Trends in Microbiology

CellPress

Review

Diversity and function of arthropod endosymbiont toxins

Jonathan H. Massey[1] and Irene L.G. Newton[1,*]

Bacterial endosymbionts induce dramatic phenotypes in their arthropod hosts, including cytoplasmic incompatibility, feminization, parthenogenesis, male killing, parasitoid defense, and pathogen blocking. The molecular mechanisms underlying these effects remain largely unknown but recent evidence suggests that protein toxins secreted by the endosymbionts play a role. Here, we describe the diversity and function of endosymbiont proteins with homology to known bacterial toxins. We focus on maternally transmitted endosymbionts belonging to the *Wolbachia*, *Rickettsia*, *Arsenophonus*, *Hamiltonella*, *Spiroplasma*, and *Cardinium* genera because of their ability to induce the above phenotypes. We identify at least 16 distinct toxin families with diverse enzymatic activities, including AMPylases, nucleases, proteases, and glycosyltransferases. Notably, several annotated toxins contain domains with homology to eukaryotic proteins, suggesting that arthropod endosymbionts mimic host biochemistry to manipulate host physiology, similar to bacterial pathogens.

Highlights

Arthropod endosymbiont genomes contain multiple proteins with homology to bacterial protein toxins.

Among *Wolbachia*, *Rickettsia*, *Arsenophonus*, *Hamiltonella*, *Spiroplasma*, and *Cardinium* genera, we identify at least 16 distinct toxin families.

Several toxins contain multiple domains, share homology with eukaryotic proteins, and are encoded by bacteriophage genomes.

Toxins known to induce male killing, cytoplasmic incompatibility, or parasitoid defense target host DNA or RNA.

Arthropod endosymbionts might induce many of their host phenotypes via toxin secretion.

Endosymbiont diversity

Arthropods are the largest and most diverse animal phylum on Earth [1]. As such, they are hosts to diverse groups of bacterial endosymbionts that have evolved multiple strategies to infect and persist inside their intracellular environments. Maternally transmitted endosymbionts belonging to the genera ***Wolbachia***, ***Rickettsia***, ***Arsenophonus***, ***Hamiltonella***, ***Spiroplasma***, and ***Cardinium*** (see Glossary) are of particular interest because of their ability to induce dramatic reproductive or defensive phenotypes in their arthropod hosts (Figure 1). These phenotypes include cytoplasmic incompatibility (CI), feminization, male killing, parthenogenesis, parasitoid defense, and pathogen blocking (Box 1). For decades, biologists have wanted to understand how endosymbionts cause these effects with the hope of applying these strategies to combat vector-borne diseases and agricultural pests [2]. But, to date, few case studies have identified a mechanism [3–5]. In each, proteins resembling toxins encoded by the endosymbiont genomes were shown to recapitulate a reproductive or defensive phenotype when expressed in an endosymbiont-free host, generating new interests in endosymbiont toxins for their ability to manipulate arthropod biology. Here, we explore what endosymbiont toxins are, categorize their diversity and function, and predict their potential mode of action in arthropod hosts.

Endosymbiont toxin families

Bacterial genomes encode an extremely rich variety of protein toxins that function during infection and competition [6]. These toxins are best studied in pathogen systems in which the microbes rely on such 'virulence factors' to invade their eukaryotic hosts, manipulate signaling pathways, silence immune responses, and regulate gene expression – often leading to disease symptoms [7]. This arms race between host and microbe is a breeding ground for toxin evolution and has led to some of the most potent toxins on Earth. Recent whole-genome sequencing projects have also uncovered homologous toxin systems in arthropod endosymbionts [8]. These findings

[1]Department of Biology, Indiana University, Bloomington, IN, USA

*Correspondence: irnewton@indiana.edu (I.L.G. Newton).

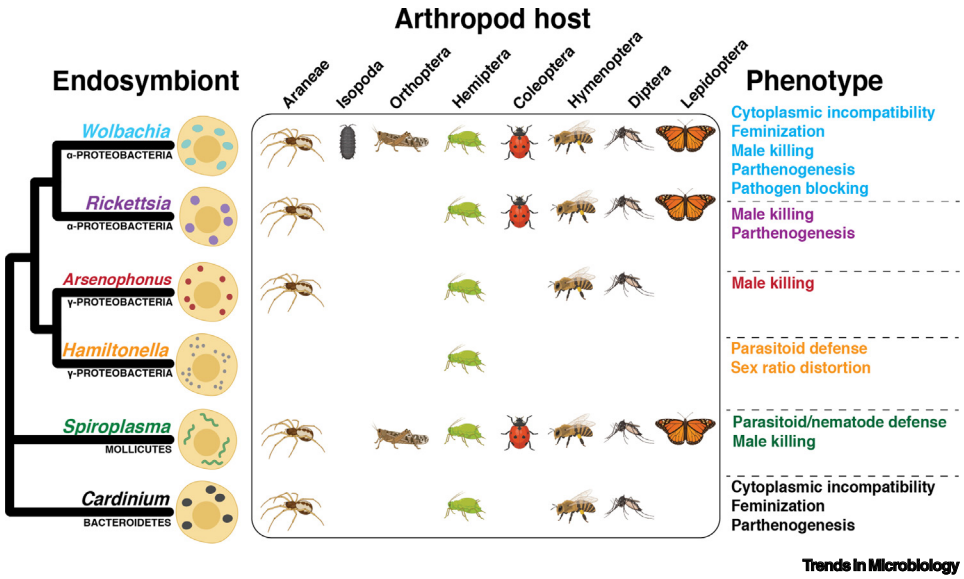

Figure 1. Diversity of arthropod endosymbiont infections and their consequences. Maternally transmitted endosymbionts belonging to the genera *Wolbachia*, *Rickettsia*, *Arsenophonus*, *Hamiltonella*, *Spiroplasma*, and *Cardinium* infect a diverse set of arthropod orders. The cladogram depicting relatedness among endosymbionts is based on alignment of 16S rRNA sequences described in [97]. Images of arthropods are shown for each endosymbiont where the presence of species belonging to these endosymbiont genera was confirmed by PCR [98]. Phenotypes in blue, purple, red, orange, green, and black summarize reproductive or resistance traits induced by *Wolbachia*, *Rickettsia*, *Arsenophonus*, *Hamiltonella*, *Spiroplasma*, and *Cardinium* species, respectively [84,88,96,99,100]. However, there is considerable phenotypic variation among different endosymbiont–host associations within a particular group. For example, not all *Wolbachia* species or strains induce cytoplasmic incompatibility in each arthropod order depicted, and different *Wolbachia* strains infecting the same host can produce different phenotypes.

have led to the hypothesis that, like their pathogenic cousins, endosymbionts also rely on protein toxin systems to infect their eukaryotic hosts and manipulate host physiology.

Here, we identify and characterize the diversity of putative protein toxins among six arthropod endosymbiont genera (*Wolbachia*, *Rickettsia*, *Arsenophonus*, *Hamiltonella*, *Spiroplasma*, and *Cardinium*). We focus on these six because of their (i) established impact on arthropod reproductive/defensive phenotypes, (ii) overlap in arthropod host orders, and (iii) available genome sequences (Figure 1 and Box 1). Variation among these six genera also encompasses a wide range of divergence times, between 1 and 3 billion years, and includes endosymbionts infecting most arthropod orders (e.g., *Wolbachia*) or only one (e.g., *Hamiltonella*) [9]. We focus on proteins containing significant homology to known toxins from other model systems, although their specific toxic phenotype in the arthropod may not yet be known. We find at least 16 toxin families (Table 1 and Figure 2), with some previously annotated and some not. Our survey is far from exhaustive; instead, we focused on these 16 toxin families because of their potential to impact arthropod biology. Some of the identified endosymbiont proteins are **polymorphic toxins**, and comprise multiple domains, and may diversify via recombination within or between species. Most of the endosymbiont toxins are highly conserved with nonendosymbiotic bacteria, and no toxins are ubiquitously present among all endosymbiont genera (Table 1 and Figure 2). We also find that certain endosymbiont genera possess more toxins than others; however, some of these results are likely due to ascertainment bias (e.g., there are many more sequenced *Wolbachia* genomes than *Cardinium* genomes). The toxin

Glossary

***Arsenophonus*:** a genus of opportunistic, facultatively intracellular gammaproteobacteria related to the other insect-associated enterics such as *Regiella* and *Hamiltonella*. The genus was first described with *A. nasoniae,* or the 'son killer' of the parasitic wasp *Nasonia vitripennis* due to the male-killing phenotype. Subsequently, related sequences and microbes have been discovered in other arthropods, including bees and ants. The model system for this genus, *A. nasoniae,* is maternally transmitted and colonizes the female venom/calyx fluid, facilitating its transfer during oviposition. Unlike other insect endosymbionts, *Arsenophonus* can be cultured in axenic media [80,81].

