extent as *DptB* deletion in *D. melanogaster*, indicating the contribution of *DptB* in this lineage is independent of host genetic background (Fig. 5B). Overall, our linear mixed-model suggests that ~87% of variation in susceptibility to *A. sicerae* can be explained by just *DptB* presence/absence and host ecology as fixed effects (marginal R^2 = .866).

These survival data establish that the *Diptericin* specificity we observed in *D. melanogaster* applies across *Drosophila* species separated by ~50 million years of evolution. The importance of AMP-microbe specificity remains true, whether variation is

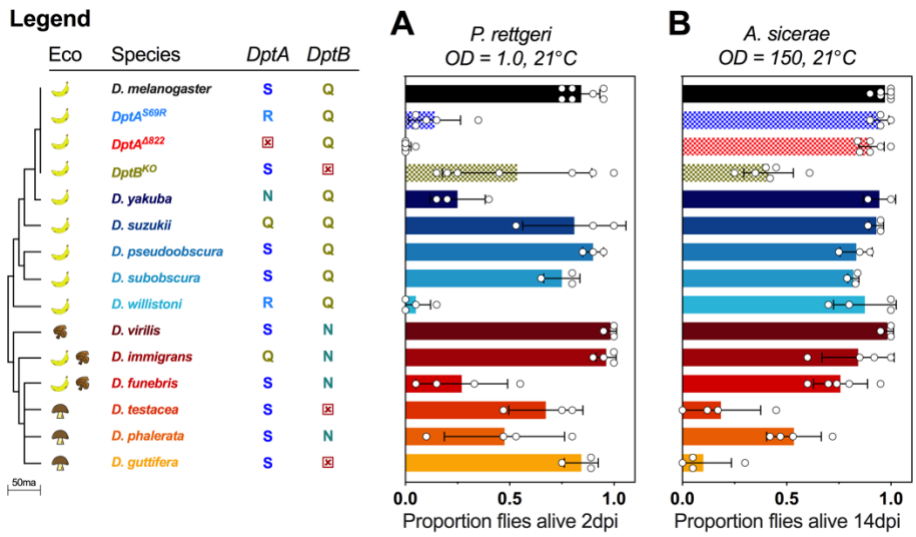


Figure 5: Diptericins predict pathogen-specific survival across the Drosophila genus. *Host phylogeny, ecology, and Diptericin complement is shown on the left. No species succumbed to clean injury within the first two weeks after infection (Fig. 5supp2). A) Susceptibility to P. rettgeri infection varies across species, with variation in survival largely explained by DptA allele and copy number, particularly within subgenus Sophophora species (D. melanogaster and blue-shaded species). B) Susceptibility to infection by A. sicerae is predicted by presence/absence of DptB, but mushroom-feeding flies (D. testacea, D. phalerata, D. guttifera) had a higher baseline susceptibility to A. sicerae infection independent of DptB loss.*

present as polymorphisms or as gene presence/absence. Our results show that single AMPs explain much of how these species can resist infection by microbes associated with their ecological microbiome. In conclusion, the host immune repertoire can adapt to the presence of new and ecologically-relevant microbes through evolution of specialized AMPs as weapons to combat specific microbes.

Discussion

While the evolution of immune genes is often interpreted with host phylogeny as a principle explanatory factor [7,46], host ecology can be a stronger driver of microbiome community structure [47]. Early studies of immune evolution suggested that AMPs were mostly generalist peptides with redundant function, suggesting AMP variation was not due to adaptive evolution [3,4]. Indeed, many studies on immune adaptation focus on broad-level effects, or identify factors specific to a single species. For instance, loss of entire immune pathways has been described for arthropods living in 'clean' or protected ecologies associated with plant-feeding [48]. The coevolution of hosts with microbial symbionts can rely on AMP-microbe dynamics maintaining a delicate balance [49], or on hosts relying on symbionts for defence [50]. Hosts can also have behavioural or ecological adaptations, preferring food resources that reduce the risk of parasitism [51]. However, despite a rich literature on immunity-microbiome interactions, the evolutionary logic explaining why the host genome encodes its particular immune effector repertoire has been difficult to approach experimentally.

Here we identify how specific effectors evolved to constrain the growth of specific microbes, finding that host ecology, not host phylogeny, makes sense of effector-microbe specificity. This contributes to a larger body of literature showing immune genes are among the fastest-evolving genes in the genome [7,52,53]. We also reveal how previously-noted bursts of positive selection in an AMP family correlate with host ecological changes, and identify the ecological microbes driving the evolution of those AMPs. The two *D. melanogaster Diptericin* genes also provide a textbook example of how gene duplication can promote immune novelty, equipping the host with extra copies of immune tools that can be adapted to specific pathogen pressures. Moving forward, it is now clear that AMPs can have strong and specific impacts on survival, and that even single residue polymorphisms can drastically impact infection outcomes [17,21,22]. One goal of infection biology is to try to identify risk factors for susceptibility present in individuals and populations. Our study reveals that characterizing the function of single effectors, interpreted through an evolution-ecology framework, can help explain how and why variation in infection susceptibility occurs within and between species.

The fly *Diptericin* repertoire reflects the presence of relevant microbes in that species' ecology. Conversely, loss or pseudogenization of *Diptericins* is observed when the microbes they target are no longer present in their environment. In a sense, this means that some AMPs seen in the genomes of these animals are, effectively, vestigial: immune genes evolved to fight microbes the extant host rarely encounters (e.g. *DptB* in *D. phalerata*). While other mechanisms of defence assuredly contribute to resistance, *Diptericins* have evolved recurrently as the fly genome's solution to control specific bacteria. Given our findings, we propose a model of AMP-microbiome evolution that includes gene duplication, sequence convergence, and gene loss, informed by the host ecology and the associated microbiome (Fig. 6). In doing so, we explain one part of why various species have the particular repertoire of AMPs that they do. This ecology-focused model of AMP-microbiome evolution provides a framework to understand how host immune systems rapidly adapt to the suite of microbes associated with a new ecological niche. These findings are likely of broad relevance to immune evolution in other animals.

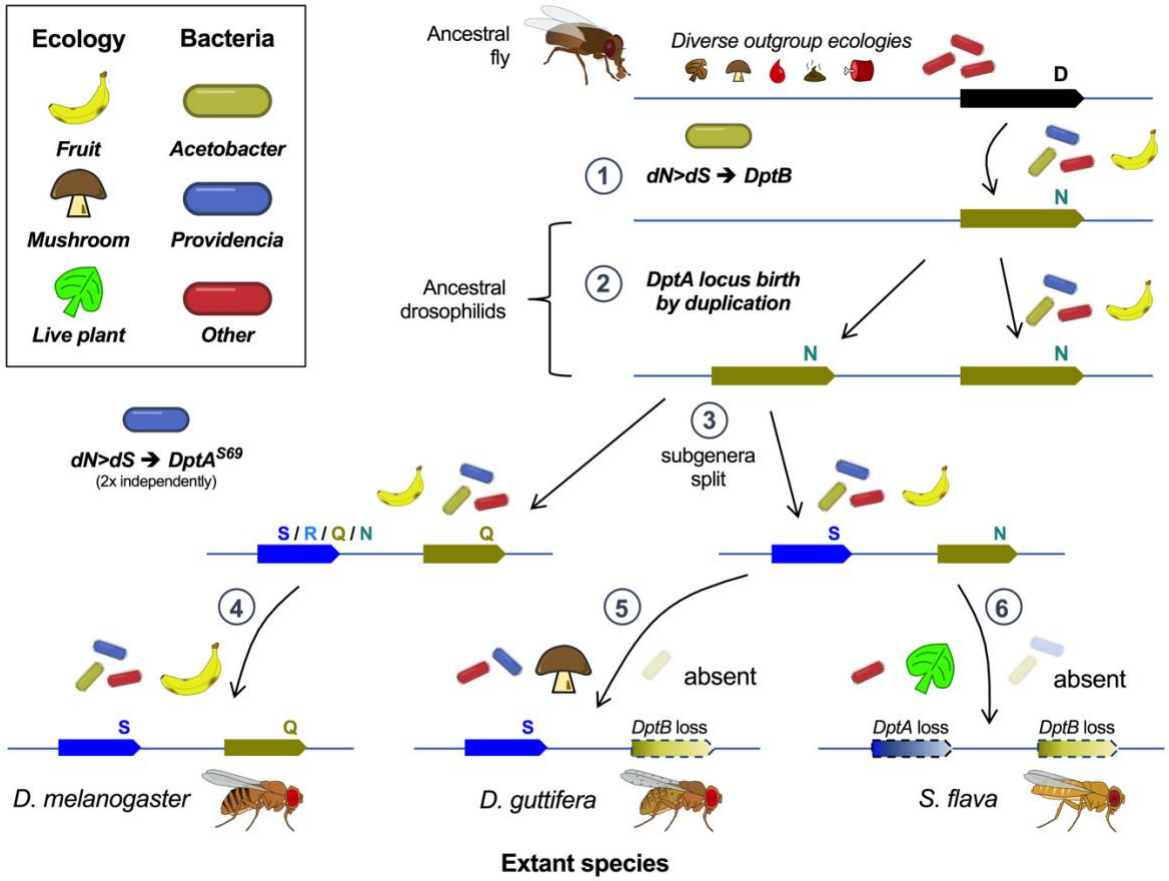


Figure 6: AMP evolution explained using Diptericins and a microbial ecology framework. *Timeline of Diptericin evolution:* ***1)*** *The drosophilid ancestor fed on fruit and was consequently exposed to Acetobacter bacteria. DptB-like sequence evolved rapidly (dN>dS) to control this novel microbe.* ***2)*** *A duplication of DptB gave rise to the DptA locus.* ***3)*** *The addition of a second Diptericin gene permitted evolutionary tinkering to control another relevant microbe: P. rettgeri, including convergent evolution of the critical S69 residue.* ***4)*** *In extant species like D. melanogaster, host ecology remains associated with both Acetobacter and Providencia. These bacteria continually select for maintenance of both genes.* ***5)*** *In mushroom-feeding species like D. guttifera, Providencia remains a threat, but mushroom ecology lacks Acetobacter. Consequently, selection is relaxed on DptB, leading to pseudogenization.* ***6)*** *In leaf-mining species like S. flava, both Acetobacter and Providencia are absent from the microbiome. As a consequence, selection is relaxed on both Diptericin genes.* ***We propose the following framework:*** *rapid evolution can reflect AMPs being tailored to control new ecologically-relevant microbes. AMP deletions are caused by further shifts in microbial ecology, where "vestigial AMPs" derived for use against specific microbes experience relaxed selection, leading to pseudogenization.*

Acknowledgements

We would like to thank Noah Whiteman, Jess Aguilar, Diler Haji, and Vincent Martinson for consultation regarding microbiome patterns informing this study's interpretation. We would also like to thank JP Boquete for transgenic fly injections. This work was funded by SNSF Postdoc.Mobility Grant number 211082 awarded to Mark Hanson, and Sinergia grant CRSII5_186397 and Novartis Foundation 532114 awarded to Bruno Lemaitre.

Materials and Methods

See supplementary text.

Bibliography

1. Daskin JH, Alford RA. Context-dependent symbioses and their potential roles in wildlife diseases. Proc Biol Sci. 2012;279: 1457–1465. doi:10.1098/rspb.2011.2276

2. Hacquard S, Spaepen S, Garrido-Oter R, Schulze-Lefert P. Interplay Between Innate Immunity and the Plant Microbiota. Annu Rev Phytopathol. 2017;55: 565–589. doi:10.1146/annurev-phyto-080516-035623

3. Hanson MA, Lemaitre B. New insights on Drosophila antimicrobial peptide function in host defense and beyond. Curr Opin Immunol. 2020;62: 22–30. doi:10.1016/j.coi.2019.11.008

4. Lazzaro BP, Zasloff M, Rolff J. Antimicrobial peptides: Application informed by evolution. Science. 2020;368. doi:10.1126/science.aau5480

5. Mookherjee N, Anderson MA, Haagsman HP, Davidson DJ. Antimicrobial host defence peptides: functions and clinical potential. Nat Rev Drug Discov. 2020. doi:10.1038/s41573-019-0058-8

6. Bosch TCG, Zasloff M. Antimicrobial Peptides—or How Our Ancestors Learned to Control the Microbiome. mBio. 2021 [cited 30 Sep 2021]. doi:10.1128/mBio.01847-21

7. Sackton TB, Lazzaro BP, Schlenke T a, Evans JD, Hultmark D, Clark AG. Dynamic evolution of the innate immune system in Drosophila. Nature genetics. 2007;39: 1461–1468. doi:10.1038/ng.2007.60

8. Lemaitre B, Hoffmann J. The Host Defense of Drosophila Melanogaster. Annual Review of Immunology. 2007;25: 697–743. doi:10.1146/annurev.immunol.25.022106.141615

9. Lin SJH, Cohen LB, Wasserman SA. Effector specificity and function in Drosophila innate immunity: Getting AMPed and dropping Boms. PLoS Pathog. 2020;16: e1008480. doi:10.1371/journal.ppat.1008480

10. Imler J-L, Bulet P. Antimicrobial peptides in Drosophila: structures, activities and gene regulation. Chemical immunology and allergy. 2005;86: 1–21. doi:10.1159/000086648

11. Hanson MA, Dostálová 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. doi:10.7554/elife.44341

12. 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. doi:10.3389/fimmu.2020.00009

13. Clemmons AW, Lindsay SA, Wasserman SA. An Effector Peptide Family Required for Drosophila Toll-Mediated Immunity. PLoS Pathogens. 2015;11. doi:10.1371/journal.ppat.1004876

14. 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. EMBO Reports. 2022 [cited 9 Nov 2022]. doi:10.15252/embr.202256036

15. Hanson MA, Cohen LB, Marra A, Iatsenko I, Wasserman SA, Lemaitre B. The Drosophila Baramicin polypeptide gene protects against fungal infection. PLoS Pathog. 2021;17: e1009846. doi:10.1371/journal.ppat.1009846

16. Huang J, Lou Y, Liu J, Bulet P, Jiao R, Hoffmann JA, et al. The BaramicinA gene is required at several steps of the host defense against Enterococcus faecalis and Metarhizium robertsii in a septic wound infection model in Drosophila melanogaster. bioRxiv; 2020 Nov. doi:10.1101/2020.11.23.394809

17. Hanson MA, Kondo S, Lemaitre B. *Drosophila* immunity: the *Drosocin* gene encodes two host defence peptides with pathogen-specific roles. Proc R Soc B. 2022;289: 20220773. doi:10.1098/rspb.2022.0773

18. Iatsenko I, Boquete JP, Lemaitre B. Microbiota-Derived Lactate Activates Production of Reactive Oxygen Species by the Intestinal NADPH Oxidase Nox and Shortens Drosophila Lifespan. Immunity. 2018;49: 929-942.e5. doi:10.1016/j.immuni.2018.09.017

19. Shaka M, Arias-Rojas A, Hrdina A, Frahm D, Iatsenko I. Lipopolysaccharide -mediated resistance to host antimicrobial peptides and hemocyte-derived reactive-oxygen species are the major Providencia alcalifaciens virulence factors in Drosophila melanogaster. Oswald E, editor. PLoS Pathog. 2022;18: e1010825. doi:10.1371/journal.ppat.1010825

