Title

Evolution of insect innate immunity through domestication of bacterial toxins

Journal

Proceedings of the National Academy of Sciences of the United States of America, 120(16)

ISSN

0027-8424

Authors

Verster, Kirsten I
Cinege, Gyöngyi
Lipinszki, Zoltán
et al.

Publication Date

2023-04-18

DOI

10.1073/pnas.2218334120

Peer reviewed

PNAS RESEARCH ARTICLE | EVOLUTION OPEN ACCESS

Evolution of insect innate immunity through domestication of bacterial toxins

Kirsten I. Verster[a,1], Gyöngyi Cinege[b,1], Zoltán Lipinszki[c], Lilla B. Magyar[b,d], Éva Kurucz[b], Rebecca L. Tarnopol[e], Edit Ábrahám[c], Zsuzsanna Darula[f,g], Marianthi Karageorgi[h], Josephine A. Tamsil[i], Saron M. Akalu[a], István Andó[b,2], and Noah K. Whiteman[a,i,j,2]

Edited by Nancy Moran, The University of Texas at Austin, Austin, TX; received October 27, 2022; accepted March 1, 2023

Toxin cargo genes are often horizontally transferred by phages between bacterial species and are known to play an important role in the evolution of bacterial pathogenesis. Here, we show how these same genes have been horizontally transferred from phage or bacteria to animals and have resulted in novel adaptations. We discovered that two widespread bacterial genes encoding toxins of animal cells, *cytolethal distending toxin subunit B* (*cdtB*) and *apoptosis-inducing protein of 56 kDa* (*aip56*), were captured by insect genomes through horizontal gene transfer from bacteria or phages. To study the function of these genes in insects, we focused on *Drosophila ananassae* as a model. In the *D. ananassae* subgroup species, *cdtB* and *aip56* are present as singular (*cdtB*) or fused copies (*cdtB::aip56*) on the second chromosome. We found that *cdtB* and *aip56* genes and encoded proteins were expressed by immune cells, some proteins were localized to the wasp embryo's serosa, and their expression increased following parasitoid wasp infection. Species of the *ananassae* subgroup are highly resistant to parasitoid wasps, and we observed that *D. ananassae* lines carrying null mutations in *cdtB* and *aip56* toxin genes were more susceptible to parasitoids than the wild type. We conclude that toxin cargo genes were captured by these insects millions of years ago and integrated as novel modules into their innate immune system. These modules now represent components of a heretofore undescribed defense response and are important for resistance to parasitoid wasps. Phage or bacterially derived eukaryotic toxin genes serve as macromutations that can spur the instantaneous evolution of novelty in animals.

horizontal gene transfer | toxins | immunology | adaptation | parasitoid

Significance

Several disease-causing bacteria produce toxins that damage host cells by triggering preprogrammed cell death. Two such bacterial toxins are called cytolethal distending toxin B and apoptosis-inducing protein of 56 kDa. We discovered that diverse insect species coopted the two bacterial genes encoding each cytotoxin through a phenomenon called horizontal gene transfer (HGT). HGT occurs when a gene from one organism is inserted into the genome of another and then is stably inherited across generations. We found that the two bacterial toxin genes were captured by an ancestral fruit fly ~21 Mya and are important for resistance against parasitoid wasps, which are principal enemies of fruit flies. These horizontally transferred genes now contribute to the fly's immune system.

Horizontal gene transfer (HGT) events from microbes to arthropods are widespread and have been shown to underlie functional adaptation in animals. For example, genes encoding a carotenoid pathway were horizontally acquired by aphids from fungi (1), and the coffee berry borer beetle acquired a mannanase-encoding gene from bacteria (2), among myriad other cases of HGT from microbial donors to diverse arthropod taxa (3, 4).

Phage-derived or associated prophage elements in bacteria were not previously known to be sources of macromutations in animals, though they are recognized as important players in bacterial evolution. In particular, phage toxin cargo genes can confer instantaneous, large fitness gains to bacterial cells. For example, temperate phages carrying toxin cargo genes (5) can become stably inherited in bacteria as prophages that encode diverse virulence factors including botulinum, cholera, diphtheria, and Shiga toxins (6). Here, we report the discovery of a macromutation arising from a phage or bacterial to animal HGT event, which spurred adaptation to parasitoid wasp infection.

Animals have benefited from the expression of phage-vectored cytotoxin genes when expressed in microbes (7, 8). Temperate, lambdoid phages of *Hamiltonella defensa* endosymbiotic bacteria called APSE protect sap-feeding insects from parasitoid wasps (9–11). This protective mutualism relies on diverse toxin cassettes carried by APSE phages. The genome sequences of three parasitoid wasp-protective APSE phages of *H. defensa* (APSE-2, APSE-6, and APSE-7) contain a homolog of *cytolethal distending toxin subunit B* (*cdtB*) (11) in a toxin cassette. CdtB is the active (A) subunit of cytolethal distending toxin (CDT) complex, a widespread AB_2 exotoxin of animal cells (12). CdtB is canonically a homolog of mammalian DNase I that produces single-stranded DNA breaks and triggers G_2/M arrest, causing cellular distention and apoptosis (12). The CDT holotoxin typically includes the binding (B) subunits CdtA and CdtC, but neither homolog is in APSE phages. It has been established that lambdoid phages can mediate HGT of CDT genes among bacteria (13).

Immediately downstream of *cdtB* in APSE-2 and some APSE-7 phage genomes is a homolog of *apoptosis-inducing protein of 56 kDa* (*aip56*) encoding the C terminus (B domain) of AIP56. AIP56 is an AB toxin (in which A is the active component and B is the binding component of the toxin) that was first isolated from the mariculture pathogen

Author contributions: K.I.V., G.C., Z.L., L.B.M., É.K., R.L.T., E.Á., Z.D., M.K., J.A.T., S.M.A., I.A., and N.K.W. designed research; performed research; contributed new reagents/analytic tools; analyzed data; and wrote the paper.

Competing interest statement: The authors have organizational affiliations to disclose. N.K.W. is on the Board of Directors of the Genetics Society of America.

This article is a PNAS Direct Submission.

[1]K.I.V. and G.C. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: ando@brc.hu or whiteman@berkeley.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2218334120/-/DCSupplemental.

Published April 10, 2023.

Photobacterium damselae subsp. *piscicida* (14). Its proteolytic activity arises from the N terminus (A domain), which cleaves NF-kB p65 and leads to apoptosis by interfering with the regulation of antiapoptotic genes (15). The A domain has signature HEXXH motif typical of most zinc metallopeptidases (16). The B domain is linked to the A domain via a disulfide bridge and facilitates cellular entry of the A domain to the target cell (15).

Previously, we found that both *cdtB* and *aip56* homologs were horizontally transferred to several insect lineages (17, 18), and we inferred three independent HGT events to the Drosophilidae. To determine the potential function and adaptive value of these insect-encoded toxins, here we studied *Drosophila ananassae* as an exemplar species because its chromosome 2L encodes *cdtB* and *aip56*, and it is amenable to genome editing (19). Furthermore, *D. ananassae* subgroup species have evolved a robust parasitoid defense mechanism distinct from that of the model organism *Drosophila melanogaster*. Generally, *D. ananassae* shows higher defense against parasitoid wasps than *D. melanogaster* (20, 21). It is known that in *D. melanogaster*, parasitoids are encapsulated with the involvement of a subset of blood cells, named lamellocytes, which are flattened cells important for the antiparasitoid response. A melanization reaction involving production of the lamellocyte-specific prophenoloxidase 3 (PPO3) contributes to parasitoid death (22). However, the gene encoding PPO3 is absent from the genomes of parasitoid-resistant *ananassae* subgroup species (23), and melanization of the capsule never occurs (21). Rather, multinucleated giant hemocytes (MGHs) encapsulate the parasitoids, suggesting that in these species an unknown, highly efficient killing mechanism must be present. We hypothesized that *D. ananassae* deploys CdtB and AIP56 proteins against parasitoid wasps through elements of the innate immune system.

Results

The Genes Encoding CdtB and a Noncatalytic Domain of AIP56 Were Horizontally Inherited from Endosymbiotic Bacteria or Phages in Insects. We previously found that *cdtB* and *aip56* were horizontally transferred to several insect lineages from bacterial endosymbionts or their phages (17, 18). Here, we report additional cases to Diptera, Hemiptera, Hymenoptera, Lepidoptera, and Thysanoptera, and the identities of the closest living relatives to the extant insect copies (Fig. 1 and *SI Appendix*, Fig. S1). Despite many independent HGT events to insects from prokaryotes, all insect CdtB proteins formed a single clade with those from *H. defensa* and APSE-2, APSE-6, and APSE-7 phages. We found similar results for the B domain of AIP56 proteins in insects (henceforth "AIP56"), which form a clade of proteins with AIP56 from APSE-2 and APSE-7 phages and insect-associated bacterial endosymbionts including *H. defensa* and *Arsenophonus* spp. Based on this evidence, we inferred that insect *cdtB* and *aip56* genes originated in ancestors of APSE phages that infected bacterial endosymbionts of arthropods, or from the bacteria themselves.

