Georgia Journal of Science

Volume 80 *No. 2 Scholarly Contributions from the Membership and Others* | Article 8

2022

Elevated Substitution Rates Among Wolbachia-infected Mosquito Species Results in Apparent Phylogenetic Discordance.

James E. Russell
Georgia Gwinnett College, jrussell@ggc.edu

Michael Saum
Georgia Gwinnett College

Rebekah Williams
Georgia Gwinnett College

Elevated Substitution Rates Among Wolbachia-infected Mosquito Species Results in Apparent Phylogenetic Discordance.

Acknowledgements

We thank professors Wendy Dustman, Rebekah Ward, Fengjie Sun, and the Lawrenceville Science Tavern for useful discussion on research results.

Elevated Substitution Rates Among *Wolbachia*-infected Mosquito Species Results in Apparent Phylogenetic Discordance.

James E. Russell*
Michael Saum
Rebekah Williams
School of Science and Technology
Georgia Gwinnett College
Lawrenceville, GA, 30019
*corresponding author
jrussell@ggc.edu

ABSTRACT

As one of the most widely distributed bacterial cytoplasmic symbionts on earth, *Wolbachia pipientis* Hertig serves as a model organism for the understanding of host-symbiont interactions. Many mosquito species are infected with *Wolbachia* strains that induce a form of reproductive manipulation called cytoplasmic incompatibility, in which infected females gain a reproductive advantage over uninfected females in mixed infection populations. The selective advantage of cytoplasmic incompatibility often results in a population sweep of *Wolbachia* and co-transmitted mitochondrial genomes. Mitochondrial evolution and phylogenetic inferences drawn from mitochondrial gene sequences are thus potentially compromised by reproductive manipulating symbionts, like *Wolbachia*. Our initial analyses of phylogenetic patterns among collected *Wolbachia*-infected and uninfected mosquito species suggested significant *Wolbachia*-induced effects on mitochondrial evolutionary patterns. Discordant mitochondrial and nuclear phylogenies among *Aedes* and *Culex* species were associated with infections status, with a distinct mitochondrial clade of infected *Aedes* and *Culex* species, separate from uninfected species of the genera. Statistical analyses of molecular substitution among infected and uninfected sequence samples revealed elevated rates of substitution for the mitochondrial sequences of the discordant infected *Aedes/Culex* clade. Subsequent phylogenetic analyses suggested that the observed discordance could be attributed to long-branch attraction effects associated with elevated rates of substitution. Our results highlight the impact cytoplasmic selection can have on phylogenetic inference in limited sample sets with *Wolbachia*-infected and uninfected species.

Keywords: *Wolbachia*, mosquito, mitochondria, cytoplasmic incompatibility

INTRODUCTION

Wolbachia is an intracellular maternally inherited bacterium that frequently acts as a reproductive parasite by manipulating host biology in such a way that its future transmission is increased regardless of cost to the host (Stouthamer et al. 1999, Werren et al. 2008). While in some arthropods *Wolbachia* biases population sex ratios in favor of transmitting females, mosquitoes and other arthropods are affected by strains of *Wolbachia* that induce the process of cytoplasmic incompatibility. Cytoplasmic incompatibility (CI) *Wolbachia* prevents uninfected female mosquitoes from producing viable offspring with infected males while *Wolbachia*-positive females can reproduce with any male regardless of the male's infection status, thus giving infected females a reproductive advantage (Binnington and Hoffman 1989, O'Neill et al. 1992). *Wolbachia* strain specificity and superinfection status (individuals infected with more than one *Wolbachia* strain) can also result in relative reproductive advantages for infected females (Sinkins et al. 1995). The well-documented selective advantage of *Wolbachia*-infected females has resulted in rapid infection frequency increases among *Drosophila simulans* populations of Australia (Kriesner et al. 2013) and California (Turelli and Hoffmann 1991, Turelli and Hoffmann 1995) and many other species and populations (Rasgon and Scott 2003, Schmidt et al. 2017).