Bacteriophage APSE: a temperate phage, found in *Hamiltonella defensa*, the lysogeny of which is correlated with protection of the *Hamiltonella* insect host against parasitoid wasps. As with bacteriophage WO, toxins encoded by the APSE genome are implicated in this protective phenotype [65,70].

Bacteriophage WO: a temperate phage, found in *Wolbachia*, for which the encapsidated genome encodes many putative toxins referred to as the 'eukaryotic association module'. In the most well studied systems, the phage also encodes the *Cif* loci implicated in the reproductive manipulation by *Wolbachia* termed cytoplasmic incompatibility [37].

***Cardinium*:** bacteria of this genus (in the phylum Bacteroidetes) are known to infect a diversity of arthropods, from spider mites to parasitic wasps, and induce a swathe of reproductive manipulations in arthropods similar to those produced by *Wolbachia,* including feminization, parthenogenesis, and cytoplasmic incompatibility [82].

***Hamiltonella*:** discovered for its defensive phenotype in aphids, this gammaproteobacterial endosymbiont is primarily maternally transmitted and is found in the ovaries, somatic tissues (including bacteriocytes), and hemolymph of its aphid and whitefly hosts. It is a facultative symbiont, however, and is found sporadically across all aphid lineages and populations [83,84].

Horizontal gene transfer (HGT): movement of genetic material between cellular lineages, often facilitated by mobile elements such as plasmids and phages.

Box 1. Phenotypic consequences of arthropod endosymbiont infections

The endosymbionts highlighted in this review are infamous for causing specific reproductive or parasite resistance phenotypes in their hosts. All of the modifications benefit infected females, and, as these symbionts are maternally transmitted, the manipulations facilitate their spread in host populations (Figure I). These modifications include the following. (A) Cytoplasmic incompatibility (CI), whereby sperm from infected males is modified in such a way that it can only be rescued by a compatible infection(s) in the fertilized egg. (B) Feminization, in which all the fertilized offspring from infected mothers develop as female, even if chromosomally male. (C) Male killing, which results in a reduction in total brood size and a skew towards female offspring from infected mothers. (D) Parthenogenesis, which allows infected females to produce all-female offspring without fertilization. (E) Parasitoid defense, whereby arthropod hosts infected with certain endosymbiont strains are less likely to succumb to parasitoid wasp infections. (F) Pathogen blocking, whereby arthropod hosts infected with certain *Wolbachia* strains are less likely to succumb to RNA virus infection. (A–F) are reviewed elsewhere [35,84,88,99,100].

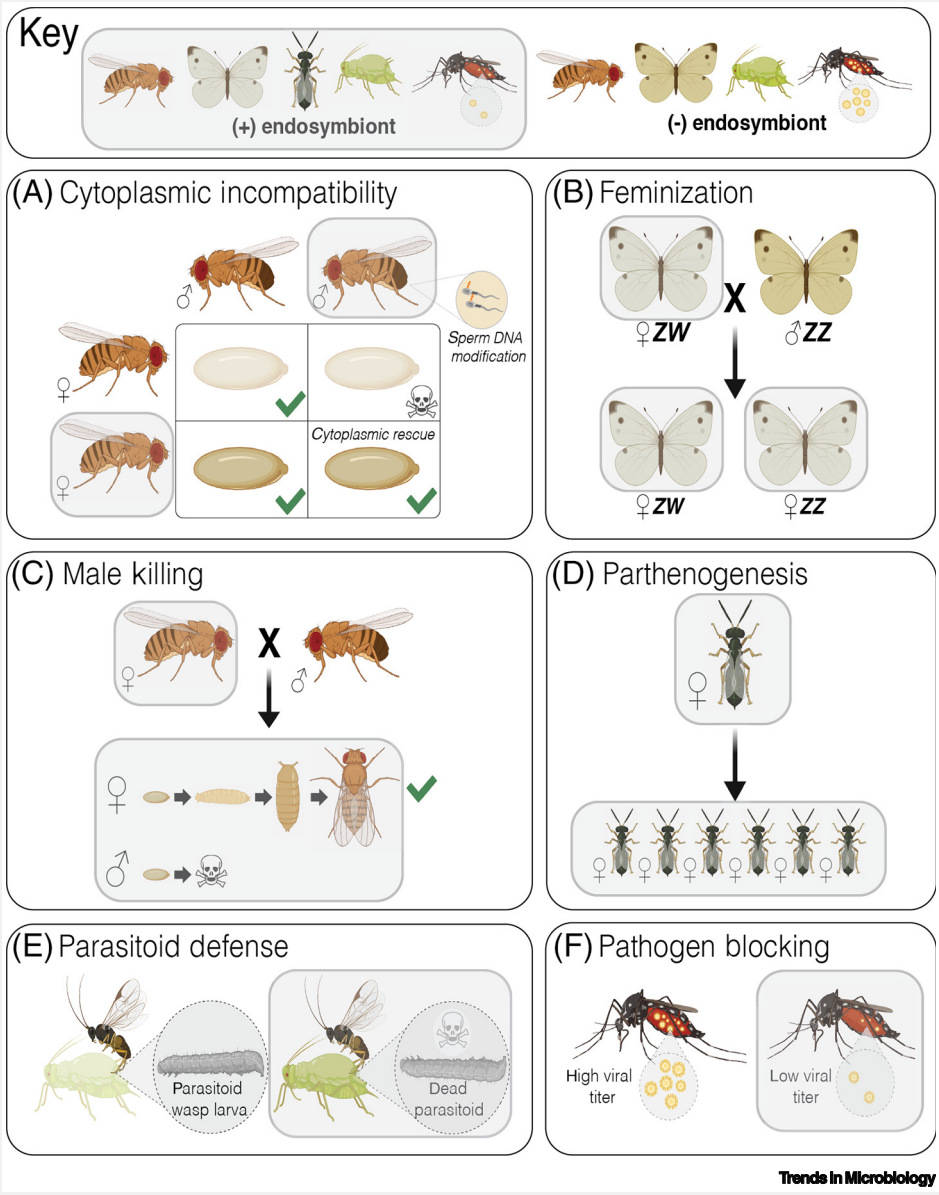

Figure I. Phenotypes induced by arthropod endosymbionts range from reproductive manipulations (A–D) to defense against parasites and pathogens (E,F).

Polymorphic toxins: multidomain effector proteins secreted by diverse secretion systems into host cells by infecting bacteria. These toxins are characterized by their diversification through recombination and domain swapping, and by their association with horizontally transferred elements [85].

***Rickettsia*:** although this genus of alphaproteobacteria is primarily known for including human pathogens, such as the causative agents of Rocky Mountain spotted fever and typhus [86], some of these bacteria may be restricted to arthropods [87]. *Rickettsia* is obligately intracellular and is often found in association with diverse arthropods such as flies, wasps, mites, beetles, and others. Like *Wolbachia, Rickettsia* often colonizes the reproductive tract of its hosts where it induces phenotypes such as male killing and parthenogenesis [88].

***Spiroplasma*:** a genus of helical mollicutes that form associations with many arthropod lineages and some plants. They are primarily extracellular in their hosts and are often found in the digestive tracts of insects. Certain strains, such as *Spiroplasma poulsonii,* grow abundantly in the hemolymph where they secrete ribosome-inactivating proteins (RIPs) that protect the host from parasitoid wasps. That same *Spiroplasma* strain is also infamous for inducing male killing in its insect hosts [89]. Some *Spiroplasma* symbioses protect the arthropod hosts from fungal or nematode infection, although it is not clear if the mechanism is conserved. *Spiroplasma* can be maternally transmitted in some hosts, in which the microbe colonizes the oocyte late during oogenesis [35,90].

Type II toxin–antitoxin (TA) systems: first discovered for their association with mobile genetic elements, these systems are comprised of a single toxin and cognate antitoxin. Classically, the TA systems promote the inheritance of selfish genetic elements as the toxin protein is more stable than the antitoxin, linking the production of the antitoxin to the presence of the genetic entity (postsegregational killing). TA systems are also involved in the bacterial stress response (or persister formation) and resistance to phage infection [18,19].

***Wolbachia*:** a genus of ubiquitous, alphaproteobacterial bacteria found to infect the cells of arthropods (40–60% of insect species) and nematodes [2]. *Wolbachia* generally colonizes the reproductive tract of its hosts (but is also

families compose a diversity of structure–function relationships (reviewed below), including adenylyltransferases (FIC toxins), nucleases (type II toxin–antitoxin systems), pore-forming toxins (TcA/TcB–TcC, ETX/MTX2, and latrotoxin-CTD), deubiquitinases (Spaid-like and CifB), ankyrin-repeats (Spaid-like and latrotoxin-CTD), and rRNA glycosylases (RIPs) (Table 1 and Figures 2 and 3).

found in some somatic tissues) and is primarily maternally transmitted. *Wolbachia* is known for its ability to manipulate arthropod reproduction in a variety of ways (Figure 1) and for limiting RNA virus replication in the host [91].