The Gene Order of *cdtB* and *aip56* in *D. ananassae* Subgroup Species Suggests HGT from an APSE-Like Sequence from a Phage or Bacterial Ancestor, Followed by Gene Fusion and Duplication. We found a single copy of *cdtB* (LOC26513850) upstream of two tandem *cdtB::aip56* fusion genes (LOC116654562 and

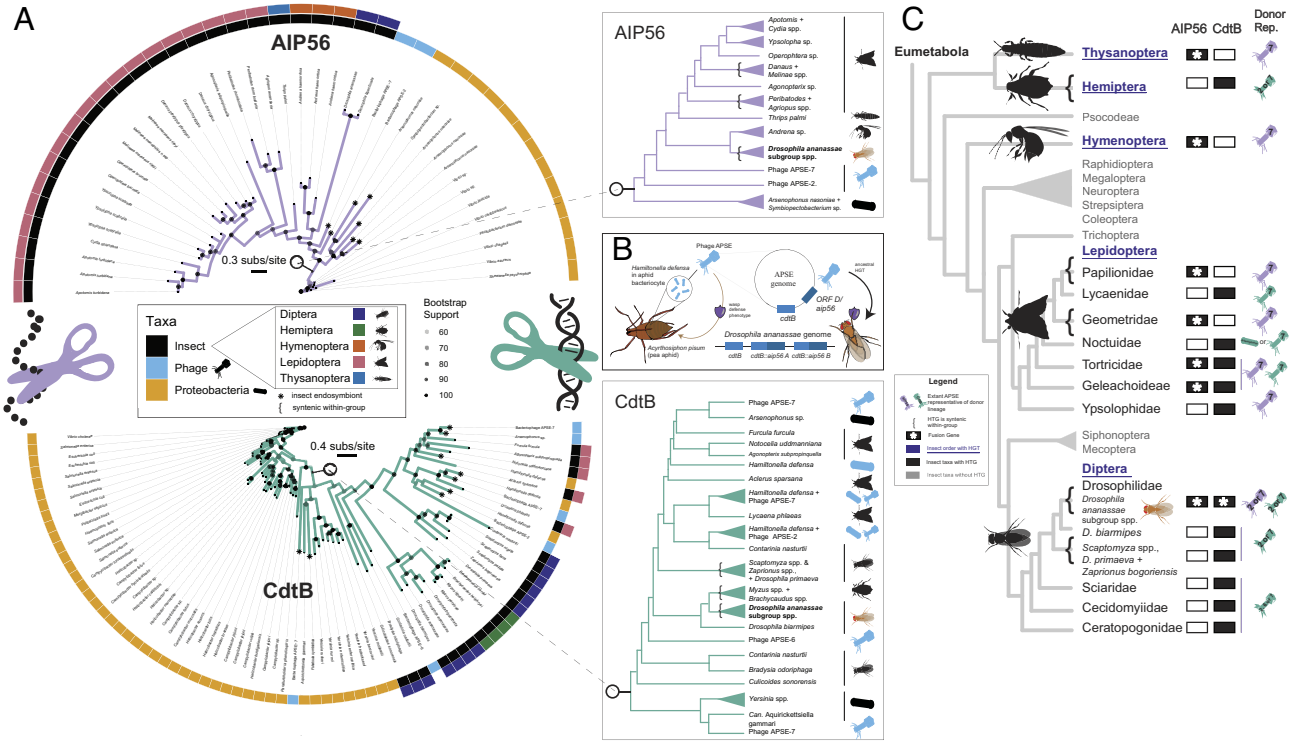

Fig. 1. *cdtB* and *aip56* genes were horizontally transferred to insects from phages or their bacterial endosymbiont hosts. (*A*) Maximum likelihood phylogenies of AIP56 and CdtB proteins show that insect-encoded horizontally transferred gene (HTG) sequences are nested within phage and endosymbiotic bacteria clades of each protein, suggesting ancestral HGT from prokaryotes or phages to insects. Several cases of HGT-derived genes are syntenic within species groups (*SI Appendix*, Tables S1 and S4), suggesting that HGT was followed by extended vertical transmission. (*B*) Synteny between *cdtB* and *aip56* in the APSE phage and *D. ananassae* subgroup species genomes suggests that ancestral HGT occurred from APSE phages or their bacterial lysogens, which may contribute to the efficient wasp resistance of *D. ananassae*. (*C*) Insect species tree highlighting the insect orders and taxa which horizontally acquired *cdtB* and/or *aip56* genes. Shown are extant representatives of the possible donor lineage.

LOC116654561, referred to as *cdtB::aip56 A* and *cdtB::aip56 B,* respectively) in the *D. ananassae* reference genome assembly.

Strikingly, this tandem arrangement maintains the gene order found in the extant APSE-2 phage genome sequence, where *aip56* is located immediately downstream of *cdtB*, consistent with ancient HGT from an APSE-like ancestral sequence in phages or bacteria (Fig. 1*B*) to a *D. ananassae* ancestor. This may then have been followed by fusion of *cdtB* with *aip56* and duplication of the resultant fusion gene, though the order of these events is unclear. All the three genes have introns, and all reference genome assemblies of the other *ananassae* species carried syntenic copies of these genes (*SI Appendix,* Table S1). Thus, HGT occurred once in this lineage based on maximum parsimony, prior to the diversification of the *D. ananassae* species subgroup. *D. ananassae* is the only species in the subgroup that has a duplication of *cdtB::aip56.* Essential DNase residues are conserved in *D. ananassae* CdtB and CdtB::AIP56 proteins (*SI Appendix*, Fig. S2), and we found that purifying selection has acted on residues encoded by *cdtB* and *cdtB::aip56* in the *D. ananassae* subgroup lineage since the HGT event (*SI Appendix*, Fig. S3 and Table S2). Thus, the function of the encoded CdtB toxins appears to remain optimized for DNase activity, although their exact organismal function is unclear.

Horizontally Transferred *cdtB* and *cdtB*::*aip56* Fusion Gene and Protein Expression Are Induced after Infection with Parasitoid Wasps. We conducted gene expression studies across development of *D. ananassae* with or without *Leptopilina boulardi* figitid parasitoid wasp infection. Single-copy *cdtB* expression was highest in the fly embryo, gradually decreased throughout larval development, and was moderately elevated after parasitoid attack (Fig. 2*A*). In contrast, *cdtB::aip56* fusion copies were marginally expressed during embryogenesis, which increased during the larval stages, and then showed a strong induction following *L. boulardi* attack in larvae (Fig. 2*A*). Expression of all the three genes was enriched in fat body and blood cells post wasp infection, consistent with a humoral innate immune role (*SI Appendix*, Fig. S4). We found that expression of these genes is not induced indiscriminately following injury but is tied specifically to wasp infection (*SI Appendix,* Fig. S5).

Next, we generated a panel of monoclonal antibodies against recombinant single-copy CdtB and CdtB::AIP56 A/B fusion proteins for immunological studies. Antibody specificity was confirmed on recombinant proteins (*SI Appendix*, Fig. S6) and biological samples by western blot (WB), indirect immunofluorescence (IIF) assay, and liquid chromatography–tandem mass spectrometry (LC-MS/MS) (Fig. 2 *B* and *C* and *SI Appendix*, Figs. S6–S9 and Table S3). Protein expression analysis during development of naïve and infected *D. ananassae* (Fig. 2*B*) mirrored the gene expression profiles (Fig. 2*A*).

IIF experiments that targeted single-copy CdtB revealed expression in embryonic macrophages (Fig. 2*C*). Consistent with the WB data, we did not detect CdtB in naïve or infected larval blood samples, or on parasitoids (*SI Appendix*, Figs. S7*B* and S8 *B–D*), but in whole larvae, the protein was weakly induced following wasp parasitization.

WB analysis of the CdtB::AIP56 A/B fusion proteins revealed their expression in larval, pupal, and adult stages, which was elevated after parasitoid wasp infection. Under nonreducing conditions, the antibodies reacted at positions corresponding both to the complete, fused proteins and those of the individual CdtB and AIP56 domains (Fig. 2*B*). However, under reducing conditions, the majority of the proteins were detected at positions corresponding to the individual CdtB and AIP56 components (*SI Appendix*, Fig. S9), suggesting cleavage of disulfide bonds linking them together. We observed the same pattern in *L. boulardi* wasp larvae extracted from the hemocoel of *D. ananassae*, suggesting that the fusion proteins bind the wasp larvae in the unreduced, processed state (Fig. 3). Structural predictions of CdtB::AIP56 further corroborate that the CdtB component is distinct from that of AIP56 (*SI Appendix*, Fig. S10).

