cifB-transcript levels largely explain cytoplasmic incompatibility variation across divergent Wolbachia

J. Dylan Shropshire, Emily Hamant, William R. Conner and Brandon S. Cooper

Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA
Department of Biological Sciences, Lehigh University, Bethlehem, PA 18015, USA
*To whom correspondence should be addressed. Email: shropshirejd@lehigh.edu

Abstract

Divergent hosts often associate with intracellular microbes that influence their fitness. Maternally transmitted Wolbachia bacteria are the most common of these endosymbionts, due largely to cytoplasmic incompatibility (CI) that kills uninfected embryos fertilized by Wolbachia-infected males. Closely related infections in females rescue CI, providing a relative fitness advantage that drives Wolbachia to high frequencies. One prophage-associated gene (cifA) governs rescue, and two contribute to CI (cifA and cifB), but CI strength ranges from very strong to very weak for unknown reasons. Here, we investigate CI-strength variation and its mechanistic underpinnings in a phylogenetic context across 20 million years (MY) of Wolbachia evolution in Drosophila hosts diverged up to 50 MY. These Wolbachia encode diverse Cif proteins (100% to 7.4% pairwise similarity), and AlphaFold structural analyses suggest that CifB sequence similarities do not predict structural similarities. We demonstrate that cifB-transcript levels in testes explain CI strength across all but two focal systems. Despite phylogenetic discordance among cifs and the bulk of the Wolbachia genome, closely related Wolbachia tend to cause similar CI strengths and transcribe cifB at similar levels. This indicates that other non-cif regions of the Wolbachia genome modulate cif-transcript levels. CI strength also increases with the length of the host’s larval life stage, presumably due to prolonged cif action. Our findings reveal that cifB-transcript levels largely explain CI strength, while highlighting other covariates. Elucidating CI’s mechanism contributes to our understanding of Wolbachia spread in natural systems and to improving the efficacy of CI-based biocontrol of arboviruses and agricultural pests globally.

Keywords: Endosymbiosis, host–microbe interactions, reproductive parasitism, Drosophila, Wolbachia

Significance Statement:

Host–microbe endosymbioses are the most intimate species interactions, and among all endosymbionts, Wolbachia bacteria are the most common. Wolbachia prevalence in nature stems largely from its ability to hijack host reproduction to spread. While some Wolbachia feminize or kill males, cytoplasmic incompatibility (CI) is the most common manipulation in Wolbachia’s arsenal. CI strength varies widely among strains, and we report that testes-transcript levels of a single CI gene (cifB) largely explains this variation across 20 million years of Wolbachia divergence. We also report other factors that contribute to and modulate CI strength. Results reveal predictors of CI-strength variation across divergent Wolbachia–host systems, which is crucial to understanding and expanding CI-based biocontrol of arboviruses and agricultural pests on multiple continents.

Introduction

Endosymbioses are intimate associations where microbes live inside the cells of other organisms (1). This type of interaction led to the evolution of mitochondria and chloroplasts, and thus all eukaryotic life on Earth (2). Maternally transmitted Wolbachia are the most common endosymbionts, infecting over half of insect species (3). While some Wolbachia are required for host survival (4, 5), many manipulate host reproduction to spread to high frequencies (6–11).

Reproductive manipulations include male-killing, feminization, and parthenogenesis (12). However, cytoplasmic incompatibility (CI) is by far the most common manipulation (Fig. 1A), occurring in at least 10 arthropod orders (13). CI kills uninfected eggs fertilized by Wolbachia-infected males. Infected embryos are rescued from CI, providing them a relative fitness advantage that promotes Wolbachia spread to high frequencies in host populations (10, 11). Two genes associated with Wolbachia’s prophage WO—an integrated temperate bacteriophage in the Wolbachia genome—govern CI and rescue: CI factors A and B (cifA and cifB). In transgenic expression systems, cifB causes CI (14–20), and cifA rescues CI (16, 21). In some systems, co-expression of cifA and cifB in the testes is necessary for CI induction (15, 16, 20). Notably, strong CI directly enables biocontrol programs to transform naturally uninfected mosquito populations with pathogen-blocking Wolbachia and reduce population sizes of disease vectors and agricultural pests globally (22–24).

CI strength is highly variable within and among Wolbachia strains, ranging from complete to statistically insignificant concentrations in natural populations. This variation in CI strength influences Wolbachia persistence and spread, and CI is a promising weapon for biocontrol (16, 21). However, the drivers of CI strength across strains are not well understood

Competing Interest: The authors declare no competing interests.

Received: May 9, 2022. Accepted: June 30, 2022

PNAS Nexus, 2022, 1, 1–11
https://doi.org/10.1093/pnasnexus/pgac099
Advance access publication date: 28 June 2022

Research Report

Significance Statement:

PNAS Nexus

© The Author(s) 2022. Published by Oxford University Press on behalf of the National Academy of Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
CI strength varies widely among Wolbachia–Drosophila associations

Since CI results in embryonic lethality, we measured embryonic hatching and defined CI strength as $s_h$, where $s_h$ is equal to 1—CI-cross hatch rate/mean hatch of compatible crosses (10, 39). Estimates of $s_h$ were made using 3-day-old males that cause relatively strong CI (25, 40). Compatible crosses did not significantly differ in egg hatch for any strain ($P > 0.05$; Supplementary Fig. S1A). Only $w$Tei (95% BCA interval of $s_h = −8% to 31%$; $P = 0.11$) and $w$Tsa ($s_h = −5% to 25%$; $P = 0.55$; Fig. 1D) did not cause statistically significant CI. $w$Tei is known to cause very weak CI, although statistically nonsignificant hatch reductions are common and depend on both Wolbachia and host variation (6). The remaining eight strains

