Rapid turnover of pathogen-blocking *Wolbachia* and their incompatibility loci

J. Dylan Shropshire[1,2*], William R. Conner[1], Dan Vanderpool[3],
Ary A. Hoffmann[4], Michael Turelli[5*], and Brandon S. Cooper[1*]

Affiliations: [1]Division of Biological Sciences, University of Montana, Missoula, Montana, USA, [2]Department of Biological Sciences, Lehigh University, Bethlehem, Pennsylvania, USA, [3]Forest Service, National Genomics Center for Wildlife and Fish Conservation, Missoula, Montana, USA, [4]Pest and Environmental Adaptation Research Group, Bio21 Institute and the School of BioSciences, The University of Melbourne, Parkville, Australia, [5]Department of Evolution and Ecology, University of California, Davis, California, USA

ORCiD: 0000-0003-4221-2178 (JDS), 0000-0001-9407-6038 (WRC), 0000-0002-6856-5636 (DV), 0000-0001-9497-7645 (AAH), 0000-0003-1188-9856 (MT), 0000-0002-8269-7731 (BSC)

***Correspondence to:** shropshirejd@lehigh.edu (JDS),mturelli@ucdavis.edu (MT), brandon.cooper@umontana.edu (BSC)

Classification: Biological Sciences, Evolution

Keywords: cytoplasmic incompatibility, *Drosophila,* endosymbiosis, horizontal transmission, host switching, insertion sequence elements

Abstract

At least half of all insect species carry maternally inherited *Wolbachia* alphaproteobacteria, making *Wolbachia* the most common endosymbionts in nature. *Wolbachia* spread to high frequencies is often due to cytoplasmic incompatibility (CI), a *Wolbachia*-induced sperm modification that kills embryos without *Wolbachia*. Several CI-causing *Wolbachia* variants, including *w*Mel from *Drosophila melanogaster*, also block viruses. Establishing pathogen-blocking *w*Mel in natural *Aedes aegypti* mosquito populations has reduced dengue disease incidence, with one study reporting about 85% reduction when *w*Mel frequency is high. However, *w*Mel transinfection establishment is challenging in many environments, highlighting the importance of identifying CI-causing *Wolbachia* variants that stably persist in diverse hosts and habitats. We demonstrate that *w*Mel-like variants have naturally established in widely distributed holometabolous dipteran and hymenopteran insects that diverged approximately 350 million years ago, with *w*Mel variants spreading rapidly among these hosts over only the last 100,000 years. *Wolbachia* genomes contain prophages that encode CI-causing operons (*cifs*). These *cifs* move among *Wolbachia* genomes – with and without prophages – even more rapidly than *Wolbachia* move among insect hosts. Our results shed light on how rapid host switching and horizontal gene transfer contribute to *Wolbachia* and *cif* diversity in nature. The diverse *w*Mel variants we report here from hosts present in different climates offer many new options for broadening *Wolbachia*-based biocontrol of diseases and pests.

Significance Statement

Wolbachia bacteria transinfected into *Aedes aegypti* mosquitoes (in particular the *w*Mel *Wolbachia* strain) are being widely released to suppress the transmission of dengue and other insect-vectored arboviruses. The bacterial strain *w*Mel originated from *Drosophila melanogaster*, but very little is known about its distribution and spread to other drosophilids and distantly related insects. Here we show that there has been a rapid rate of spread of *Wolbachia* strains closely related to *w*Mel not just into other flies but also into wasps, indicating a large pool of *w*Mel variants that can be employed in vector-borne disease control. *Wolbachia* prophage genes essential to this spread have evolved particularly rapidly.

Main Text

Maternally transmitted *Wolbachia* bacteria were first discovered in the ovaries of the mosquito *Culex pipiens* (*1*) and are now recognized as the most common endosymbiont of insects, occurring in about half of all species (*2*). *Wolbachia*'s success stems largely from its ability to affect host reproduction to favor *Wolbachia* transmission. For example, cytoplasmic incompatibility (CI) kills *Wolbachia*-free embryos fertilized by *Wolbachia*-carrying males, driving the endosymbiont to high frequencies. Model *Drosophila* systems have facilitated understanding of *Wolbachia* population biology and CI. The first two *Wolbachia* variants discovered in *Drosophila* species were *w*Ri, found in southern California *D. simulans* (*3*) and *w*Mel, found in Australian *D. melanogaster* (*4*). Both variants rapidly spread worldwide within these invasive hosts (*5–7*). *w*Ri now occurs at relatively stable frequencies (~93%) in almost all characterized *D. simulans* populations (*8*), while the frequencies of *w*Mel and closely related variants vary significantly among host populations (*9–11*).

CI enables the suppression of vector and pest populations and the transformation of natural vector populations with pathogen-blocking *Wolbachia*. When *w*Mel *Wolbachia* are moved from *D. melanogaster* to *Aedes aegypti*–species diverged 250 million years ago (MYA)–CI associated with *w*Mel can drive the endosymbiont to high frequencies where it can alleviate human disease burden (*12*), with one study reporting an 86% reduction in dengue (*13*). While artificial *w*Mel transfer between divergent Diptera has led to successful disease control (*13*), not all attempted *w*Mel introductions have succeeded (*14*, *15*). This raises the issue of whether alternative *w*Mel variants might spread in these situations, but the diversity of *w*Mel variants and the timescale of their host-switching remain uncertain. By comparing phylograms and chronograms for host and *Wolbachia* genomes, we document that variants closely related to *w*Mel have naturally transfered among dipteran and hymenopteran hosts 100 million years (MY) more diverged than donor *D. melanogaster* and recipient *Ae. aegypti*. These host invasions are accompanied by even faster turnover of CI-causing operons (*cifs*) via both phage-dependent and phage-independent transfer among *Wolbachia* genomes. Our results offer insights for further broadening applications of *w*Mel variants and their diverse *cifs*.

Rapid spread of *w*Mel-like *Wolbachia* Across Hosts Diverged 350 MYA

To illustrate the timescales of *Wolbachia* movement across host species and of *cifs* among *Wolbachia* genomes, we focus on *Wolbachia* closely related to *w*Mel. We use "*w*Mel-like" to designate the clade examined here (*16*), comparing the results to "*w*Ri-like" *Wolbachia* (*17*) (Table S1). Although the clade boundaries are arbitrary, our conclusions concerning rapid movement of closely related *Wolbachia* among distantly related hosts—and rapid turnover of *Wovirus* and *cifs* within those *Wolbachia* genomes—rest on robustly estimated chronograms. When pervasive recombination and horizontal exchange were initially described among diverse *Wolbachia* (*18*, *19*), it was conjectured that recombination precluded reliable bifurcating phylogenies and chronograms for these mosaic genomes. However, subsequent analyses cf.(*20*) show that despite rapid and frequent movements of CI-determining loci, phages, and other genetic elements, large portions of the *w*Mel-like and *w*Ri-like genomes show no significant evidence of recombination (Supplementary Discussion). We use these apparently quasi-clonal genomic regions for our phylogenetic and divergence-time analyses and show that previous analyses which did not explicitly control for recombination are robust (Fig. S1; Supplementary Discussion) (*17*, *21*).

We consider 20 *w*Mel-like *Wolbachia* whose host species include dipteran and hymenopteran hosts, with a most recent common ancestor (MRCA) about 350 MYA (Fig. 1A-C) (*22*). Using all single-copy genes of equal length with little evidence of past recombination (Supplementary Discussion), the bulk of these *Wolbachia* genomes diverged only about 128 KYA (95% credible interval: 51 – 263 KYA) (Fig. 1A). Fig. 1B shows an approximate chronogram for the most diverged hosts: a wasp, *Diachasma alloeum*; a stalk-eyed fly, *Sphyracephala brevicornis*; and *D. teissieri* as a representative drosophilid. (The divergence time of the insect orders Diptera, which includes the families Drosophilidae and Diopoidea (stalk-eyed flies), and Hymenoptera is comparable to the crown age of all extant tetrapods, ~373 MY (*23*). In contrast, the *w*Mel-like *Wolbachia* in these most-diverged hosts, denoted *w*Dal, *w*Sbr and *w*Mel, respectively, diverged only about 94 KYA (95% CI: 38 – 189 KYA) (Fig. 1A). We also report *w*Mel-like *Wolbachia* in 18 drosophilids (Fig. 1C), including a variant, *w*Zts, in tropical *Zaprionus tscasi* that is now the closest known relative of *w*Mel in *D. melanogaster*. *Zaprionus* flies are members of the *Drosophila* subgenus that diverged from the *Sophophora* subgenus about 47 MYA (95% CI: 43.9 – 49.9 MYA), highlighting

*w*Mel proliferation across species that span the entire paraphyletic genus also named *Drosophila* (*24*), similar to *w*Ri (*17*). We discuss additional inferences that can be made when more than one *Wolbachia* sequence from these hosts and others become available and provide a correction of prior claims concerning the *Wolbachia* in *D. suzukii* and *D. subpulchrella* based on species misidentification (Fig. S2, Supplementary Discussion). The complete set of hosts and their *w*Mel-like *Wolbachia* are provided in Table S1.