Though our WB data showed fusion CdtB A/B in blood samples (*SI Appendix*, Fig. S6*B*), and in the isolated parasitoid wasps (Fig. 3*A*), they were not detected using IIF (*SI Appendix*, Fig. S8 *F–H*), suggesting the masking of CdtB A/B toxin epitopes in vivo. AIP56 A/B components were detected by both WB and IIF in the hemolymph clot, in MGHs, spherical blood cells, and microvesicles (*SI Appendix*, Figs. S6*B* and S8 *J* and *L*), confirming the presence of the fusion proteins and the epitope hiding of their CdtB domain. Moreover, AIP56 A/B components were also detected in precipitates localized on parasitoids, and in a single layer of outer epithelial cells, the serosa of the wasp (Fig. 3*B* and *SI Appendix,* Fig. S8*K* and Movie S1).

***cdtB* and *cdtB*::*aip56* Mutants Are Susceptible to Parasitoid Wasp Infection.** We next generated a panel of backcrossed, homozygous null mutant lines of *D. ananassae* (*SI Appendix*, Fig. S11) whose genotypes represented several combinations of loss-of-function (LOF) mutations in *cdtB* and *cdtB::aip56* genes, including a triple null mutant for all the three genes. Wild-type and mutant flies were then exposed to gravid females of three *Leptopilina* wasp species used to dissect mechanisms of innate immunity in *Drosophila* (20, 21), *L. boulardi, L. heterotoma,* and *L. victoriae*. *L. boulardi* is the least virulent and specialist of the broader *melanogaster* group, which includes *D. ananassae. L. heterotoma* is a virulent generalist that attacks species inside and outside the *melanogaster* group. Finally, *L. victoriae* is the most virulent, specialized on the *ananassae* species subgroup (24). It attacks *D. ananassae* in its native range in South and Southeast Asia and the Indo and South Pacific (further descriptions of the wasp and fly ranges can be found in *SI Appendix*, *Extended Discussion*).

We found that the *cdtB* and *cdtB::aip56* mutant *D. ananassae* lines were more susceptible to wasp infections (Fig. 4) than the wild type. We observed that some *L. boulardi* adults emerged in mutants lacking both *cdtB::aip56 A* and *cdtB::aip56 B*, even though *L. boulardi* does not normally complete development in *D. ananassae* (21). Mutant *D. ananassae* lines were more susceptible to wasp infections (Fig. 4) as the number of LOF mutations in *cdtB* and *cdtB::aip56* genes increased. In particular, the triple mutant *D. ananassae* lines had significantly lower host survival rates when attacked by each of the three wasp species compared to wild-type flies and had significantly higher rates of emergences in each of the wasp species. Overall, these experiments show that the three horizontally transferred genes (HTGs) are essential for wild-type resistance against *Leptopilina* wasps in *D. ananassae*.

Discussion

Here, we report the repeated horizontal transfer of cytotoxin genes derived from an APSE-like ancestral sequence in phages or bacteria to insect genomes. We dissect their potential adaptive value in a recipient lineage, the fruit fly *D. ananassae*, where gene, protein, and knockout data are each consistent with an immune role. These same horizontally transferred cytotoxins are adaptive to host insects when expressed in bacterial intermediaries (e.g., in the case of *H. defensa* and its APSE phages) because they provide protection from parasitoid wasp attacks. Our knockout studies in *D. ananassae* suggest that this is also true when homologs

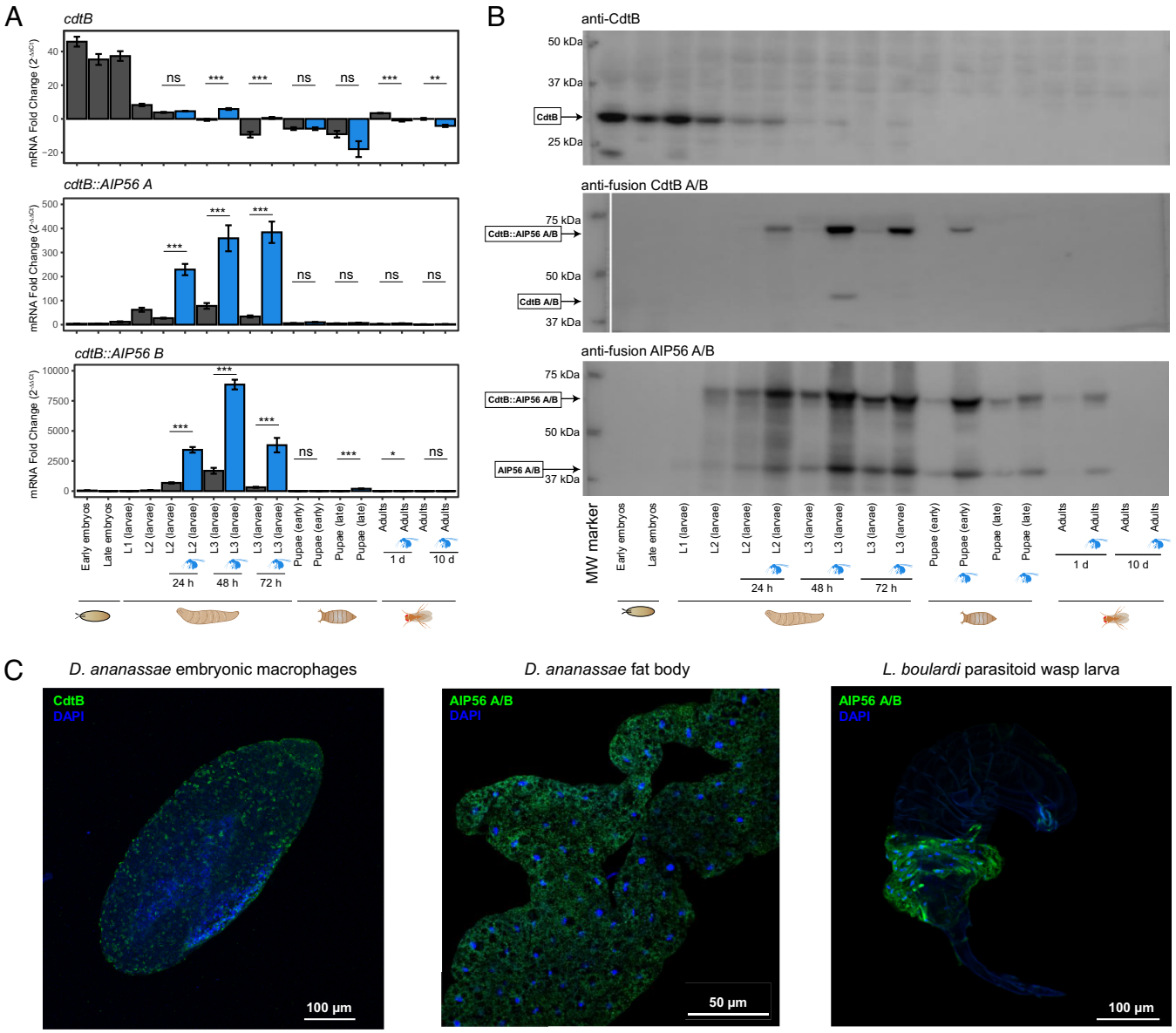

Fig. 2. Horizontally transferred *cdtB* and *cdtB::aip56* fusion gene and protein expression are induced after infection with endoparasitoid wasps. (*A*) Expression of *cdtB, cdtB::aip56 A*, and *cdtB::aip56 B* genes varies across *D. ananassae* developmental stages and is induced after *L. boulardi* infections (labeled in blue). L1, L2, and L3 correspond to first, second, and third larval developmental stages, respectively. 24 h, 48 h, and 72 h correspond to hours following wasp infection or control treatment. The error bars indicate the SE of four to six independent data points. Student's *t* test was used for statistical analysis of gene expression profiles between naïve and wasp-induced larval samples. *P*-values: <0.05 = *; <0.01 = **; <0.001 = ***; ns = not significant. ΔΔCt was calculated by normalizing ΔCt against those of 10-d-old naïve adult samples. (*B*) WB analysis of single-copy CdtB and fusion proteins, following the same developmental series and parasitization regimes as in (*A*). Protein levels mirror those of messenger RNA (mRNA) (*A*). (*C*) Expression of CdtB and AIP56 A/B proteins detected by IIF analysis.

of these cytotoxins are encoded in the genomes of the host insects themselves.