Results and discussion

Wolbachia phylogeny and host association

Our 10 focal Wolbachia include $w$Mel-like variants ($w$Mel of D. melanogaster and $w$Tei of D. teissieri), $w$Ri-like variants ($w$Ri of D. simulans, $w$Ano of D. anomalata, $w$Aur of D. auraria, and $w$Tri of D. triauraria), and several other Group-A strains ($w$Ha of D. simulans, $w$Boc of D. bokii, $w$Bai of D. baimai, and $w$Tsa of D. tsacasi) that infect divergent Drosophila host species. According to a Bayesian chronogram with absolute age estimates (156 genes; 130,359 bp), Group-A strains diverged from Group-B Wolbachia like $w$Mau in D. mauritiana 9.9 to 54 MY ago (Fig. 1B), agreeing with past estimates (37). Our 10 Group-A strains encompass 3.6 to 19.2 MY of evolution (Fig. 1B). These Wolbachia infect nine Drosophila species diverged 10 to 50 MY (Fig. 1C) (38).
**Shropshire et al.**

**Fig. 2.** CifA and CifB variation. (A) Similarity matrix displaying pairwise amino acid similarity for CifA proteins ranging from 100% to 15.4%, displaying considerable genetic diversity. (B) CifA (left) and CifB (right) protein architecture with HHpred annotations with $P > 80\%$ are displayed as large colored boxes. CifA had no confident annotations. The PD-(D/E)XK pair was the only feature represented in all CifB structures. (C) AlphaFold structures of CifA and CifB’s PD-(D/E)XK pair from four divergent Cif Types. Proteins are colored by pLDDT—a metric of structure confidence. pLDDT scores range from 0 to 100 where pLDDT > 90 is high confidence, 70 > pLDDT > 50 is low confidence, and pLDDT < 50 is very low confidence.

were classified into weak or strong, where strong-CI inducers had less than 20% egg hatch on average. wMel ($s_h = 9\%$ to 25%; $P = 3.5E-09$) and wBoc ($s_h = 7\%$ to 32%; $P = 2.4E-3$; Fig. 1D) caused weak but significant CI, while wRi ($s_h = 90\%$ to 95%; $P = 1.4E-14$), wAno ($s_h = 100\%$ to 100%; $P = 1.4E-8$), wAur ($s_h = 45\%$ to 93%; $P = 2.2E-4$), wTh ($s_h = 90\%$ to 96%; $P = 9.9E-12$), wHa ($s_h = 77\%$ to 93%; $P = 1.4E-12$), and wBai ($s_h = 92\%$ to 98%; $P = 7.4E-06$) caused strong CI (Fig. 1D).

**Cif proteins are highly diverged, and CifB evolves faster than CifA.**

Cifs are classified into five phylogenetic clades (Types 1 to 5) (41), each thought to contribute to CI (13, 18, 20). All focal Wolbachia genomes contained one to five CifA and CifB pairs, including at least one pair of Type 1 cifs in each genome (cif[T1]; Supplementary Fig. S1B). Type 2 (wRi and wAno), Type 3 (wTsa), Type 4 (wTei), and Type 5 (wTri) were also represented across strains (Supplementary Fig. S1B). CifA ($N = 21$) and CifB ($N = 19$) protein sequences were aligned, and pairwise identity was calculated for shared sites. CifA protein sequences ranged from 100–15.4% identity (Fig. 2A), where CifA[T1] copies shared between 100% and 60.8% identity with other CifA[T1] proteins and between 45.2% and 18.2% identity with CifA[T2–T5] proteins (Fig. 2A). Conversely, CifB sequences ranged from 100% to 7.4% identity, where CifB[T1] shared between 100% and 29.7% identity with other CifB[T1] proteins and between 26% and 7.4% identity with CifB[T2–T5] proteins (Supplementary Fig. S1C).

We annotated functional domains using HHpred (42). CifA had no hits above 80% probability—our threshold for accepting annotations (Fig. 2B; Supplementary Table S2). The lack of confident annotations for CifA indicates that CifA either has a novel function not currently represented in the database or, more likely, that it is too diverged from database proteins to be confidently annotated. However, CifB had six confident domain predictions (Supplementary Table S2): PD-(D/E)XK nuclease, Ulp1 protease or peptidase, Avirulence protein, OTU-like cysteine protease, Ankyrin/DNA-binding, and Latrotoxin (Fig. 2B). The last three domains were unique to CifB[T5], and the avirulence protein domain was unique to CifB[T1–T2]. The Ulp1 protease domain, known to be a functional deubiquitinase (14), was present in all CifB[T1] proteins and absent in other CifB Types. Deubiquitination activity does not appear to be essential for CI induction (32). Despite considerable sequence variation (Supplementary...
Fig. S1C), a pair of PD-(D/E)XK-nuclease domains was present in all CifB proteins—CifB<sub>T3</sub> and T4<sub>l</sub> are functional nucleases (18, 19), and CifB<sub>T3</sub> is a DNase with unknown catalytic residues (43). We conclude that Cifs are divergent in both sequence and domains but with several shared features. Indeed, the representation of the PD-(D/E)XK pair in all CifB Types suggests it is functionally important (13, 43) for CI and/or another unknown phenotype.