20. Zanchi C, Johnston PR, Rolff J. Evolution of defence cocktails: Antimicrobial peptide combinations reduce mortality and persistent infection. Molecular Ecology. 2017;26: 5334–5343. doi:10.1111/mec.14267

21. Unckless RL, Howick VM, Lazzaro BP. Convergent Balancing Selection on an Antimicrobial Peptide in Drosophila. Current Biology. 2016;26: 257–262. doi:10.1016/j.cub.2015.11.063

22. Myers AN, Lawhon SD, Diesel AB, Bradley CW, Rodrigues Hoffmann A, Murphy WJ, et al. An ancient haplotype containing antimicrobial peptide gene variants is associated with severe fungal skin disease in Persian cats. Clark LA, editor. PLoS Genet. 2022;18: e1010062. doi:10.1371/journal.pgen.1010062

23. Unckless RL, Lazzaro BP. The potential for adaptive maintenance of diversity in insect antimicrobial peptides. Philosophical Transactions of the Royal Society B: Biological Sciences. 2016;371: 20150291. doi:10.1098/rstb.2015.0291

24. Halldórsdóttir K, Árnason E. Trans-species polymorphism at antimicrobial innate immunity cathelicidin genes of Atlantic cod and related species. PeerJ. 2015;3: e976. doi:10.7717/peerj.976

25. Hellgren O, Sheldon BC. Locus-specific protocol for nine different innate immune genes (antimicrobial peptides: β-defensins) across passerine bird species reveals within-species coding variation and a case of trans-species polymorphisms. Molecular Ecology Resources. 2011;11: 686–692. doi:10.1111/j.1755-0998.2011.02995.x

26. Linzmeier RM, Ganz T. Human defensin gene copy number polymorphisms: Comprehensive analysis of independent variation in α- and β-defensin regions at 8p22-p23. Genomics. 2005;86: 423–430. doi:10.1016/j.ygeno.2005.06.003

27. Smith KJ, Gwyer Findlay E. Expression of antimicrobial host defence peptides in the central nervous system during health and disease. Discovery Immunology. 2022; kyac003. doi:10.1093/discim/kyac003

28. Marra A, Hanson MA, Kondo S, Erkosar B, Lemaitre B. *Drosophila* antimicrobial peptides and lysozymes regulate gut microbiota composition and abundance. Microbiology; 2021 Mar. doi:10.1101/2021.03.19.436153

29. Dimarcq JL, Keppi E, Dunbar B, Lambert J, Reichhart JM, Hoffmann D, et al. Insect immunity. Purification and characterization of a family of novel inducible antibacterial proteins from immunized larvae of the dipteran Phormia terranovae and complete amino-acid sequence of the predominant member, diptericin A. Eur J Biochem. 1988;171: 17–22. doi:10.1111/j.1432-1033.1988.tb13752.x

30. Hanson MA, Lemaitre B, Unckless RL. Dynamic Evolution of Antimicrobial Peptides Underscores Trade-Offs Between Immunity and Ecological Fitness. Frontiers in Immunology. 2019;10: 2620. doi:doi: 10.3389/fimmu.2019.02620

31. Parisot N, Vargas-Chávez C, Goubert C, Baa-Puyoulet P, Balmand S, Beranger L, et al. The transposable element-rich genome of the cereal pest Sitophilus oryzae. BMC Biol. 2021;19: 241. doi:10.1186/s12915-021-01158-2

32. Hanson MA, Hamilton PT, Perlman SJ. Immune genes and divergent antimicrobial peptides in flies of the subgenus Drosophila. BMC evolutionary biology. 2016;16: 228. doi:10.1186/s12862-016-0805-y

33. Suvorov A, Kim BY, Wang J, Armstrong EE, Peede D, D'Agostino ERR, et al. Widespread introgression across a phylogeny of 155 Drosophila genomes. Current Biology. 2021; S0960982221014962. doi:10.1016/j.cub.2021.10.052

34. Bing X, Gerlach J, Loeb G, Buchon N. Nutrient-Dependent Impact of Microbes on Drosophila suzukii Development. mBio. 2018;9: e02199-17. doi:10.1128/mBio.02199-17

35. Pais IS, Valente RS, Sporniak M, Teixeira L. Drosophila melanogaster establishes a species-specific mutualistic interaction with stable gut-colonizing bacteria. Moran N, editor. PLoS Biol. 2018;16: e2005710. doi:10.1371/journal.pbio.2005710

36. Martinez-Sañudo I, Simonato M, Squartini A, Mori N, Marri L, Mazzon L. Metagenomic analysis reveals changes of the Drosophila suzukii microbiota in the newly colonized regions. Insect Sci. 2018;25: 833–846. doi:10.1111/1744-7917.12458

37. Chen J-S, Tsaur S-C, Ting C-T, Fang S. Dietary Utilization Drives the Differentiation of Gut Bacterial Communities between Specialist and Generalist Drosophilid Flies. Chaston JM, editor. Microbiol Spectr. 2022; e01418-22. doi:10.1128/spectrum.01418-22

38. Martinson VG, Douglas AE, Jaenike J. Community structure of the gut microbiota in sympatric species of wild Drosophila. Ecol Lett. 2017;20: 629–639. doi:10.1111/ele.12761

39. Bost A, Martinson VG, Franzenburg S, Adair KL, Albasi A, Wells MT, et al. Functional variation in the gut microbiome of wild *Drosophila* populations. Mol Ecol. 2018;27: 2834–2845. doi:10.1111/mec.14728

40. Goldman-Huertas B, Mitchell RF, Lapoint RT, Faucher CP, Hildebrand JG, Whiteman NK. Evolution of herbivory in Drosophilidae linked to loss of behaviors, antennal responses, odorant receptors, and ancestral diet. Proceedings of the National Academy of Sciences of the United States of America. 2015. doi:10.1073/pnas.1424656112

41. O'Connor TK, Humphrey PT, Lapoint RT, Whiteman NK, O'Grady PM. Microbial interactions and the ecology and evolution of Hawaiian Drosophilidae. Front Microbiol. 2014;5: 616. doi:10.3389/fmicb.2014.00616

42. Copley RR, Totrov M, Linnell J, Field S, Ragoussis J, Udalova IA. Functional conservation of Rel binding sites in drosophilid genomes. Genome Research. 2007;17: 1327–1335. doi:10.1101/gr.6490707

43. De Cock M, Virgilio M, Vandamme P, Bourtzis K, De Meyer M, Willems A. Comparative Microbiomics of Tephritid Frugivorous Pests (Diptera: Tephritidae) From the Field: A Tale of High Variability Across and Within Species. Front Microbiol. 2020;11: 1890. doi:10.3389/fmicb.2020.01890

44. Kounatidis I, Crotti E, Sapountzis P, Sacchi L, Rizzi A, Chouaia B, et al. Acetobacter tropicalis is a major symbiont of the olive fruit fly (Bactrocera oleae). Appl Environ Microbiol. 2009;75: 3281–3288. doi:10.1128/AEM.02933-08

45. Mazzon L, Piscedda A, Simonato M, Martinez-Sañudo I, Squartini A, Girolami V. Presence of specific symbiotic bacteria in flies of the subfamily Tephritinae (Diptera Tephritidae) and their phylogenetic relationships: proposal of "Candidatus Stammerula tephritidis." Int J Syst Evol Microbiol. 2008;58: 1277–1287. doi:10.1099/ijs.0.65287-0

46. Imrie RM, Roberts KE, Longdon B. Between virus correlations in the outcome of infection across host species: evidence of virus genotype by host species interactions. Evolutionary Biology; 2021 Feb. doi:10.1101/2021.02.16.431403

47. Moraitou M, Forsythe A, Fellows Yates JA, Brealey JC, Warinner C, Guschanski K. Ecology, not host phylogeny, shapes the oral microbiome in closely related species. Falush D, editor. Molecular Biology and Evolution. 2022; msac263. doi:10.1093/molbev/msac263

48. Gerardo NM, Altincicek B, Anselme C, Atamian H, Barribeau SM, de Vos M, et al. Immunity and other defenses in pea aphids, Acyrthosiphon pisum. Genome Biology. 2010;11. doi:10.1186/gb-2010-11-2-r21

49. Login FH, Balmand S, Vallier A, Vincent-Monégat C, Vigneron A, Weiss-Gayet M, et al. Antimicrobial peptides keep insect endosymbionts under control. Science. 2011;334: 362–365. doi:10.1126/science.1209728

50. Vorburger C, Perlman SJ. The role of defensive symbionts in host-parasite coevolution. Biol Rev Camb Philos Soc. 2018;93: 1747–1764. doi:10.1111/brv.12417

51. Salazar-Jaramillo L, Paspati A, Van De Zande L, Vermeulen CJ, Schwander T, Wertheim B. Evolution of a cellular immune response in Drosophila: A phenotypic and genomic comparative analysis. Genome Biology and Evolution. 2014;6: 273–289. doi:10.1093/gbe/evu012

52. Shultz AJ, Sackton TB. Immune genes are hotspots of shared positive selection across birds and mammals. eLife. 2019;8: e41815. doi:10.7554/eLife.41815

53. Patil A, Hughes AL, Zhang G. Rapid evolution and diversification of mammalian α-defensins as revealed by comparative analysis of rodent and primate genes. Physiological Genomics. 2004;20: 1–11. doi:10.1152/physiolgenomics.00150.2004

Figure captions

Figure 2 supp1: *Acetobacter sp. ML04.1* is a strain of *Acetobacter sicerae*. Maximum likelihood phylogeny (100 bootstraps) made using concatenated sequences of *16S, DnaK, GroEL,* and *RpoB* genes from sequenced strains and Sanger sequences collected for this study. Alignment file available as File S2. We sequenced the genome of *Acetobacter sp. ML04.1*, and designate it *Acetobacter sicerae* strain "BELCH", for Berra Erkosar (BE) who originally isolated this bacterial species from the food medium of flies reared in the city of Lasanne (L) Switzerland (CH). Sequencing methods and statistics are available in the supplementary materials and methods, and raw sequence reads are available online (GenBank accession: PRJNA873675).

Figure 2 supp2: *Acetobacter sicerae* is avirulent upon oral infection. A) *Acetobacter sicerae* is not virulent when administered orally to $DptA^{S69}$ wild-type, $DptA^{S69R}$, $DptA^{\Delta 822}$, $DptB^{KO}$, *ΔAMP14* and Rel^{E20} flies previously cleared of the microbiota by bleaching, and reared on antibiotic-medium until emergence, whereupon they were given *A. sicerae*-inoculated food. The reduced lifespan of Rel^{E20} flies is also observed in absence of infection. B-C) Silencing *DptB* ubiquitously by RNAi (genotype: *Act>DptB-IR*) but not *AttA* or *DptA* induces both the susceptibility (B) and bloating phenotype (C) after systemic infection with *A. sicerae.*

Figure 2 supp3: *DptB*-deficient flies display a marked susceptibility to several *Acetobacter* species. A-F) Survival of $DptA^{S69}$ wild-type, $DptA^{\Delta 822}$, $DptB^{KO}$, Dpt^{SK1}, *ΔAMP8*, *ΔAMP8*,Dpt^{SK1}, *ΔAMP14,* and Rel^{E20} flies at 25°C was monitored upon septic injury with various *Acetobacter* species. Flies deficient for *DptB* by either $DptB^{KO}$ or Dpt^{SK1} mutation (Dpt^{SK1} present in *ΔAMP14*) were specifically susceptible to infection by *A. sicerae* (A), *A. indonesiensis* (B), and *A. orientalis* (C). For *A. tropicalis*, both *Diptericins* and other AMPs contribute to resistance, evidenced by survival of *DptA/DptB* mutants, *ΔAMP8* compared to *ΔAMP8*, Dpt^{SK1} flies. Of note, $DptB^{KO}$ flies, but not *ΔAMP8,* shows the *Acetobacter*-induced bloating after *A. tropicalis* infection (Fig. 2supp4D). Infection with *A. aceti* (G) and *A. pomorum* (H) at 25°C did not results in marked mortality, even in Rel^{E20} flies lacking the Imd pathway. Interestingly, *A. aceti* infection causes bloating in all genotypes tested, with levels comparable even between wild-type $DptA^{S69}$ and Rel^{E20} flies. This suggests the bloating phenotype can be induced independent of *DptB-Acetobacter* specificity.

Figure 2 supp4: Prevalence of the bloating phenotype in *AMP* deficient flies upon systemic infection with various *Acetobacter* species. The prevalence of the bloating phenotype of $DptA^{S69}$ wild-type, $DptA^{\Delta 822}$, $DptB^{KO}$, Dpt^{SK1}, *ΔAMP8*, *ΔAMP8*,Dpt^{SK1}, *ΔAMP14,* and Rel^{E20} flies at 25°C was monitored upon septic injury with various *Acetobacter* species. Flies lacking *DptB* display a marked increase in the bloating phenotype upon systemic infection with *A. sicerae* (A), *A. indonesiensis* (B), *A. orientalis (C) and A. tropicalis (D).* Meanwhile *A. aceti* infection (E) induces bloating in ~25% of flies regardless of genotype, although it is relatively avirulent in terms of fly mortality (Fig. 2supp3E). In one of three replicate experiments, bloating was induced by *A. pomorum* in *DptB*-deficient flies while in the two other replicates, bloating was minimal, even in *ΔAMP14* and Rel^{E20} flies. Bloating was monitored in flies daily, however we only show bloating at the time point prior to onset of major mortality for each *Acetobacter* species, which differ in virulence. We will note here that *DptB*-deficient flies that were not bloated at the time point shown often ultimately became bloated prior to mortality at a later time point.

Figure 3supp1: Previous infection GWAS did not assess *DptB* in defence against *P. rettgeri* as few polymorphisms of consequence are present in the study population, the *Drosophila* Genetic Reference Panel (DGRP). There is only one mutation in *DptB* in the DGRP that is likely to cause loss-of-function (2R_14755267_SNP), present in only one strain (DGRP 223), which would

not be possible to detect as significant through GWAS analysis. Furthermore, despite a clear importance of *DptB* against *A. sicerae* (Fig. 2), a future GWAS study using *A. sicerae* would likely fail to find *DptB* as a gene of importance against this microbe, given the very few strains with non-synonymous changes in the DptB G domain (active domain). This result emphasizes how future studies in other animals will benefit from both forward genetic approaches (such as GWAS), and reverse genetic approaches (e.g. through systematic gene disruption), if the goal is to reveal AMP-microbe specificities. Here we note that, prior to this study, neither Unckless et al. [21] nor Hanson et al. [11] assessed *DptB* in defence against *P. rettgeri*.

Figure 3 supp2: *DptA* expression levels are lower in *DptB^{KO}* flies compared to wild-type. The expression of *DptA* is specifically reduced in *DptB^{KO}* flies, which could be due to a cis-genetic effect of the *DptB deficiency* on *DptA* inducibility. The Imd response, evidenced by induction of *Drosocin (Dro)* and *Attacin B* (*AttB)*, is fully-induced in *DptB^{KO}* flies. Significance levels: ns = not significant, * = $p < .05$, ** = $p < .01$. Error bars represent standard deviation.