Rapid Introgressive Transfer of *Wolbachia* Between Some Closely Related Species

Many obligate mutualistic endosymbionts like the *Wolbachia* in filarial nematodes (*25*) and *Buchnera* in aphids (*26*) are acquired cladogenically. In contrast, all but one of these *w*Mel-like *Wolbachia* must have been acquired through introgression or non-sexual horizontal transmission. Introgression is plausible only between the most closely related drosophilid species (indicated by the colored triangles in Fig. 1C). Joint analysis of mtDNA and *Wolbachia* sequences implies that the three-species *yakuba* clade (*D. teissieri*, (*D. yakuba*, *D. santomea*)) first acquired *Wolbachia* by horizontal transmission from an unknown host, then transferred it within the clade through hybridization and introgression (*27*). *D. yakuba* hybridizes with its endemic sister species *D. santomea* on the island of São Tomé (*10*, *28*, *29*), and with *D. teissieri* on the edges of forests on the nearby island of Bioko (*30*). *Wolbachia* and mtDNA chronograms are generally concordant for these three hosts and indicate more recent common ancestors than for the bulk of the host nuclear genomes (*27*). *Z. taronus* co-occurs on São Tomé, particularly with *D. santomea* at high altitudes, but diverged from the *D. yakuba* triad about 47 MYA (Fig. 1C), making introgression impossible. Yet, its *Wolbachia* (*w*Zta) diverged from the *w*Yak-clade *Wolbachia* only about 14–84 KY. These data illustrate horizontal *Wolbachia* transfer between distantly related species with overlapping ranges and habitats.

Other possible examples of introgressive *Wolbachia* acquisition involve two species pairs in the *D. montium* subgroup, (*D. seguyi*, *D. malagassya*) and (*D. bocqueti*, *D.* sp. aff. *chauvacae*), whose *Wolbachia* diverged on the order of 10 KY. Both host pairs appear as sister species (*31*), so introgressive *Wolbachia* transfer is plausible. However, the mtDNA third-position coding sites differ by 0.53% and 1.15% respectively, corresponding to divergence times on the order of 100 KY or longer (*32*), inconsistent with introgressive acquisition of the corresponding *Wolbachia*. For the more distantly related pairs (*Z. taronus*, *Z. tsacasi*) and (*D. borealis*, *D. incompta*), the mtDNA third-site differences of 19.3% and 30% respectively, decisively preclude introgressive *Wolbachia* transfer.

***w*Mel-like *Wolbachia* Hosts are Diverse and Cytoplasmic Incompatibility is Common**

Notably, the hosts of these *w*Mel-like *Wolbachia* are extraordinarily diverse in terms of ecology and geography. They range from cosmopolitan human-associated species (*D. melanogaster*, *D. simulans* and *S. pallida*) to endemics restricted to small oceanic islands (*D. santomea* and *D. arawakana*). The drosophilids include one that breeds and feeds on flowers (*D. incompta*), a mushroom specialist (*D. recens*), and classic generalists (e.g., *D. melanogaster* and *D. simulans*). As expected, hosts with closely related *Wolbachia* co-occur somewhere – or did in the recent past. For instance, *w*Au was previously observed in *D. simulans* in Florida and Ecuador (*6*). Thus, before *w*Au was displaced by *w*Ri, *w*Au-carrying *D. simulans* probably co-occurred with *D. tropicalis* found only in Central and South America and Caribbean islands, which harbors *w*Tro, sister to *w*Au in Fig.1C. Although the wasp *Diachasma alloeum* parasitizes the tephritid *Rhagoletis pomonella*, none of *R. pomonella*'s several *Wolbachia* seem to be *w*Mel-like (*33*).

We analyzed the phylogenetic distribution of CI—and the *cif* genes underlying it (*34–37*). CI contributes to *Wolbachia* incidence among and frequency within species (*38*), as well as *Wolbachia* persistence within both natural and transinfected hosts (*13*, *39*, *40*). Of our 20 *w*Mel- and 8 *w*Ri-like *Wolbachia*, all but 11 *w*Mel-like strains have been tested for CI. For 2 of these 11 *w*Mel-like strains (*w*Seg in *D. seguyi* and *w*Bocq in *D. bocqueti*), stocks were available for us to test for CI. Comparing hatch rates of embryos produced by the CI cross to hatch rates of embryos from the three compatible crosses, we found that *w*Seg and *w*Bocq cause relatively strong CI, with relative hatch rates from incompatible crosses less than half that from compatible crosses. Including these two new CI systems, 8 *w*Mel-like and 6 *w*Ri-like *Wolbachia* in our study cause CI (Fig. 2A; see Table S1 for references). CI strength varies greatly among these systems and others (*41*), but may also vary within systems as a function of male

age (*42*), environmental factors (*34*), and host backgrounds (*10*, *34*, *42*, *43*). As an example, *w*Mel causes strong CI in *D. simulans* (*44*) and *Ae. aegypti* (*12*) backgrounds, but generally weak CI in its source host, highlighting *w*Mel's ability to cause strong CI in hosts that have presumably not evolved to suppress *w*Mel-induced CI. Additional large-scale tests for CI are needed for *Wolbachia* that reportedly do not cause CI, but maintain *cifs*, as these strains may cause CI under some conditions.

cif Genes and Proteins are Highly Diverse Among *w*Mel-like *Wolbachia*

CI is caused by two-gene *cif* operons, with paternal expression of *cifB* (and occasionally *cifA*) killing embryos unless a complementary *cifA* copy is expressed maternally (*35–37*, *45–49*). *cifs* are generally associated with *Wovirus* bacteriophages that are themselves subdivided into four groups typed using serine recombinase (sr) alleles (sr1WO–sr4WO), with three harboring *cifs* (sr1WO-sr3WO) (*50*). *Wolbachia*-encoded *cifs* span five described clades called Types ($cif_{[T1]}$–$cif_{[T5]}$) (*51*), and individual *Wolbachia* genomes often include multiple *cif* operons (Fig. 2B) (*51*, *52*). Excluding non-CI-inducing *w*Mel-like *w*Au and *w*Tro (*6*, *53*), all other *Wolbachia* in our analyses encode between one and three *cif* operons from four of the five described *cif* types: all genomes except *w*Sbr contain a $cif_{[T1]}$ operon, eight contain a $cif_{[T2]}$ operon, and $cif_{[T4]}$ and $cif_{[T5]}$ operons each appear in four. *w*Mel-like and *w*Ri-like *Wolbachia* do not encode $cif_{[T3]}$ operons (Fig. 2B). Even within *cif* types, there is significant variation in protein sequence, length, and domain composition (Fig. S3A-C; Supplementary Discussion). An exceptional example is $CifB_{wZta[T1-2]}$ (i.e., the second copy of $CifB_{[T1]}$ found in *w*Zta). It is 357 amino acids longer than the next largest $CifB_{[T1]}$ and shares only 51% to 39% sequence identity with other $CifB_{[T1]}$ proteins (Table S2). Despite its considerable divergence, *w*Zta's CifB is more similar to other $CifB_{[T1]}$ proteins in our dataset than to any other putatively functional CifB Types, and retains a pair of PD-(D/E)XK nuclease domains ubiquitous among CifB proteins.