FIC family toxins

FIC (filamentation induced by cAMP) proteins were first discovered in *Escherichia coli* for their proposed function in cell division [10]. Under high cAMP and temperature conditions, loss-of-function *fic* mutants showed significantly increased filamentation phenotypes relative to controls. Subsequently, the first crystal structure of an FIC protein was solved in 2006 from *Helicobacter pylori* (Protein Data Bank ID 2F6S), and FIC domains were soon after shown to have adenylyltransferase enzymatic activity [11,12]. The conserved His within the FIC motif HXFX(D/E)(G/A)N(G/K)RXXR is necessary for the catalytic modification of protein targets by addition of adenosine monophosphate groups (AMPylation or adenylylation) [13]. FIC proteins are found across many different bacterial clades, and some FIC domain-containing proteins are secreted by pathogens into host cells, where they modify host target proteins [14]. For example, in *Vibrio parahaemolyticus* and *Histophilus somni*, FIC proteins induce cytotoxicity upon expression in eukaryotic host cells by catalyzing AMPylation of host target proteins, inducing a dramatic collapse of the actin cytoskeleton [11,12]. We found FIC domain-containing proteins to be the most prevalent protein toxin family among arthropod endosymbiont genera (Table 1 and Figure 2), with some *Wolbachia* genomes containing up to three copies of FIC enzymes [15]. The influence of host actin modification on *Wolbachia* titer [16,17] leads us to speculate that arthropod endosymbionts might rely on actin-modifying toxins like FICs to manipulate the host actin environment for proper invasion and/or transmission.

Type II toxin–antitoxin (TA) systems

Type II TA systems were originally discovered for their function in 'addiction' plasmids that drive the maintenance of toxin-containing mobile elements in bacteria, but they have also been found on bacterial chromosomes [18], where they may regulate bacterial dormancy or inhibit phage propagation by slowing cell growth [19]. These TA systems are highly diverse and prevalent across both the bacterial and archaeal domains of life [20]. Multiple type II TA systems are present among the endosymbiont genomes (Figure 2). Type II toxins are defined by their direct interaction with a cognate antitoxin that mediates inhibition. Many of the type II TA systems we identified belong to toxin families that show endoribonuclease activities, which disrupt protein translation by binding to ribosomal subunits and cleaving rRNA. These include RelE, YoeB, BrnT, ChpB, and VapC toxin families (Table 1 and Figure 2). On the *Wolbachia pipientis* strain *w*Mel chromosome, for example, multiple copies of RelE/RelB TA genes (WD_0122–WD_0127) are encoded in an array. Overexpression of a type II TA protein toxin (e.g., RelE) in the absence of its antitoxin partner (e.g., RelB), results in cell death, which likely explains why type II TA systems are typically encoded as a single operon containing both the toxin and antitoxin genes [21]. However, the cellular roles of TA systems, at physiologically relevant expression levels, is likely more complex [22]. We found that members of the RelE/ParE type II TA toxin superfamily were always accompanied by their partner antitoxins RelB/ParD in *Wolbachia*, *Rickettsia*, *Arsenophonus*, and *Hamiltonella* genomes, but we did not always identify antitoxin partners for RatA/RatB, BrnT/BrnA, ChpB/ChpS, or VapC/VapB, suggesting either incomplete genome assemblies or the presence of some other mechanism underlying antitoxin activity in these species. The function of type II TA systems in arthropod endosymbiont biology remains unknown, but it is tempting to speculate that these TA systems play some role in modulating bacterial growth during host association [23] or altering phage propagation between strains coinfecting the same host. Alternatively, it is

Table 1. Arthropod endosymbiont protein toxins

Toxin	Function	Present in species of	Example accession ID	Structural homologs [92,93] >80% probability [94]	Amino acid length (domain coordinates) [93]	Mobile element-associated
FIC family	Catalyze the transfer of AMP moieties to host target proteins, disrupting their activity	*Wolbachia*	WP_010082532.1	FIC (3EQX)	360 (18–358)	(–)
		Rickettsia	WP_202069843.1	FIC (3EQX)	358 (18–357)	?
		Hamiltonella	WP_072239752.1	FIC (3EQX)	176 (1–175)	?
		Spiroplasma	WP_094049572.1	FIC (3EQX)	361 (16–360)	?
		Cardinium	WP_144086617.1	FIC (3EQX)	374 (14–372)	?
RelE/ParE	Type II TA endoribonuclease/ DNA-gyrase inhibitor	*Wolbachia*	WP_047758703.1	RelE (2KHE)	90 (1–86)	(–)
		Rickettsia	WP_041077629.1	RelE (2KHE)	87 (1–85)	?
		Arsenophonus	WP_026823134.1	RelE (2KHE)	88 (1–88)	?
		Hamiltonella	WP_174889601.1	RelE (2KHE)	88 (1–88)	?
RatA	Type II TA inhibits translation by blocking 50S association with 30S subunits	*Wolbachia*	WP_182309761.1	Polyketide cyclase (3TL1)	153 (1–150)	?
		Rickettsia	WP_012150424.1	Polyketide cyclase (3TL1)	146 (145)	?
		Arsenophonus	WP_071846843.1	Polyketide cyclase (3TL1)	146 (1–144)	?
		Hamiltonella	WP_095034427.1	Polyketide cyclase (3TL1)	144 (1–143)	?
TcB–TcC	Component of the Tc holotoxin complex that folds into a cocoon-like structure before translocating into host cells	*Wolbachia*	WP_010962649.1	TcdB2, TccC3 (6SUF)	2843 (1765–2223)	(+) WO-Island [37]
		Arsenophonus	WP_180558610.1	TcdB2, TccC3 (4O9X)	1486 (2–1485)	?
		Hamiltonella	ACJ10121.1	TcdB2, TccC3 (6SUP)	1683 (652–1375)	APSE [70]
TcA	Component of the Tc holotoxin complex, which forms a pentameric pore that translocates TcB–TcC toxins in the host cells	*Wolbachia*	WP_136132830.1	TcdA1 (6RW6)	1136 (198–1013)	?
		Arsenophonus	WP_180558609.1	TcdA1 (6RW6)	2458 (2–2456)	?
		Hamiltonella	WP_174889479.1	TcdA1 (6RW6)	2521 (7–2520)	?
BrnT	Type II TA endoribonuclease	*Rickettsia*	WP_041078225.1	BrnT (3U97)	99 (1–90)	?
		Hamiltonella	WP_015873100.1	BrnT (3U97)	92 (1–83)	?
Spaid-like	Male killing toxin thought to interact with the male-specific lethal complex in *Drosophila*	*Wolbachia*	WP_010962723.1	DARPin (6MOK)	966 (142–449)	(+) WOMelB [37]
				OTU (6W9R)	966 (754–943)	
		Spiroplasma	WP_105629072.1	DARPin (6MOK)	1065 (33–331)	Plasmid pSMSRO [95]
				OTU (6W9R)CdtB	1065 (449–503)	
CdtB	Active subunit of cytolethal distending toxin complex showing DNase activity	*Arsenophonus*	WP_129109532.1	(1SR4)	306 (39–306)	?
		Hamiltonella	WP_100096556.1	CdtB (1SR4)	329 (49–329)	APSE [70]
Glycosylating toxin TcdA	Glycosylates members of the Ras superfamily of small GTPases, disrupting host cell signaling pathways	*Spiroplasma*	WP_047791727.1	Glucosyltransferase (2VK9)	674 (1–496)	?
		Cardinium	WP_184891138.1	Glucosyltransferase (5UQM)	766 (35–581)	?
RIPs	Inhibit eukaryotic protein synthesis by binding to 28S ribosomal RNA and cleaving an adenine in the sarcin–ricin loop structure	*Hamiltonella*	NP_050968.1	Shiga-like toxin B (1C4Q)	102 (28–100)	APSE [70]
		Spiroplasma	WP_040094559.1	Agglutinin-1 RIP (2ZR1)	448 (208–413)	?
ChpB	Type II TA endoribonuclease	*Hamiltonella*	WP_015873809.1	Kid toxin protein (1M1F)	113 (4–113)	?

(continued on next page)

Table 1. (continued)

Toxin	Function	Present in species of	Example accession ID	Structural homologs [92,93] >80% probability [94]	Amino acid length (domain coordinates) [93]	Mobile element-associated
CifA/CifB	CifB toxins show deubiquitylase and nuclease activity, whereas CifA function is unknown	*Wolbachia*	WP_012481787.1/ WP_010962722.1	No clear homology/ nuclease (PF08011.12), Ulp protease (6UPS)	NA 1166 (266–378; 599–733) 1166 (875–1928)	(+) WOMelB [37] (+) WOMelB [37]
ETX/MTX2	β-pore-forming toxins with receptor-binding activity that facilitates *C. perfringens* infections in livestock	*Spiroplasma*	WP_047791882.1	Epsilon toxin (3ZJX)	292 (91–276)	Plasmid pSneo [96]
Latrotoxin-CTD	Pore-forming toxins that cause paralysis in black widow spider prey	*Wolbachia*	WP_182365036.1	OTU (6W9R) Nuclease (PF08011.12) Ankyrin-repeat (5Y4D) LatrotoxinCTD (PF15658.7)	2833 (52–189) 2833 (464–593; 849–986) 2833 (2066–2550) 2833 (2654–2833)	(+) WO-Island [37]
VapC	Type II TA endoribonuclease	*Rickettsia*	WP_008580287.1	Ribonuclease VapC (6NKL)	132 (3–130)	?
Botulinum toxin	Virulence factor that causes botulism by blocking nerve signaling through cleavage of SNARE proteins	*Arsenophonus*	WP_026823090.1	*Clostridium* P47 protein (6EKT)	425 (3–425)	?

possible that some arthropod endosymbionts secrete type II TA toxins to manipulate their hosts, as expression of the toxin in a eukaryotic context is toxic to the cell [24].