Figure 3 supp3: Silencing *DptA,* but not *DptB,* by RNAi causes a severe susceptibility to *P. rettgeri*. A) Survival data from Fig. 3C including additional control genotypes (RNAi constructs without the GAL4 driver). Total experiment numbers per genotype (n^{exp}) are shown. B) RT-qPCR at 7hpi in flies infected by *P. rettgeri*, monitoring the expression of *DptA*, *DptB*, *Drososin* (*Dro*) and *Attacin B (AttB)* in flies over-expressing *DptA-IR* and *DptB-IR* confirms gene-specific knockdown. Silencing *DptA* and *DptB* has no effect on the expression of *Dro* and *AttB* indicating that *Diptericin* silencing does not impact the Imd response. Significance levels: ns = not significant, * = $p < .05$, ** = $p < .01$. Error bars represent standard deviation.

Figure 4 supp1: Alignment of the active Diptericin domain (G domain) across species. Blue background = DptA of subgenus Sophophora, Red background = DptA of subgenus Drosophila ("DptC" clade in [32]), olive background = DptB of Drosophilidae, purple background = DptB-like genes of Tephritidae. Cladogram on left is the result of a Neighbour-joining phylogenetic tree (10,000 bootstraps). Tephritid Dpt genes cluster paraphyletically within the drosophilid *DptB* clade, indicating convergent evolution towards a DptB-like gene structure more broadly, alongside parallel evolution of the Q/N polymorphism. Of note: *Scaptomyza* species that lack *Relish* binding sites in their *Dpt* gene promoters (Fig. 4supp4) also show unique mutations in highly-conserved motifs: for instance, the universal motif GGPYGN just upstream of the S69R polymorphic site is found across the phylogeny, but has a non-synonymous change to encode GGPYGD in *S. flava* and *S. graminum DptC*, possibly suggesting relaxed selection permitting changes to an otherwise essential Diptericin domain motif.

Figure 4 supp2: Tracing the evolution of the Diptericin polymorphic site across Diptera. The ancestral D-encoding codon of Diptericin was GAT, which enables a one-step mutation to AAT encoding N, which is likely the ancestral codon in both *DptB*-like genes of Tephritidae and *DptB* of Drosophilidae. The subgenus Sophohora serine is ancestrally encoded by AGC/AGT, suggesting a single mutation could have caused an AAT>AGT asparagine-serine polymorphism, ultimately settling on the AGC serine codon in the Sophophora ancestor. However, the subgenus Drosophila serine is encoded by an ancestral TCA codon, indicating it emerged by convergent evolution. Moreover, the Sophophora AGC codon is one mutational step away from either asparagine or arginine, whereas the subgenus Drosophila TCA codon requires a minimum of 2 mutations to reach any other common polymorphism residue. Consequently, little variation is seen in the subgenus Drosophila 1-exon *DptA* locus gene, whereas a high degree of variation is seen in the subgenus Sophophora 1-exon *DptA* between serine and arginine (e.g. *D. melanogaster, D. willistoni*), asparagine (e.g. *D. yakuba*), and glutamine (*D. suzukii, D. ananassae*).

Figure 4 supp3: summary of microbiome studies informing interpretation of Figure 4. Species are listed by phylogeny. Microbe presence/absence is annotated according to terms outlined at bottom. Both the species itself, and the closest-related species with a microbiome study are given as some inferences are made by using microbiome studies in close relatives, whereas *DptA/DptB* evolution is screened specifically in species with sequenced genomes.

Figure 4 supp4: Analysis of *Diptericin* loci suggests pseudogenization of *Diptericin* genes in *Scaptomyza* compared to related flies. Light blue annotations indicate 250bp of sequence upstream of the transcription start site of each *Diptericin* gene, where *Relish* binding motifs are expected. These subgenus Drosophila species have had additional *DptA* locus duplication events (*DptA* named *"DptC"* here, as these genes are part of the *DptC* clade of *Diptericins* [32]). Related to the main text, "*DptA1"* = *"DptC1"*. In the two leaf-mining *Scaptomyza* species, *DptC1* has been pseudogenized by a premature stop codon. *Scaptomyza pallida* likely shared a leaf-mining ancestor, but now feeds on decaying vegetation and mushroom [41], likely exposing it to *Providencia*, which should maintain selective pressure to keep *DptA*. Accordingly, its *DptC1* gene retains an intact coding sequence. However, in all three *Scaptomyza* species, *DptB* appears to have lost Relish binding sites ("Rel" annotations), which are needed for immune inducibility. The Relish binding motif used here (GGRDNNHHBS, generated by Copley et al. [42]) finds abundant *Relish* binding sites in the promoter of the *AttC* AMP gene across *Scaptomyza* species (Figure 4supp5), indicating Scaptomyza still uses the same general motif as other species for Rel-mediated immune regulation. Two of the three *Scaptomyza* species have also lost all *Relish* binding sites in the promoter of *DptC2*, suggesting selection to maintain this gene may also have relaxed in the *Scaptomyza* ancestor, beginning a possible pseudogenization process in some species.

Figure 4 supp5: The Relish binding motif "GGRDNNHHBS" is observed upstream of the *AttC* AMP genes of *Scaptomyza* species. Scanning the *AttC* gene promoter confirms the same Relish binding motif used by other species for AMP induction is abundant in other Rel-regulated AMP genes of *Scaptomyza*. Thus, the absence of Relish binding sites upstream of Scaptomyza *Diptericins* described in Fig. 4supp4 likely affects their expression, consistent with pseudogenization.

Figure 5 supp1: Use of 21°C reduces *A. sicerae* virulence compared to 25°C. Even at this temperature with reduced pathogen virulence, $DptB^{KO}$ susceptibility mirrors that of $\Delta AMP14$ and Rel^{E20} flies. This further emphasizes the high importance of *DptB* as the principle effector explaining Imd-mediated defence against this microbe.

Figure 5 supp2: Clean injury controls of different fly species. No fly species succumbs to injury alone in a marked way, confirming that susceptibility to infections in Fig. 5 reflect bacteria-specific immune competence. Differential survival upon infection by *P. rettgeri* or *A. sicerae* per species in Fig. 5 also suggests no specific susceptibility to injury alone in any species tested.

Figure 5 supp3: Survival curves for experiments summarized in Figure 5. *Acetobacter sicerae* experiments were monitored until day 21 to allow for complete mortality of the most susceptible fly strains/species given reduced virulence at this temperature (Fig. 5supp1).

Fig 1

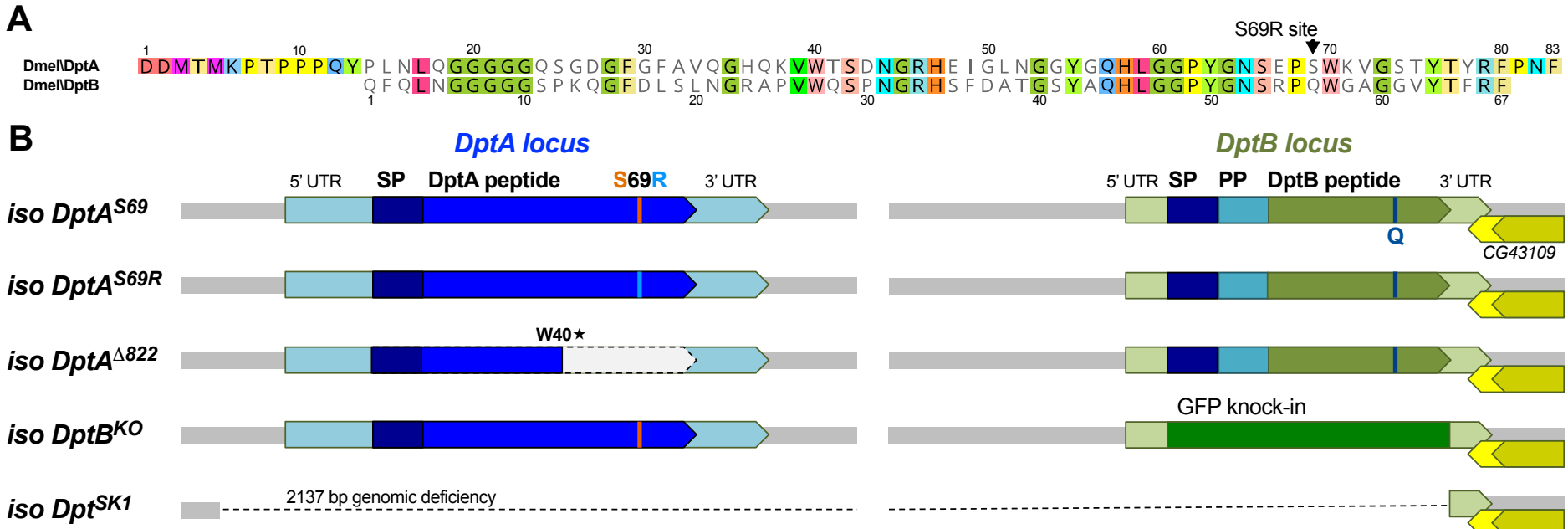
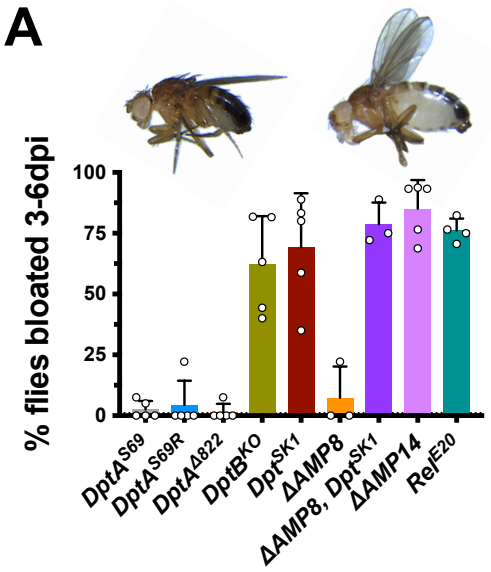
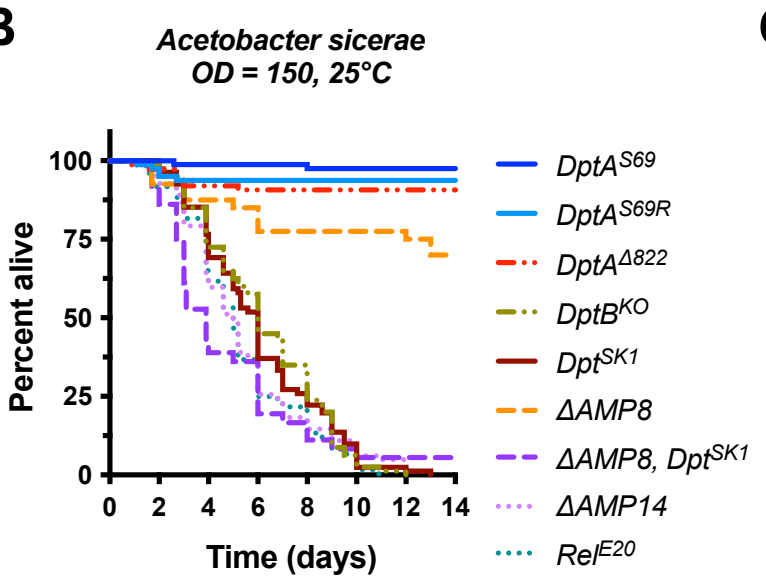

A
S69R site
Dmel\DptA
Dmel\DptB
B
DptA locus
DptB locus
5' UTR
SP
DptA peptide
S69R
3' UTR
PP
DptB peptide
Q
CG43109
iso DptA^S69
iso DptA^S69R
iso DptA^Δ822
W40★
iso DptB^KO
GFP knock-in
iso Dpt^SK1
2137 bp genomic deficiency

Fig 2


A
% flies bloated 3-6dpi
100
75
50
25
0
DptA^S69
DptA^S69R
DptA^Δ822
DptB^KO
Dpt^SK1
ΔAMP8
ΔAMP8, Dpt^SK1
ΔAMP14
Rel^E20


B
Acetobacter sicerae
OD = 150, 25°C
Percent alive
100
75
50
25
0
0 2 4 6 8 10 12 14
Time (days)
DptA^S69
DptA^S69R
DptA^Δ822
DptB^KO
Dpt^SK1
ΔAMP8
ΔAMP8, Dpt^SK1
ΔAMP14
Rel^E20

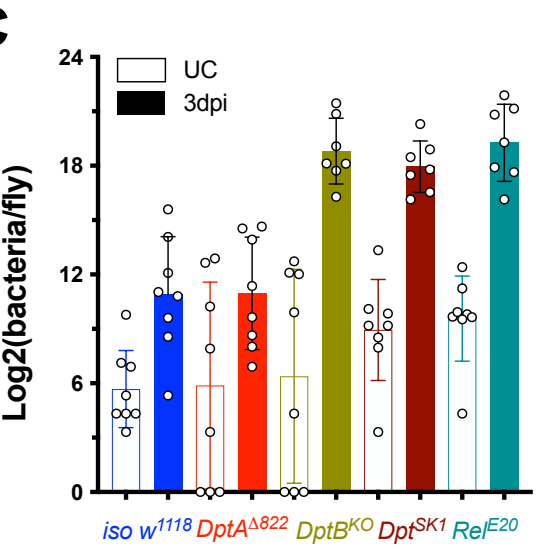

C
UC
3dpi
Log2(bacteria/fly)
24
18
12
6
0
iso w^1118
DptA^Δ822
DptB^KO
Dpt^SK1
Rel^E20