The combination of phenotypically–derived selection for the spread of *Wolbachia* and linkage disequilibrium between maternally-inherited *Wolbachia* and mitochondrial DNA results in indirect selection on mtDNA among infected populations. (Hurst and Jiggins 2005). Since infected females typically co-transmit mtDNA with *Wolbachia* the rapid spread of *Wolbachia* can result in a mitochondrial selective sweep (Richardson et al. 2012, Shoemaker et al. 2004, Turelli et al. 1992). *Wolbachia*-mediated selective sweeps have been implicated in the reduced mitochondrial genetic diversity of *Drosophila recens* Wheeler (Shoemaker et al. 2004), *Eurema hecabe* Linnaeus (Narita et al. 2006), *Eupristina verticillata* Waterston (Sun et al. 2010), *Coenonympha tullia* Müller (Kodandaramaiah et al. 2013), *Cydia fagiglandana* Zeller (Avtziz et al. 2014), and the mosquito species *Aedes albopictus* Skuse (Armbruster et al. 2003) and *Culex pipiens* Linnaeus (Atyame et al. 2011). The evolutionary impact of selective sweeps and reduced genetic variation is a limitation on the adaptive potential of mitochondrial gene pools resulting from reduced effective population size and elevated patterns of non-synonymous substitution (Cariou et al. 2017, Shoemaker et al. 2004). This is of particular relevance for disease vector mitigation efforts among mosquito species using *Wolbachia* as a gene drive mechanism (Hoffmann et al. 2011, Ruang-areerate and Kittayapong 2006, Xi et al. 2005).

Wolbachia is primarily sustained through vertical or mother-to-offspring transmission, though in many species horizontal transmission of *Wolbachia* has been implied from the isolation of similar strains in unrelated host species and from discrepancies between evolutionary histories of endosymbionts and hosts (Kikuchi and Fukatsu 2003, Schilthuizen and Stouthamer 1997, Werren et al. 1995, Zhou et al. 1998). Violation of the assumption of infected host species-*Wolbachia* coevolution together with indirect selection on mtDNA, can complicate phylogenetic analyses using mitochondrial molecular markers, like the *CO1* gene region (Hurst and Jiggins 2005). Reports of discordance between mitochondrial and nuclear phylogenies have been

attributed to divergent mitochondrial clades within species complexes associated with hybrid introgression of mitochondria and subsequent *Wolbachia*-mediated selective sweeps (Dyer et al. 2011, Jiggins 2003, Narita et al., 2006, Whitworth et al. 2007, Turelli et al. 2018). The utility of the *CO1* gene region as a reliable marker for species designation and phylogenetic analyses has been brought into question by such findings (Hurst and Jiggins 2005). However, Smith et al. (2012) found no evidence of mitochondrial *CO1* barcoding taxon misidentification associated with *Wolbachia* infection, noting the potential for such and recommending *Wolbachia* screening in DNA barcoding studies.

The aim of this study was to investigate the evolutionary impact of *Wolbachia* infections on mosquito species by analyzing the phylogenetic structure of the mosquito mitochondrial (*CO1* barcoding region), nuclear (*ITS2* gene region), and *Wolbachia* MLST gene regions *coxA* and *gatB* (Baldo et al. 2006). Violation of an *a priori* assumption of mitochondrial phylogenetic concordance among sampled species resulted in a post-hoc test of the hypothesis that *Wolbachia*-induced selection resulted in accelerated rates of evolution for mitochondrial gene sequences among infected species.

MATERIALS & METHODS

Collection and Identification of Specimens

Adult mosquitoes were collected during July-September of 2016 from four different locations across metro Atlanta, GA, USA and from May-August of 2017 from five different locations across metro Atlanta, using human landing catch and CO_2-baited CDC light traps (John W. Hock Co., Gainesville, FL) charged with 95% EtOH. Specimens were identified morphologically using dichotomous taxonomic keys (Burkett-Cadena, 2013) and were grouped by genus and/or species and location, then stored in 100% ethanol in a -40 °C freezer until DNA extraction was performed.