Pore-forming toxins

Pore-forming toxins, which are the largest class of bacterial toxins, are primarily used as virulence factors by pathogens during infection [25]. Pore-forming toxins create holes in target cell membranes during pathogen or predator attack, allowing new virulence factors to enter host cells during infection [26,27]. We confirm previous evidence that the genes encoding pore-forming toxins are present among multiple arthropod endosymbiont species' genomes (Figures 2 and 3). We identified genes encoding components of the insecticidal Tc toxin complex in *Wolbachia*, *Arsenophonus*, and *Hamiltonella* genomes, ETX/MTX2 in *Spiroplasma* genomes, and latrotoxin in *Wolbachia* genomes (Table 1 and Figure 2).

One example of pore-forming toxins, found primarily in pathogenic microbes [28], is the Tc toxin complex. These toxins were first studied in insect entomopathogens, such as *Photorhabdus luminescens*, where secretion of the Tc toxin complex in the host gut during infection creates large cavities in the epithelium that lead to host death [29]. Recent crystal structures from the *P. luminescens* Tc complex illustrate how the holotoxin (one of the largest known) forms and functions [30,31]. The pore apparatus is made up of a pentameric ring of TcA proteins that open and close under different pH environments. The TcA ring functions as a translocation channel for TcB–TcC fusion proteins during infection. Before translocation, TcB–TcC folds into a cocoon structure, encapsulating its toxic residue via rearrangement hotspot (Rhs)-repeat motifs on the C-terminal end (Figures 2 and 3). Together, the Tc holotoxin (TcA + TcB–TcC) interacts with host cell membranes to deliver toxins into the host cytosol [30,31]. The impact of the Tc toxin complex on *Wolbachia*, *Arsenophonus*, or *Hamiltonella* biology remains unclear. One clue, however, stems from a TcB–TcC (WD_0513) homolog that resides on the **bacteriophage**

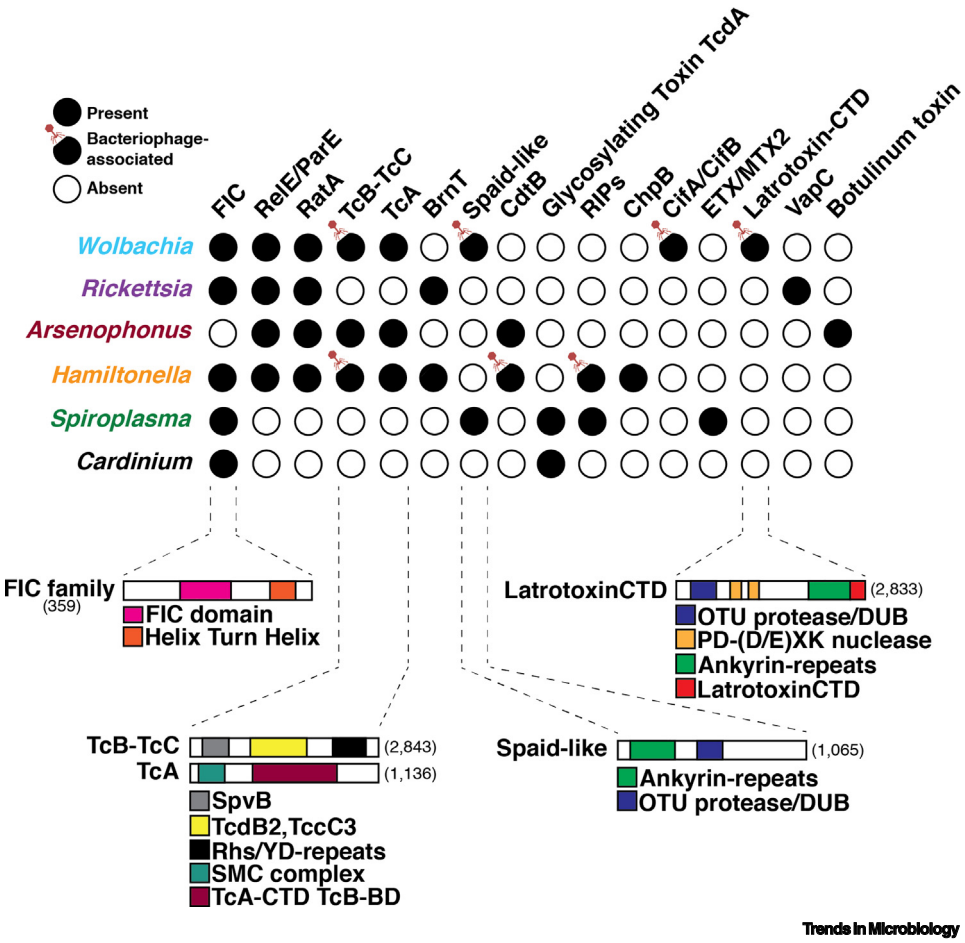

Figure 2. Diversity of protein toxin families identified in each endosymbiont genus. To identify protein toxins in each endosymbiont genus we used protein homolog queries (BLASTp; E < 1e-4) from known toxins in pathogenic bacteria as well as known homologs from toxins that cause specific toxic phenotypes in certain arthropod–endosymbiont interactions. HHPred was then used to identify domain architectures. Accession IDs for each protein toxin are referenced in Table 1. Black-filled circles denote the presence of the toxin in at least one endosymbiont species for each genus. Red bacteriophage icons denote toxins that are located within bacteriophage regions in the endosymbiont genome (Table 1). Open circles denote the absence of the toxin according to multiple failed BLASTp searches using protein queries of toxin homologs. Protein domain arrangements are shown below certain toxins to illustrate multidomain architectures (e.g., polymorphic toxins).

WO island in some *W. pipientis* genomes (Table 1). Increased copy number of this gene, along with seven others, is associated with a higher *W. pipientis* titer [32], suggesting that components of pore-forming toxins might facilitate arthropod endosymbiont infections.

The pore-forming epsilon (ETX)/mosquitocidal (MTX2) toxins were discovered in *Clostridium perfringens* as virulence factors underlying the rapidly fatal disease enterotoxemia in sheep, goat, and cattle [33]. Secretion of ETX/MTX2 by *C. perfringens* causes edema in major organs, eventually leading to severe neurological symptoms and death [34]. ETX/MTX2 is thought to be secreted initially as a prototoxin. It then undergoes proteolytic cleavage and oligomerization before forming a mature heptameric complex that inserts into the host cell membrane [33]. Multiple *Spiroplasma* species' genomes encode ETX/MTX2-like toxin homologs of varying lengths [35,36].