Fig 3

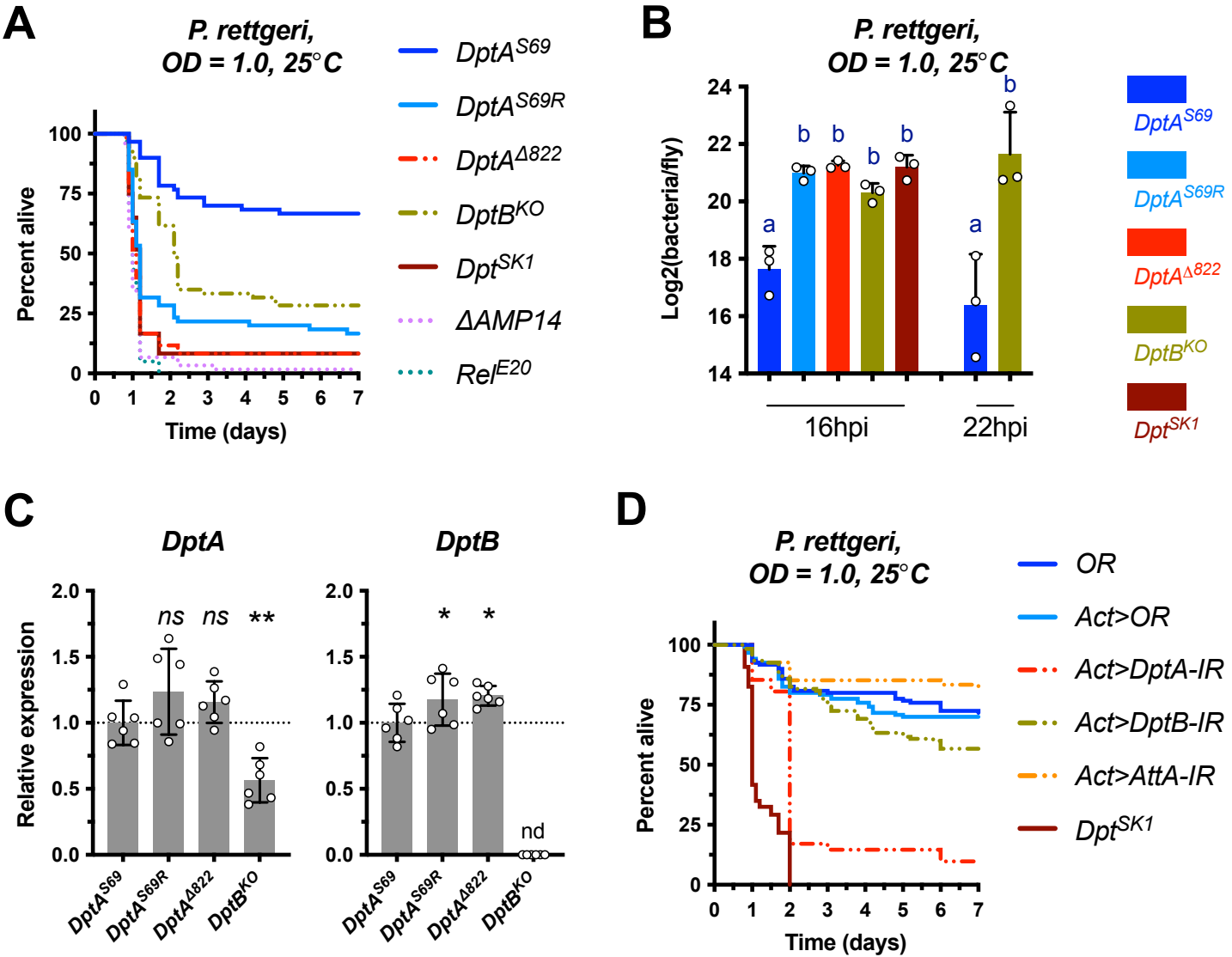

A
P. rettgeri,
OD = 1.0, 25°C
Percent alive
Time (days)
DptA^S69
DptA^S69R
DptA^Δ822
DptB^KO
Dpt^SK1
ΔAMP14
Rel^E20
B
P. rettgeri,
OD = 1.0, 25°C
Log2(bacteria/fly)
16hpi
22hpi
DptA^S69
DptA^S69R
DptA^Δ822
DptB^KO
Dpt^SK1
C
DptA
DptB
Relative expression
ns
ns
**
*
*
nd
DptA^S69
DptA^S69R
DptA^Δ822
DptB^KO
D
P. rettgeri,
OD = 1.0, 25°C
Percent alive
Time (days)
OR
Act>OR
Act>DptA-IR
Act>DptB-IR
Act>AttA-IR
Dpt^SK1

Fig 4

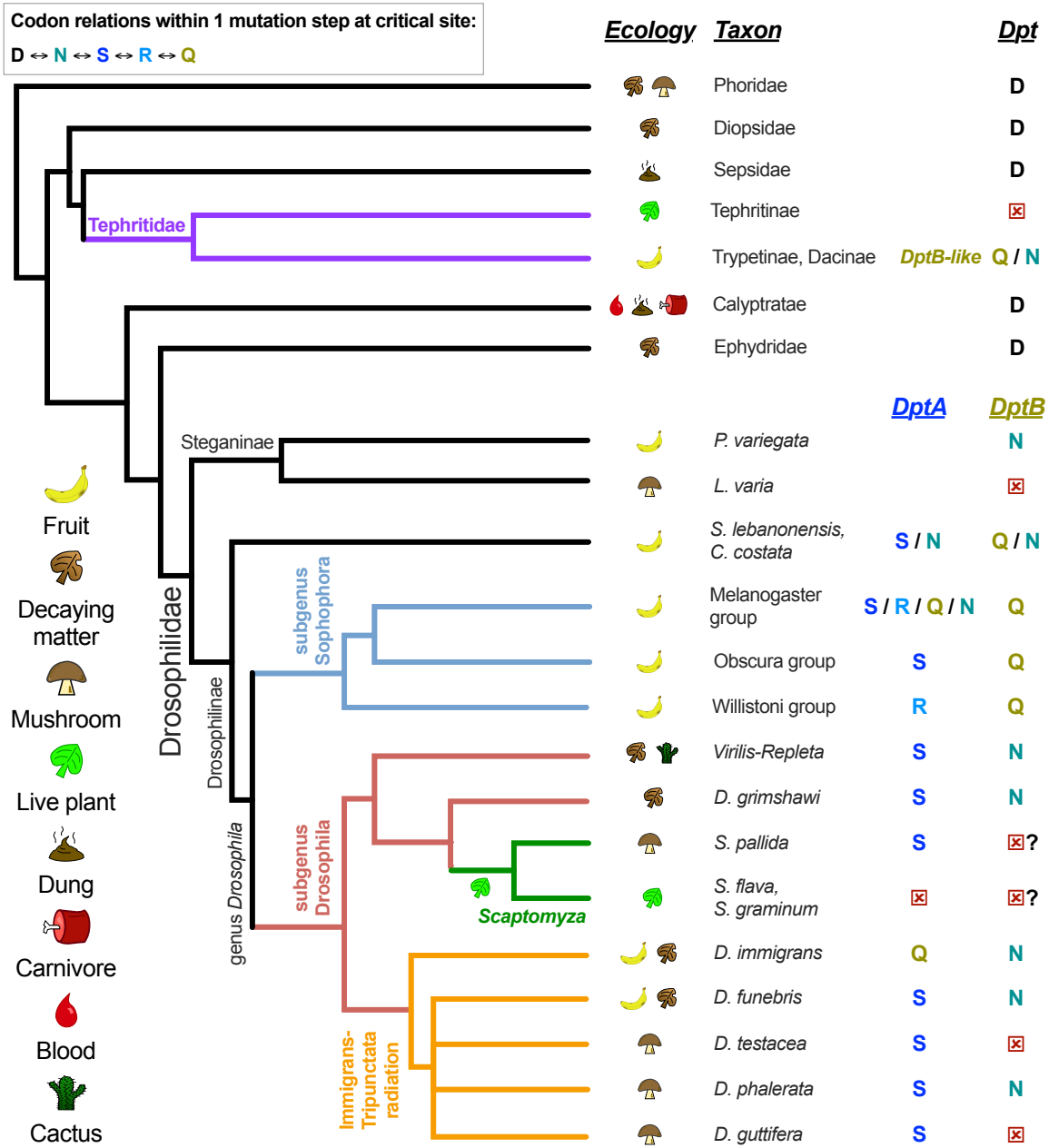
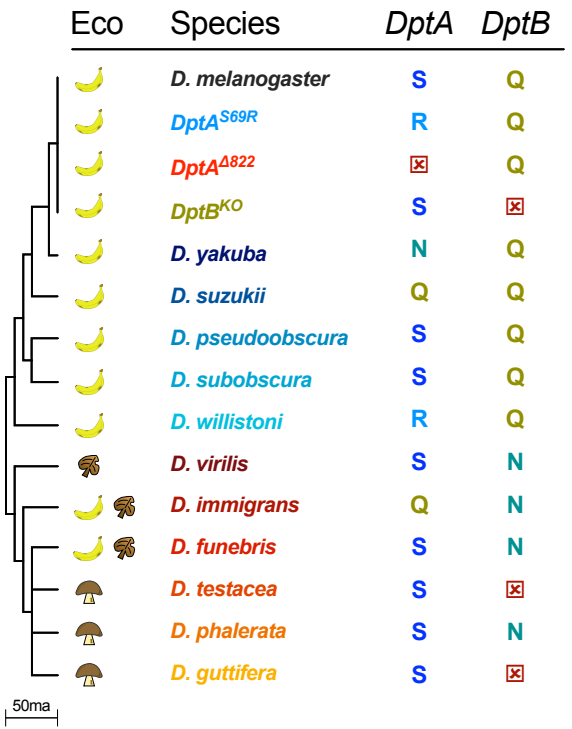
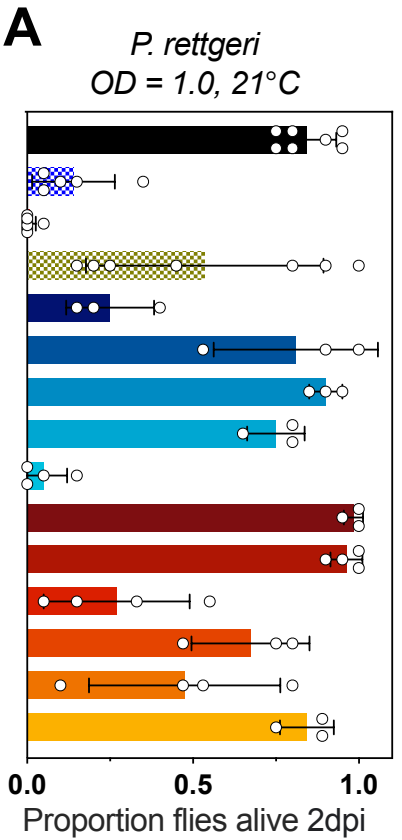

Codon relations within 1 mutation step at critical site:
D ↔ N ↔ S ↔ R ↔ Q
Ecology
Taxon
Dpt
Phoridae
D
Diopsidae
D
Sepsidae
D
Tephritidae
Tephritinae
Trypetinae, Dacinae
DptB-like
Q / N
Calyptratae
D
Ephydridae
D
DptA
DptB
Steganinae
P. variegata
N
L. varia
S. lebanonensis,
C. costata
S / N
Q / N
Melanogaster
group
S / R / Q / N
Q
subgenus
Sophophora
Obscura group
S
Q
Willistoni group
R
Q
Drosophilidae
Drosophilinae
Virilis-Repleta
S
N
D. grimshawi
S
N
S. pallida
S
?
genus Drosophila
subgenus
Drosophila
S. flava,
S. graminum
?
Scaptomyza
D. immigrans
Q
N
D. funebris
S
N
D. testacea
S
Immigrans-
Tripunctata
radiation
D. phalerata
S
N
D. guttifera
S
Fruit
Decaying
matter
Mushroom
Live plant
Dung
Carnivore
Blood
Cactus

Fig 5

Legend


Eco
Species
DptA
DptB
D. melanogaster S Q
DptA S69R R Q
DptA Δ822 Q
DptB KO S
D. yakuba N Q
D. suzukii Q Q
D. pseudoobscura S Q
D. subobscura S Q
D. willistoni R Q
D. virilis S N
D. immigrans Q N
D. funebris S N
D. testacea S
D. phalerata S N
D. guttifera S
50ma


A
P. rettgeri
OD = 1.0, 21°C
0.0
0.5
1.0
Proportion flies alive 2dpi

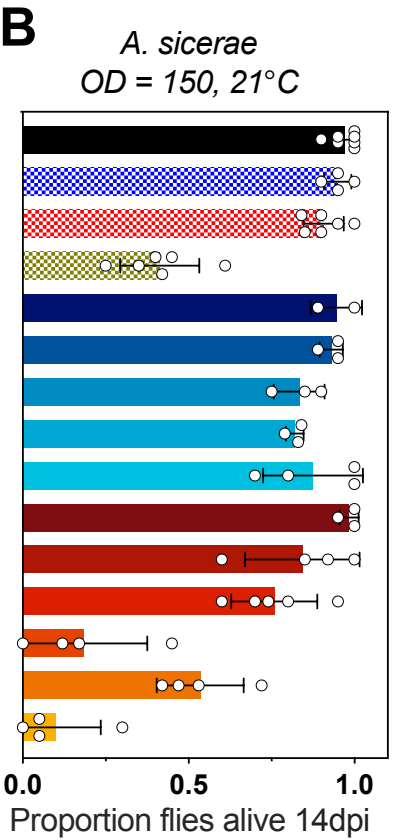

B
A. sicerae
OD = 150, 21°C
0.0
0.5
1.0
Proportion flies alive 14dpi

Fig 6

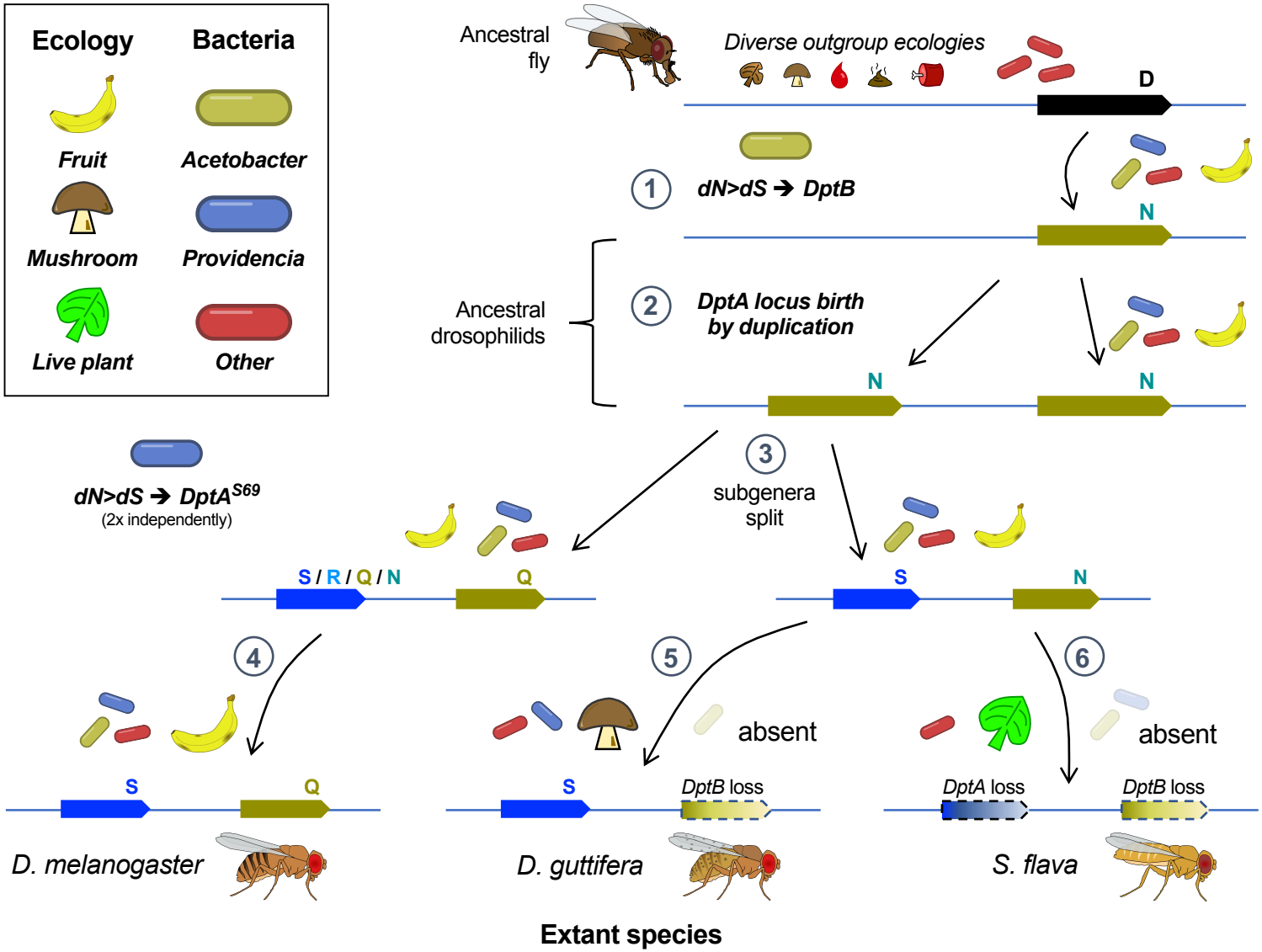

Ecology
Bacteria
Fruit
Acetobacter
Mushroom
Providencia
Live plant
Other
Ancestral fly
Diverse outgroup ecologies
D
1
dN>dS → DptB
N
Ancestral drosophilids
2
DptA locus birth by duplication
N
N
3
subgenera split
dN>dS → DptA^S69
(2x independently)
S / R / Q / N
Q
S
N
4
5
6
absent
absent
S
Q
S
DptB loss
DptA loss
DptB loss
D. melanogaster
D. guttifera
S. flava
Extant species