CifB sequence and structural similarities do not covary

Theory predicts that CI-causing genes like cifB are under weak purifying selection since they are expressed in males that do not transmit the infection, while rescue genes like cifA should be preserved (36). Indeed, genomic and phenotypic analyses of putatively pseudogenized Cifs suggest that CifB has a higher rate of putative pseudogenization than CifA (37, 41). For example, non-CI-causing wMau has a putatively pseudogenized CifB, intact CifA, and can rescue CI induced by closely related strains (37, 44). Consistent with these results, our CifA sequences (95% CI of the median = 57.5% to 63.9% identity) were significantly more similar than were CifB sequences (95% CI = 41.9% to 55.0%) according to a Wilcoxon matched-pairs signed-rank test (P < 0.0001). However, this result may also be explained by relaxed selection due to structural flexibility in the protein.

To test this hypothesis, we generated AlphaFold (45) CifA structures, and CifB structures surrounding the shared PD-(D/E)XK pair (CifB<sub>nuc</sub>), for representative Cif proteins from Types 1, 2, 4, and 5 (Fig. 2C). Type 3 was excluded because Type 3 wTsa was putatively pseudogenized (see below), and only the PD-(D/E)XK-nuclease pair of CifB (CifB<sub>nuc</sub>) was analyzed as it was shared across all Types. AlphaFold confidence, measured as a Local Distance Difference Test (<i>pLDDT</i>) score, suggests confident structural predictions (<i>pLDDT</i> = between 74.9 and 86.9; Supplementary Fig. S2; Supplementary Table S3). However, the C-terminus of all CifA protein predictions were low confidence and structurally disordered in the wMel, wRi, and wTsa proteins (Fig. 2C; Supplementary Fig. S2A–D). Low-confidence structures can be an indicator of conserved but represented by highly divergent sequences (47).

Cif pseudogenization does not explain non-CI-inducing strains

Two of our Wolbachia (wTei and wTsa) failed to cause statistically significant embryonic mortality in the CI crosses. We tested the hypothesis that non-CI inducers had pseudogenized Cif proteins (37, 41) by searching for premature stop codons that truncate the proteins (Fig. 2B). While wTei appeared to have normal Cif proteins, wRi, wHa, wBoc, and wTsa encoded at least one Cif copy that was putatively pseudogenized. Three CifA copies are truncated (wHa[T1], wBoc[T1], and wTsa[T3]), and four CifB copies are truncated (wRi[T1–1], wRi[T1–2], wHa[T1], and wTsa[T3]). Alternative upstream start codons exist for all but the pseudogenized CifB<sub>T2</sub> and T4<sub>l</sub>, which has a transposon inserted after the 348th nucleotide. We conclude that statistically nonsignificant wTei and wTsa CI cannot be explained by pseudogenization alone since wTei’s Cifs appear normal and wTsa has four putatively intact Cif<sub>T1</sub> pairs. Prior genomic and functional analyses of wYak-clade Wolbachia support our wTei result (6, 34). However, sequence variation may still contribute to inefficiencies in protein function. Indeed, transgenic expression and biochemical assays revealed that a single Valine to Leucine substitution in CifB<sub>wYak[T1]</sub>, a very close homolog of CifB<sub>bno[wT1]</sub>, significantly inhibits enzymatic activity relative to sister wMel (34). Finally, wBoc was the only strain that lacked an intact Cif pair. Two hypotheses could explain why wBoc induced weak CI despite pseudogenized Cifs. First, expression of the upstream coding sequence may be sufficient to yield a phenotype. Second, since complete circularized genomes were only available for wMel, wRi, and wHa, it is plausible that the wBoc genome contains additional unknown cif copies that our analyses did not detect.

Wolbachia densities in testes do not fully explain CI strength

The bacterial density hypothesis predicts that CI strength covaries with Wolbachia densities in testes (13), though there are exceptions where strong-CI-inducing strains have unusually low densities (50). We tested the bacterial density hypothesis by measuring Wolbachia densities in whole-testes extracts via qPCR (Fig. 3A) and performing categorical and qualitative correlation analyses with CI strength.

First, we predicted that strong-, weak-, and non-CI-inducing Wolbachia would have higher, similar, and lower Wolbachia densities in testes than model wMel that caused weak CI. However, only one strong-CI-inducing strain, wRi, occurred at statistically higher densities in testes than wMel (Fig. 3A). Three of the other strong-CI-inducing strains (wAur, wTri, and wHa) had similar densities, and two had lower densities (wBai and wAno) compared to wMel-infected testes. Contrary to our prediction, the other weak-CI line, wBoc, also occurred at lower densities in testes than wMel. Finally, non-CI-causing wTei and wTsa had similar and lower densities than wMel, respectively (Fig. 3A). Intriguingly, despite diverging from wRi on the order of thousands of years ago (Fig. 1B), strong-CI-inducing wAno and wAur occur at lower densities in their hosts, relative to wRi in D. simulans. In summary, only wRi and wTsa fit our predictions for categorical assignment (Table 1), which suggests that sequence similarity is not a good indicator of structural similarity. This is consistent with findings that PD-(D/E)XK domains are structurally conserved but represented by highly divergent sequences (47).
We next tested the bacterial density hypothesis using a phylogenetic generalized least squares (PGLS) regression that accounts for phylogenetic signal and dependencies within the residuals of our comparisons. Wolbachia densities in testes had a statistically nonsignificant relationship with CI strength among our focal Wolbachia ($R^2 = 0.03$, $P = 0.62$; Supplementary Table S4). However, we also subset our data by removing non-CI inducers ($w$Tei and $w$Tsa) and/or low-density strains ($w$Ano, $w$Boc, $w$Bai, and $w$Tsa), since they seemingly yielded stronger CI than expected from their densities. Removal of low-density Wolbachia revealed a strong positive relationship between Wolbachia densities and CI strength ($R^2 = 0.99$, $P = 2.56E−05$; Supplementary Table S4). Similarly, removal of non-CI strains recovered a strong correlation between densities and CI strength ($R^2 = 0.99$, $P = 9.74E−07$; Supplementary Table S4). We conclude that Wolbachia densities in the testes cannot fully explain CI-strength variation among our ten systems, but it trends with CI strength among a subset of strains. Notably, while Wolbachia densities measured via qPCR from whole-testes extracts is a poor predictor of CI strength across systems, it remains plausible that Wolbachia density in specific stages of spermatogenesis is a strong predictor (51).