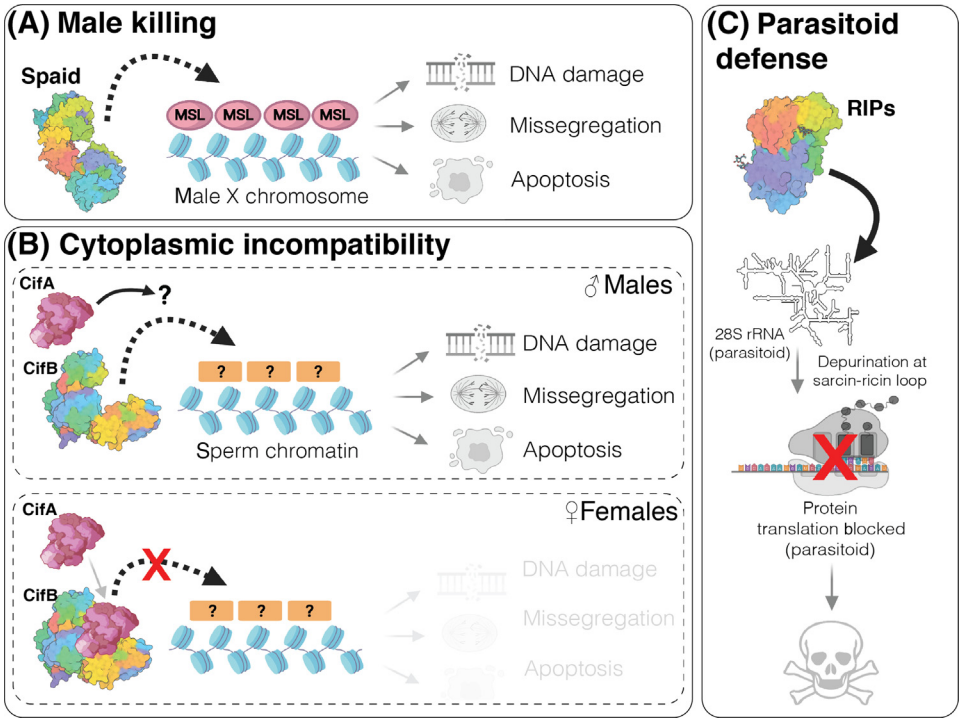

Figure 3. Models illustrating hypothesized modes of action for toxins capable of inducing arthropod phenotypes. For each panel, broken arrows denote direct or indirect interactions, where current evidence is insufficient to distinguish. Unbroken arrows denote direct interactions. (A) Spaid (*Spiroplasma poulsonii* androcidin) overexpression in *S. poulsonii*-free *Drosophila melanogaster* hosts induces potent male killing [4]. Previous work found that genes encoding proteins in the male-specific lethal (MSL) dosage compensation complex in *D. melanogaster* are required, in part, for *S. poulsonii*-induced male killing [101]. Spaid colocalizes with MSL1 in *D. melanogaster* cells, and ectopic expression of MSL2 with Spaid in females is sufficient to induce cell death [4]. Spaid thus appears to cause male killing by interfering with dosage compensation on the male X chromosome, leading to apparent DNA damage, chromatin bridge formation, and mis-segregation during cell division – eventually causing death. (B) In *D. melanogaster* hosts, coexpression of (cytoplasmic incompatibility factor) *cifA* and *cifB* genes in males induces cytoplasmic incompatibility (CI) in embryos from matings with *Wolbachia*-free females [3,5,46,47]. Females infected with *Wolbachia* or overexpressing *cifA* can rescue embryos from CI, and biochemical assays show that CifA binds directly with CifB. CifA and CifB coexpression in male sperm appears to recapitulate *Wolbachia*-induced CI effects, causing chromatin bridge formation, mis-segregation, and lethality in embryos from matings with CifA-free mothers. (C) *D. melanogaster* and *Drosophila neotestacea* infected with *Spiroplasma* that expresses ribosome-inactivating toxins (RIPs) are resistant to wasp parasitoid attacks. This appears to be caused in part by RIP-mediated cleavage of an adenosine residue in the sarcin–ricin loop (depurination) of wasp 28S rRNA [61]. Cleavage at this residue likely inhibits protein translation (causing death) in wasp larvae infecting the fly host.

In *Spiroplasma poulsonii* strain sHy, that infects *Drosophila hydei*, an ETX/MTX2-like homolog also contains an OTU (ovarian tumor) deubiquitinase domain [36], whereas we identified smaller versions of ETX/MTX2-like homologs in *Spiroplasma eriocheiris* and *Spiroplasma ixodetis*. The function of ETX/MTX2-like toxins in *Spiroplasma*–arthropod interactions remains unknown.

Latrotoxins are pore-forming proteins that are also present in the genomes of *Wolbachia* species [37]. The structure and function of latrotoxins have been studied for their importance in the black widow spider (*Latrodectus* spp.), which injects latrotoxins into its invertebrate prey, causing paralysis [38]. In *Wolbachia*, several large protein toxins possess a latrotoxin C-terminal domain (CTD) (Figure 2) [37]. Like TcB–TcC proteins, these proteins also contain multiple conserved motifs such as *Salmonella* plasmid Virulence B (SpvB) domains, ankyrin-repeats, Rhs-repeats,

deaminase domains, and programmed cell death (PCD) NACHT domains. In spiders, latrotoxin-CTD is proteolytically cleaved during cell release, forming the mature latrotoxin after apparent disintegration of the source cell [39,40]. In *Wolbachia*, latrotoxin-CTD domains are also preceded by furin cleavage sites, suggesting a similar mechanism for toxin cleavage and maturation in these endosymbionts [37]. Interestingly, in *Wolbachia* strain *w*Mel, an apparent cotranscribed operon (WD_0512, WD_0513, and WD_0514) [41] encodes not only a latrotoxin-CTD gene but also TcB–TcC (see above) and ankyrin-repeat genes, respectively. Gene copy number of this region (along with five other linked genes) is positively correlated with *Wolbachia* virulence in *Drosophila melanogaster* [32], suggesting a potential function of multiple pore-forming toxins in endosymbiont overgrowth and pathogenesis. The potency of pore-forming toxins in persistent, maternally transmitted infections presents an exciting puzzle to solve. Are these toxins less potent in arthropod endosymbionts, and if not, how are they regulated during apparently beneficial endosymbiont infections?

Deubiquitinating enzymes (DUBs)

Exemplifying their important role in pathogenesis, recent evidence has highlighted the presence of diverse DUBs in viruses and bacteria, where these DUBs suppress host ubiquitin-mediated immune responses during infection [42,43]. DUBs are specialized proteases that remove ubiquitin (Ub) from protein substrates in eukaryotes, affecting protein degradation, localization, function, or protein–protein interactions [44,45]. As shown previously [3–5], we also identified DUB domains present in *Wolbachia* and *Spiroplasma* protein toxins that function in host reproductive manipulations (Figures 2 and 3). The DUB domain in the *Wolbachia* cytoplasmic incompatibility factor CifB, for example, is required for its toxic effect on male sperm chromatin in *D. melanogaster* (Box 1 and Figure 3). Inactivating the catalytic residue in the DUB domain renders the CifB toxin harmless [3]. The protein CifA, encoded in the same operon as CifB, must be coexpressed with CifB in heterologous assays to induce CI in male hosts; in females, however, expression of CifA is capable of suppressing CifB toxicity and rescuing CI, fully recapitulating the *Wolbachia* CI phenotype (Box 1 and Figure 3) [3,46,47] CifA is capable of binding CifB directly, which might underlie its antitoxin activity [3]. The fact that CifA and CifB activities depend on each other mirrors a type II TA system, with the male host as the toxin target instead of a bacterial competitor [3,48].

A DUB is also partly responsible for the male killing effects of *S. poulsonii* in *D. melanogaster* (Figure 3). The toxin Spaid (*S. poulsonii* androcidin) was recently identified as the major effector underlying *S. poulsonii* potent male-killing phenotypes (Box 1 and Figure 3) [4]. Expression of Spaid in the absence of *S. poulsonii* infection causes male death at the second instar larval stage of development but has no effect on female survival. Spaid contains four ankyrin repeat (AR) motifs and an OTU DUB domain (Figure 2). Deletion of the AR motifs abolishes Spaid toxicity, allowing adult males to survive. Deletion of the OTU domain allowed males to survive until pupation, suggesting that overexpression of Spaid without an OTU domain is still toxic at later stages of development; this toxicity, however, might be an artifact of artificially high expression levels in transgenic flies [4]. Deletion of the OTU domain also seemed to inhibit Spaid's ability to localize to the host cell nucleus, where it is hypothesized to damage DNA in males by interacting with the male-specific lethal (MSL) complex on the X chromosome (Figure 3). Interestingly, in *Wolbachia* strain *w*Mel, WD_0633 also possesses ARs and an OTU domain, giving it a 'Spaid-like' domain architecture (Table 1), but heterologous expression in *D. melanogaster* did not induce male killing [49]. Similarly, in *Rickettsia felis*, a plasmid-encoded toxin contains ARs and an OTU domain together, suggesting that diverse endosymbionts might converge on this toxic protein architecture [50]. In endosymbionts notorious for manipulating arthropod reproduction, therefore, DUB domains in Spaid and CifB appear critical for endosymbiont-induced host toxicity at DNA in male flies. In eukaryotes, DUBs play important roles in the cell DNA damage response by

regulating enzymes involved in DNA synthesis and repair [45]. We speculate that endosymbionts rely on DUBs to interfere with these pathways to manipulate host reproduction in sex-specific ways.

Ankyrin repeats (ARs)

It was previously thought that ARs are primarily restricted to eukaryotic proteins, but recent evidence suggests that they are also encoded by the genomes of pathogenic bacteria and viruses [51,52]. ARs are one of the most common domains found in eukaryotic proteins. They were first discovered as repeat sequences in yeast and *D. melanogaster* [53] and are characterized by a repeating 33 amino acid motif (containing the conserved N-terminal residues G-X-TPLHLA) that folds into a helix–loop–helix–β-hairpin/loop structure [54]. ARs containing genes are especially prevalent in the genomes of the endosymbiont *Wolbachia*. Four percent of the *w*Pip (*Wolbachia* strain of *Culex pipiens*) genome, for example, encodes AR genes (60/1386 coding sequences) – the highest reported for any prokaryote [55].