cif Operons Turn Over Rapidly Among Closely Related Wolbachia

As previously noted, the phylogenies of *cifs* and the *Wolbachia* that contain them are discordant (*27*, *35*, *51*); but how rapidly do *cifs* turnover among *Wolbachia* genomes? In the *w*SYTZ clade (*w*SYT plus *w*Zta), *w*SYT genomes contain only one $cif_{[T1]}$ operon but *w*Zta contains two, indicating a gain or loss in the last 14,000–84,000 years (Fig. 2). The level of divergence between *w*Zta's two $cif_{[T1]}$ operons rules out duplication as an origin of the second pair (Fig. 2C). Turnover is not restricted to particular *cif* types, as exemplified by the *w*SYTZ clade acquiring $cif_{[T4]}$ operons since it diverged from (*w*Mel, *w*Zts) 28–132 KYA; moreover, in each *Wolbachia*, the $cif_{[T4]}$ operons occur in the same phage as the $cif_{[T1]}$ operon. *w*Yak and *w*San diverged from *w*Tei only 1.5–10 KYA, yet they have a second $cif_{[T4]}$ operon (*54*). The trio ((*w*Ara, *w*Spa), *w*Sbr), which diverged 37–181 KYA, also have several notable changes. *w*Ara contains two $cif_{[T1]}$ operons, while its sister *w*Spa contains a $cif_{[T1]}$ operon and a $cif_{[T2]}$ operon. The outgroup in this trio, *w*Sbr, does not contain a $cif_{[T1]}$ operon, but instead contains $cif_{[T5]}$ and $cif_{[T2]}$ operons. The trio ((*w*Ara, *w*Spa), *w*Sbr) illustrates turnover of $cif_{[T1]}$, $cif_{[T2]}$ and $cif_{[T5]}$ operons within 28–132 KY. The variants (*w*Inc, (*w*Bor, (*w*Au, *w*Tro))), which diverged 36–179 KYA, indicate turnover, loss and gain of *cifs*: both *w*Inc and *w*Bor contain a $cif_{[T1]}$ operon, only *w*Bor contains a $cif_{[T5]}$ operon, while sister *w*Au and *w*Tro contain no *cifs*. Finally, the *w*Ri-like clade, which diverged 5–27 KYA (*17*), displays comparably rapid *cif* turnover. Most dramatically, the pair (*w*Aur, *w*Tri), which diverged only 13–3400 years ago (Supplementary Discussion), differ in that only *w*Tri contains a $cif_{[T5]}$ operon. Overall, diverse *cif* operons move among closely related *Wolbachia* even more rapidly than these *Wolbachia* move among distantly related hosts.

We further illustrate rapid *cif* turnover by focusing on homologs of $cifA_{[T1]}$. $cif_{[T1]}$ variants are the most common *cifs* among our *w*Mel-like and *w*Ri-like *Wolbachia*. We estimate the *w*Mel-like and *w*Ri-like clades diverged about 450 K–1.1 MYA (SI). Among *w*Mel-like variants, 11 contain very similar $cifA_{[T1]}$ on a polytomy closely related to $cifA_{[T1]}$ of the *w*Ri-like variants (Fig. 2C). The *w*Ri-like clade, with MRCA 5–27 KYA, contains very similar $cifA_{[T1]}$ alleles, suggesting that the $cif_{[T1]}$ homologs are codiverging with these *w*Ri-like *Wolbachia*. In contrast, the $cifA_{[T1]}$ alleles of the *w*Mel-like variants *w*Dal, *w*Ara, *w*Bor and *w*Inc are more similar to the $cifA_{[T1]}$ of *w*Ri-like *w*Aur and *w*Tri (99.8–99.2% aa identity) than to the $cifA_{[T1]}$ alleles found in the 11 *w*Mel-like *Wolbachia* in the polytomy (Fig. 2C). This suggests a recent transfer of $cif_{[T1]}$ operons between the *w*Mel-like and *w*Ri-like clades, or from a third unknown source. Closely related $cifA_{[T1]}$ variants are in *w*Mel-like *w*Dal, *w*Ara, *w*Spa, and *w*Zta. The grouping of *w*Zta and *w*Dal again reflects the flux of $cif_{[T1]}$ loci across *w*Mel-like *Wolbachia*.

Wovirus Turnover Does Not Fully Explain *cif* Movement

Two additional aspects of our data on *cif* transfer are worth emphasizing. First, the *Wovirus* bacteriophages that contain *cifs* turnover among *Wolbachia* genomes (Supplementary Discussion) on the order of 1–100 KY. An exceptional case involves the *w*Bocq genome that contains an sr3WO *Wovirus* that is absent from the genome of sister *w*Ach. This implies gain or loss of this *Wovirus* by (*w*Bocq, *w*Ach) in the last 3.3–25 KYA. Second, while *cifs* clearly transfer along with the *Wovirus* that carry them, *cifs* also move among divergent *Wovirus* classes (i.e., phage-independent *cif* turnover) (*27*)cf.(*54*). This interchange is documented in two ways: closely related *cifs* are found in distantly related phages and distantly related *cifs* are found in closely related phages (Fig. 3, and Supplementary Discussion). We demonstrate this by comparing phylograms of the serine recombinase genes (sr) of sr3WO *Wovirus* (sr3) to phylograms of $cifA_{[T1]}$ alleles associated with them. The ten sr3WO *Wovirus* in *w*Ri-like *Wolbachia* have identical sr3 alleles and are closely related to sr3 alleles found in several *w*Mel-like *Wolbachia*. These sr3 alleles are more distantly related to several other sr3 alleles that include a copy found in *w*Mel from *D. melanogaster*. In contrast, almost all $cifA_{[T1]}$ alleles associated with *w*Mel-like sr3WO *Wovirus* are very closely related to $cifA_{[T1]}$ alleles associated with *w*Ri-like sr3WO *Wovirus* (Fig. 3). This generalizes phage-independent *cif* transfer among divergent *Wolbachia* as first documented for Type IV loci in *w*Yak by Cooper et al. (2019) (*27*) and later misinterpreted by Baião et al. (2021) (*54*) (Supplementary Discussion). Higher quality assemblies for known donor and recipient *Wolbachia* will be essential for establishing the relative role of insertion sequence (IS) elements (*27*) and other factors like plasmids in this transfer. We hypothesize a critical role for IS elements in the phage-independent *cif* transfer we document, as supported by IS elements flanking the majority of *cifs* in our analyses (Table S3).

Selection Acts to Preserve *cifA* and Nuclease Domains within *cifB*

What is the fate of these *cifs*? Theory predicts that once *Wolbachia* are established in a host species, natural selection does not act to maintain CI but does act to maintain resistance to CI (*55*, *56*). Consistent with these predictions and prior observations (*21*, *51*), Fig. 2B shows that putative pseudogenization (i.e., truncation) is more common for *cifB* than for *cifA* (see also Fig. S4A,B). Still, as noted by Beckmann et al. (2021) (*57*), CI is incredibly common despite weak selection on the phenotype (*34*). This paradoxical prevalence of CI across *Wolbachia* lineages can be explained in part by clade selection in which CI-causing *Wolbachia* lineages are more likely to be transmitted to new host species because they typically have higher frequencies within host species and persist longer than do non-CI causing *Wolbachia* (*38*). However, CifB also contributes to alternative functions that include *Wolbachia* titer regulation through interactions with host autophagy (*58*). This suggests that Cif moonlighting could plausibly contribute to preservation of particular *cifs*.

To assess patterns of selection across Cifs, we calculated the ratio of non-synonymous (d_N) to synonymous substitutions (d_S) for each Cif protein, using a 3-dimensional spherical sliding window across the length of AlphaFold-derived Cif structures (Fig. 4A, Fig. S3, Fig. S4C,D, Movie 1). Fig. 4B shows that $CifA_{[T1]}$ proteins are more similar to one another in terms of both sequence identity and structural similarity than $CifB_{[T1]}$ proteins from the same pairs. As predicted, CifA of Types 1 and 2 had lower d_N/d_S ratios than did CifB of the same type (Fig. 4, S3D-E) (Supplementary Discussion), consistent with purifying selection maintaining CifA. Putatively pseudogenized CifA and CifB proteins have higher d_N/d_S ratios than do intact proteins (e.g., CifA $d_N/d_S \sim 1$; Fig. 4E-F), further supporting the presumption that in-frame stop codons interfere with Cif function (Supplementary Discussion). In contrast, Types 3 and 4 CifA and CifB have comparable d_N/d_S ratios within each type, which could plausibly stem from pleiotropy or loss-of-function (Fig. S3D-E). $CifA_{[T1]}$ and $CifB_{[T1]}$ binding residues and $CifB_{[T1]}$'s Deubiquitinase domain have d_N/d_S ratios comparable to non-domain associated residues. However, the two $CifB_{[T1]}$ nuclease domains both have lower d_N/d_S ratios than do other residues (Fig. 4C-D, G-H). Indeed, across Cif Types, CifB's first nuclease domain has lower d_N/d_S ratios than non-domain associated residues (Fig. S3F-H). While CifB's nuclease activity may not always contribute to observed CI expression (*59*), selection may still act to maintain other nuclease-associated features (e.g., DNA binding) and contribute to CifB's association with chromatin restructuring (Supplementary Discussion) (*60*, *61*). Interestingly, our data reveal that sites with d_N/d_S ratios above 1, consistent with positive selection, are primarily located within regions of the protein that are not domain-associated (Fig. 4D). Many of these sites appear at the surface, aligning with the expectation that surface residues are key to host-microbe interactions, potentially facilitating interactions between Cif and host proteins.