**cifβ-transcript levels in testes largely explain CI strength**

The Wolbachia-density hypothesis is based on the presumption that CI-gene expression positively covaries with Wolbachia abundance in host tissues (13). Thus, we tested if cifA- and/or cifβ-transcript levels in whole-testes extracts covaried with CI strength using RT-qPCR. As with Wolbachia densities, we tested if cif-transcript levels and CI strength covaried via categorical and/or qualitative analyses. Relative to weak-CI-causing $w$Mel, we predicted that strong-, weak-, and non-CI strains would have higher, similar, or lower cif-transcript levels, respectively.

Regarding cifβ-transcript levels, seven of nine strains matched our predicted transcript levels in testes relative to cifB$_{w$Mel$[T1]}$. Five strong-CI strains ($w$Ri, $w$Ano, $w$Aur, $w$Tri, and $w$Ha) had copies with only higher transcript levels, the weak-CI $w$Boc had similar transcript levels, and the non-CI-inducing $w$Tsa had significantly lower transcript levels (Fig. 3C). Of the remaining two strains, strong-CI $w$Bai only had copies with significantly lower transcript levels, and $w$Tei that did not cause statistically significant CI had a copy with lower transcript levels (cifB$_{w$Mel$[T1]}$) and one with higher transcript levels (cifB$_{w$Mel$[T1]}$; Fig. 3C). Finally, PGLS regression revealed a significant positive relationship between cifB$_{w$Mel$[T1]}$-transcript levels and CI strength when low-density Wolbachia ($R^2 = 0.33$, $P = 0.04$) and non-CI-inducing strains were removed from the analysis ($R^2 = 0.89$, $P = 4.66E−03$; Supplementary Table S4). A single cifB copy (cifB$_{w$Mel$[T1]}$) was selected for analysis from each strain since Type 1 is the only cifB Type maintained across all strains.

Conversely, only five strains matched our predicted transcript levels pattern relative to cifA$_{w$Mel$[T1]}$ in testes: Three strong-CI strains ($w$Ri, $w$Ano, and $w$Tri) had at least one cifA copy with higher transcript levels, the weak-CI strain $w$Boc had similar transcript levels, and the strong-CI strain $w$Ha had copies with similar and/or lower transcript levels that likely amount to higher combined transcript levels (Fig. 3B). The remaining five strains had higher ($w$Tei and $w$Tsa) and lower ($w$Ano and $w$Bai) transcript levels than predicted based on their CI strength (Fig. 3B). A PGLS regression revealed no significant relationship between cifA$_{w$Mel$[T1]}$-transcript levels and CI strength (Supplementary Table S4). In summary, categorical analysis of cifA-transcript levels relative to $w$Mel suggests that CI strength can be explained by cifA-transcript levels in the testes of $w$Ri, $w$Ano, $w$Tri, $w$Boc, and $w$Ha but not $w$Tei, $w$Ano, $w$Bai, and $w$Tsa (Table 1). The quantitative analysis supports the disconnect between cifA$_{w$Mel$[T1]}$-transcript levels and CI strength but
does not account for the combinatorial impact of expressing multiple cifA copies. We also tested for a relationship between cif-
copy number and CI strength and found no significant relation-
ship ($R^2 = 0.11; P = 0.76$).

In our analyses reveal for the first time that cifB-transcript
levels in tests largely explains CI-strength variation. However,
two strains are interesting exceptions: uTei did not cause sta-
tistically significant CI, yet cifBwT1-transcript levels were rela-
tively high, and uBai caused very strong CI, but cifBwBaiT1-1 and
cifBwBaiT1-2-transcript levels were much lower than expected. We
propose four non-exclusive hypotheses for these patterns. First,
we measured transcript levels as a proxy for translation, but the
two can be decoupled (52). Thus, uTei may produce fewer proteins
per transcript than other strains, and uBai may produce more proteins per transcript. Second, we measured Wolbachia densities and cif-transcript levels from full-tests extracts, but Wolbachia localization within spermatogenesis is linked to CI-strength vari-
ation (51). It is plausible that uTei is poorly localized for CI in-
duction while uBai is optimally localized. Third, theory and em-
pirical studies agree that hosts develop resistance to CI (6, 36,
40). Since most of the strains investigated were from different
species, it is plausible—indeed, perhaps likely—that host factors
modulate CI-strength variation. Future work aimed at determin-
ing the contributions of host factors to CI strength and the rela-
tionship between cifB-transcript levels and CI strength will be
particularly useful. Finally, cif-genetic diversity may contribute
to variation in CI efficiencies. This hypothesis appears particularly
likely for uTei since a single amino acid substitution between CifBwMelT1 and a CifBwTeiT1 homolog in wYak inhibits transgenic
enzymatic activity and CI penetrance (34). CifBwBaiT1-2 uniquely
encodes a domain with structural homology to the cysteine pro-
tease avirulence protein AvrPphB of Pseudomonas (Fig. 2B). AvrP-
phB cleaves host kinases and inhibits Arabidopsis defenses against
Pseudomonas colonization (53). Thus, the CifBwBaiT1-2 avirulence
domain may interfere with host CI-resistance mechanisms (54).
However, the molecular function of the AvrPphB-like residues re-
tained in CifBwBaiT1-2 and the means of CI suppression are un-
known. Future work will be necessary to determine the relative
contributions of translation, localization, and genetic variation to
CI caused by these extraordinary strains.