The only ARs so far directly implicated in endosymbiont toxin activity, based on specific biological assays, are those present in Spaid from *S. poulsonii* (discussed above). Despite their prevalence, little is known about AR function in *Wolbachia*. Heterologous expression of AR genes in *D. melanogaster* eggs did not induce reproductive manipulations [56], though some ARs appear sex-specifically expressed [55]. In certain pathogenic bacteria, however, ARs are critical for toxin activity. In *Legionella pneumophila*, for example, the toxic effector AnkX uses ARs to bind specifically to the human GTPase Rab1b and a FIC domain to transfer a phosphocholine (PC) moiety [57,58]. A crystal structure of AnkX in complex with Rab1b showed its 13 ARs forming a cup-like structure around Rab1b, where different AR residues interact specifically with different Rab1b-exposed domains [58]. Purified versions of AnkX with truncated ARs showed severely reduced phosphocholination rates, confirming their necessity in AnkX binding of Rab1b substrates [58]. Given bacterial ARs' capacity to interact with eukaryotic protein domains, the prevalence of ARs in *Wolbachia* genomes (and in association with bacteriophage WO [37]) hints at their potential to be secreted effectors. Future studies focused on whether *Wolbachia* secrete ARs into the host and whether these ARs bind host proteins will be informative in exploring their activity in endosymbiont–arthropod interactions.

Ribosome-inactivating proteins (RIPs)

RIP toxins are produced by fungi, plants, and bacteria; pathogenic bacteria use them as a virulence factor, while in eukaryotes, they are primarily defensive. Bacterial pathogens secrete RIPs and, after binding to the host cell membrane, RIP toxins are endocytosed, during which the catalytic subunit A is cleaved into fragments A1 and A2 in the endoplasmic reticulum [59]. After translocation to the cytosol, the A1 fragment binds a region of the host 28S ribosomal subunit called the sarcin–ricin loop (SRL) and cleaves a displayed adenosine residue (depurination) necessary for protein synthesis. Cleavage of this residue (rRNA glycosylation) blocks protein synthesis, eventually killing the host. We confirmed previous evidence of RIP toxins in both *Hamiltonella* and *Spiroplasma* genomes (Figure 2) [60]. Genomes from *Spiroplasma* contained genes with homology to the RIP 'Shiga' toxin A1 fragment, whereas genomes from *Hamiltonella* only contained genes with homology to the B subunit (Table 1 and Figure 2).

Some evidence points to endosymbiont RIP toxins playing important roles in insect defense from wasp parasitoids as well as parasitic nematodes. In *Drosophila*, RIPs were first implicated in protection against the nematode *Howardula aoronymphium*, which parasitizes *Drosophila neotestacea* hosts, resulting in sterility. When *D. neotestacea* is also infected with *Spiroplasma*, it is protected from *H. aoronymphium*-induced sterility. RIPs derived from *Spiroplasma* were shown to be able to depurinate *H. aoronymphium* rRNA *in vitro*, and the same depurination

signatures were detected from *H. aoronymphium* rRNA extracted from *Spiroplasma*-infected *D. neotestacea* [60]. These results suggest that *Spiroplasma* RIPs protect *D. neotestacea* hosts from sterilization by inhibiting protein synthesis in *H. aoronymphium*. Surprisingly, *D. neotestacea* rRNA did not show significant signatures of depurination during coinfection, indicating either localized delivery of RIPs to *H. aoronymphium* via *Spiroplasma* secretion or some RIP target selection mechanism [60]. Similarly, *D. melanogaster* strains infected with *S. poulsonii* containing RIP toxins are more likely to resist wasp attacks [61]. RNA isolated from wasps that parasitize *D. melanogaster* hosts infected with *S. poulsonii* shows signatures of depurination (as do the *D. melanogaster* rRNAs), strongly suggesting that *S. poulsonii* RIPs are active during infection. RIP toxins, therefore, are potent protein synthesis inhibitors that are associated with arthropod defense in multiple parasitism scenarios (Figure 3). In the case of *D. melanogaster* and *Spiroplasma*, the RIPs were also toxic to the flies [61,62], suggesting a tradeoff between toxicity and defense during wasp parasitism.

Aphid hosts of *Hamiltonella* are also vulnerable to attack by multiple endoparasitic wasp species. The pea aphid *Acyrthosiphon pisum* is less likely to succumb to parasitism if it also harbors the endosymbiont *Hamiltonella defensa*. But not all *H. defensa* strains confer resistance. Instead, strains infected with the **bacteriophage APSE** are more likely to provide protection to their aphid hosts [63]. The effect of APSE on parasitoid resistance is correlated with the presence of RIP-like toxins (as well as CdtB and YD-repeat toxins) in the phage genome, suggesting that phage-encoded toxins offer protection to aphid hosts [64,65]. Interestingly, different APSE variants themselves exhibit different degrees of protection against parasitoids. In particular, APSE strain APSE3, which harbors a YD-repeat toxin, offers high levels of protection when experimentally transferred to a phage-free strain of *H. defensa* [66]. APSE-encoded toxins other than RIPs, therefore, might underlie APSE mediated parasitoid defense.

The impact of bacteriophage on the diversity of arthropod endosymbiont toxins

Several of the toxins we describe are encoded on mobile genetic elements in their endosymbiont genomes. These include TcB–TcC, Spaid-like, CdtB, RIPs, CifA/CifB, ETX/MTX2, and Iatrotoxin-CTD in *Wolbachia*, *Hamiltonella*, and *Spiroplasma* (Table 1 and Figure 2). This result is striking. Mobile elements are rare in endosymbionts as opportunities for interaction and genetic exchange dwindle in host cellular environments [67]. Arthropod endosymbionts like *Wolbachia*, *Hamiltonella*, and *Spiroplasma* are outliers, therefore, since they harbor multiple mobile elements including bacteriophage, transposable elements, and plasmids [35,36,67–69]. The prevalence of potent toxin genes on mobile elements in these endosymbiont genera has important evolutionary consequences that we discuss below.

Mobile elements, such as bacteriophage, can impact the evolution of the toxins they encode as well as interactions between endosymbiont and host. Homology between phage APSE of *H. defensa* and phage pSOG3 of *Sodalis glossinidius*, for example, suggests that phages are not only horizontally exchanged but also recombined between different coinfecting endosymbiont species [65,70]. The same has been observed for bacteriophage WO in *Wolbachia* strains that coinfect *Drosophila* and *Nasonia* species [71,72]. Phage-mediated **horizontal gene transfer (HGT)** and subsequent recombination of DNA between endosymbionts, therefore, could be a major source of toxin diversification as toxin genes move from one host to another. Phage-mediated HGT of toxins between endosymbionts can also impact arms races between host and endosymbiont since endosymbionts that acquire toxins might show increased virulence relative to toxin-free endosymbionts [73,74]. This arms race itself is likely ripe for toxin evolution as new mutations that increase toxin potency will be

rapidly selected (although potentially at the cost of host fitness). Interestingly, recent evidence has demonstrated the transfer of an endosymbiont-derived toxin (CdtB) into eukaryotic hosts, likely via APSE or *H. defensa*, presenting yet another mode of toxin diversification [75,76]. Future work dedicated to understanding how endosymbiont toxins evolve and move across endosymbiont/host clades will provide clarity on their impact in host–endosymbiont interactions and evolution.

Concluding remarks

Endosymbionts interact with their hosts along a spectrum of fitness consequences. The spectrum ranges from completely antagonistic interactions, in which endosymbionts are parasites at the cost of their host fitness, to completely mutualistic interactions, in which endosymbionts are cooperators and benefit host fitness [77]. In *Wolbachia pipientis*, for example, two strains (*w*Mel and *w*MelPop) infecting *D. melanogaster* fall at opposite ends of this spectrum, with *w*Mel benefiting its host by providing protection from viral pathogens, and *w*MelPop harming its host by over multiplying – though a high *w*MelPop titer also confers increased pathogen protection [78]. These effects are likely due to genetic differences between the endosymbiont strains: *w*MelPop shows a significantly higher copy number (up to 13) of an array of eight genes (termed the Octomom region) than *w*Mel [32]. At least two of the genes in Octomom encode proteins with homology to toxin domains: the pore-forming toxins Iatrotoxin-CTD (WD_0512) and TcB-TcC (WD_0513). Both WD_0512 and WD_0513 are expressed in *D. melanogaster* during infection [79] but their function/toxicity remains unclear. We speculate that endosymbionts rely on such protein toxins to walk along this parasitism–mutualism spectrum during coevolution with their hosts, inducing arthropod phenotypes (e.g., CI) to different degrees. Understanding the structure and function of these individual toxins opens up the possibility of understanding how and why different endosymbiont strains and species harm or help their arthropod hosts while shaping their biology (see Outstanding questions).

Outstanding questions

Do arthropod endosymbiont proteins with homology to bacterial pathogen toxins retain their toxic activity in endosymbiont–arthropod interactions?

Are predicted endosymbiont toxins secreted into host cells during infection?

How do prophages within endosymbiont genomes impact the evolution of predicted toxin effectors?

Does toxin evolution predict whether endosymbionts behave as mutualists or parasites with their arthropod hosts?