The *H. defensa* endosymbionts impose fitness costs on sap-feeding insects when parasitization rates are low (25), a trade-off that may explain why endosymbionts and phages are frequently lost and regained in natural populations (11). Our study shows how the reach of this widespread defensive mutualism has been extended through HGT from phages or prophages to insects directly. In the case of *D. ananassae*, the bacterial intermediary that is housed as an intracellular mutualist and associated costs of this symbiosis have been obviated through HGT. This is a biological example of the economic principle of disintermediation or "cutting out the middleman." However, there are likely pleiotropic costs associated with expressing a eukaryotic cytotoxin in a eukaryotic cell as well. It remains to be understood how such costs have been mitigated during the domestication process in *D. ananassae*.

We suggested that *D. ananassae cdtB* and *aip56* genes may have been inherited sequences from APSE-like phage ancestors or bacteria ~21 Mya (17), a hypothesis strengthened by the observed synteny between these genes in the extant APSE-2 and *D. ananassae* subgroup species genomes (Fig. 1*A*). However, due to the vast expanses of time involved, the shuttling of toxin gene cassettes between taxa via HGT, a lack of sampling, and temporal biases, we are unlikely to ever know the specific provenance of these genes. We hypothesize that they came from insect endosymbiotic bacteria or their phages, due to the increased likelihood of HGT between spatially overlapping organisms (18) and the large representation of insect endosymbionts in the CdtB and AIP56 phylogenies (*SI Appendix*, Fig. S1). One possible candidate could be an ancestor of *Arsenophonus* spp., a clade of insect intracellular symbionts which has been found to acquire APSE-like gene cassettes through HGT (11). Following the initial HGT from a microbe or phage

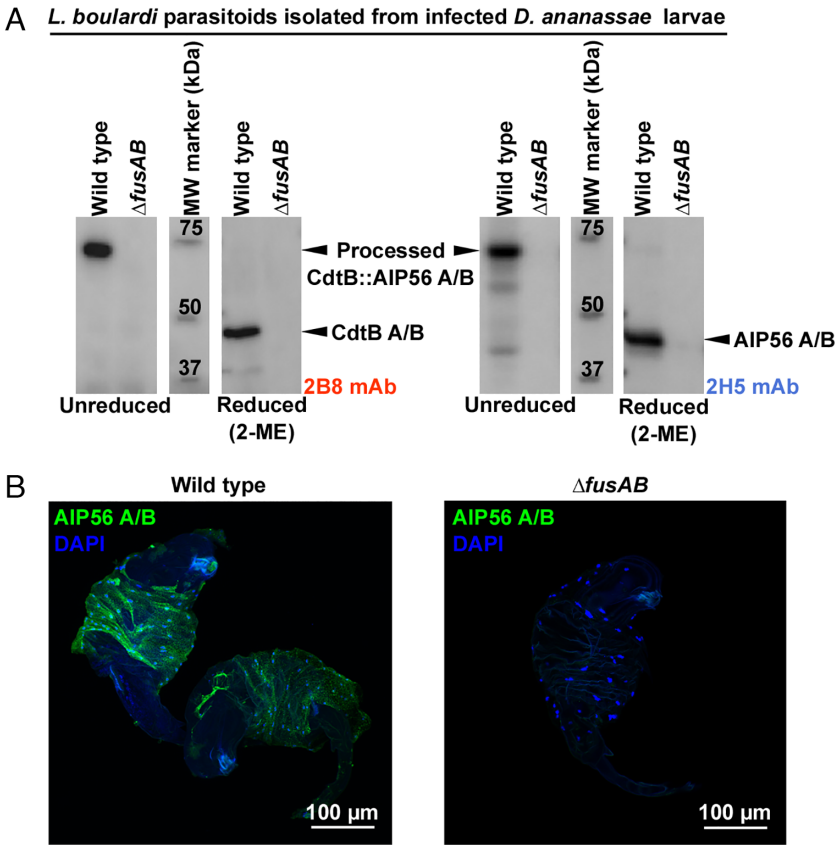

Fig. 3. Processed CdtB::AIP56 A/B proteins are present in parasitoid wasps isolated from infected *D. ananassae* larvae. (*A*) The CdtB::AIP56 A/B could be reduced by 2-ME, which revealed that disulfide bonds link the CdtB and the AIP56 subunits of the processed fusion proteins. Twenty microgram protein was loaded per lane. (*B*) AIP56 A/B proteins in parasitoid wasp larvae isolated from wild-type and mutant *D. ananassae.*

to a eukaryotic genome, the gene has to undergo a process of domestication, wherein it is incorporated into existing genetic networks or pathways (26), and may acquire other eukaryotic motifs such as introns. For example, *cdtB* and *cdtB::aip56* genes in the *D. ananassae* lineage have acquired eukaryotic motifs such as untranslated regions (UTRs), TATA boxes, polyA signals, and cleavage sites (17), in addition to having been integrated into mainstream immune mechanisms, as we show here.

The pathogenicity of CdtB in other contexts is facilitated by additional subunits, such as CdtA and CdtC in several bacterial species [*Campylobacter jejuni* and *Helicobacter hepaticus,* among others (12, 27), or PltA and PltB in *Salmonella* Typhi (28)]. However, CdtB can function as a DNase by itself. Some studies have suggested that horizontal transfer is more likely to occur in genes that, like *cdtB*, can serve as self-contained units, i.e., those that are not part of protein complexes or complicated pathways (29). This could explain the abundance of *cdtB* HGT events across the insects as a whole (Fig. 1).

Given that the canonical proteolytic A domain of AIP56 is missing from all insects, and most insect-associated phages or

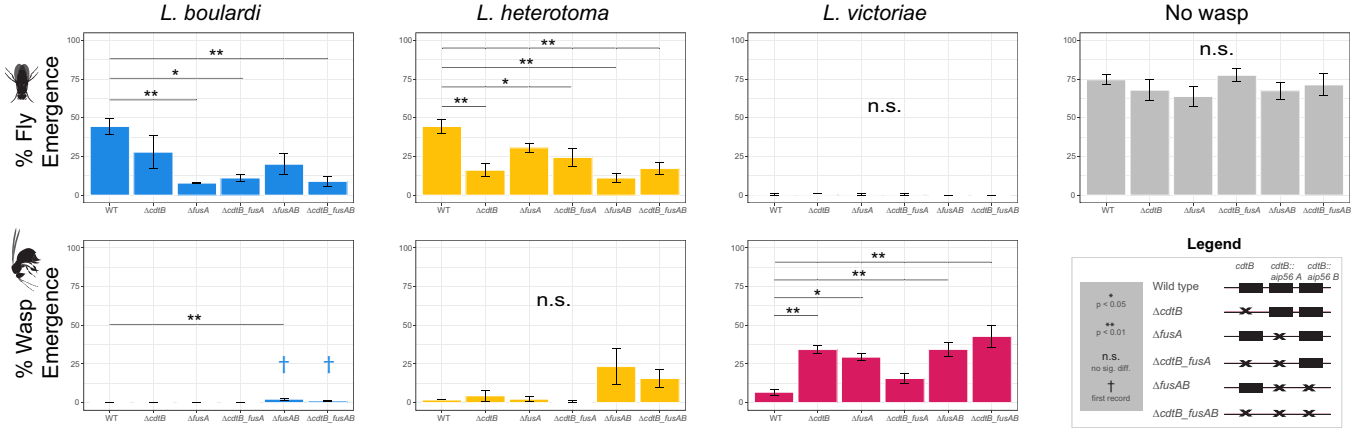

Fig. 4. *D. ananassae* mutants are less likely to survive parasitization from *L. boulardi*, *L. heterotoma,* and *L. victoriae* than wild-type flies. *Leptopilina* spp. wasps are more likely to eclose when developing in *D. ananassae* mutants. The error bars indicate the SEM from three to four independent data points. Tukey's post-hoc test was used to compare means and assess statistical significance of differences; for clarity, only differences between mutant lines and wild type are shown. *P*-values: <0.05 = *; <0.01 = **; ns = not significant. Results of all pair-wise comparisons between all mutant lines are in *SI Appendix,* Table S6.

bacteria, other encoded genes may have served a similar catalytic function after fusing to the nonproteolytic AIP56 B domain. These include CdtB in *D. ananassae* or SltxA in lepidopterans and hymenopterans (*SI Appendix, Extended Discussion*). This hypothesis is strengthened by our finding that *D. ananassae* CdtB is connected to AIP56 B domain via disulfide bridges, a phenomenon characteristic of bacterial AB toxins (15). Other examples of bacterial CdtB sequences linked to noncanonical "B" subunits (e.g., Shiga-like toxin) have been reported (30). Such multiplexed gene fusions are common and have been shown to contribute to the evolution of multidomain proteins in both bacteria and eukaryotes (31). We hypothesize that, similar to its function in the pathogenic *P. damselae* subsp. *piscicida*, the B domain of AIP56 facilitates cellular entry of the catalytic subunit into the target cells. Though both *cdtB* and *aip56* cooccur in some APSE phages, the actual fusion of the two genes appears to be unique to insects.