### Strong CI and cifBwT1-transcript levels show evidence of Wolbachia phylogenetic signal

We first calculated Fritz and Purvis’ D statistic (55), which assesses
phylogenetic signals of binary traits. $D = 1$ represents a randomly
distributed trait across the phylogeny, while $D = 0$ suggests the
trait is clumped as if evolving under Brownian motion (i.e. phy-
logenetic signal). $P_1$ and $P_2$ indicate the probability that the $D$
value differs from $D = 1$ and $D = 0$, respectively. Causing strong CI
($D = -0.70; P_1 = 0.01; P_2 = 0.78$; Supplementary Fig. S3A) and
high-cifBwT1 transcription ($D = -0.26; P_1 = 0.02; P_2 = 0.62$; Sup-
plementary Fig. S3B) showed evidence of phylogenetic clumping.
Conversely, causing CI, having low-Wolbachia densities, or express-
ing high levels of a cifA[T1] copy was randomly distributed across
the phylogeny (Supplementary Fig. S3C–E). However, Geiger simu-
lations suggest that the inclusion of more Wolbachia-infected sys-
tems may help differentiate our $D$ values for low-Wolbachia den-
sity, cifA[T1]-transcript levels, and CI induction from randomness
(Supplementary Fig. S3), although the sample sizes required are
notably unreasonable. For non-proportional data (i.e. CI-strength
measures), we also calculated the maximum-likelihood value of
Pagel’s lambda ($\lambda$) (56), using the continuous measures of our
traits ($\lambda = 0$ indicates no phylogenetic signal). This approach
supported the finding that cifBwT1-transcript levels have phylogenetic
signal ($\lambda = 0.88; P = 0.03$), and that cifA[T1]-transcript levels ($\lambda =
0.68; P = 0.73$), and Wolbachia density ($\lambda = 0.84; P = 0.12$) do not.in
summary, these results suggest that Wolbachia that cause strong
CI and have high cifBwT1 transcription tend to be closely related.
We hypothesize that non-cif regions of the Wolbachia genome
modulate cifB-transcript levels. Future work is required to deter-
mine if this is via active modulation of cifB expression or the in-
direct effect of loci that regulate Wolbachia abundance in host
hierarchies.

### Larval, but not pupal, development time positively covaries with CI strength

Cardinium-induced CI strength (33) positively covaries with the
length of Encarsia wasp pupation time. Conversely, longer
development time covaries with weaker CI in uMel-infected
D. melanogaster (28). Here, we test for the first time whether
Wolbachia-induced CI strength across numerous Wolbachia-
Drosophila systems covaries with larval (Supplementary Fig. S4A),
pupal (Supplementary Fig. S4B), and/or egg-to-adult development
times (Supplementary Fig. S4C). A PGLS regression revealed a
significant positive relationship between CI strength and the
length of the larval life stage ($R^2 = 0.47; P = 0.041$), but not with
the length of pupation ($R^2 = 0; P = 0.98$) or with overall egg-
to-adult development time ($R^2 = 0.21; P = 0.22$; Supplementary

---

**Table 1. Summary of CI-strength, Wolbachia-density, and cif-transcript level results.**

<table>
<thead>
<tr>
<th>Wolbachia of host</th>
<th>CI strength</th>
<th>Wolbachia density</th>
<th>cifA level</th>
<th>cifB level</th>
</tr>
</thead>
<tbody>
<tr>
<td>uMel of D. melanoagaeter</td>
<td>Weak</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>uTei of D. tessereri</td>
<td>None</td>
<td>Medium</td>
<td>Medium/Medium</td>
<td>Low/High</td>
</tr>
<tr>
<td>uRi of D. simulans</td>
<td>Strong</td>
<td>High</td>
<td>High</td>
<td>High/High</td>
</tr>
<tr>
<td>uAno of D. anomala</td>
<td>Strong</td>
<td>Low</td>
<td>High/Medium</td>
<td>High/High</td>
</tr>
<tr>
<td>uAur of D. araria</td>
<td>Medium</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>uTri of D. triaria</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>uHa of D. simulans</td>
<td>Strong</td>
<td>Medium</td>
<td>Low/High*</td>
<td>Low/High</td>
</tr>
<tr>
<td>uBoc of D. bokci</td>
<td>Weak</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>uBai of D. baimai</td>
<td>Strong</td>
<td>Low</td>
<td>Medium</td>
<td>Low/Low</td>
</tr>
<tr>
<td>uTsa of D. tsacasi</td>
<td>None</td>
<td>Medium</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

Note: High, medium, and low density/cif transcription is defined as higher, similar, and lower than the weak CI-causing uMel. We predicted that Wolbachia density/cif transcription would be higher, similar, and lower than uMel for strong, weak, and non-CI inducers. Measures consistent with these predictions are bolded.

*Co-expression of uHa’s two cifA copies may yield higher total expression than cifAvMel[T1], but neither copy alone had higher transcript levels.
Table S5; Supplementary Fig. S4D–F). Removing low-density Wolbachia strains resulted in an even stronger relationship between larval development and CI strength ($R^2 = 0.86$, $P = 0.02$) and revealed a weak yet statistically significant relationship with egg-to-adult development time ($R^2 = 0.78$, $P = 0.048$; Supplementary Table S5; Supplementary Fig. S4G–I), likely driven by the larval pattern. In summary, we find that larval development time is a predictor of Wolbachia-induced CI across systems. We present two hypotheses for this correlation. First, as has been proposed for Cardinium (57, 58), longer development may enable CI products to accumulate or provide more time to act on their host targets in specific life stages. Second, development time may covary with other contributors of CI-strength variation, including the accessibility of host targets. Future investigations of the developmental restrictions of CI will be necessary to determine the mechanistic basis of the relationship between host development time and CI strength. Moreover, while these data support a relationship between development time and CI strength across systems, it remains to be determined if slow and fast developing males from the same Wolbachia–Drosophila system cause different CI strengths.