Acknowledgments

We thank members of the Newton laboratory for their helpful discussions. Special thanks to Danny Rice for advice on structural homology detection. This work was supported by awards NIH R01AI144430 and NSF BII 2022049 (to I.L.G.N.).

Declaration of interests

There are no interests to declare.

References

1. Giribet, G. and Edgecombe, G.D. (2012) Reevaluating the arthropod tree of life. *Annu. Rev. Entomol.* 57, 167–186
2. Kaur, R. *et al.* (2021) Living in the endosymbiotic world of *Wolbachia*: a centennial review. *Cell Host Microbe* 29, 879–893
3. Beckmann, J.F. *et al.* (2017) A *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat. Microbiol.* 2, 17007
4. Harumoto, T. and Lemaitre, B. (2018) Male-killing toxin in a bacterial symbiont of *Drosophila*. *Nature* 557, 252–255
5. LePage, D.P. *et al.* (2017) Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. *Nature* 543, 243–247
6. Alouf, J.E. *et al.* (2005) *The Comprehensive Sourcebook of Bacterial Protein Toxins*, Elsevier
7. Schmitt, C.K. *et al.* (1999) Bacterial toxins: friends or foes? *Emerg. Infect. Dis.* 5, 224–234
8. Zhang, D. *et al.* (2012) Polymorphic toxin systems: Comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. *Biol. Direct* 7, 18
9. Kumar, S. *et al.* (2017) TimeTree: a resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* 34, 1812–1819
10. Kawamukai, M. *et al.* (1989) Nucleotide sequences of fic and fic-1 genes involved in cell filamentation induced by cyclic AMP in *Escherichia coli*. *J. Bacteriol.* 171, 4525–4529
11. Worby, C.A. *et al.* (2009) The fic domain: regulation of cell signaling by adenylylation. *Mol. Cell* 34, 93–103
12. Yarbrough, M.L. *et al.* (2009) AMPylation of Rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. *Science* 323, 269–272
13. Xiao, J. *et al.* (2010) Structural basis of Fic-mediated adenylylation. *Nat. Struct. Mol. Biol.* 17, 1004–1010
14. Veyron, S. *et al.* (2018) FIC proteins: from bacteria to humans and back again. *Pathog. Dis.* 76, fty012
15. Ellegaard, K.M. *et al.* (2013) Comparative genomics of *Wolbachia* and the bacterial species concept. *PLoS Genet.* 9, e1003381
16. Newton, I.L.G. *et al.* (2015) *Wolbachia* utilize host actin for efficient maternal transmission in *Drosophila melanogaster*. *PLoS Pathog.* 11, e1004798
17. Sheehan, K.B. *et al.* (2016) Identification and characterization of a candidate *Wolbachia pipientis* type IV effector that interacts with the actin cytoskeleton. *mBio* 7, e00622-1
18. Harms, A. *et al.* (2018) Toxins, targets, and triggers: an overview of toxin–antitoxin biology. *Mol. Cell* 70, 768–784

19. Dy, R.L. *et al.* (2014) Remarkable mechanisms in microbes to resist phage infections. *Annu. Rev. Virol.* 1, 307–331
20. Leplae, R. *et al.* (2011) Diversity of bacterial type II toxin–antitoxin systems: a comprehensive search and functional analysis of novel families. *Nucleic Acids Res.* 39, 5513–5525
21. Gotfredsen, M. and Gerdes, K. (1998) The *Escherichia coli* relBE genes belong to a new toxin-antitoxin gene family. *Mol. Microbiol.* 29, 1065–1076
22. Song, S. and Wood, T.K. (2018) Post-segregational killing and phage inhibition are not mediated by cell death through toxin/antitoxin systems. *Front. Microbiol.* 9, 814
23. Stårsta, M. *et al.* (2020) RHS-elements function as type II toxin–antitoxin modules that regulate intra-macrophage replication of *Salmonella* Typhimurium. *PLoS Genet.* 16, e1008607
24. Audoly, G. *et al.* (2011) Effect of rickettsial toxin VapC on its eukaryotic host. *PLoS One* 6, e26528
25. Dal Peraro, M. and van der Goot, F.G. (2016) Pore-forming toxins: ancient, but never really out of fashion. *Nat. Rev. Microbiol.* 14, 77–92
26. Gilbert, R.J.C. (2002) Pore-forming toxins. *Cell. Mol. Life Sci.* 59, 832–844
27. Meusch, D. *et al.* (2014) Mechanism of Tc toxin action revealed in molecular detail. *Nature* 508, 61–65
28. Song, N. *et al.* (2021) Genome-wide dissection reveals diverse pathogenic roles of bacterial Tc toxins. *PLoS Pathog.* 17, e1009102
29. Bowen, D. *et al.* (1998) Insecticidal toxins from the bacterium *Photorhabdus luminescens*. *Science* 280, 2129–2132
30. Gatsogiannis, C. *et al.* (2013) A syringe-like injection mechanism in *Photorhabdus luminescens* toxins. *Nature* 495, 520–523
31. Gatsogiannis, C. *et al.* (2016) Membrane insertion of a Tc toxin in near-atomic detail. *Nat. Struct. Mol. Biol.* 23, 884–890
32. Chrostek, E. and Teixeira, L. (2015) Mutualism breakdown by amplification of *Wolbachia* genes. *PLoS Biol.* 13, e1002065
33. Lacomel, C.J. *et al.* (2021) Branching out the aerolysin, ETX/MTX-2 and Toxin_10 family of pore forming proteins. *J. Invertebr. Pathol.* Published online March 26, 2021. https://doi.org/10.1016/j.jip.2021.107570
34. Finnie, J.W. *et al.* (1999) Neuronal damage produced in rat brains by *Clostridium perfringens* type D epsilon toxin. *J. Comp. Pathol.* 120, 415–420
35. Ballinger, M.J. and Perlman, S.J. (2019) The defensive *Spiroplasma*. *Curr. Opin. Insect Sci.* 32, 36–41
36. Gerth, M. *et al.* (2021) Rapid molecular evolution of *Spiroplasma* symbionts of *Drosophila*. *Microb. Genom.* 7 Published online February 16, 2021. https://doi.org/10.1099/mgen.0.000503
37. Bordenstein, S.R. and Bordenstein, S.R. (2016) Eukaryotic association module in phage WO genomes from *Wolbachia*. *Nat. Commun.* 7, 13155
38. Orlova, E.V. *et al.* (2000) Structure of α-latrotoxin oligomers reveals that divalent cation-dependent tetramers form membrane pores. *Nat. Struct. Biol.* 7, 48–53
39. Dulubova, I.E. *et al.* (1996) Cloning and structure of delta-latroinsectotoxin, a novel insect-specific member of the latrotoxin family: functional expression requires C-terminal truncation. *J. Biol. Chem.* 271, 7535–7543
40. Rohou, A. *et al.* (2007) Insecticidal toxins from black widow spider venom. *Toxicon* 49, 531–549
41. Iturbe-Ormaetxe, I. *et al.* (2005) Distribution, expression, and motif variability of ankyrin domain genes in *Wolbachia pipientis*. *J. Bacteriol.* 187, 5136–5145
42. Kubori, T. *et al.* (2019) Emerging insights into bacterial deubiquitinases. *Curr. Opin. Microbiol.* 47, 14–19
43. Schubert, A.F. *et al.* (2020) Identification and characterization of diverse OTU deubiquitinases in bacteria. *EMBO J.* 39, e105127
44. Komander, D. (2010) Mechanism, specificity and structure of the deubiquitinases. *Subcell. Biochem.* 54, 69–87
45. Komander, D. *et al.* (2009) Breaking the chains: structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* 10, 550–563
46. Shropshire, J.D. and Bordenstein, S.R. (2019) Two-By-One model of cytoplasmic incompatibility: Synthetic recapitulation by transgenic expression of cifA and cifB in *Drosophila*. *PLoS Genet.* 15, e1008221
47. Shropshire, J.D. *et al.* (2018) One prophage WO gene rescues cytoplasmic incompatibility in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 115, 4987–4991
48. Chen, H. *et al.* (2020) The biochemistry of cytoplasmic incompatibility caused by endosymbiotic bacteria. *Genes* 11, 852
49. Perlmutter, J.I. *et al.* (2020) Transgenic testing does not support a role for additional candidate genes in *Wolbachia* male killing or cytoplasmic incompatibility. *mSystems* 5, e00658-19
50. Gillespie, J.J. *et al.* (2018) A tangled web: origins of reproductive parasitism. *Genome Biol. Evol.* 10, 2292–2309
51. Al-Khodor, S. *et al.* (2010) Functional diversity of ankyrin repeats in microbial proteins. *Trends Microbiol.* 18, 132–139
52. Pan, X. *et al.* (2008) Ankyrin repeat proteins comprise a diverse family of bacterial type IV effectors. *Science* 320, 1651–1654
53. Breeden, L. and Nasmyth, K. (1987) Similarity between cell-cycle genes of budding yeast and fission yeast and the Notch gene of *Drosophila*. *Nature* 329, 651–654
54. Sedgwick, S.G. and Smerdon, S.J. (1999) The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biochem. Sci.* 24, 311–316
55. Walker, T. *et al.* (2007) Ankyrin repeat domain-encoding genes in the wPip strain of *Wolbachia* from the *Culex pipiens* group. *BMC Biol.* 5, 39
56. Yamada, R. *et al.* (2011) Functional test of the influence of *Wolbachia* genes on cytoplasmic incompatibility expression in *Drosophila melanogaster*. *Insect Mol. Biol.* 20, 75–85
57. Campanacci, V. *et al.* (2013) Structure of the *Legionella* effector AnkX reveals the mechanism of phosphocholine transfer by the FIC domain. *EMBO J.* 32, 1469–1477
58. Ernst, S. *et al.* (2020) *Legionella* effector AnkX displaces the switch II region for Rab1b phosphocholination. *Sci. Adv.* 6, eaaz8041
59. Sandvig, K. (2005) The Shiga toxins: properties and action on cells. In *The Comprehensive Sourcebook of Bacterial Protein Toxins* (3rd edn) (Alouf, J.E. *et al.*, eds), pp. 310–322, Academic Press
60. Hamilton, P.T. *et al.* (2016) A ribosome-inactivating protein in a *Drosophila* defensive symbiont. *Proc. Natl. Acad. Sci. U. S. A.* 113, 350–355
61. Ballinger, M.J. and Perlman, S.J. (2017) Generality of toxins in defensive symbiosis: ribosome-inactivating proteins and defense against parasitic wasps in *Drosophila*. *PLoS Pathog.* 13, e1006431
62. Garcia-Arraez, M.G. *et al.* (2019) Functional analysis of RIP toxins from the *Drosophila* endosymbiont *Spiroplasma poulsonii*. *BMC Microbiol.* 19, 46
63. Oliver, K.M. *et al.* (2009) Bacteriophages encode factors required for protection in a symbiotic mutualism. *Science* 325, 992–994
64. Oliver, K.M. and Perlman, S.J. (2020) Toxin-mediated protection against natural enemies by insect defensive symbionts. *Adv. Insect Physiol.* 58, 277–316
65. Rouïl, J. *et al.* (2020) The protector within: comparative genomics of APSE phages across aphids reveals rampant recombination and diverse toxin arsenals. *Genome Biol. Evol.* 12, 878–889
66. Brandt, J.W. *et al.* (2017) Culture of an aphid heritable symbiont demonstrates its direct role in defence against parasitoids. *Proc. Biol. Sci.* 284, 20171925
67. Newton, I.L.G. and Bordenstein, S.R. (2011) Correlations between bacterial ecology and mobile DNA. *Curr. Microbiol.* 62, 198–208
68. Klasson, L. *et al.* (2008) Genome evolution of *Wolbachia* strain w Pip from the *Culex pipiens* group. *Mol. Biol. Evol.* 25, 1877–1887
69. Reveillaud, J. *et al.* (2019) The *Wolbachia* mobilome in *Culex pipiens* includes a putative plasmid. *Nat. Commun.* 10, 1051
70. Degnan, P.H. and Moran, N.A. (2008) Diverse phage-encoded toxins in a protective insect endosymbiont. *Appl. Environ. Microbiol.* 74, 6782–6791