Epitope masking in the CdtB domain of the fusion proteins, under in vivo conditions, suggests that the eukaryotic genotoxin is enveloped by antitoxin-like molecules (possibly involving AIP56) to avoid cellular damage of the eukaryotic host. We showed that AIP56 was detected on the wasp serosa and, given that AIP56 is linked to CdtB via disulfide bonds, we conclude that fusion CdtB may target the serosa as well. The serosa plays an essential role in nutrient uptake and pathogenesis during the development of diverse parasitoid wasps (32). In some wasp species, this membrane gives rise to single cells called teratocytes that disassociate, invade the host tissues, and suppress the innate immune response (32). We speculate that AIP56 A/B components may bind to a parasitoid wasp-specific receptor, which facilitates entry of the fusion CdtB components into the serosa. This intoxicated serosal layer could be associated with parasitoid death (Fig. 3*B* and *SI Appendix,* Fig. S8*K* and Movie S1). Further research is needed to characterize the mechanism of action of these toxins in the parasitoid and host.

Parasitization assays showed that not only CdtB::AIP56 fusion proteins were involved in the parasitoid defense response (Fig. 4), but also the single-copy CdtB, which was not detected on parasitoid wasps isolated from the host (*SI Appendix*, Fig. S8 *C* and *D*), possibly due to an indirect role in the defense reaction. It is known that embryonic macrophages, where CdtB is expressed, are cells involved in remodeling of the nerve cord that deposit extracellular matrix proteins and retaining of innate immune functions through development (33). We therefore hypothesize that CdtB plays an indirect role in defense via an unknown signaling pathway and may also have a pleiotropic developmental function. As CdtB exhibits potent phosphatidylinositol-3,4,5-triphosphate phosphatase activity (34), small fold induction for CdtB might be able to trigger large changes in expression of other elements, which could also be involved in the protection mechanisms against parasitoids.

The fitness advantage conferred by *cdtB* and *cdtB::aip56* genes was large when wild-type flies were compared to the toxin gene knockout fly lines for each of the three endoparasitoid wasps with varying immune strategies. Following the parasitoid infection, both the host and the invader start a complex and species-dependent defense mechanism involving virus-like particles originating from the venom injected during oviposition (20, 35). The outcome of these elaborate defense mechanisms, which are an emergent property of the host–parasite interaction (36), is reflected both in fly and wasp emergence rates, to varying degrees. Our results show that CdtB and CdtB::AIP56 are important components of the parasitoid immune response in *D. ananassae*, but their absence is not sufficient to eradicate it altogether (Fig. 4).

Critically, the HGT-derived cytotoxins serve to bolster the immune system of insects like *D. ananassae* against parasitoid wasp attacks. Mortality rates from endoparasitoids can exceed 50% in natural drosophilid populations (20), and thus new traits that enhance protection have the potential to be strongly favored by natural selection. In contrast to the diversified effector repertoires and adaptive immune systems of chordates, insects have relied instead on a highly conserved effector pool coupled with an eclectic but effective defense strategy that includes the use of toxins provided by other organisms. This includes dietary toxins such as ethanol that protect *D. melanogaster* against parasitoid wasp attack (37), and Ribosome Inactivating Protein (RIP) toxins encoded by *Spiroplasma* endosymbionts that protect mushroom-feeding drosophilids from both nematodes and parasitoid wasps (38, 39). A third strategy that has evolved, as we have shown here, is through cytotoxins ancestrally directed at animal cells that have been gained through ancient HGT events and provide resistance against parasitoid wasp attack.

Despite the initial outcrossing of the mutant fly lines, and the similar emergence rates of mutant and wild-type flies in the absence of parasitoid pressure, we cannot eliminate the possibility that nontarget, pleiotropic effects or off-target mutations led to deficiencies in these LOF mutant lines that were brought forward under stress. Future studies are needed that focus on phenotyping of multiple LOF mutant alleles, on corroborating these results with genetic rescue in *D. ananassae,* and gain-of-function experiments in *D. melanogaster*.

It is dogma that phage-derived macromutations result in the instantaneous origin of new virulence factors in prokaryotes. Our study demonstrates that some of the same toxin cargo genes are utilized as macromutations by animals too, resulting in a repeated pattern of the evolution of novelty across domains of life.

Materials and Methods

Identification of CdtB and AIP56 in Insects. To search for cases of HGT of *cdtB* and *aip56* in eukaryotes, we used HGT screening methods as described previously (18), with modifications. AIP56 and CdtB proteins from APSE phages and insects were queried against publicly available, translated eukaryotic genomes on National Center for Biotechnology Information (NCBI) GenBank, Darwin Tree of Life project (www.darwintreeoflife.org), and individual aphid genome sequences. Sequences with E-values <0.0001 were retained for analysis and manually validated for evidence of contamination (more details in *SI Appendix, Supplementary Methods*). A summary of the new insect sequences is shown in *SI Appendix,* Table S4.

TBLASTN was used to identify the coordinates of *cdtB* and *cdtB::aip56* genes in *Drosophila ananassae* subgroup species from published genomes (*SI Appendix, Extended Methods*). The two genes are found in all sequenced genomes in this group, are syntenic, and are between 400 and 550 bp of each other, with *cdtB* upstream of *cdtB::aip56*. *D. ananassae* appears to be the only species with two copies of *cdtB::aip56*, which was validated with PCR and Sanger sequencing (*SI Appendix,* Table S5).

Generation of AIP56 and CdtB Toxin Phylogenies. The multidomain nature of AIP56, its fusion to other encoded proteins, high divergence, and prior misannotations complicated previous attempts to create informative AIP56 phylogenies (18). To overcome these confounds, we queried AIP56 B domain sequences from four query species: *Arsenophonus nasoniae* (WP_051297188.1: 346-470), *Photobacterium damselae* subsp. *piscicida* (WP_012954632.1:354-479), APSE-2 (ACJ10170.1), and *D. ananassae* (XP_017099943.1: 477-649) against NCBI GenBank and extracted the top 100 "aligned amino acid sequences" per search, removing duplicates. CD-HIT at 90% similarity was used to prune redundant sequences in the main figure. Sequences were aligned in MAFFT v7.450 (40, 41), and protein topologies were inferred using maximum likelihood as implemented in W-IQ-TREE (http://iqtree.cibiv.univie.ac.at/) (42) using the best-fit model as assessed by BIC in ModelFinder. Nodes with 50% bootstrap support were collapsed using the di2multi function in ape v5.4 (43). Phylogenies

were visualized and annotated using ggtree v. 2.5.0.991 (44). For CdtB, similar methods were used, except with a 0.8 cutoff for CD-HIT and the following query sequences: XP_031641203 (*Contarinia nasturtii*); XP_014760894 (*D. ananassae*); XP_022163116 (*Myzus persicae*); QDF82162 (*Scaptomyza flava*); CAB3623624 (APSE-7); CAB3775476 (APSE-2); WP_100096556 (*Hamiltonella defensa*); and *Lycaena phlaeas* sequence (scaffold CAJOSV010000055.1, translated nt 1772611-1773549). The full trees for each protein can be seen in *SI Appendix*, Fig. S1.

Alignment and Structural Predictions of CdtB and AIP56. Representative CdtB sequences were aligned in MAFFT (*SI Appendix*, Fig. S2). We found that previously annotated residues corresponding to DNA binding, metal ion binding, and enzyme catalysis were highly conserved between prokaryotes, phage, and insects (17) (*SI Appendix*, Fig. S2), suggesting conserved DNase activity. AIP56, and in particular the B domain, is not as well characterized as CdtB, and thus the alignment is not informative as to conserved function. *D. ananassae* CdtB and CdtB::AIP56 sequences were submitted to the AlphaFold v2.1.0 colab notebook (https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb) (45). We used AlphaFold because previous attempts to predict AIP56 structure have failed on other software such as Phyre2 (18, 45), and AlphaFold predicts protein structures via a machine learning approach even when no similar structure is known (45). PDB files for top-ranked predictions are in Datasets S1–S3.

Measuring Evolutionary Constraint on *cdtB* and *cdtB*::*aip56* in *D. ananassae* Subgroup Species. Ka/Ks analyses were calculated using UniBCCS online tool, and then confirmed with HyPhy aBSREL (46). Selection at specific codons was tested using FUBAR (47) using default settings. For more details, see *SI Appendix*, *Extended Methods*.