Conclusions

By leveraging CI diversity across 20 MY of Wolbachia and 50 MY of host evolution, we discovered that cifB-transcript levels largely explain CI-strength variation. This result confirms long-held predictions and enables future in-depth mechanistic investigation of the dynamic relationships between transcript levels, density, and localization on CI strength. Our analyses reveal five additional impactful findings. First, our focal Wolbachia encode considerable Cif-sequence variation, with some strains having as little as 7.4% pairwise similarity. Second, CifB sequence evolves faster than CifA, but the similarity in the protein structures of CifB nuclease pairs does not covary with CifB sequence similarities. Third, Wolbachia densities and cifA-transcript levels in testes fail to explain CI strength broadly. Fourth, CI tends to be stronger when larval development time is longer. Finally, cifB and strong-CI expression are similar among closely related Wolbachia, suggesting that other regions of the Wolbachia genome modulate cif-transcript levels, either actively or indirectly via effects on Wolbachia abundance in host tissues. This latter finding is particularly notable given that cifs are associated with prophage-WO regions that tend to be phylogenetically discordant from the bulk of the Wolbachia genome. Taken together, these results inform the mechanism and evolution of CI, bringing us closer to developing a complete understanding of CI-strength variation and how it governs Wolbachia’s standing as the world’s most common endosymbiont.

Materials and methods

Wolbachia chronogram

Genomes used are listed in Supplementary Table S6 (6, 37, 38, 59–61). We used Prokka v. 1.11 to annotate genomes, then extracted all single-copy genes that were the same length in each genome for phylogenetic analyses. 156 genes met these criteria. The genes were aligned with MAFFT v. 7 (62), then concatenated. RevBayes v. 1.1.1 was used for all phylogenetic analyses (63). We generated a Wolbachia absolute chronogram using the same GTR+I model as (38), except that we used a relaxed clock with a branch rate prior of $\Gamma(7,7)$ for each branch, normalized to a mean of 1 across all branches. Briefly, we assumed a constant-rate sampled-birth-death process, which has four parameters: speciation rate, extinction rate, sampling probability, and the age of the root. We used previously reported speciation and extinction rate estimates based on empirical information (38). Specifically, we previously specified empirical lognormal hyperpriors to determine the means of these distributions so that the prior expected number of species under the birth–death process was equal to the known number of species in the group. The sampling fraction, $\rho$, was set to 0.1. To estimate absolute node ages, we used prior estimates of substitutions per site per year (64). We refer readers to (38) for a detailed description of these methods.

Host phylogeny

We constructed a host phylogeny with 20 nuclear genes (aconitase, aldolase, bicoicid, ebony, enolase, esc, g6pdh, glyp, glys, ninaE, pepck, pgk, pgm, pic, ptc, tpi, transaldolase, white, wingless, and yellow). Coding sequences for these genes in D. melanogaster and D. simulans were obtained from FlyBase. Orthologs in D. teissieri (65), D. auraria and D. triauraria (66), D. biimai, D. bocki, and D. tsacasi (67), and D. anomala (38) were obtained with BLAST using the D. melanogaster sequences. The resulting sequences were aligned with MAFFT v. 7 (62) and hand curated to remove fragments of introns that could cause frameshifts.

We used these genes to estimate a phylogeny using RevBayes v. 1.1.1, following (38). We used a GTR+I+F model with four rate categories, partitioning by gene and codon position. Each partition had an independent rate multiplier with prior $\Gamma(1,1)$ [i.e. Exp(1)], as well as stationary frequencies and exchangeability rates drawn from flat, symmetrical Dirichlet distributions [i.e. Dirichlet(1,1,1,…)]. The model used a uniform prior over all possible topologies. Branch lengths were drawn from a flat, symmetrical Dirichlet distribution and thus summed to 1. Since the expected number of substitutions along a branch equals the branch length times the rate multiplier, the expected number of substitutions across the entire tree for a partition is equal to the partition’s rate multiplier. Four independent runs were performed, and all converged to the same topology. For additional details on the priors and their justifications, consult (38).

Fly lines, care, and maintenance

Fly lines used are listed in Supplementary Table S6. Uninfected flies were derived via tetracycline treatment in prior studies (15, 68). Tetracycline cleared lines were used in experiments over a year after treatment, avoiding the effects of antibiotic treatment on mitochondria metabolism and density (69). Infection status was confirmed with PCR of the Wolbachia surface protein gene. An arthropod-specific 28S rDNA was amplified as a control for DNA quality (6, 37). DNA was extracted for symbiont checks using a squish buffer protocol, described in (25). Flies were reared in vials with 10 mL of food made of cornmeal, dry corn syrup, malt extract, inactive yeast, soy flour, and agar (25). Fly stocks were maintained at 23°C before and during experiments. Flies were anesthetized using CO2 for virgin collections and dissections. During hatch-rate assays, flies were mouth aspirated between vials.