Thus, while theory predicts that selection does not act on the CI phenotype (*55*), selection on alternative CifB functions may plausibly delay mutational disruption of *cifB* (*57*).

Conclusions

Our findings confirm that non-sexual horizontal *Wolbachia* acquisition—and introgressive transfer between close relatives—commonly occurs on the order of 10–100 KY. These conclusions are robust to uncertainties about *Wolbachia* divergence times. Among recently diverged *Wolbachia*, distantly related *cif* operons rapidly turnover via phage-mediated and phage-independent transfer. The commonness of non-sexually acquired *Wolbachia* and *cif* transfer among *Wolbachia* genomes indicates that opportunities for horizontal transfer must be common. Because many insects carry vertically transmitted *Wolbachia* (*2*), we expect that ephemeral somatic infections are common cf.(*62*). Future analyses should focus on understanding the ecology of non-sexual *Wolbachia* transfer and the cellular-genetic basis of successful non-sexual *Wolbachia* establishment (or not) in new hosts, as well as on *Wovirus/cif* transfer between *Wolbachia* genomes.

Wolbachia offer a practical mechanism for mitigating human diseases. While *w*Mel introductions into *Ae. aegypti* have been very effective where they have spread to high frequencies (*13*, *39*), alternative *Wolbachia*-based interventions are needed. For example, *w*Mel has been lost in some release locations (*14*, *15*), particularly under extremely hot conditions. Temperature can affect CI strength (*43*, *63*), rates of imperfect *Wolbachia* transmission (*11*, *64*), and *Wolbachia* effects on host behaviors (*16*). Identifying strong-CI-causing and virus-blocking *Wolbachia* from the tropics could facilitate *Wolbachia* biocontrol, and *w*Mel-like variants that naturally associate with tropical host species are obvious candidates (*49*, *65*). Our study introduces a comprehensive panel of closely related *w*Mel-like *Wolbachia* that exhibit diverse ecology, geography and *cif* profiles. The discovery of a tropical *Wolbachia* variant (*w*Zts), sister to *w*Mel in *D. melanogaster*, may be immediately useful for applications. While *w*Zts CI has not yet been tested, all six *Z. tscasi* sampled in nature carry *w*Zts (producing 0.61 as the 95% lower bound for *w*Zts frequency), consistent with strong CI. This versatility contributes to moving towards more customized and environment-specific *Wolbachia* applications.

References:

1. J. Yen, A. Barr, Etiological agent of cytoplasmic incompatibility in *Culex pipiens*. *J. Invertebr. Pathol.* **22**, 242–250 (1973).
2. L. A. Weinert, E. V. Araujo-Jnr, M. Z. Ahmed, J. J. Welch, The incidence of bacterial endosymbionts in terrestrial arthropods. *Proc. R. Soc. B-Biol. Sci.* **282**, 20150249 (2015).
3. A. Hoffmann, M. Turelli, G. Simmons, Unidirectional incompatibility between populations of *Drosophila simulans*. *Evolution* **40**, 692–701 (1986).
4. A. Hoffmann, Partial cytoplasmic incompatibility between 2 Australian populations of *Drosophila* melanogaster. *Entomol. Exp. Appl.* **48**, 61–67 (1988).
5. M. Turelli, A. Hoffmann, Rapid spread of an inherited incompatibility factor in California *Drosophila*. *Nature* **353**, 440–442 (1991).
6. M. Turelli, A. Hoffmann, Cytoplasmic incompatibility in *Drosophila simulans* - Dynamics and parameter estimates from natural-populations. *Genetics* **140**, 1319–1338 (1995).
7. M. F. Richardson, L. A. Weinert, J. J. Welch, R. S. Linheiro, M. M. Magwire, F. M. Jiggins, C. M. Bergman, Population genomics of the *Wolbachia* endosymbiont in *Drosophila melanogaster*. *PLOS Genet.* **8**, e1003129 (2012).
8. P. Kriesner, A. A. Hoffmann, S. F. Lee, M. Turelli, A. R. Weeks, Rapid sequential spread of two *Wolbachia* variants in *Drosophila simulans*. *PLOS Pathog.* **9**, e1003607 (2013).
9. P. Kriesner, W. R. Conner, A. R. Weeks, M. Turelli, A. A. Hoffmann, Persistence of a *Wolbachia* infection frequency cline in *Drosophila melanogaster* and the possible role of reproductive dormancy. *Evolution* **70**, 979–997 (2016).
10. B. S. Cooper, P. S. Ginsberg, M. Turelli, D. R. Matute, *Wolbachia* in the *Drosophila yakuba* complex: Pervasive frequency variation and weak cytoplasmic incompatibility, but no apparent effect on reproductive isolation. *Genetics* **205**, 333 (2017).
11. M. T. J. Hague, J. D. Shropshire, C. N. Caldwell, J. P. Statz, K. A. Stanek, W. R. Conner, B. S. Cooper, Temperature effects on cellular host-microbe interactions explain continent-wide endosymbiont prevalence. *Curr. Biol.* **32**, 878 (2022).
12. T. Walker, P. H. Johnson, L. A. Moreira, I. Iturbe-Ormaetxe, F. D. Frentiu, C. J. McMeniman, Y. S. Leong, Y. Dong, J. Axford, P. Kriesner, A. L. Lloyd, S. A. Ritchie, S. L. O'Neill, A. A. Hoffmann, The *w*Mel *Wolbachia* strain blocks dengue and invades caged *Aedes aegypti* populations. *Nature* **476**, 450-U101 (2011).
13. A. Utarini, C. Indriani, R. A. Ahmad, W. Tantowijoyo, E. Arguni, M. R. Ansari, E. Supriyati, D. S. Wardana, Y. Meitika, I. Ernesia, I. Nurhayati, E. Prabowo, B. Andari, B. R. Green, L. Hodgson, Z. Cutcher, E. Rances, P. A. Ryan, S. L. O'Neill, S. M. Dufault, S. K. Tanamas, N. P. Jewell, K. L. Anders, C. P. Simmons, Efficacy of *Wolbachia*-infected mosquito deployments for the control of dengue. *N. Engl. J. Med.* **384**, 2177–2186 (2021).
14. J. S. Moledo Gesto, S. B. Pinto, F. B. Stehling Dias, J. Peixoto, G. Costa, S. Kutcher, J. Montgomery, B. R. Green, K. L. Anders, P. A. Ryan, C. P. Simmons, S. L. O'Neill, L. A. Moreira, Large-scale deployment and establishment of *Wolbachia* into the *Aedes aegypti* population in Rio de Janeiro, Brazil. *Front. Microbiol.* **12**, 711107 (2021).
15. N. T. Hien, D. D. Anh, N. H. Le, N. T. Yen, T. V. Phong, V. S. Nam, T. N. Duong, N. B. Nguyen, D. T. T. Huong, L. Q. Hung, C. N. T. Trinh, N. V. Hoang, V. Q. Mai, L. T. Nghia, N. T. Dong, L. H. Tho, S. Kutcher, T. P. Hurst, J. L. Montgomery, M. Woolfit, E. Rances, P. Kyrylos, K. L Anders, L. Nguyen, J. Brown-Kenyon, A. Caird, B. J. McLean, I. Iturbe-Ormaetxe, S. A. Ritchie, S. L. O'Neill, P. A. Ryan, Environmental factors influence the local establishment of *Wolbachia* in *Aedes aegypti* mosquitoes in two small communities in central Vietnam. *Gates Open Res.* **5**, 147 (2021).
16. M. T. J. Hague, H. Mavengere, D. R. Matute, B. S. Cooper, Environmental and genetic contributions to imperfect *w*Mel-like *Wolbachia* transmission and frequency variation. *Genetics* **215**, 1117–1132 (2020).
17. M. Turelli, B. S. Cooper, K. M. Richardson, P. S. Ginsberg, B. Peckenpaugh, C. X. Antelope, K. J. Kim, M. R. May, A. Abrieux, D. A. Wilson, M. J. Bronski, B. R. Moore, J.-J. Gao, M. B. Eisen, J. C. Chiu, W. R. Conner, A. A. Hoffmann, Rapid global spread of *w*Ri-like *Wolbachia* across multiple *Drosophila*. *Curr. Biol.* **28**, 963 (2018).
18. L. Baldo, S. Bordenstein, J. J. Wernegreen, J. H. Werren, Widespread recombination throughout *Wolbachia*