71. Bordenstein, S.R. and Wernegreen, J.J. (2004) Bacteriophage flux in endosymbionts (*Wolbachia*): infection frequency, lateral transfer, and recombination rates. *Mol. Biol. Evol.* 21, 1981–1991
72. Kent, B.N. *et al.* (2011) Complete bacteriophage transfer in a bacterial endosymbiont (*Wolbachia*) determined by targeted genome capture. *Genome Biol. Evol.* 3, 209–218
73. Harrison, E. and Brockhurst, M.A. (2017) Ecological and evolutionary benefits of temperate phage: what does or doesn't kill you makes you stronger. *Bioessays* 39, 1700112
74. Ma, W. *et al.* (2006) Type III effector diversification via both pathoadaptation and horizontal transfer in response to a coevolutionary arms race. *PLoS Genet.* 2, e209
75. Verster, K.I. *et al.* (2021) Horizontal transfer of microbial toxin genes to gall midge genomes. *bioRxiv* Published online February 4, 2021. https://doi.org/10.1101/2021.02.03.429655
76. Verster, K.I. *et al.* (2019) Horizontal transfer of bacterial cytolethal distending toxin B genes to insects. *Mol. Biol. Evol.* 36, 2105–2110
77. Wernegreen, J.J. (2004) Endosymbiosis: lessons in conflict resolution. *PLoS Biol.* 2, E68
78. Chrostek, E. *et al.* (2013) *Wolbachia* variants induce differential protection to viruses in *Drosophila melanogaster*: a phenotypic and phylogenomic analysis. *PLoS Genet.* 9, e1003896
79. Gutzwiller, F. *et al.* (2015) Dynamics of *Wolbachia pipientis* gene expression across the *Drosophila melanogaster* life cycle. *G3* 5, 2843–2856
80. Hypsa, V. and Dale, C. (1997) *In vitro* culture and phylogenetic analysis of 'Candidatus Arsenophonus triatominarum', an intracellular bacterium from the triatomine bug, *Triatoma infestans*. *Int. J. Syst. Bacteriol.* 47, 1140–1144
81. Nadal-Jimenez, P. *et al.* (2019) Genetic manipulation allows *in vivo* tracking of the life cycle of the son-killer symbiont, *Arsenophonus nasoniae*, and reveals patterns of host invasion, tropism and pathology. *Environ. Microbiol.* 21, 3172–3182
82. Zchori-Fein, E. and Perlman, S.J. (2004) Distribution of the bacterial symbiont *Cardinium* in arthropods. *Mol. Ecol.* 13, 2009–2016
83. Degnan, P.H. *et al.* (2009) *Hamiltonella defensa*, genome evolution of protective bacterial endosymbiont from pathogenic ancestors. *Proc. Natl. Acad. Sci. U. S. A.* 106, 9063–9068
84. Moran, N.A. *et al.* (2005) The players in a mutualistic symbiosis: insects, bacteria, viruses, and virulence genes. *Proc. Natl. Acad. Sci. U. S. A.* 102, 16919–16926
85. Ruhe, Z.C. *et al.* (2020) Polymorphic toxins and their immunity proteins: diversity, evolution, and mechanisms of delivery. *Annu. Rev. Microbiol.* 74, 497–520
86. Azad, A.F. and Beard, C.B. (1998) Rickettsial pathogens and their arthropod vectors. *Emerg. Infect. Dis.* 4, 179–186
87. Weinert, L.A. *et al.* (2009) Evolution and diversity of *Rickettsia* bacteria. *BMC Biol.* 7, 6
88. Perlman, S.J. *et al.* (2006) The emerging diversity of *Rickettsia*. *Proc. R. Soc. B Biol. Sci.* 273, 2097–2106
89. Montenegro, H. *et al.* (2005) Male-killing *Spiroplasma* naturally infecting *Drosophila melanogaster*. *Insect Mol. Biol.* 14, 281–287
90. Herren, J.K. *et al.* (2014) Insect endosymbiont proliferation is limited by lipid availability. *eLife* 3, e02964
91. Lindsey, A.R.I. *et al.* (2018) Conflict in the intracellular lives of endosymbionts and viruses: a mechanistic look at *Wolbachia*-mediated pathogen-blocking. *Viruses* 10, 141
92. Bateman, A. *et al.* (2002) The Pfam protein families database. *Nucleic Acids Res.* 30, 276–280
93. Berman, H.M. *et al.* (2000) The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242
94. Zimmermann, L. *et al.* (2018) A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* 430, 2237–2243
95. Masson, F. *et al.* (2018) *In vitro* culture of the insect endosymbiont *Spiroplasma poulsonii* highlights bacterial genes involved in host–symbiont interaction. *mBio* 9, e00024-18
96. Ballinger, M.J. *et al.* (2019) Toxin and genome evolution in a *Drosophila* defensive symbiosis. *Genome Biol. Evol.* 11, 253–262
97. Toju, H. and Fukatsu, T. (2011) Diversity and infection prevalence of endosymbionts in natural populations of the chestnut weevil: relevance of local climate and host plants. *Mol. Ecol.* 20, 853–868
98. Duron, O. *et al.* (2008) The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC Biol.* 6, 27
99. Gotoh, T. *et al.* (2007) *Cardinium* symbionts cause cytoplasmic incompatibility in spider mites. *Heredity* 98, 13–20
100. Landmann, F. (2019) The *Wolbachia* endosymbionts. *Microbiol. Spectr.* 7 https://doi.org/10.1128/microbiolspec.BAI-0018-2019
101. Veneti, Z. *et al.* (2005) A functional dosage compensation complex required for male killing in *Drosophila*. *Science* 307, 1461–1463