Hatch-rate and egg-lay assays

Four crosses were performed for each strain (Fig. 1A). The one exception was wAur-infected D. auraria, where only crosses with infected males were performed as uninfected males were sickly and failed to mate in pairs. Virgin 6–8-d-old females were added to vials containing a spoon filled with fly food—one fly was added per vial. Food for these assays was the same as standard rearing.
media but with blue food coloring and 1% extra agar. Fresh yeast paste (3:2 water to yeast) was smeared on the food. After 4.5 h of acclimation, a single 3-d-old virgin male was added to each vial and then incubated overnight. In the morning, flies were aspirated into new vials with a fresh spoon. Vials were incubated for another 24 h before flies were removed via aspirating. Eggs were counted on spoons immediately after flies were removed, and the remaining unhatched eggs were counted after 48 h. The hatch rate was calculated as the proportion of eggs hatched per spoon in 48 h. CI strength ($s_c$) was measured for each sample as 1—(CI-cross hatch rate/mean hatch of compatible crosses) (10). 95% BCa confidence intervals and means of CI strength were calculated using boot in R (70). We used Mann–Whitney U tests in R to determine if hatching differed between CI and compatible crosses for each strain. A Kruskal–Wallis followed by Dunn’s test was performed to determine if any cross types differed. Samples with fewer than ten eggs laid were excluded from hatch-rate analyses.

CI strength is amenable to variation in male age (25, 40), rearing density (28), paternal grandmother age (71), and temperature (72). To enable comparison among different Wolbachia–host associations, we controlled these factors. Virgin males were aged for 3 d before pairing with females since younger D. scosca and D. bocki males failed to mate and lay. Flies were given only bocki 3 d before pairing with females since younger citations, we controlled these factors. Virgin males were aged for (72). To enable comparison among different (74, 75). Second, the HHpred webserver (42) was used to identify between Cif sequence and structural similarity was assessed conducted twice—switching the order of template and exper-

Cif sequence and structure analyses

cif sequences were retrieved from the genomes of all strains used in this study with BlastP. Representative cifA and cifB genes from each of the five phylogenetic Types of cif genes were used as query sequences, based on previous assignments (41). cifA_{Dm}{T1}, cifB_{Bft}{T2}, cifA_{BF}{T4}, cifA_{Bf}{T4}, cifB_{BF}{T4}, cifB_{Bf}{T4}, and cifB_{Bft}{T2} were assigned a Type designation based on their closest hit from the query sequences and the Type of their nearby cifA or cifB partner. Glimmer 3 was used to identify the coding region surrounding each blast hit in Genieous Prime and then translated (73).

Three methods were used to characterize Cif proteins further. First, MUSCLE alignments for CifA and CifB were generated in Geneious Prime (73) and were used to create heatmaps of pairwise protein similarity in R using the corplot package (74, 75). Second, the HHpred webserver (42) was used to identify protein domains. Databases used were SCOPe70_2.07, Pfam-A_v33.1, COG_KOG_v1.0, and SMART_v6.0. Annotations with 80wererecorded. If multiple hits were retrieved for a single region, only the highest probability annotation was taken and used for annotation.

Finally, ColabFold (76), an AlphaFold2 (45) Google Collaboratory notebook, was used to generate tertiary structures for CifA and CifB’s PD-(D/E)XXK domains. Full CifB proteins were not generated due to restrictions on protein length and computational power. Tertiary structure similarity was determined using the Zhang-lab TM-score webserver (48). Each pairwise comparison was conducted twice—switching the order of template and experimental structures—and the average was taken. The relationship between Cif sequence and structural similarity was assessed using a Wilcoxon matched-pairs signed-rank test in Graphpad Prism 8.

Tissue collection and nucleotide purification

Siblings from hatch-rate assays were collected for DNA and RNA extractions for Wolbachia-density and cif-transcript-level assays. All tissue was collected the day after crossing for the hatch-rate experiment. Virgin males were CO₂ anesthetized, and testes were dissected in chilled phosphate-buffered saline. Five pairs of testes were placed into a single tube for DNA extractions and stored at −80°C until processing. The tissue was homogenized, and DNA was extracted with the DNeasy Blood and Tissue kit (Qiagen). For RNA extractions, 15 pairs of testes were placed into a single tube containing 200 μL of Trizol and four 3 mm glass beads. Samples were then homogenized using a TissueLyser II (Qiagen) at 25 Hz for 2 min, centrifuged, and stored at −80°C until processing.

RNA samples were thawed, 200 μL of additional Trizol was added, and tissue was further homogenized at 25 Hz for 2 ml. RNA was extracted using the Direct-Zol RNA Miniprep kit (Zymo Research) following the manufacturer’s recommendations, but with an added wash step. On-column DNase treatment was not performed. The “rigorous” treatment protocol from the DNA-free kit (Ambion) was used to degrade DNA in RNA samples. Samples were confirmed DNA-free using PCR and gel electrophoresis for an arthropod-specific 28S rDNA (6, 37). The Qubit RNA HS Assay Kit (Invitrogen) was used to measure RNA concentration. Samples were diluted to 20 ng/μL, 2 μL of TATAA Universal RNA Spike I (tatabiocenter) was added to each sample, and 16 μL of the sample was converted to cDNA using SuperScript IV VILO Master Mix (Invitrogen).

Relative-abundance and gene-transcript-level assays

Abundance and transcript-level assays were performed on testes nucleotide extracts. Wolbachia density was measured via qPCR as the relative abundance of the Wolbachia gene ftsZ and the single-copy gene mid1, which is conserved across the Drosophila genus (25, 77). cif-A- and cif-B-transcript levels were measured via RT-qPCR as the relative abundance of cif target to a spike-in control (described below). Since csfA are highly diverse, eight cifA primer sets and ten cifB primer sets were designed to capture the transcript levels of all variants encoded by our 10 focal Wolbachia. cifB primers were designed using Primer3 in Geneious Prime (73) (Supplementary Table S7). All primers were efficient for relevant hosts and Wolbachia, except Type 3 wtZs primers, which were not adequately tested due to insufficient transcript levels.