genomes. *Mol. Biol. Evol.* **23**, 437–449 (2006).
19. F. M. Jiggins, J. H. G. von der Schulenburg, G. D. D. Hurst, M. E. N. Majerus, Recombination confounds interpretations of *Wolbachia* evolution. *Proc. R. Soc. Lond. B Biol. Sci.* **268**, 1423–1427 (2001).
20. X. Wang, X. Xiong, W. Cao, C. Zhang, J. H. Werren, X. Wang, Phylogenomic analysis of *Wolbachia* strains reveals patterns of genome evolution and recombination. *Genome Biol. Evol.* **12**, 2508–2520 (2020).
21. M. K. Meany, W. R. Conner, S. Richter, J. A. Bailey, M. Turelli, B. S. Cooper, Loss of cytoplasmic incompatibility and minimal fecundity effects explain relatively low *Wolbachia* frequencies in *Drosophila mauritiana*. *Evolution* **73**, 1278–1295 (2019).
22. Y.-H. Wang, M. S. Engel, J. A. Rafael, H.-Y. Wu, D. Rédei, Q. Xie, G. Wang, X.-G. Liu, W.-J. Bu, Fossil record of stem groups employed in evaluating the chronogram of insects (Arthropoda: Hexapoda). *Sci. Rep.* **6**, 38939 (2016).
23. T. R. Simões, S. E. Pierce, Sustained high rates of morphological evolution during the rise of tetrapods. *Nat. Ecol. Evol.* **5**, 1403–1414 (2021).
24. A. Suvorov, B. Y. Kim, J. Wang, E. E. Armstrong, D. Peede, E. R. R. D'Agostino, D. K. Price, P. Waddell, M. Lang, V. Courtier-Orgogozo, J. R. David, D. Petrov, D. R. Matute, D. R. Schrider, A. A. Comeault, Widespread introgression across a phylogeny of 155 *Drosophila* genomes. *Curr. Biol. CB* **32**, 111-123.e5 (2022).
25. F. Comandatore, D. Sassera, M. Montagna, S. Kumar, G. Koutsovoulos, G. Thomas, C. Repton, S. A. Babayan, N. Gray, R. Cordaux, A. Darby, B. Makepeace, M. Blaxter, Phylogenomics and analysis of shared genes suggest a single transition to mutualism in *Wolbachia* of nematodes. *Genome Biol. Evol.* **5**, 1668–1674 (2013).
26. P. Baumann, C. Lai, L. Baumann, D. Rouhbakhsh, N. Moran, M. Clark, Mutualistic associations of aphids and prokaryotes - Biology of the genus *Buchnera*. *Appl. Environ. Microbiol.* **61**, 1–7 (1995).
27. B. S. Cooper, D. Vanderpool, W. R. Conner, D. R. Matute, M. Turelli, *Wolbachia* acquisition by *Drosophila yakuba*-clade hosts and transfer of incompatibility loci between distantly related *Wolbachia*. *Genetics* **212**, 1399–1419 (2019).
28. D. Lachaise, M. Harry, M. Solignac, F. Lemeunier, V. Benassi, M. L. Cariou, Evolutionary novelties in islands: *Drosophila santomea*, a new *melanogaster* sister species from São Tomé. *Proc. R. Soc. B-Biol. Sci.* **267**, 1487–1495 (2000).
29. A. A. Comeault, A. Venkat, D. R. Matute, Correlated evolution of male and female reproductive traits drive a cascading effect of reinforcement in *Drosophila yakuba*. *Proc. R. Soc. B-Biol. Sci.* **283**, 20160730 (2016).
30. B. S. Cooper, A. Sedghifar, W. T. Nash, A. A. Comeault, D. R. Matute, A maladaptive combination of traits contributes to the maintenance of a *Drosophila* hybrid zone. *Curr. Biol.* **28**, 2940 (2018).
31. W. R. Conner, E. K. Delaney, M. J. Bronski, P. S. Ginsberg, T. B. Wheeler, K. M. Richardson, B. Peckenpaugh, K. J. Kim, M. Watada, A. A. Hoffmann, M. B. Eisen, A. Kopp, B. S. Cooper, M. Turelli, A phylogeny for the *Drosophila montium* species group: A model clade for comparative analyses. *Mol. Phylogenet. Evol.* **158**, 107061 (2021).
32. S. Y. W. Ho, M. J. Phillips, A. Cooper, A. J. Drummond, Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol.* **22**, 1561–1568 (2005).
33. H. Schuler, W. Arthofer, M. Riegler, C. Bertheau, S. Krumböck, K. Köppler, H. Vogt, L. A. F. Teixeira, C. Stauffer, Multiple *Wolbachia* infections in *Rhagoletis pomonella*. *Entomol. Exp. Appl.* **139**, 138–144 (2011).
34. J. D. Shropshire, B. Leigh, S. R. Bordenstein, Symbiont-mediated cytoplasmic incompatibility: What have we learned in 50 years? *eLife* **9**, e61989 (2020).
35. D. P. LePage, J. A. Metcalf, S. R. Bordenstein, J. On, J. I. Perlmutter, J. D. Shropshire, E. M. Layton, L. J. Funkhouser-Jones, J. F. Beckmann, S. R. Bordenstein, Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. *Nature* **543**, 243 (2017).
36. J. F. Beckmann, J. A. Ronau, M. Hochstrasser, A *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat. Microbiol.* **2**, 17007 (2017).
37. J. D. Shropshire, S. R. Bordenstein, Two-By-One model of cytoplasmic incompatibility: Synthetic recapitulation by transgenic expression of *cifA* and *cifB* in *Drosophila*. *PLOS Genet.* **15**, e1008221 (2019).
38. M. Turelli, A. Katznelson, P. S. Ginsberg, Why *Wolbachia*-induced cytoplasmic incompatibility is so common. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2211637119 (2022).
39. A. A. Hoffmann, B. L. Montgomery, J. Popovici, I. Iturbe-Ormaetxe, P. H. Johnson, F. Muzzi, M. Greenfield,