Fig 2supp1

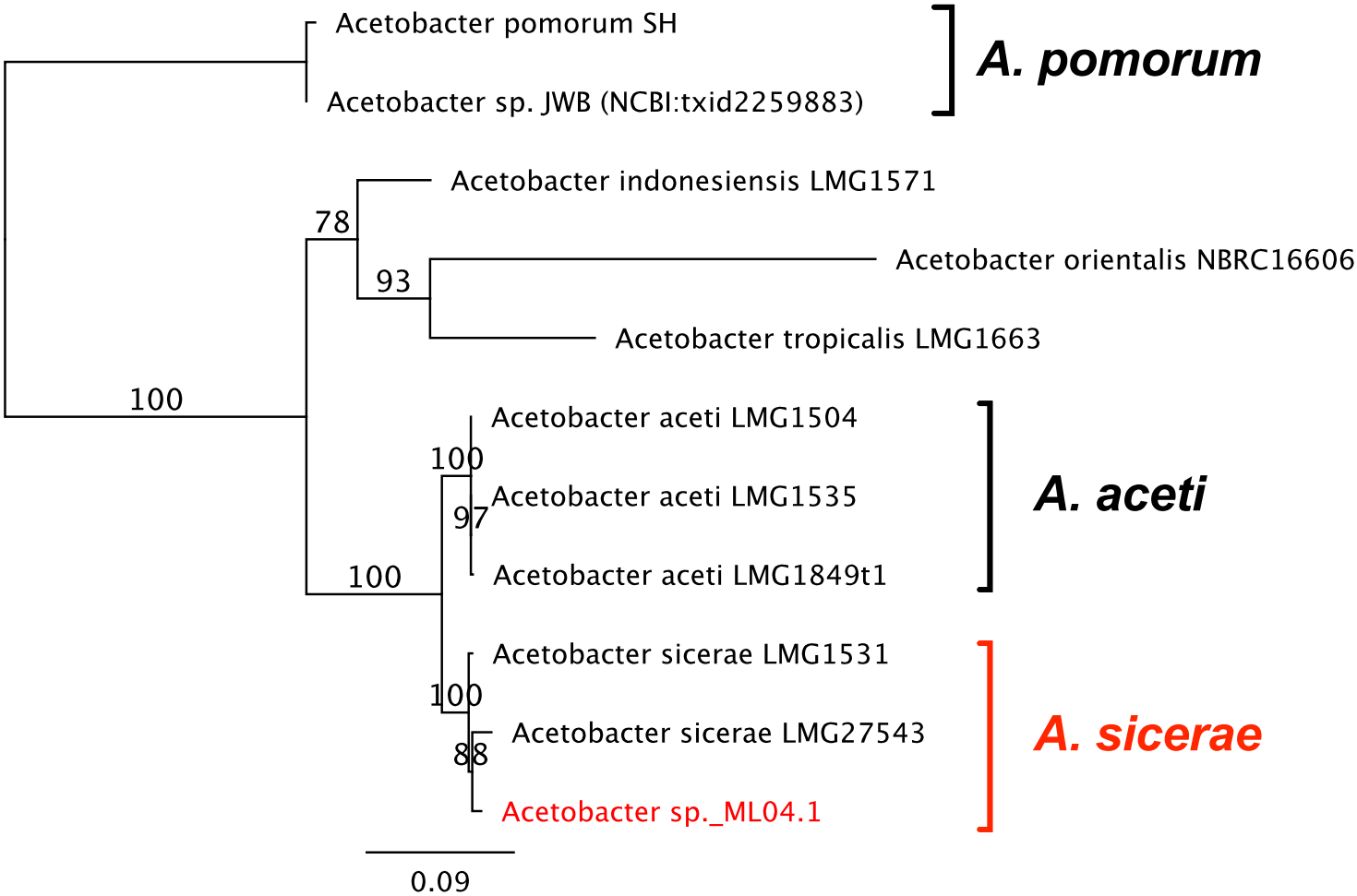
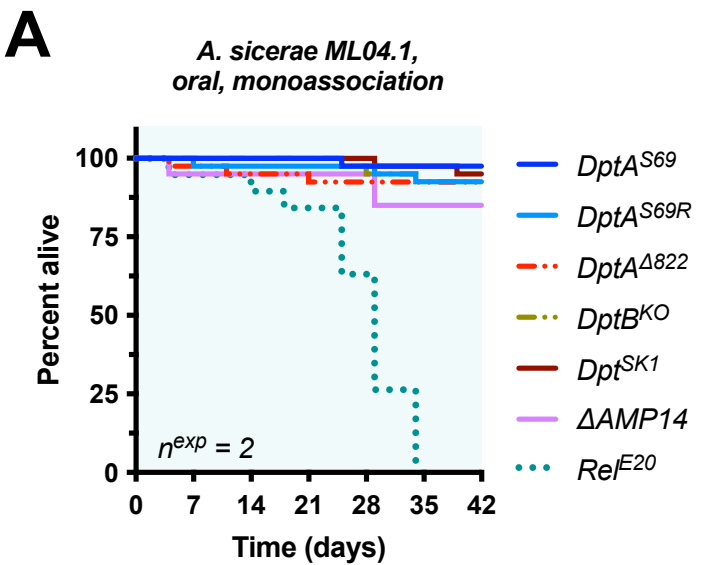
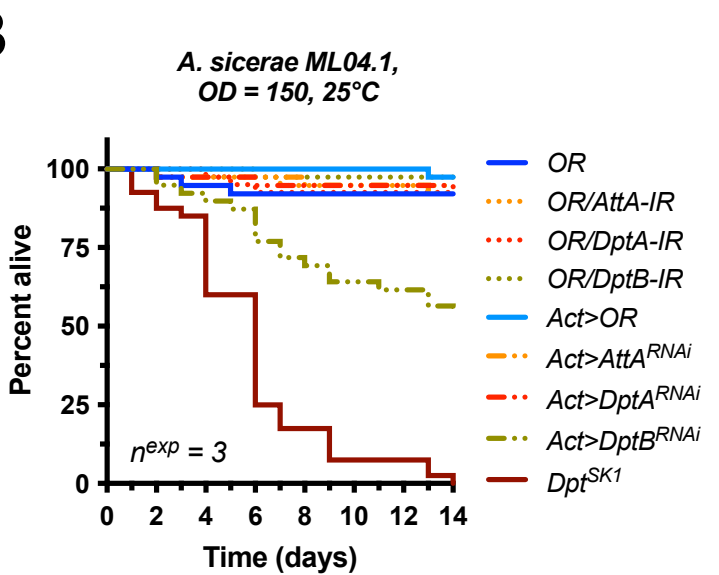

Acetobacter pomorum SH
Acetobacter sp. JWB (NCBI:txid2259883)
A. pomorum
Acetobacter indonesiensis LMG1571
78
93
Acetobacter orientalis NBRC16606
Acetobacter tropicalis LMG1663
100
Acetobacter aceti LMG1504
100
Acetobacter aceti LMG1535
97
Acetobacter aceti LMG1849t1
A. aceti
100
Acetobacter sicerae LMG1531
100
Acetobacter sicerae LMG27543
88
Acetobacter sp._ML04.1
A. sicerae
0.09

Fig 2supp2


A
A. sicerae ML04.1,
oral, monoassociation
Percent alive
100
75
50
25
0
0
7
14
21
28
35
42
Time (days)
n^exp = 2
DptA^S69
DptA^S69R
DptA^Δ822
DptB^KO
Dpt^SK1
ΔAMP14
Rel^E20


B
A. sicerae ML04.1,
OD = 150, 25°C
Percent alive
100
75
50
25
0
0
2
4
6
8
10
12
14
Time (days)
n^exp = 3
OR
OR/AttA-IR
OR/DptA-IR
OR/DptB-IR
Act>OR
Act>AttA^RNAi
Act>DptA^RNAi
Act>DptB^RNAi
Dpt^SK1

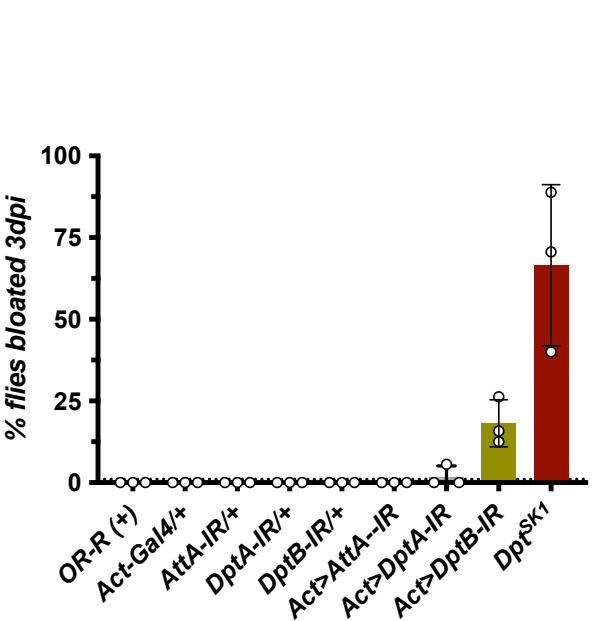

C
% flies bloated 3dpi
100
75
50
25
0
OR-R (+)
Act-Gal4/+
AttA-IR/+
DptA-IR/+
DptB-IR/+
Act>AttA--IR
Act>DptA-IR
Act>DptB-IR
Dpt^SK1

Fig 2supp3

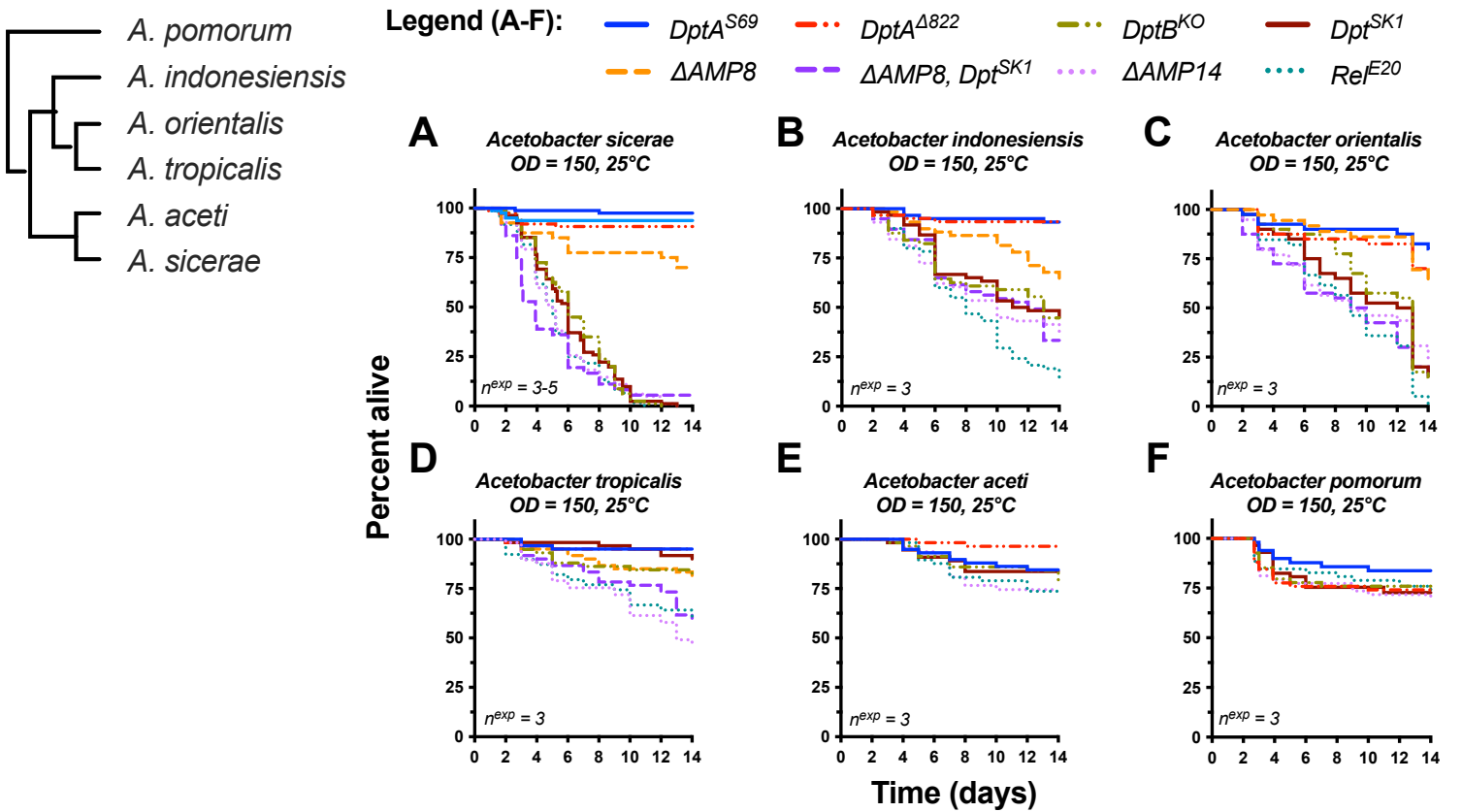

A. pomorum
A. indonesiensis
A. orientalis
A. tropicalis
A. aceti
A. sicerae
Legend (A-F):
DptA^S69
DptA^Δ822
DptB^KO
Dpt^SK1
ΔAMP8
ΔAMP8, Dpt^SK1
ΔAMP14
Rel^E20
A
Acetobacter sicerae
OD = 150, 25°C
n^exp = 3-5
B
Acetobacter indonesiensis
OD = 150, 25°C
n^exp = 3
C
Acetobacter orientalis
OD = 150, 25°C
n^exp = 3
D
Acetobacter tropicalis
OD = 150, 25°C
n^exp = 3
E
Acetobacter aceti
OD = 150, 25°C
n^exp = 3
F
Acetobacter pomorum
OD = 150, 25°C
n^exp = 3
Percent alive
Time (days)