Insect Stocks and Culturing. *D. ananassae* wild type (14024-0371.13), obtained from UC San Diego *Drosophila* Stock Center, was kept at 25 °C on standard yeast-cornmeal food. The *Leptopilina boulardi* G486, *L. victoriae* LvUNK, and *L. heterotoma* Lh14 were a gift from Todd Schlenke (University of Arizona) and maintained on *D. melanogaster* Oregon R hosts.

Generation of Wasp-Induced Samples for RNA and Protein Assays. To generate *D. ananassae* larvae that were infected with parasitoid wasps for gene and protein expression experiments, a total of 70 staged, second instar *D. ananassae* larvae were exposed to 17 mated female *L. boulardi* G486 parasitoid wasps for 6 h. Parasitized fly larvae were then selected by screening under a dissecting microscope 24 h following wasp induction based on the presence of the melanized site of oviposition on the cuticle. The selected wasp-infected larvae were placed in vials with standard yeast-cornmeal food for further culturing until the required developmental stage for harvesting.

qRT-PCR. RNA samples were prepared by RNeasy mini kit (Qiagen) according to the manufacturer's instructions. For each sample, 25 µL complementary DNA (cDNA) was generated from 1,000 ng RNA using the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific) with the Oligo(dT)18 Primer. PerfeCTa SYBER Green SuperMix (Quanta bio) and 2 µL 10 times diluted cDNA were used for qRT-PCR. A Rotor-Gene Q (Qiagen) qPCR platform was used with the following reaction conditions: 95 °C 2 min, 45 cycles at 95 °C for 10 s, 57 °C for 45 s, and 72 °C for 15 s. Two biological replicates were used, each generated from a minimum of 30 individuals. In each qRT-PCR experiment, two to three technical replicates were applied. To interpret gene expression levels, the ΔΔCt method was used. Cycle threshold (Ct) values of the *cdtB* and *cdtB::aip56* genes were normalized to those of the housekeeping *GF23239* gene (LOC6505882). Further normalization was done to the respective sample in the experiment (labeled in the figure legends). $2^{-\Delta\Delta Ct}$ values are presented on graphs, which show the messenger RNA (mRNA) fold change of the respective sample compared to that considered as reference. Data analysis was performed with Rotor-Gene Q Series Software and Q-Rex 1.0. Student's *t* tests and Tukey's post hoc tests were used for further statistical analysis between samples. Primers are listed in *SI Appendix*, Table S5.

DNA Constructs for Recombinant Protein Production. We generated untagged *cdtB* from the pET-His6-TEV-cdtB [2B-T] plasmid (17) by standard mutagenesis. *cdtB A* (nt 61-840 of the *cdtB aip56 A* CDS), *cdtB B* (nt 61-840 of the *cdtB aip56 B* CDS), *aip56 A* (nt 841-1941 of the *cdtB aip56 A* CDS), and *aip56 B* (nt 841-1953 of the *cdtB aip56 B* CDS) were PCR amplified [using Q5 High-fidelity DNA polymerase (NEB)] from coding DNA sequences reverse transcribed from mRNA prepared from *L. boulardi* G486 infected wild-type *D. ananassae* larvae. To generate His6-tagged *aip56-A* and *aip56-B* or His6-TEV-tagged *cdtB A* and *cdtB B*, coding sequences were cloned into the pET-DUET1 (Merck-Millipore) or pDEST17 (Thermo Fisher Scientific) plasmids, respectively (without the predicted N-terminal signal peptides). All constructs were validated by DNA sequencing. Sequences of oligonucleotide primers are provided in *SI Appendix*, Table S5.

Recombinant Protein (Antigen) Expression and Purification. CdtB, His6::TEV::CdtB A (CdtB A), His6::TEV::CdtB B (CdtB B), His6::AIP56 A (AIP56 A), and His6::AIP56 B (AIP56 B) were expressed in SixPack *Escherichia coli* strain (48) as follows: cells were grown in 500 mL standard Luria-Bertani broth to A_{600}~0.6 and expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 20 h at 16 °C. The cells were harvested by centrifugation and lysed by sonication. CdtB was purified under native conditions with immobilized metal affinity chromatography (IMAC) (using Ni-Sepharose 6 Fast Flow resin, Cytiva), according to the manufacturer and further purified on HiLoad 16/600 Superdex 75 size-exclusion chromatography column (Cytiva) preequilibrated with sterile phosphate-buffered saline (PBS). CdtB A, CdtB B, AIP56 A, and AIP56 B were purified under denaturing conditions with IMAC (using Ni-Sepharose 6 Fast Flow resin, Cytiva), according to the manufacturer. Proteins were renatured on beads by washing with a linear gradient of 8-1 M urea in 100 mM sodium phosphate and 10 mM tris hydroxymethyl aminomethane (Tris), pH 8.0. Proteins were eluted by boiling the beads in sterile PBS supplemented with 0.5% SDS. Purified proteins were analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) (purity >95%) and sequence validated by mass spectrometry before using them for immunization and antibody production.

Hybridoma Production and Screen for Specific Antibodies. BALB/C mice were immunized three times at 3-wk intervals using 1 µg recombinant peptide. For hybridoma generation, the spleen cells were fused with Sp2/0 myeloma cells using polyethylene glycol (PEG 1450). Hybridomas were selected and maintained by Kohler and Milstein's hybridoma technology (1976). Screening was done by standard enzyme-linked immunosorbent assay (ELISA), and plates (Corning) were coated with 100 ng/mL recombinant proteins. Sheep anti-Mouse immunoglobulin G (IgG), Horseradish peroxidase-linked whole antibody (GE Healthcare, UK) (1:10,000), and o-Phenylenediamine (Sigma-Aldrich) were used for detection. The supernatants were further characterized by WB using recombinant proteins and larval extracts. Selected cell lines, 3G9 (anti-CdtB), 2B8 (anti-CdtB A/B), and 2H5 (anti-AIP56 A/B), were subcloned by limiting dilution, and their supernatants were used throughout.

Preparation of Tissue Extracts. Tissues, hemocytes, and parasitoids were isolated in Schneider's medium (Lonza) supplemented with 5% fetal bovine serum (Gibco) and 1 nM 1-phenyl 2-thiourea (Sigma) (CSM). The isolated parasitoids were washed three times in *Drosophila* Ringer solution, centrifuged for 20 s, 700 × g at 4 °C, and the pellets were harvested for protein analysis. The extracts were prepared in lysis buffer (50 mM Tris HCl, pH 8.0; 150 mM NaCl; 1.0% Nonidet P-40; 0.5% sodium deoxycholate; 0.1% SDS) supplemented with protease inhibitor mixture (Boehringer Mannheim) and 1 mM phenyl-methylsulfonyl fluoride, for immunoprecipitation and in sample buffer (250 mM Tris pH = 6.8, 35% glycerol, 0.75 mg/mL Bromophenol blue, 9.2% SDS) for WBs, using a homogenizer, and incubated for 1 h on ice. For WB, samples were boiled 5 min and centrifuged at 18,000 × g for 5 min. Protein concentrations were determined by Amido Black assay. To generate reduced conditions, 5% 2-Mercaptoethanol (2-ME) was added.

WB Analysis and Immunoprecipitation. Crude protein extracts were run on 10% or 12% SDS PAGE, blotted to polyvinylidene difluoride (PVDF) membrane (Millipore), blocked with 5% nonfat milk in Tris-buffered saline (TBS) (10 mM Tris pH 7.5, 150 mM NaCl), and incubated with primary antibodies (polyclonal serum 1:10,000 dilution, hybridoma supernatant 1:1 dilution in RPMI Medium1640 (gibco) containing 5% fetal bovine serum (FBS)) for 1 h. Membranes were washed three times (10 min each) with TBS containing 0.1% Tween 20, incubated with Polyclonal Goat Anti-Mouse Immunoglobulins/horseradish peroxidase (HRP) (Dako) (1:10,000 diluted in TBS containing 0.1% Tween 20 and 1% bovine serum albumin (BSA)), and washed three times. Reactions were visualized with Immobilon Western Chemiluminescent HRP Substrate (Merck).

Immunoprecipitation was carried out with the 2H5 antibody, cross-linked with dimethyl pimelimidate dihydrochloride (Sigma) to Protein G Sepharose beads (Cytiva). The samples were run on 10% SDS-PAGE under reducing conditions; the silver-stained bands were cut out and subjected to mass spectrometric identification (LC-MS/MS).