In RT-qPCR, Cq values are commonly normalized to an endogeneously expressed gene to control for variation in RNA quality, reverse transcript levels, amplification, and pipetting. Endogenous controls must be consistently expressed across treatment groups to be valid. A single gene is unlikely to meet these criteria since we analyze transcript levels across nine Drosophila species. As such, we opted to normalize our qPCR results against an RNA spike-in control added to each sample after RNA dilution and prior to reverse transcript levels (described above). The spike-in control accounts for variation in reverse transcript levels, amplification, and pipetting that occurs after the spike-in is added to the sample and is considered equivalent or better than endogenous controls in some cases (78–80).

Samples were tested in triplicate using Powerup SYBR Green Master Mix (Applied Biosystems). Abundance FC was calculated as $2^{-\Delta\Delta Ct}$ of ftsZ and mid1 relative to a random uMel sample.
Transcript levels FC was calculated as $2^{-\Delta\Delta C_t}$ of the cif target relative to the spike-in control compared to the average transcript levels of four reference samples with four unique cif copies: cifuMel[T1], ciftuBai[T2], cifuAur[T3], and cifuTri[T5]. FDR-corrected pairwise t tests were used to determine if Wolbachia density differed between systems. Analyses excluded samples with a Cq standard deviation exceeding 0.4 between triplicate measures.

**Development assays**

We measured the developmental timing of fly species used in our study, in parallel. Stock bottles were cleared of flies, and all flies were collected after three days and held in a vial with fresh food for 24 h. Ten non-virgin Wolbachia-infected females were moved to a vial with a spoon with food, as described above for hatch-rate assays. Vials were monitored to determine the number of eggs laid and hatched every 3 h between 8 AM and 8 PM. After 20 to 30 eggs were laid, females were removed from vials, and the spoon was transferred to a vial containing 10 mL of fresh food. After eggs hatched, pupation was recorded every 6 h and adult emergence were recorded once daily. A vial was monitored until all adults emerged or there was no emergence for three days. Larval, pupal, and egg-to-adult development times were calculated as the time between the first egg hatch and the last larva pupation, the first pupation and the last adult emergence, and the first egg lay and last adult emergence per vial, respectively.

**Correlation analyses**

We tested whether Wolbachia density, cif-transcript levels, or development time correlated with CI strength. Each Wolbachia-host pair represented a single sample in each analysis. PGLS regressions were used to test for relationships between BCAs means of CI strength and the means of the traits above (density and transcript levels data were log2 transformed before taking the average) using caper in R (81). Analyses were performed with and without maximum-likelihood correction of $\lambda$ and Akaike information criterion values were generated for each model. In all cases, correcting for $\lambda$ yielded a preferred model, which we have reported.

**Phylogenetic signal**

We investigated whether any of these traits exhibit phylogenetic signals using our *Wolbachia* phylogeny and data for CI strength, Wolbachia density, and cif-transcript levels. For cif-transcript levels, only Type 1 cifs were included in analyses, and only a single allele was selected from strains with multiple Type 1 copies (cifuMel[T1] and cifuBai[T1-2]). Two methods were used to investigate phylogenetic signals.

First, we calculated Fritz and Purvis’ $D$ statistic using the caper R package, which estimates phylogenetic signal based on binary traits (55, 81). Binary traits were assigned as follows. Wolbachia density was low for uAno, uBai, uBoc, and uTsa and high for other strains. cifA-transcript levels were high for uAno, uRi, and uTri and low for other strains. cifB-transcript levels were high for uAno, uAur, uHa, uRi, uTei, and uTri and low for other strains. CI strength was treated as binary by either comparing strong CI (uAno, uAur, uBai, uHa, uRi, uTei) vs. all other strains or CI (strong-CI strains plus uMel and uBoc) vs. non-CI strains. The $D$ statistic compares the observed $D$ value to alternative $D$ values generated with simulated data based on a random phylogenetic pattern and Brownian motion using 1,000 permutations each. To investigate if larger phylogenies improve $D$ statistic estimation, we used the Geiger package in R (82) to simulate 100 permutations of the $D$ statistic for trees with 10, 25, 50, and 100 taxa ($\lambda = 1$ or 0) given the prevalence of our binary traits.

Next, we calculated Pagel’s lambda ($\lambda$), using the phytools package in R, which estimates phylogenetic signal based on continuous traits (56, 83). A likelihood ratio test was used to compare our fitted value of $\lambda$ to a model assuming no phylogenetic signal. Pagel’s $\lambda$ was not calculated for CI since CI data is proportional. We used the mean of log2-transformed data for relative abundance and transcript levels.

**Figure generation**

Figures were created using GGPlot2, and figure aesthetics were edited in Affinity Designer 1.8 (Serif Europe, Nottingham, UK).

**Acknowledgments**

We thank Michael Hague for assistance with phylogenetic analyses and Tim Wheeler for assistance in the laboratory. Comments by members of the Cooper lab and four anonymous reviewers improved earlier versions of the manuscript.

**Supplementary material**

Supplementary material is available at PNAS Nexus online.

**Funding**

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

**Authors’ contributions**

JDS: Conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, supervision, validation, writing—original draft, and writing—review and editing. EH: Investigation, validation, and writing—review and editing; WRC: Investigation, formal analysis, and writing—review and editing; and BSC: Conceptualization, funding acquisition, methodology, project administration, supervision, visualization, writing—original draft, and writing—review and editing.

**Data availability**

All data are made publicly available in the supplement of this manuscript.

**References**