Fig 2supp4

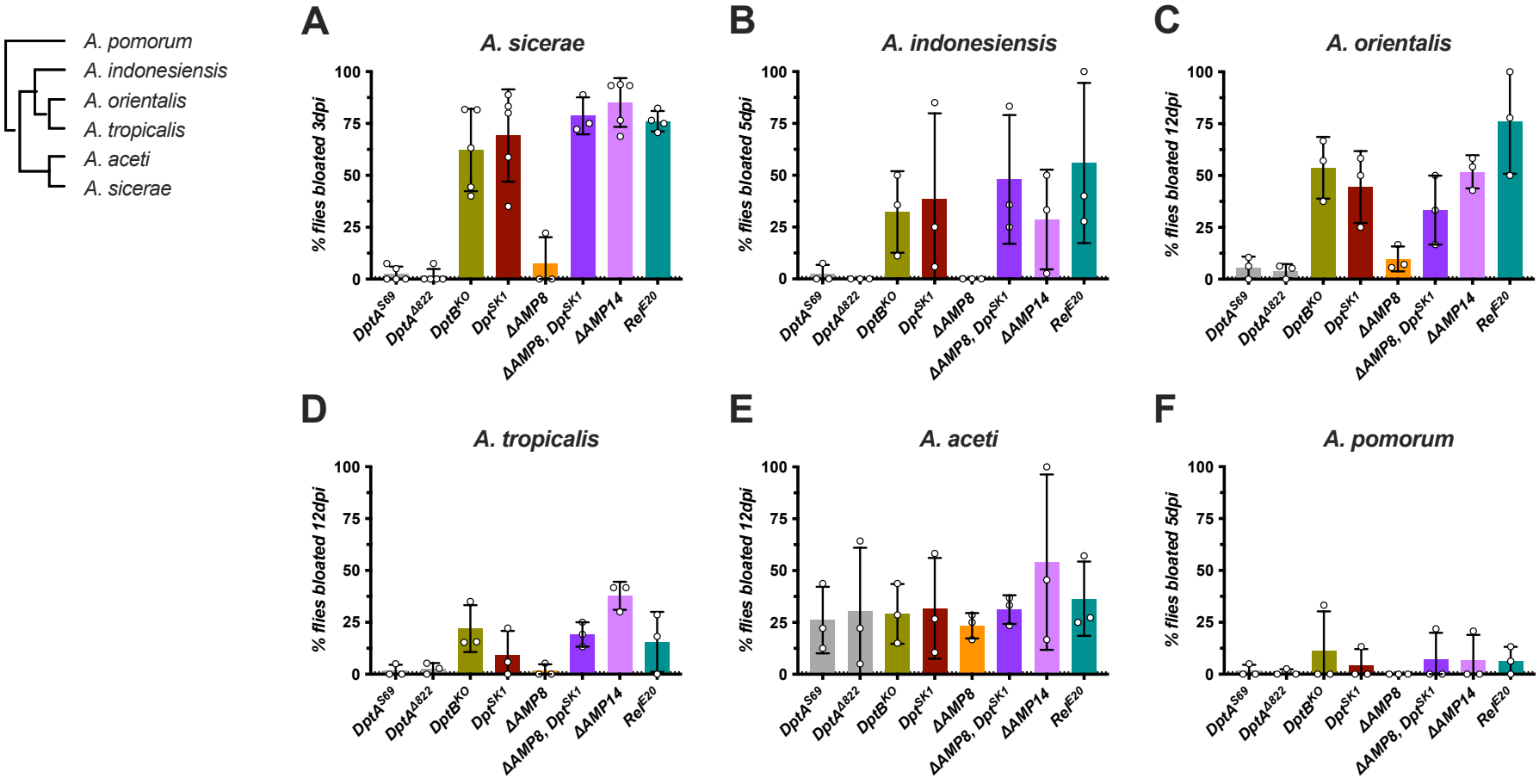

A. pomorum
A. indonesiensis
A. orientalis
A. tropicalis
A. aceti
A. sicerae
A
A. sicerae
% flies bloated 3dpi
B
A. indonesiensis
% flies bloated 5dpi
C
A. orientalis
% flies bloated 12dpi
D
A. tropicalis
% flies bloated 12dpi
E
A. aceti
% flies bloated 12dpi
F
A. pomorum
% flies bloated 5dpi
0
25
50
75
100
DptA^S69
DptA^Δ22
DptB^KO
Dpt^SK1
ΔAMP8
ΔAMP8, Dpt^SK1
ΔAMP14
Rel^E20

Fig3supp1

Allele	Region	Status	Allele	Allele frequency	Codon	Residue
2R_14754932_SNP	Signal peptide	Dmel_R6 reference allele	T	96.0%	CAT	H
		Alternate allele	A	3.4%	CAA	Q
		Other/unknown	-	0.5%	CA?	H/Q
2R_14754939_SNP	Signal peptide	Dmel_R6 reference allele	G	3.4%	GCT	A
		Alternate allele	T	81.0%	TCT	S
		Other/unknown	-	15.6%	?CT	A/S/P/T
2R_14754945_SNP	Signal peptide	Dmel_R6 reference allele	C	98.5%	CTT	L
		Alternate allele	T	0.5%	TTT	F
		Other/unknown	-	1.0%	?TT	L/F/I/V
2R_14755006_SNP	Propeptide	Dmel_R6 reference allele	C	27.3%	CCC	P
		Alternate allele	T	67.3%	CTC	L
		Other/unknown	-	5.4%	C?C	P/L/H/R
2R_14755147_SNP	2nd residue of putative G domain	Dmel_R6 reference allele	C	25.9%	TTC	F
		Alternate allele	G	68.3%	TTG	L
		Other/unknown	-	5.9%	TT?	F/L
2R_14755153_SNP	4th residue of putative G domain	Dmel_R6 reference allele	G	99.0%	TTG	L
		Alternate allele	T	0.5%	TTT	F
		Other/unknown	-	0.5%	TT?	L/F
2R_14755267_SNP	**G domain**	**Dmel_R6 reference allele**	**T**	**99.5%**	**TAT**	**Y**
Found exclusively in DGRP 223		**Alternate allele**	**A**	**0.5%**	**TAA**	**STOP**
		Other/unknown	**-**	**0.0%**	**-**	
2R_14755287_SNP	G domain	Dmel_R6 reference allele	C	97.0%	CCC	P
		Alternate allele	G	2.0%	CGC	R
		Other/unknown	-	1.0%	C?C	P/R/L/H
2R_14755325_SNP	G domain	Dmel_R6 reference allele	G	99.0%	GTG	V
		Alternate allele	A	1.0%	ATG	M
		Other/unknown	-	0.0%		V/M/L

Fig3supp2

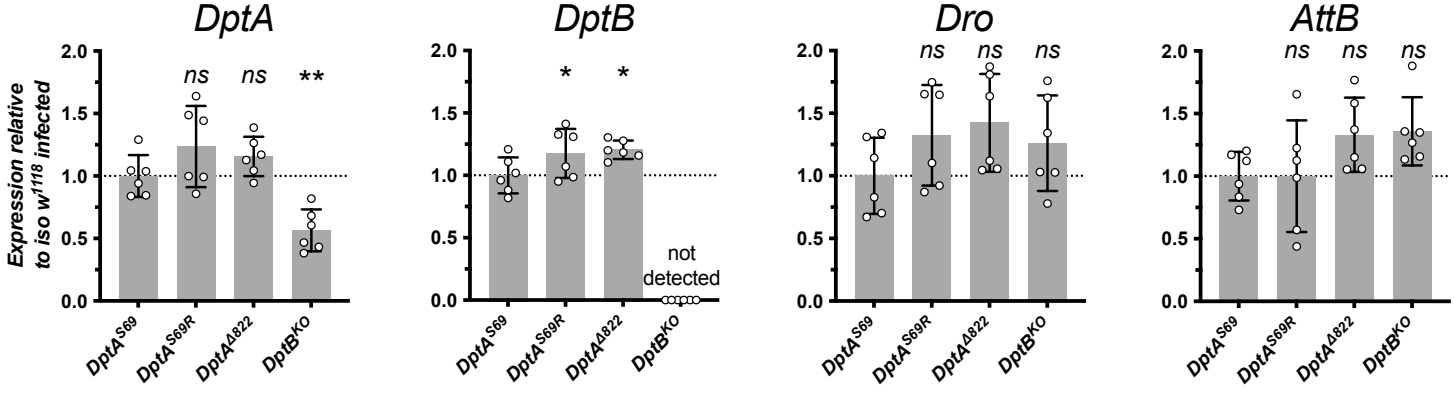

DptA
DptB
Dro
AttB
Expression relative to iso w1118 infected
2.0
1.5
1.0
0.5
0.0
ns
ns
**
*
*
ns
ns
ns
ns
ns
ns
not detected
DptA S69
DptA S69R
DptA Δ822
DptB KO

Fig3supp3

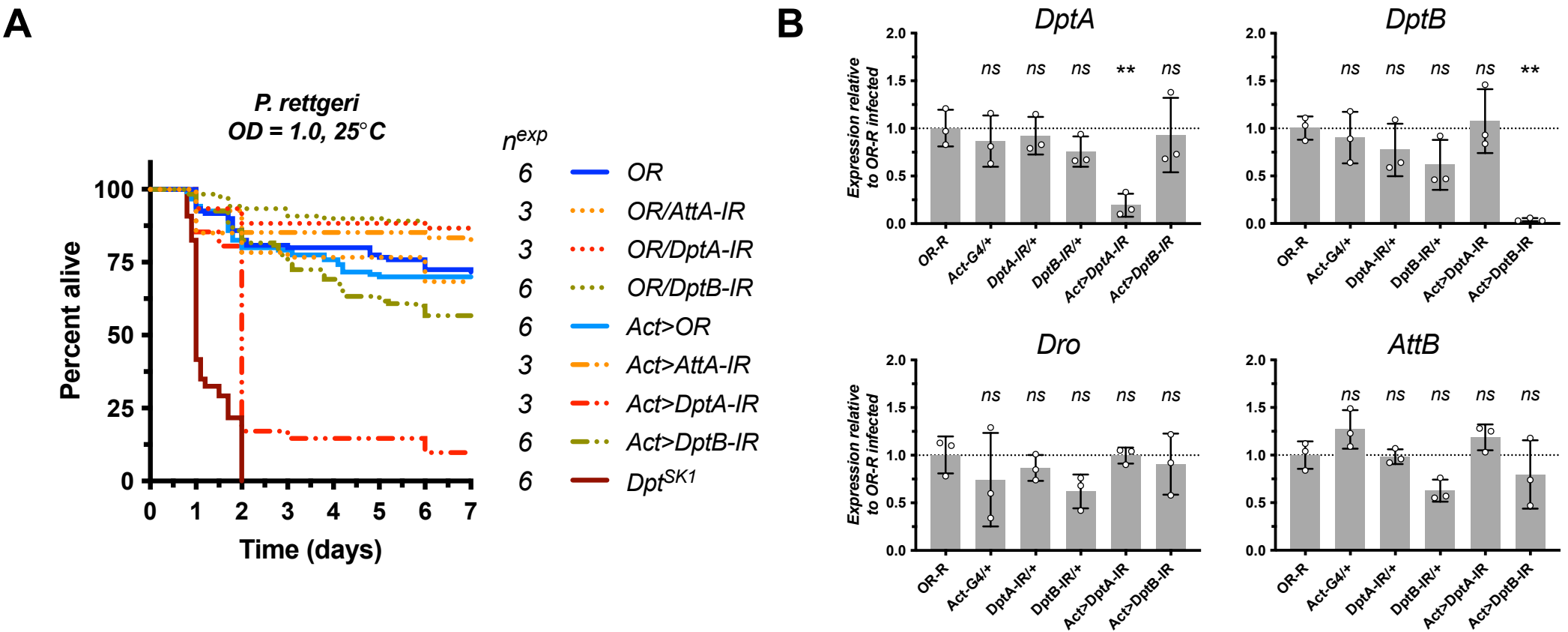

A
P. rettgeri
OD = 1.0, 25°C
Percent alive
Time (days)
n^exp
6 OR
3 OR/AttA-IR
3 OR/DptA-IR
6 OR/DptB-IR
6 Act>OR
3 Act>AttA-IR
3 Act>DptA-IR
6 Act>DptB-IR
6 Dpt^SK1
B
DptA
DptB
Dro
AttB
Expression relative to OR-R infected
ns
**
OR-R
Act-G4/+
DptA-IR/+
DptB-IR/+
Act>DptA-IR
Act>DptB-IR