Identification of CdtB:AIP56 and AIP56 Using LC-MS/MS. The proteins with an apparent ~70 kDa molecular weight showing strong reactivity with the anti-AIP56 antibodies were subjected to tryptic in-gel digestion and the resulting peptide mixtures were analyzed using a Waters MClass nHPLC-Thermo Orbitrap Elite LC-MS system in a data-dependent fashion (49). Proteins were identified using the Protein Prospector BatchTag Web software using the *Drosophila* subset of the UniProt database supplemented with the CdtB:AIP56 sequences, applying score-based acceptance criteria. The CdtB:AIP56 fusion proteins were identified with multiple peptides affording 33% and 17% sequence coverage for CdtB::AIP56 B and CdtB::AIP56 A, respectively (*SI Appendix,* Table S3).

IIF Analysis. Larvae were dissected in CSM, fat bodies and parasitoids were removed and fixed with 2% paraformaldehyde for 10 min, washed three times in PBS (5 min each), and blocked with 0.1% BSA in PBS supplemented with 0.01% Triton X-100. Hemocytes were adhered for 1 h on microscope slides in CSM, fixed with acetone for 6 min, air dried, and blocked with 0.1% BSA in PBS for 20 min. Samples were incubated with the primary antibodies for 2 h, washed three times in PBS, incubated with the Alexa Fluor 488 goat anti-mouse IgG secondary antibody (Invitrogen, 1:1,000), containing DAPI (2.5 μg/mL) for 45 min, washed three times in PBS, mounted in Fluoromount G medium, and analyzed with an epifluorescence microscope (Zeiss Axioscope 2 MOT) or with an Olympus FV1000 confocal laser scanning microscope. Embryonic samples were prepared as previously described (50).

Generation of *D. ananassae* Mutant Lines. We generated five backcrossed *D. ananassae* mutant lines with CRISPR/Cas9-mediated mutagenesis via nonhomologous end joining in a wild-type (14024-0371.13) background. The *ΔcdtB* line carries a 1,000 bp deletion in *cdtB*; the Δ*fusA* line has a 5 bp deletion in *cdtB::aip56 A*; the *ΔcdtB_fusA* line has a 1 bp frameshift deletion at *cdtB::aip56 A* in a *ΔcdtB* background; the Δ*fusAB* line encodes a 7 bp frameshift deletion in *cdtB::aip56 A,* and a 400 bp deletion in *cdtB::aip56 B*; and the *ΔcdtB_fusAB* line has a 400 bp deletion in *cdtB::aip56 A*, a 399 bp deletion in *cdtB::aip56 B*, in a *ΔcdtB* background (*SI Appendix,* Fig. S11). Mutant lines were verified with bidirectional Sanger sequencing, using primers listed in *SI Appendix,* Table S5. Sequences of altered genes can be found in GenBank accession numbers OQ302761-OQ302767. For more information on the generation of mutant lines, please see *SI Appendix, Supplementary Methods*.

Parasitization Assay. Seventy, second instar *D. ananassae* larvae were exposed to 17 female parasitoid wasps for 6 h. Forty-eight hours after the attack period, 10 *D. ananassae* larvae were dissected to assay for parasitoids. A vial was considered available for the parasitization assay when each tested larva carried at least one parasitoid (51). For each fly line, three to four (the latter for untreated controls) independent parasitization assays were carried out. The eclosed flies and wasps were counted for 15- and 30-d post-infection, respectively. Control uninfected flies were treated under identical conditions. After determining the data approximated a normal distribution using the Shapiro–Wilk test, we conducted ANOVA and post-hoc Tukey HSD tests (these results are shown in *SI Appendix,* Table S6).

Data, Materials, and Software Availability. Partial sequences of altered *D. ananassae* genes have been deposited in GenBank under accession numbers: (OQ302761 (52), OQ302762 (53), OQ302763 (54), OQ302764 (55), OQ302765 (56), OQ302766 (57), OQ302767 (58).

ACKNOWLEDGMENTS. The technical help of Olga Kovalcsik, Anita Balázs, Zsuzsánna Réthi Nagy, and Henrietta Kovács is acknowledged. Thanks to Dr. Todd Schlenke who provided parasitoid wasp specimens. We extend additional thanks to Jessica Aguilar, Susan Bernstein, Ted Chor, Dr. Doris Bachtrog, Dr. Kamalatkar Chatla, and Dr. Kevin Wei for animal husbandry support, and Dr. Nicolas Alexandre for bioinformatics assistance. K.I.V. was supported by the NSF Graduate Research Fellowship, the Philomathia Foundation Fellowship, and the NSF Postdoctoral Research Fellowship. G.C. was supported by the National Research, Development and Innovation Office (NKFI) K128762 grants from the Hungarian NSF, and the BO/00552/20/8 Bolyai János Research Scholarship from the Hungarian Academy of Sciences. Z.L. was supported by grants from the Hungarian Academy of Sciences [Bolyai János Research Scholarship (BO/00329/15) and Lendület Program Grant (LP2017-7/2017)]. L.B.M. was supported by the Doctoral School of Biology, University of Szeged. R.L.T. was supported by the NIH Genetic Dissection of Cells and Organisms Training grant (award no. 5T32GM132022-03) and the NSF Graduate Research Fellowship. The Hungarian Center of Excellence for Molecular Medicine has received funding from the European Union's Horizon 2020 Research and Innovation Program under grant agreement no. 739593. N.K.W. was supported by a grant from the National Institute of General Medical Sciences of the NIH (award no. R35GM119816). I.A. was supported by the National Research, Development and Innovation Office (NKFI) K135877 grant from the Hungarian NSF.

Author affiliations: [a]Department of Integrative Biology, University of California, Berkeley, CA 94720; [b]Innate Immunity Group, Institute of Genetics, Biological Research Centre, Eötvös Loránd Research Network, Szeged 6726, Hungary; [c]MTA SZBK Lendület Laboratory of Cell Cycle Regulation, Institute of Biochemistry, Biological Research Centre, Eötvös Loránd Research Network, Szeged 6726, Hungary; [d]Doctoral School of Biology, University of Szeged, Szeged 6720, Hungary; [e]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720; [f]Single Cell Omics Advanced Core Facility, Hungarian Centre of Excellence for Molecular Medicine, Szeged 6728, Hungary; [g]Laboratory of Proteomics Research, Biological Research Centre, Eötvös Loránd Research Network, Szeged 6726, Hungary; [h]Department of Biology, Stanford University, Palo Alto, CA 94305; [i]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720; and [j]Helen Wills Neuroscience Institute, University of California, Berkeley, CA 94720

1. N. A. Moran, T. Jarvik, Lateral transfer of genes from fungi underlies carotenoid production in aphids. *Science* **328**, 624–627 (2010).
2. R. Acuña *et al.*, Adaptive horizontal transfer of a bacterial gene to an invasive insect pest of coffee. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4197–4202 (2012).
3. N. Wybouw, Y. Pauchet, D. G. Heckel, T. Van Leeuwen, Horizontal gene transfer contributes to the evolution of arthropod herbivory. *Genome Biol. Evol.* **8**, 1785–1801 (2016).
4. F. Husnik, J. P. McCutcheon, Functional horizontal gene transfer from bacteria to eukaryotes. *Nat. Rev. Microbiol.* **16**, 67–79 (2018).
5. M. Touchon, J. A. Moura de Sousa, E. P. Rocha, Embracing the enemy: the diversification of microbial gene repertoires by phage-mediated horizontal gene transfer. *Curr. Opin. Microbiol.* **38**, 66–73 (2017).
6. H. Brüssow, C. Canchaya, W.-D. Hardt, Phages and the evolution of bacterial pathogens: From genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* **68**, 560–602, table of contents (2004).
7. K. M. Oliver, S. J. Perlman, Toxin-mediated protection against natural enemies by insect defensive symbionts. *Adv. In Insect Phys.* **58**, 277–316 (2020), 10.1016/bs.aiip.2020.03.005.
8. J. H. Massey, I. L. G. Newton, Diversity and function of arthropod endosymbiont toxins. *Trends Microbiol.* **30**, 185–198 (2022).
9. F. van der Wilk, A. M. Dullemans, M. Verbeek, J. F. J. M. van den Heuvel, Isolation and characterization of APSE-1, a bacteriophage infecting the secondary endosymbiont of *Acyrthosiphon pisum*. *Virology* **262**, 104–113 (1999).
10. K. M. Oliver, P. H. Degnan, M. S. Hunter, N. A. Moran, Bacteriophages encode factors required for protection in a symbiotic mutualism. *Science* **325**, 992–994 (2009).
11. B. M. Boyd, G. Chevignon, V. Patel, K. M. Oliver, M. R. Strand, Evolutionary genomics of APSE: a tailed phage that lysogenically converts the bacterium *Hamiltonella defensa* into a heritable protective symbiont of aphids. *Virol. J.* **18**, 219 (2021).
12. M. Lara-Tejero, J. Galan, CdtA, CdtB, and CdtC form a tripartite complex that is required for cytolethal distending toxin activity. *Infect. Immun.* **69**, 4358–4365 (2001).
13. M. Asakura *et al.*, An inducible lambdoid prophage encoding cytolethal distending toxin (Cdt-I) and a type III effector protein in enteropathogenic *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 14483–14488 (2007).
14. M. T. Silva, N. M. S. Dos Santos, A. do Vale, AIP56: A novel bacterial apoptogenic toxin. *Toxins* **2**, 905–918 (2010).
15. D. S. Silva *et al.*, The apoptogenic toxin AIP56 is a metalloprotease A-B toxin that cleaves NF-κb P65. *PLoS Pathog.* **9**, e1003128 (2013).
16. N. M. Hooper, Families of zinc metalloproteases. *FEBS Lett.* **354**, 1–6 (1994).
17. K. I. Verster *et al.*, Horizontal transfer of bacterial cytolethal distending toxin B genes to insects. *Mol. Biol. Evol.* **36**, 3–8 (2019).
18. K. I. Verster, R. L. Tarnopol, S. M. Akalu, N. K. Whiteman, Horizontal transfer of microbial toxin genes to gall midge genomes. *Genome Biol. Evol.* **13**, evab202 (2021).
19. J. E. Bubnell, C. K. S. Ulbing, P. Fernandez Begne, C. F. Aquadro, Functional divergence of the *bag-of-marbles* gene in the *Drosophila melanogaster* species group. *Mol. Biol. Evol.* **39**, msac137 (2022).
20. T. A. Schlenke, J. Morales, S. Govind, A. G. Clark, Contrasting infection strategies in generalist and specialist wasp parasitoids of *Drosophila melanogaster*. *PLoS Pathog.* **3**, 1486–1501 (2007).
21. R. Markus *et al.*, Multinucleated giant hemocytes are effector cells in cell-mediated immune responses of *Drosophila*. *J. Innate Immun.* **7**, 340–353 (2015).
22. J. P. Dudzic, S. Kondo, R. Ueda, C. M. Bergman, B. Lemaitre, *Drosophila* innate immunity: Regional and functional specialization of prophenoloxidases. *BMC Biol.* **13**, 81 (2015).
23. D. Hultmark, I. Andó, Hematopoietic plasticity mapped in *Drosophila* and other insects. *Elife* **11**, e78906 (2022).