M. Durkan, Y. S. Leong, Y. Dong, H. Cook, J. Axford, A. G. Callahan, N. Kenny, C. Omodei, E. A. McGraw, P. A. Ryan, S. A. Ritchie, M. Turelli, S. L. O'Neill, Successful establishment of *Wolbachia* in *Aedes* populations to suppress dengue transmission. *Nature* **476**, 454-U107 (2011).
40. A. Hoffmann, M. Turelli, L. Harshman, Factors affecting the distribution of cytoplasmic incompatibility in *Drosophila simulans*. *Genetics* **126**, 933–948 (1990).
41. J. D. Shropshire, E. Hamant, W. R. Conner, B. S. Cooper, *cifB*-transcript levels largely explain cytoplasmic incompatibility variation across divergent *Wolbachia*. *PNAS Nexus* **1**, pgac099 (2022).
42. J. D. Shropshire, E. Hamant, B. S. Cooper, Male age and *Wolbachia* dynamics: investigating how fast and why bacterial densities and cytoplasmic incompatibility strengths vary. *mBio* **12**, e02998-21 (2021).
43. K. T. Reynolds, A. A. Hoffmann, Male age, host effects and the weak expression or nonexpression of cytoplasmic incompatibility in *Drosophila* strains infected by maternally transmitted *Wolbachia*. *Genet. Res.* **80**, 79–87 (2002).
44. S. Zabalou, A. Apostolaki, S. Pattas, Z. Veneti, C. Paraskevopoulos, I. Livadaras, G. Markakis, T. Brissac, H. Mercot, K. Bourtzis, Multiple rescue factors within a *Wolbachia* strain. *Genetics* **178**, 2145–2160 (2008).
45. J. D. Shropshire, J. On, E. M. Layton, H. Zhou, S. R. Bordenstein, One prophage WO gene rescues cytoplasmic incompatibility in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 4987–4991 (2018).
46. K. L. Adams, D. G. Abernathy, B. C. Willett, E. K. Selland, M. A. Itoe, F. Catteruccia, *Wolbachia cifB* induces cytoplasmic incompatibility in the malaria mosquito vector. *Nat. Microbiol.* **6**, 1575–1582 (2021).
47. G. Sun, M. Zhang, H. Chen, M. Hochstrasser, The CinB nuclease from *w*No *Wolbachia* is sufficient for induction of cytoplasmic incompatibility in *Drosophila*. *mBio* **13**, e03177-21 (2022).
48. Y. Xiao, H. Chen, H. Wang, M. Zhang, X. Chen, J. M. Berk, L. Zhang, Y. Wei, W. Li, W. Cui, F. Wang, Q. Wang, C. Cui, T. Li, C. Chen, S. Ye, L. Zhang, X. Ji, J. Huang, W. Wang, Z. Wang, M. Hochstrasser, H. Yang, Structural and mechanistic insights into the complexes formed by *Wolbachia* cytoplasmic incompatibility factors. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2107699118 (2021).
49. J. D. Shropshire, R. Rosenberg, S. R. Bordenstein, The impacts of cytoplasmic incompatibility factor (*cifA* and *cifB*) genetic variation on phenotypes. *Genetics* **217**, iyaa007 (2021).
50. S. R. Bordenstein, S. R. Bordenstein, Widespread phages of endosymbionts: Phage WO genomics and the proposed taxonomic classification of Symbioviridae. *PLOS Genet.* **18**, e1010227 (2022).
51. J. Martinez, L. Klasson, J. J. Welch, F. M. Jiggins, Life and death of selfish genes: Comparative genomics reveals the dynamic evolution of cytoplasmic incompatibility. *Mol. Biol. Evol.* **38**, 2–15 (2021).
52. M. Bonneau, C. Atyame, M. Beji, F. Justy, M. Cohen-Gonsaud, M. Sicard, M. Weill, *Culex pipiens* crossing type diversity is governed by an amplified and polymorphic operon of *Wolbachia*. *Nat. Commun.* **9**, 319 (2018).
53. J. Martinez, S. Ok, S. Smith, K. Snoeck, J. P. Day, F. M. Jiggins, Should symbionts be nice or selfish? Antiviral effects of *Wolbachia* are costly but reproductive parasitism is not. *PLOS Pathog.* **11**, e1005021 (2015).
54. G. C. Baião, J. Janice, M. Galinou, L. Klasson, Comparative genomics reveals factors associated with phenotypic expression of *Wolbachia*. *Genome Biol. Evol.* **13**, evab111 (2021).
55. M. Turelli, Evolution of incompatibility-inducing microbes and their hosts. *Evolution* **48**, 1500–1513 (1994).
56. R. Haygood, M. Turelli, Evolution of incompatibility-inducing microbes in subdivided host populations. *Evolution* **63**, 432–447 (2009).
57. J. F. Beckmann, K. Van Vaerenberghe, D. E. Akwa, B. S. Cooper, A single mutation weakens symbiont-induced reproductive manipulation through reductions in deubiquitylation efficiency. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2113271118 (2021).
58. M. Deehan, W. Lin, B. Blum, A. Emili, H. Frydman, Intracellular density of *Wolbachia* is mediated by host autophagy and the bacterial cytoplasmic incompatibility gene *cifB* in a cell type-dependent manner in *Drosophila melanogaster*. *mBio* **12**, e02205-20 (2021).
59. R. Kaur, J. D. Shropshire, B. A. Leigh, S. R. Bordenstein, Nuclease proteins CifA and CifB promote spermatid DNA damage associated with symbiont-induced cytoplasmic incompatibility. *bioRxiv*, 2022.04.04.487029 (2022).
60. K. Terretaz, B. Horard, M. Weill, B. Loppin, F. Landmann, Functional analysis of *Wolbachia* Cid effectors unravels cooperative interactions to target host chromatin during replication. *PLOS Pathog.* **19**, e1011211

(2023).
61. R. Kaur, B. A. Leigh, I. T. Ritchie, S. R. Bordenstein, The Cif proteins from *Wolbachia* prophage WO modify sperm genome integrity to establish cytoplasmic incompatibility. *PLOS Biol.* **20**, e3001584 (2022).
62. S. Towett-Kirui, J. L. Morrow, S. Close, J. E. Royer, M. Riegler, Host-endoparasitoid-endosymbiont relationships: concealed Strepsiptera provide new twist to *Wolbachia* in Australian tephritid fruit flies. *Environ. Microbiol.* **23**, 5587–5604 (2021).
63. P. A. Ross, S. A. Ritchie, J. K. Axford, A. A. Hoffmann, Loss of cytoplasmic incompatibility in *Wolbachia*-infected *Aedes aegypti* under field conditions. *PLoS Negl. Trop. Dis.* **13**, e0007357 (2019).
64. P. A. Ross, I. Wiwatanaratanabutr, J. K. Axford, V. L. White, N. M. Endersby-Harshman, A. A. Hoffmann, *wolbachia* infections in *Aedes aegypti* differ markedly in their response to cyclical heat stress. *PLOS Pathog.* **13**, e1006006 (2017).
65. X. Gu, P. A. Ross, J. Rodriguez-Andres, K. L. Robinson, Q. Yang, M.-J. Lau, A. A. Hoffmann, A *w*Mel *Wolbachia* variant in *Aedes aegypti* from field-collected *Drosophila melanogaster* with increased phenotypic stability under heat stress. *Environ. Microbiol.* **24**, 2119–2135 (2022).
66. B. M. Wiegmann, M. D. Trautwein, I. S. Winkler, N. B. Barr, J.-W. Kim, C. Lambkin, M. A. Bertone, B. K. Cassel, K. M. Bayless, A. M. Heimberg, B. M. Wheeler, K. J. Peterson, T. Pape, B. J. Sinclair, J. H. Skevington, V. Blagoderov, J. Caravas, S. N. Kutty, U. Schmidt-Ott, G. E. Kampmeier, F. C. Thompson, D. A. Grimaldi, A. T. Beckenbach, G. W. Courtney, M. Friedrich, R. Meier, D. K. Yeates, Episodic radiations in the fly tree of life. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 5690–5695 (2011).

Acknowledgments:

We thank Daniel Matute for directing us to data for *Zaprionus* species and John Statz for comments that improved the manuscript.

Funding

This work was supported by National Science Foundation (NSF) CAREER (2145195) and National Institutes of Health MIRA (R35GM124701) Awards to BSC. JDS was supported by an NSF Postdoctoral Research Fellowship in Biology (DBI-2010210) and start-up funding from Lehigh University.

Author contributions

Conceptualization: JDS, MT, BSC; Data curation: WRC, DV; Formal analysis: JDS, WRC, DV, MT, BSC; Funding acquisition: JDS, BSC; Project administration: JDS, MT, BSC; Resources: BSC; Supervision: MT, BSC; Validation: JDS, WRC, DV; Visualization: JDS, DV, BSC; Writing – Original draft: JDS, MT, BSC; Writing – Review & editing: JDS, WRC, DV, AAH, MT, BSC.

Competing interests
The authors declare no competing interests.

Data and materials availability

Source data for the main and Supplementary Data figures are provided in the online version of this paper or in Dryad. Newly sequenced *Wolbachia* genomes are available under BioProject PRJNA1021588.

List of Supplementary Materials

Materials and Methods
Supplementary Discussion
Figs. S1 to S5
Tables S1 to S4
Movie S1