Fig4supp1

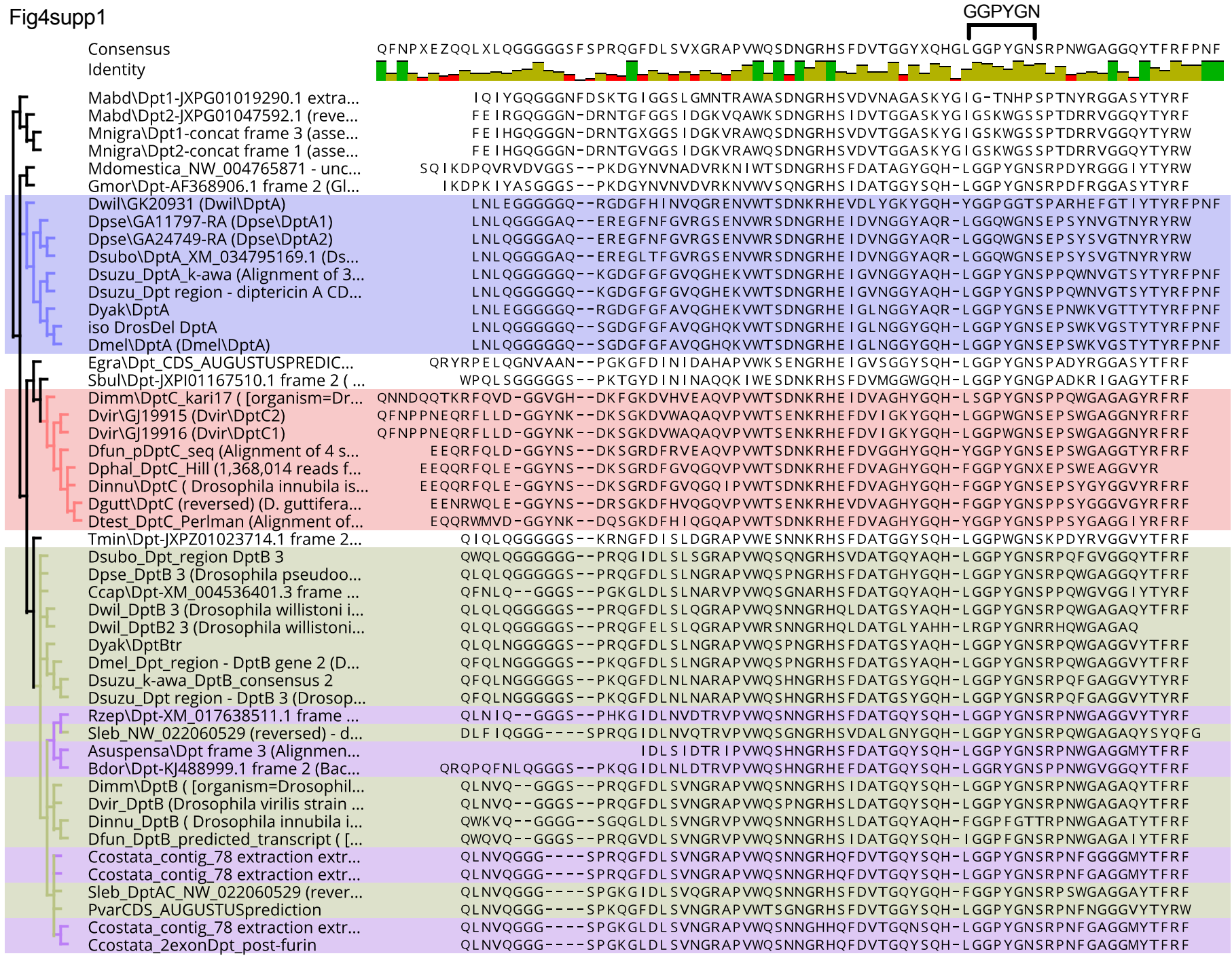

GGPYGN
Consensus
Identity
Mabd\Dpt1-JXPG01019290.1 extra...
Mabd\Dpt2-JXPG01047592.1 (reve...
Mnigra\Dpt1-concat frame 3 (asse...
Mnigra\Dpt2-concat frame 1 (asse...
Mdomestica_NW_004765871 - unc...
Gmor\Dpt-AF368906.1 frame 2 (Gl...
Dwil\GK20931 (Dwil\DptA)
Dpse\GA11797-RA (Dpse\DptA1)
Dpse\GA24749-RA (Dpse\DptA2)
Dsubo\DptA_XM_034795169.1 (Ds...
Dsuzu_DptA_k-awa (Alignment of 3...
Dsuzu_Dpt region - diptericin A CD...
Dyak\DptA
iso DrosDel DptA
Dmel\DptA (Dmel\DptA)
Egra\Dpt_CDS_AUGUSTUSPREDIC...
Sbul\Dpt-JXPI01167510.1 frame 2 (...
Dimm\DptC_kari17 ([organism=Dr...
Dvir\GJ19915 (Dvir\DptC2)
Dvir\GJ19916 (Dvir\DptC1)
Dfun_pDptC_seq (Alignment of 4 s...
Dphal_DptC_Hill (1,368,014 reads f...
Dinnu\DptC (Drosophila innubila is...
Dgutt\DptC (reversed) (D. guttifera...
Dtest_DptC_Perlman (Alignment of...
Tmin\Dpt-JXPZ01023714.1 frame 2...
Dsubo_Dpt_region DptB 3
Dpse_DptB 3 (Drosophila pseudoo...
Ccap\Dpt-XM_004536401.3 frame ...
Dwil_DptB 3 (Drosophila willistoni i...
Dwil_DptB2 3 (Drosophila willistoni...
Dyak\DptBtr
Dmel_Dpt_region - DptB gene 2 (D...
Dsuzu_k-awa_DptB_consensus 2
Dsuzu_Dpt region - DptB 3 (Drosop...
Rzep\Dpt-XM_017638511.1 frame ...
Sleb_NW_022060529 (reversed) - d...
Asuspensa\Dpt frame 3 (Alignmen...
Bdor\Dpt-KJ488999.1 frame 2 (Bac...
Dimm\DptB ([organism=Drosophil...
Dvir_DptB (Drosophila virilis strain ...
Dinnu_DptB (Drosophila innubila i...
Dfun_DptB_predicted_transcript ([...
Ccostata_contig_78 extraction extr...
Ccostata_contig_78 extraction extr...
Sleb_DptAC_NW_022060529 (rever...
PvarCDS_AUGUSTUSprediction
Ccostata_contig_78 extraction extr...
Ccostata_2exonDpt_post-furin

Fig4supp2

Codon table

TTT	Phe (F)	TCT	Ser (S)	TAT	Tyr (Y)	TGT	Cys (C)
TTC		TCC		TAC		TGC	
TTA	Leu (L)	TCA		TAA	STOP	TGA	STOP
TTG		TCG		TAG		TGG	Trp (W)
CTT	Leu (L)	CCT	Pro (P)	CAT	His (H)	CGT	Arg (R)
CTC		CCC		CAC		CGC	
CTA		CCA		CAA	Gln (Q)	CGA	
CTG		CCG		CAG		CGG	
ATT	Ile (I)	ACT	Thr (T)	AAT	Asn (N)	AGT	Ser (S)
ATC		ACC		AAC		AGC	
ATA		ACA		AAA	Lys (K)	AGA	Arg (R)
ATG	Met (M)	ACG		AAG		AGG	
GTT	Val (V)	GCT	Ala (A)	GAT	Asp (D)	GGT	Gly (G)
GTC		GCC		GAC		GGC	
GTA		GCA		GAA	Glu (E)	GGA	
GTG		GCG		GAG		GGG	

Outgroup Diptera

Family	Subfamily	Genus	species	1-exon Dpt	Codon	Dpt(s)	Codon
Phoridae	Metopininae	*Megaselia*	*abdita*	-	-	D,N	GAT, AAT
		Megaselia	*nigra*	-	-	D,D	GAT
Sepsidae	Sepsinae	*Themira*	*minor*	-	-	D,D,E	GAT,GAC,GAA
Diopsidae	Diopsinae	*Sphyracephala*	*brevicornis*	-	-	D	GAT
Tephritidae	Trypetinae	*Anastrepha*	*suspensa*	-	-	N	AAT
		Rhagoletis	*zephyria*	-	-	N	AAT
	Dacinae	*Ceratitis*	*capitata*	-	-	Q	CAA
		Bactrocera	*oleae*	-	-	N	AAT
Calliphoridae	Lucilinae	*Lucilia*	*cuprina*	-	-	D	GAT
	Chrysominae	*Protophormia*	*terranovae*	-	-	D	GAT
Sarcophagidae	Sarcophaginae	*Sarcophaga*	*bullata*	-	-	D,D,D,D,D,D,D	GAT
Glossinidae	-	*Glossina*	*morsitans*	-	-	D	GAT
Muscidae	Muscinae	*Musca*	*domestica*	-	-	D,D	GAT
Ephydridae	Ephydrinae	*Ephydra*	*gracilis*	-	-	D	GAC
Drosophilidae	Steganinae	*Phortica*	*variegata*	-	-	N	AAT
		Leucophenga	*varia*	-	-	del	n/a
	Drosophilinae	*Chymomyza*	*costata*	N, N, N	AAT	N	AAT
		Scaptodrosophila	*lebanonensis*	S	AGT	Q	CAA
		Drosophila	*busckii*	S,S,S,S,S,S,S	TCT,TCC	N,N,N	AAT
		Zaprionus	*indianus*	S	TCT	N	AAC
		Zaprionus	*capensis*	S	TCT	N	AAT

Genus *Drosophila*

Subgenus	Species group	Genus	species	1-exon Dpt (DptA)	Codon	2-exon Dpt (DptB)	Codon
Sophophora	Melanogaster	*Drosophila*	*melanogaster*	S	AGC	Q	CAG
			melanogaster	R	AGA	Q	CAG
		Drosophila	*yakuba*	N	AAC	Q	CAG
		Drosophila	*suzukii (k-awa036)*	Q	CAA	Q	CAG
		Drosophila	*ananassae*	Q	CAA	Q	CAA
	Obscura	*Drosophila*	*pseudoobscura*	S, S	AGC	Q	CAG
		Drosophila	*subobscura*	S	AGC	Q	CAG
	Willistoni	*Drosophila*	*willistoni*	R	AGA	Q, Q	CAA
Drosophila	Virilis-Repleta	*Drosophila*	*virilis*	S, S	TCA	N	AAT
		Drosophila	*mojavensis*	S,S,S	TCA,TCC	N	AAC
	Hawaiian Drosophila	*Drosophila*	*grimshawi*	S, S	TCA	N	AAT
	Immigrans	*Drosophila*	*immigrans (Longdon)*	Q	CAA	N	AAC
		Drosophila	*albomicans*	S	TCT	N	AAT
	Funebris	*Drosophila*	*funebris*	S	TCA	N	AAC
	Testacea	*Drosophila*	*testacea*	S	TCA	del	n/a
	Quinaria	*Drosophila*	*phalerata*	S	TCA	N	AAT
		Drosophila	*guttifera*	S	TCA	del	n/a

Fig4supp3

Species/Family	Species group	Closest species with similar ecology and microbiome study from wild flies	Species group of closest species	Ecology	Providencia	Acetobacter	DptA	DptB	Reference(s)
Tephritidae	Fruit feeders	*Many species sampled*	*Tephritidae*	fruit	yes	yes	n/a	Q/N	Kounatidis et al (2009), De Cock et al. (2020)
Tephritidae	Tephritinae	*Many species sampled*	*Tephritinae*	plant	no mention	no	n/a	deleted	Mazzon et al. (2008)
D. melanogaster	Melanogaster	*D. melanogaster*	*Melanogaster*	Fruit, compost	stochastic	yes	S/R	Q	Chen et al. (2022)
D. yakuba	Melanogaster	*D. melanogaster*	*Melanogaster*	Fruit, compost	stochastic	yes	N	Q	Chen et al. (2022)
D. suzukii	Melanogaster	*D. suzukii*	*Melanogaster*	Fruit	stochastic	yes	Q	Q	Chen et al. (2022)
D. pseudoobscura	Obscura	*D. subobscura*	*Obscura*	Fruit	no mention	yes	S, S	Q	Jaramillo and Castañeda (2021)
D. subobscura	Obscura	*D. subobscura*	*Obscura*	Fruit	no mention	yes	S	Q	Jaramillo and Castañeda (2021)
D. willistoni*	Willistoni	*D. willistoni**	*Willistoni*	Fruit	no mention	yes	R	Q	Wong et al. (2013), Adair et al. (2020)
D. virilis*	Virilis	*D. virilis**	*Virilis*	Flux	yes	yes	S, S	N	Wong et al. (2013)
D. mojavensis	Repleta	*D. nigrospiracula*	*Repleta*	Cactus	yes	stochastic	S, S, S	N	Martinson et al. (2018), Chen et al. (2022)
D. grimshawi	Hawaii	nd	nd	Decaying tree bark	nd	nd	S, S	N	Markow and O'Grady (2008), no data on microbiome
S. pallida	Scaptomyza	*D. neotestacea, D. falleni*	*Testacea*	Mushroom, Decaying vegetation	yes	no	S, S-pseudogene?	N-pseudogene?	Martinson et al. (2017), Bost et al. (2018)
S. graminum	Scaptomyza	*S. flava*	*Scaptomyza*	Leaf parasite	no	no	deleted, S-pseudogene?	N-pseudogene?	O' Connor et al. (2014) and personal communication Whiteman lab
S. flava	Scaptomyza	*S. flava*	*Scaptomyza*	Leaf parasite	no	no	deleted, S-pseudogene?	N-pseudogene?	O' Connor et al. (2014) and personal communication Whiteman lab
D. immigrans	Immigrans	*D. sulfurigaster, D. pallidifrons, D. rubida*	*Immigrans*	Fruit, compost	yes	yes	Q	N	Brown et al. (https://doi.org/10.1101/2021.07.28.454263)
D. funebris	Funebris	*D. sulfurigaster, D. pallidifrons, D. rubida*	*Immigrans*	Fruit, compost, mushroom	yes	yes	S	N	Brown et al. (https://doi.org/10.1101/2021.07.28.454263)
D. testacea	Testacea	*D. neotestacea*	*Testacea*	Mushroom	yes	no	S	deleted	Martinson et al. (2017), Bost et al. (2018) and personal communication Vincent Martinson
D. phalerata	Quinaria section	*D. falleni*	*Quinaria section*	Mushroom	yes	no	S	N	Martinson et al. (2017), Bost et al. (2018) and personal communication Vincent Martinson
D. guttifera	Quinaria section	*D. falleni*	*Quinaria section*	Mushroom	yes	no	S	deleted	Martinson et al. (2017), Bost et al. (2018) and personal communication Vincent Martinson

Terms used in table to bin microbe presence/absence

yes = typically present across independent samples
stochastic = sometimes absent across independent samples
no = usually absent across independent samples
no mention = the OTUs described in the manuscript could reasonably overlap with the bacteria of interest, but no specific mention of the relevant genus or family was made
nd = no data
* = only available study was in lab-reared flies, which are more likely to be dominated by Acetobacter and Lactobacillus members

Fig4supp4

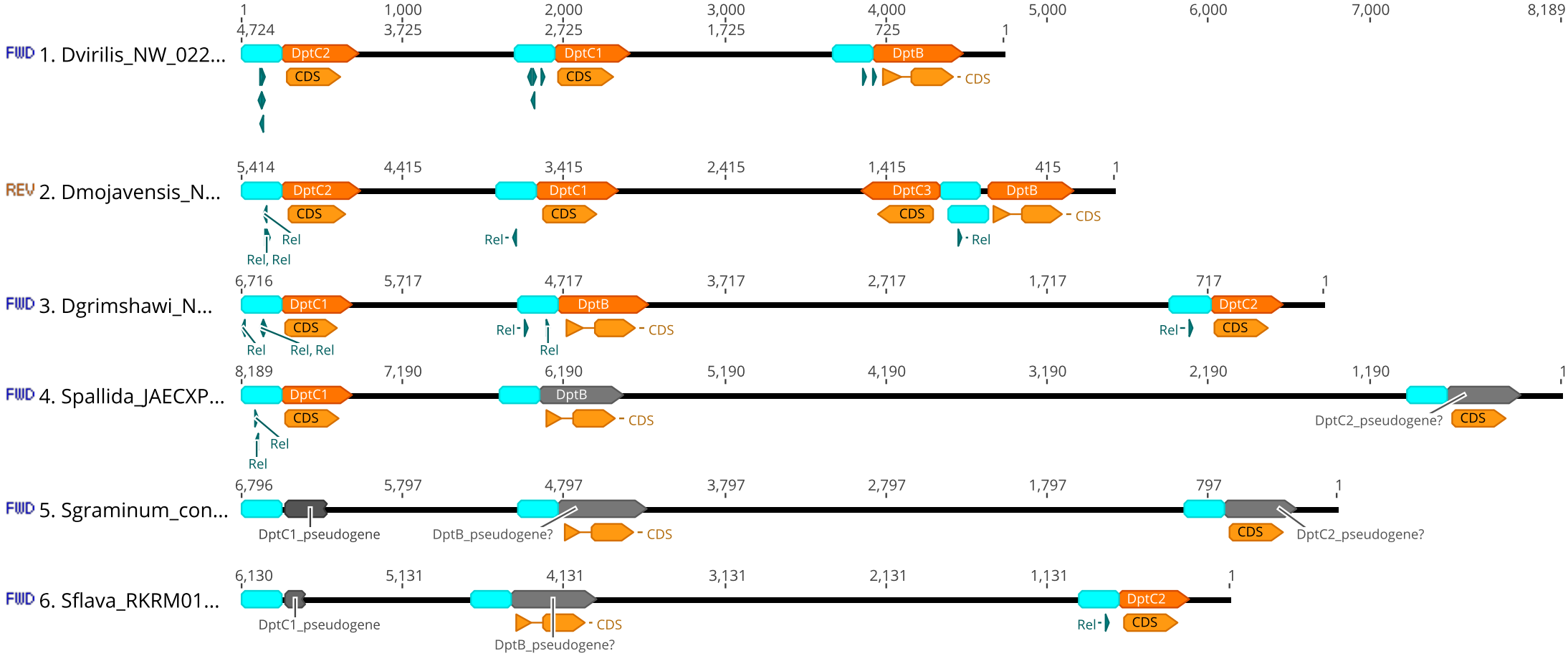

1. Dvirilis_NW_022...
FWD
DptC2
DptC1
DptB
CDS
2. Dmojavensis_N...
REV
DptC3
Rel
Rel, Rel
3. Dgrimshawi_N...
4. Spallida_JAECXP...
DptC2_pseudogene?
5. Sgraminum_con...
DptC1_pseudogene
DptB_pseudogene?
6. Sflava_RKRM01...