24. M. T. Kimura, A. Suwito, What determines host acceptance and suitability in tropical Asian *Drosophila* parasitoids? *Environ. Entomol.* **43**, 123–130 (2014).
25. D. J. Leybourne, T. A. Valentine, J. I. B. Bos, A. J. Karley, A fitness cost resulting from *Hamiltonella defensa* infection is associated with altered probing and feeding behaviour in *Rhopalosiphum padi*. *J. Exp. Biol.* **223**, jeb207936 (2020).
26. D. A. Baltrus, Exploring the costs of horizontal gene transfer. *Trends Ecol. Evol.* **28**, 489–495 (2013).
27. Z. Ge, D. B. Schauer, J. G. Fox, In vivo virulence properties of bacterial cytolethal-distending toxin. *Cell. Microbiol.* **10**, 1599–1607 (2008).
28. E. H. Mezal, D. Bae, A. A. Khan, Detection and functionality of the CdtB, PltA, and PltB from *Salmonella enterica* serovar Javiana. *Pathog. Dis.* **72**, 95–103 (2014).
29. Y. Moran, D. Fredman, P. Szczesny, M. Grynberg, U. Technau, Recurrent horizontal transfer of bacterial toxin genes to eukaryotes. *Mol. Biol. Evol.* **29**, 2223–2230 (2012).
30. X. Gao *et al.*, Evolution of host adaptation in the *Salmonella* typhoid toxin. *Nat. Microbiol.* **2**, 1592–1599 (2017).
31. S. Pasek, J. L. Risler, P. Brézellec, Gene fusion/fission is a major contributor to evolution of multi-domain bacterial proteins. *Bioinformatics* **22**, 1418–1423 (2006).
32. M. R. Strand, Teratocytes and their functions in parasitoids. *Curr. Opin. Insect Sci.* **6**, 68–73 (2014).
33. K. Comber *et al.*, A dual role for the βPS integrin myospheroid in mediating *Drosophila* embryonic macrophage migration. *J. Cell Sci.* **126**, 3475–3484 (2013).
34. G. Huang *et al.*, The active subunit of the cytolethal distending toxin, CdtB, derived from both *Haemophilus ducreyi* and *Campylobacter jejuni* exhibits potent phosphatidylinositol-3,4,5-triphosphate phosphatase activity. *Front. Cell. Infect. Microbiol.* **11**, 664221 (2021).
35. B. Wertheim, Adaptations and counter-adaptations in *Drosophila* host-parasitoid interactions: Advances in the molecular mechanisms. *Curr. Opin. Insect Sci.* **51**, 100896 (2022).
36. L. Lambrechts, S. Fellous, J. C. Koella, Coevolutionary interactions between host and parasite genotypes. *Trends Parasitol.* **22**, 12–16 (2006).
37. B. Z. Kacsoh, Z. R. Lynch, N. T. Mortimer, T. A. Schlenke, Fruit flies medicate offspring after seeing parasites. *Science* **339**, 947–950 (2013).
38. J. Jaenike, R. Unckless, S. N. Cockburn, L. M. Boelio, S. J. Perlman, Adaptation via symbiosis: Recent spread of a *Drosophila* defensive symbiont. *Science* **329**, 212–215 (2010).
39. M. J. Ballinger, S. J. Perlman, Generality of toxins in defensive symbiosis: Ribosome-inactivating proteins and defense against parasitic wasps in *Drosophila*. *PLoS Pathog.* **13**, 1–19 (2017).
40. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
41. K. Katoh, K. Misawa, K.-I. Kuma, T. Miyata, MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
42. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
43. E. Paradis, K. Schliep, ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).
44. G. Yu, Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinfor.* **69**, e96 (2020).
45. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
46. M. D. Smith *et al.*, Less is more: An adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* **32**, 1342–1353 (2015).
47. B. Murrell *et al.*, FUBAR: A fast, unconstrained bayesian approximation for inferring selection. *Mol. Biol. Evol.* **30**, 1196–1205 (2013).
48. Z. Lipinszki *et al.*, Enhancing the translational capacity of *E. coli* by resolving the codon bias. *ACS. Synth. Biol.* **7**, 2656–2664 (2018).
49. Z. Rethi-Nagy *et al.*, STABILON, a novel sequence motif that enhances the expression and accumulation of intracellular and secreted proteins. *Int. J. Mol. Sci.* **23**, 8168 (2022).
50. G. Cinege *et al.*, Genes encoding cuticular proteins are components of the *Nimrod* gene cluster in *Drosophila*. *Insect Biochem. Mol. Biol.* **87**, 45–54 (2017).
51. B. Z. Kacsoh, J. Bozler, T. A. Schlenke, A role for nematocytes in the cellular immune response of the drosophilid *Zaprionus indianus*. *Parasitology* **141**, 697–715 (2014).
52. K. I. Verster *et al.*, *Drosophila ananassae* isolate DeltacdtB_cdtB sequence. NCBI GenBank. https://www.ncbi.nlm.nih.gov/nuccore/OQ302761. Deposited 12 March 2023.
53. K. I. Verster *et al.*, *Drosophila ananassae* isolate DeltacdtB_fusA_fusionA sequence. NCBI GenBank. https://www.ncbi.nlm.nih.gov/nuccore/OQ302762. Deposited 12 March 2023.
54. K. I. Verster *et al.*, *Drosophila ananassae* isolate DeltacdtB_fusAB_fusionA sequence. NCBI GenBank. https://www.ncbi.nlm.nih.gov/nuccore/OQ302763. Deposited 12 March 2023.
55. K. I. Verster *et al.*, *Drosophila ananassae* isolate DeltafusA_fusionA sequence. NCBI GenBank. https://www.ncbi.nlm.nih.gov/nuccore/OQ302764. Deposited 12 March 2023.
56. K. I. Verster *et al.*, *Drosophila ananassae* isolate DeltafusAB_fusionA sequence. NCBI GenBank. https://www.ncbi.nlm.nih.gov/nuccore/OQ302765. Deposited 12 March 2023.
57. K. I. Verster *et al.*, *Drosophila ananassae* isolate DeltafusAB_fusionB sequence. NCBI GenBank. https://www.ncbi.nlm.nih.gov/nuccore/OQ302766. Deposited 12 March 2023.
58. K. I. Verster *et al.*, *Drosophila ananassae* isolate DeltacdtB_fusAB_fusionB sequence. NCBI GenBank. https://www.ncbi.nlm.nih.gov/nuccore/OQ302767. Deposited 12 March 2023.