Figure 1

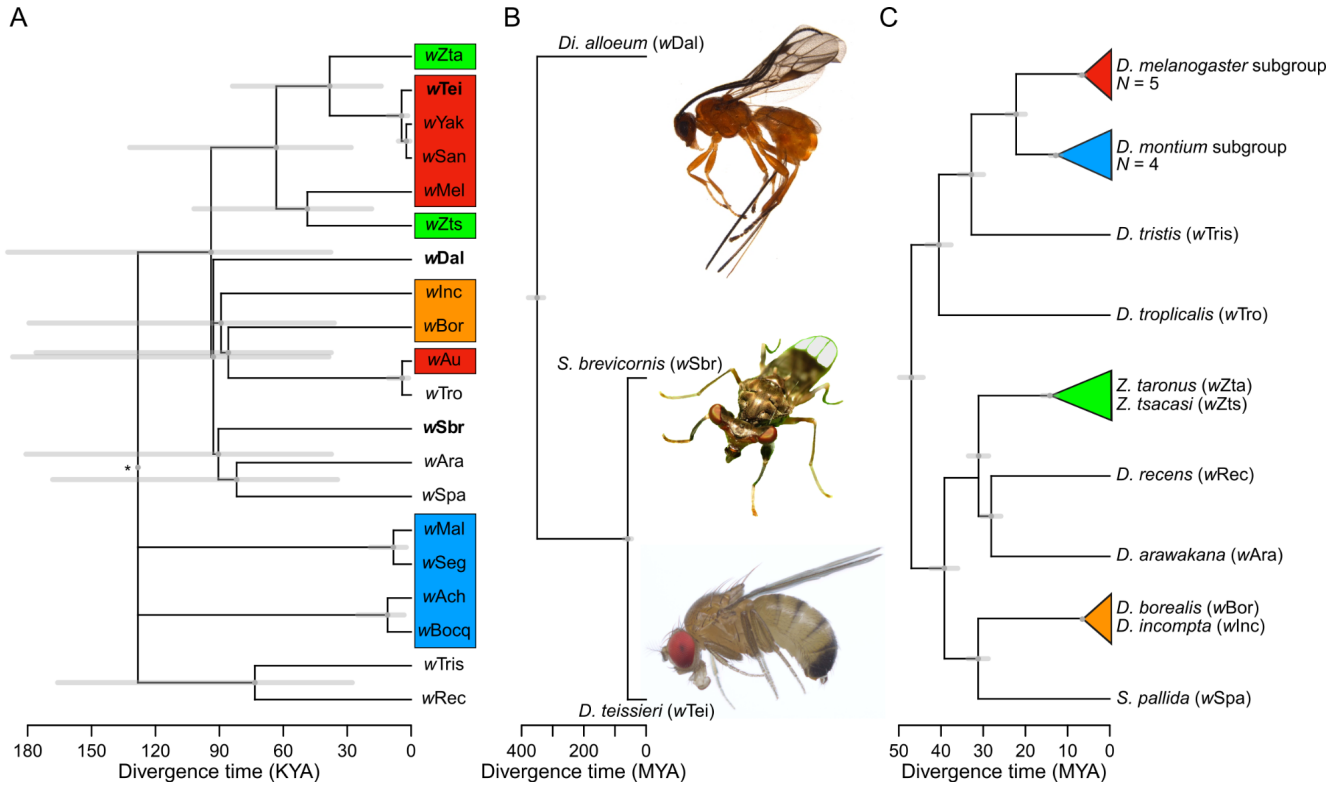

Fig. 1. *w*Mel-like *Wolbachia* that diverged approximately 94 KYA occupy insects that diverged about 350 MYA. (A) An absolute chronogram with the *Wolbachia* associated with the most distantly related hosts in bold. The colored *Wolbachia* labels match the host clades from Panel C. Horizontal gray bars are 95% credible intervals of divergence times. The crown age is 128 KY, with 95% credible interval of 263 to 51 KY*. **(B)** An approximate chronogram for the most distantly related host clades containing *w*Mel-like *Wolbachia*: Hymenoptera (*Diachasma alloeum*) and within Diptera, Diopsidea (*Spyracephala brevicornis*) and Drosophilidae (*D. teissieri* presented as a representative drosophilid). Diptera and Hymenoptera diverged about 350 MYA (378–329 MYA, Devonian–Carboniferous) and span Holometabola (Misof et al. 2014; Johnson et al. 2018). Drosophilidae and the Diopsoidea superfamily containing Diopsidea stalk-eyed flies diverged about 59 MYA based on the crown age of the Drosophilidae (47 MY) (*24*), and the crown age of Schizophora (70 MY) (*66*). **(C)** A chronogram for drosophilid hosts with node ages and approximate confidence intervals estimated from the fossil-calibrated chronogram of Suvorov et al. (2022) (*24*). Images taken by Centre for Biodiversity Genomics (*D. alloeum*), Katja Schulz (*S. brevicornis*) and Tim Wheeler (*D. teissieri*).

Figure 2

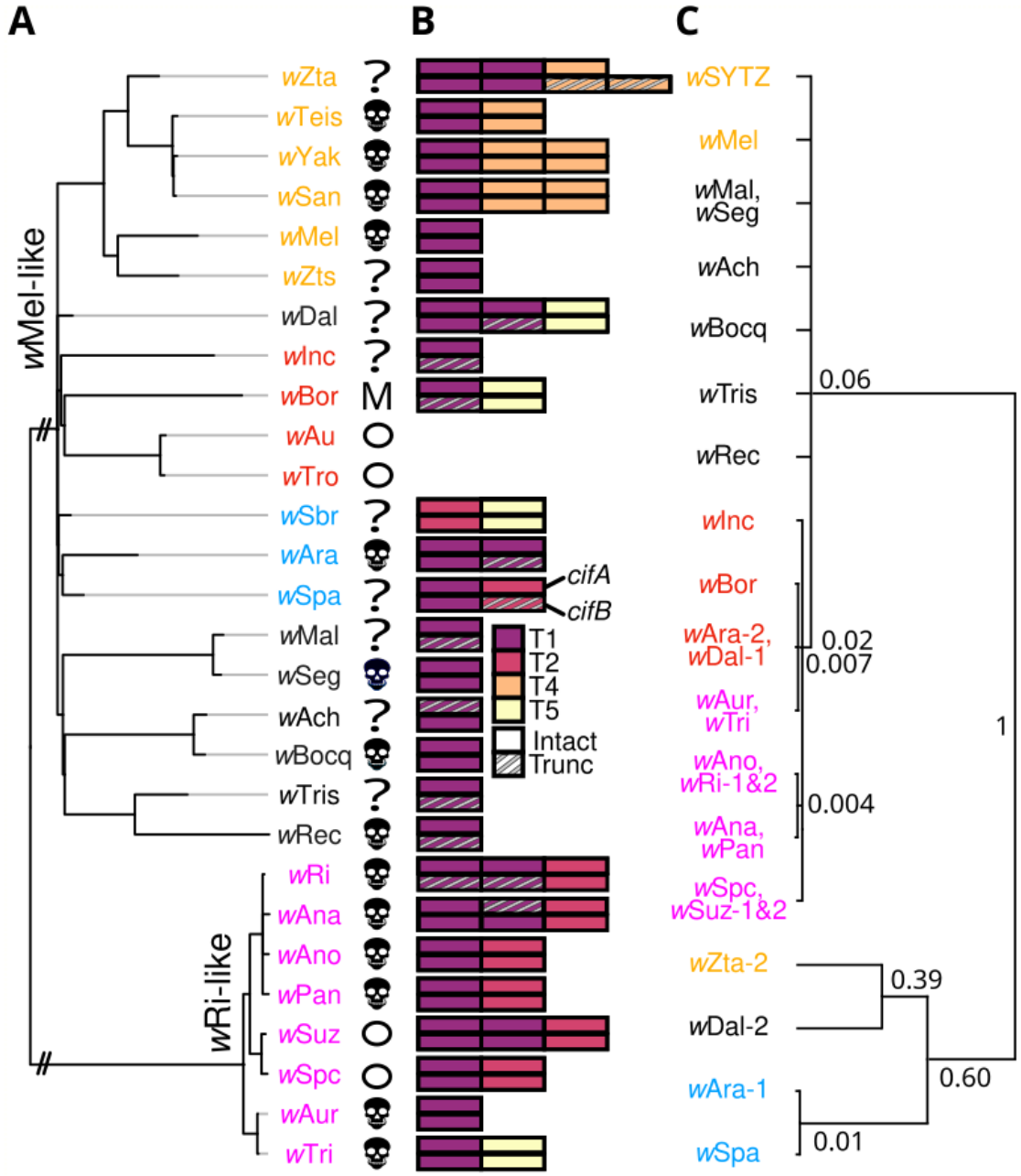

Fig. 2. Diverse *cif* operons rapidly turnover among *w*Mel-and *w*Ri-like genomes. (A) A phylogram of *w*Mel-like and *w*Ri-like *Wolbachia*, including variants that cause CI (💀), do not cause CI (circles), or have unknown CI status (?). *w*Bor does not cause CI, but it does kill males (M). The *w*Mel-like and *w*Ri-like clades diverged about 1.1 MYA (see SI). Branches leading to these clades are shortened (//) and light gray branch extensions are used to improve visualization. **(B)** These closely related Wolbachia carry four of five known *cif* operon Types (T1–T5). *cifA* (top) and *cifB* (bottom) schematics are presented with operon copies adjacent to one another. **(C)** A relative chronogram for $cifA_{[T1]}$ with node labels indicating relative ages, scaled to 1 for the most diverged. Identical sequences are collapsed into a single tip, and nodes with posterior probability < 0.95 are collapsed into polytomies. Strain labels are colored to highlight *Wolbachia-cifA* discordance.