Fig4supp5

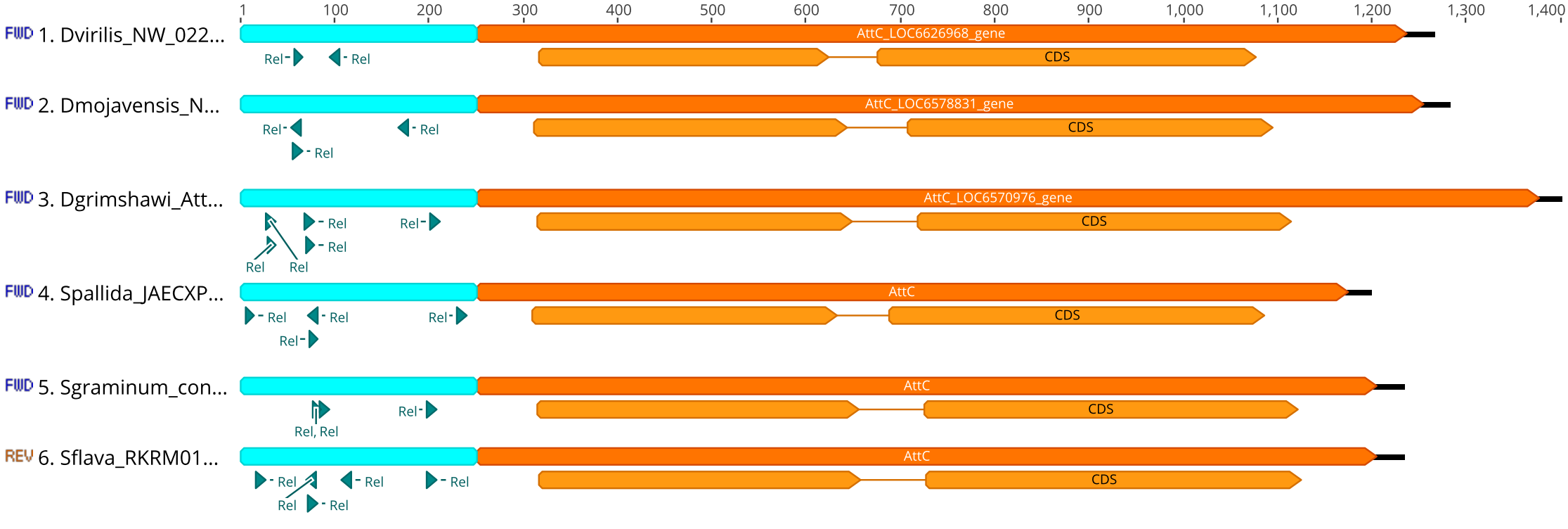

1
100
200
300
400
500
600
700
800
900
1,000
1,100
1,200
1,300
1,400
FWD 1. Dvirilis_NW_022...
AttC_LOC6626968_gene
Rel
Rel
CDS
FWD 2. Dmojavensis_N...
AttC_LOC6578831_gene
Rel
Rel
Rel
CDS
FWD 3. Dgrimshawi_Att...
AttC_LOC6570976_gene
Rel
Rel
Rel
Rel
Rel
CDS
FWD 4. Spallida_JAECXP...
AttC
Rel
Rel
Rel
Rel
CDS
FWD 5. Sgraminum_con...
AttC
Rel
Rel, Rel
CDS
REV 6. Sflava_RKRM01...
AttC
Rel
Rel
Rel
Rel
Rel
CDS

Fig5supp1

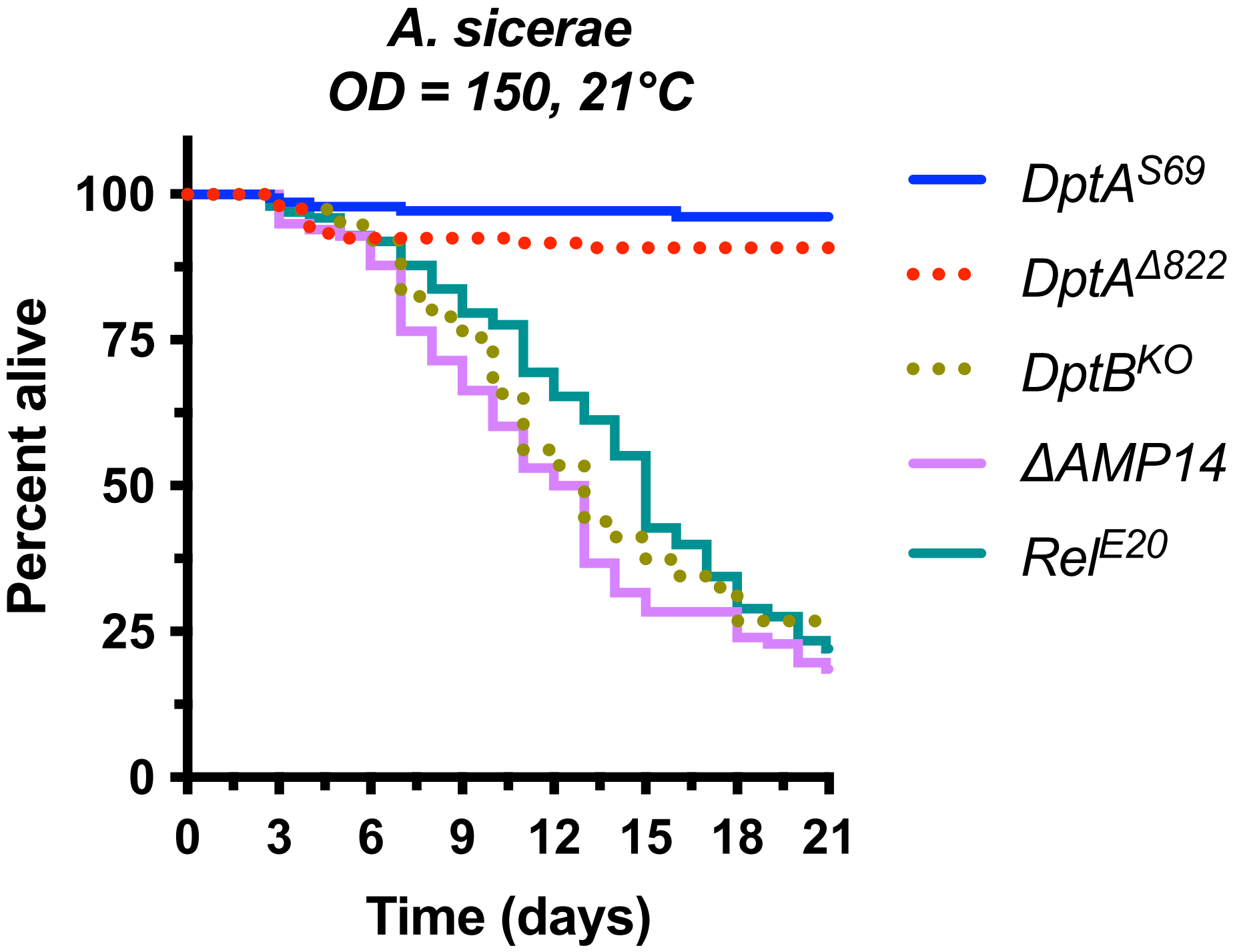

A. sicerae
OD = 150, 21°C
Percent alive
100
75
50
25
0
0
3
6
9
12
15
18
21
Time (days)
DptA^S69
DptA^Δ822
DptB^KO
ΔAMP14
Rel^E20

Fig5supp2

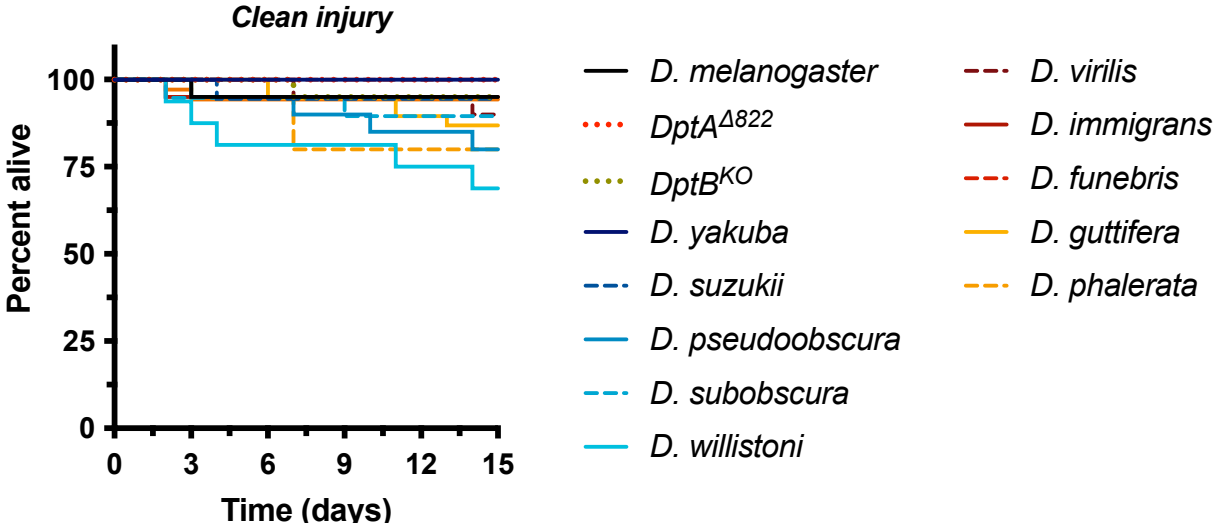
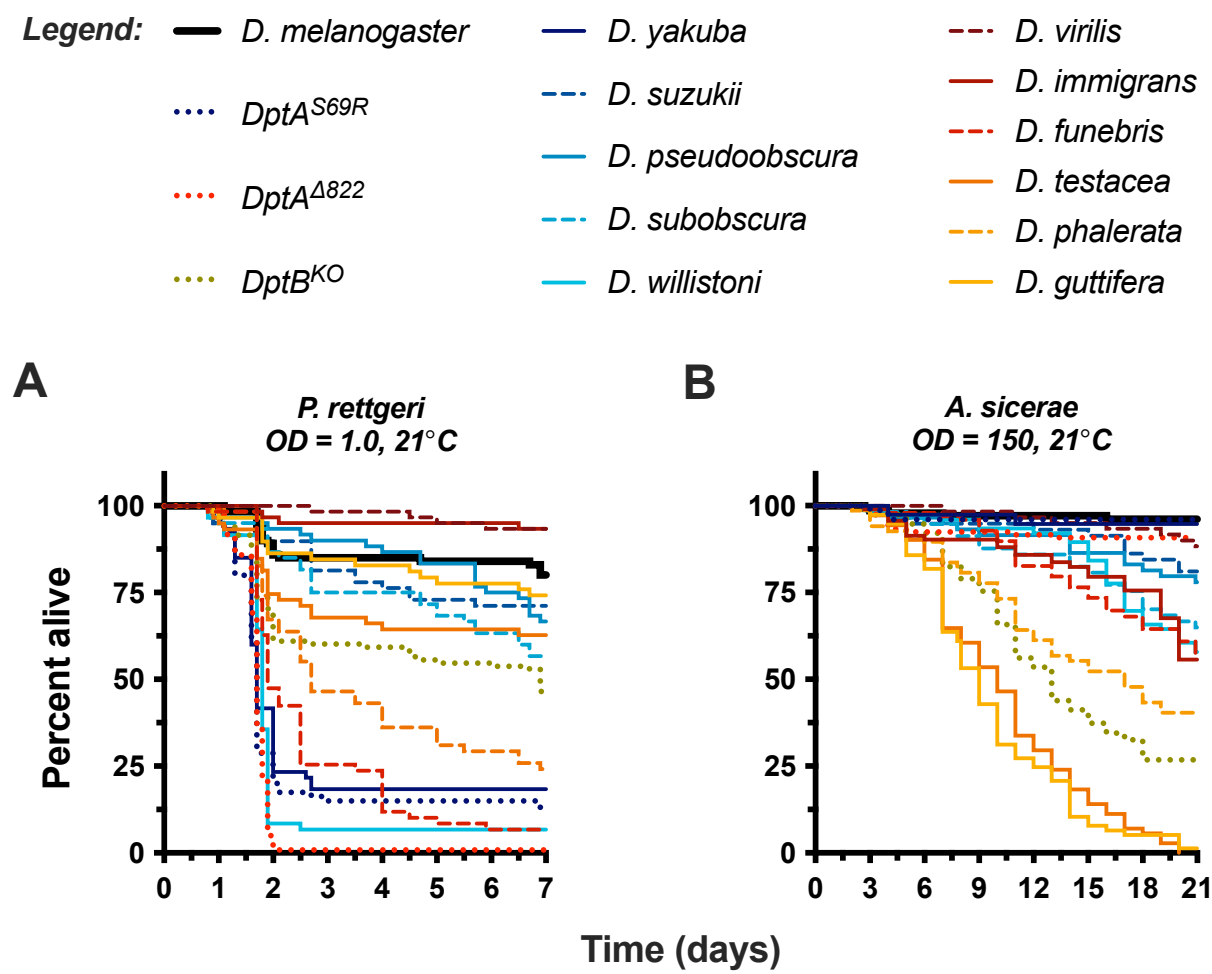

Clean injury
Percent alive
100
75
50
25
0
0
3
6
9
12
15
Time (days)
D. melanogaster
DptA$^{\Delta 822}$
DptB^{KO}
D. yakuba
D. suzukii
D. pseudoobscura
D. subobscura
D. willistoni
D. virilis
D. immigrans
D. funebris
D. guttifera
D. phalerata

Fig5supp3


Legend:
D. melanogaster
DptA^S69R
DptA^Δ822
DptB^KO
D. yakuba
D. suzukii
D. pseudoobscura
D. subobscura
D. willistoni
D. virilis
D. immigrans
D. funebris
D. testacea
D. phalerata
D. guttifera
A
P. rettgeri
OD = 1.0, 21°C
B
A. sicerae
OD = 150, 21°C
Percent alive
100
75
50
25
0
0 1 2 3 4 5 6 7
0 3 6 9 12 15 18 21
Time (days)