Figure 3

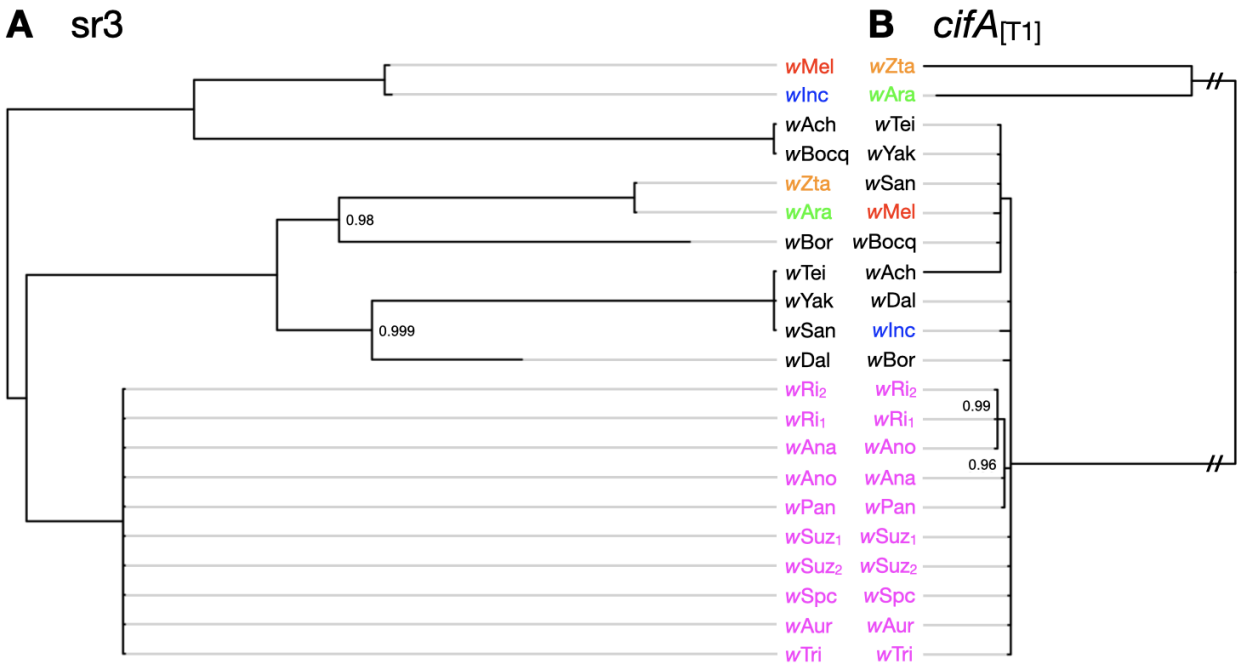

Fig. 3. Discordant phylograms for sr3 alleles of sr3WO and linked *cifA*$_{[T1]}$ alleles demonstrate phage-independent cif turnover. (A) Phylogram for sr alleles facing **(B)** phylogram for *cifA*$_{[T1]}$ alleles linked to these sr3 alleles. Branches leading to the two sets of closely related *cifA*$_{[T1]}$ alleles are shortened (//) to improve visualization. *w*Ri-like strains are shown in magenta, and focal *w*Mel-like strains are colored to highlight sr3-*cifA*$_{[T1]}$ discordance. Subscripts represent different sr3 copies within the same *Wolbachia* genome (see Table S3) in cases where multiple sr3 alleles can be associated with specific *cifA*$_{[T1]}$ copies. Subscripts presented for *cifA*$_{[T1]}$ alleles denote associated sr alleles. Light gray branch extensions are provided to simplify sr3-*cifA*$_{[T1]}$ comparisons. Nodes with posterior probability < 0.95 are collapsed into polytomies. Posterior support values appear only at nodes with support less than 1.

Figure 4

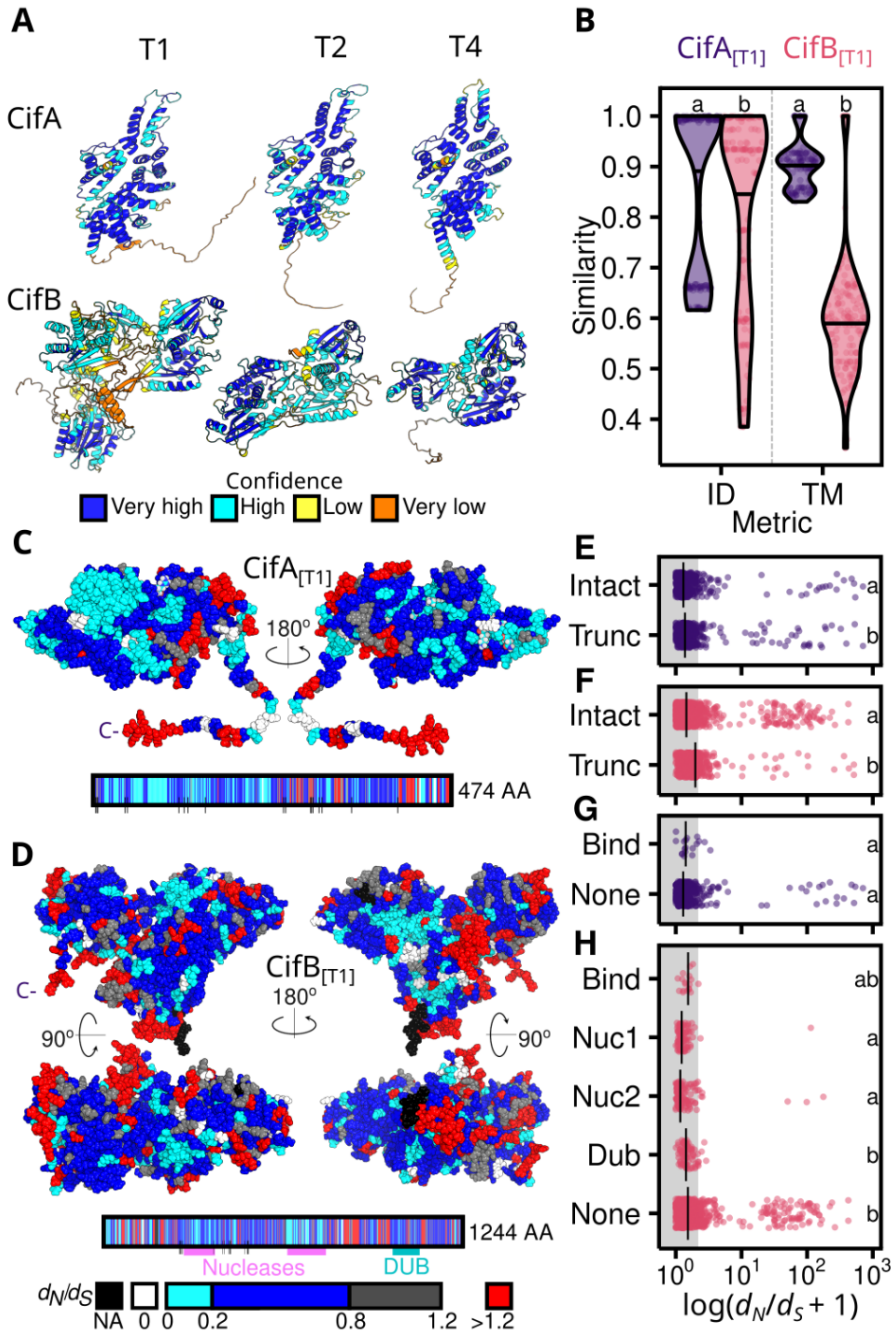

Fig. 4. Cifs are highly variable and CifA tends to have lower values of d_N/d_S than CifB. (A) Representative $Cif_{wMel[T1]}$, $Cif_{wRi[T2]}$, and $Cif_{wYak[T4]}$ AlphaFold structures colored by confidence (pLDDT) per residue. **(B)** Putatively intact $CifA_{[T1]}$ proteins are more similar than $CifB_{[T1]}$ from the same pairs ($N = 18$) in terms of both sequence identity (ID) and structural similarity (TM). d_N/d_S for **(C)** $CifA_{[T1]}$ and **(D)** $CifB_{[T1]}$ displayed on $Cif_{wMel[T1]}$ AlphaFold structures and linear schematics. Median d_N/d_S per residue is calculated using a 10 Å spherical sliding window in pairwise comparisons of $Cif_{wMel[T1]}$ to other $Cif_{[T1]}$ proteins (CifA: $N = 9$; CifB: $N = 4$). Colored boxes and black lines below schematics indicate domains and CifA-CifB binding sites, respectively. d_N/d_S is relatively lower for intact than truncated **(E)** $CifA_{[T1]}$ and **(F)** $CifB_{[T1]}$. **(G)** d_N/d_S for $CifA_{[T1]}$ binding sites tends to not differ from d_N/d_S for other residues. **(H)** d_N/d_S tends to be lower for $CifB_{[T1]}$ nuclease domains than for other residues. Shared letters within plots b and e-h represent statistically similar groups determined by a Mann-Whitney U test (2 groups) or a Kruskal-Wallis and Dunn's multiple comparison test (>2 groups). *P*-values are presented in Table S4.