DNA extraction, amplification and sequencing

DNA was extracted from whole mosquitoes using a standard Qiagen DNeasy Blood and Tissue kit (Qiagen, Valencia, CA). PCR amplifications of partial mitochondrial cytochrome oxidase subunit one (*CO1*) sequences and nuclear gene *ITS2* sequences for all specimens were conducted using universal *CO1* primers HCO and LCO (Folmer et al. 1994) and universal *ITS2* 5.8S and 28S primers (Manonmani et al. 2001) (Supp. Table 1).

Wolbachia infection was assessed for all samples by PCR amplification of 16S rDNA genes using *Wolbachia*-specific primers (O'Neill et al. 1992). To further determine *Wolbachia* diversity, samples previously amplified for 16S were subjected to a PCR assay with primers for specific amplification of *coxA* and *gatB* (Baldo et al. 2006) (Supp. Table 1). For all reactions, PCR was carried out with an initial denaturation step at 95°C for 3 min followed by 30 cycles with a denaturation step at 94°C for 30 sec, elongation at 54°C for 30 sec, and extension at 72°C for 90 sec, then a final extension step at 72°C for 5 min. PCR products were electrophoresed on agarose gel stained with SYBR Green and visualized with ultraviolet light, after which PCR products matching

target gene size completed Sanger capillary sequencing at the Georgia Genomics Facility (Athens, GA, USA) and Eurofins Genomics Facility (Louisville, KY, USA).

Data analysis

All sequences for *CO1*, *ITS2*, *coxA* and *gatB*, were viewed using FinchTV 1.4.0 (Geospiza, Inc.; Seattle, WA, USA) and CHROMAS 2.6.6 (Technelysium Pty Ltd, South Brisbane, AU). GenBank's BLASTn tool (http://www.ncbi.nlm.nih.gov/BLAST/) was used to perform similarity searches for all sequences in order to support or reject morphological identification of specimens, as well as to retrieve reference sequences for subsequent alignment in MEGA 7 using MUSCLE with default parameters. Uncalled bases and alignment inconsistencies were resolved by analysis of chromatogram files and manual manipulation of sequence alignments.

Alignments for each gene were analyzed using neighbor joining, maximum parsimony and maximum likelihood methods in MEGA 7. Tree topologies were compared among phylogenetic methods for consensus. Maximum likelihood analyses for all sequence alignments, nucleotide and amino acid, were conducted using best fit substitution models based on lowest BIC, AICc, and highest log likelihood scores in MEGA 7.0. Concatenation of *coxA* and *gatB* gene sequences were phylogenetically analyzed using *Wolbachia* MLST (https://pubmlst.org/*Wolbachia*/) sequences from *Brugia malayi* Brug as an outgroup, with *Drosophila melanogaster* Meigen and *Muscidifurax uniraptor* Girault and Sanders as representative *Wolbachia* supergroup A samples and *Drosophila simulans* Sturtevant and *Trichogramma deion* Pinto and Oatman as representative *Wolbachia* supergroup B samples.

DNA polymorphism and population genetic analyses of neutral evolution were conducted using Tajima's D and Fu and Li's tests in DnaSP v6 (Rozas et al. 2017). Between-species analysis of molecular evolution was conducted through analyses of pairwise sequences comparisons, using *Drosophlia neotestacea* Grimaldi, James and Jaenike as an outgroup, with Tajima's Relative Rate test in MEGA and chi square analysis of unique nucleotide differences in R package rr.test. A one-sided *t*-test on the unique nucleotide differences (d) between infected and uninfected individuals in select *Aedes* and *Culex* genera was conducted with the following parameters: Null Hypothesis H0: $\mu_d \leq 15$, Alternative Hypothesis H1: $\mu_d > 15$; where μ_d is the species population mean of the differences (d) between unique nucleotides for Seq A(Infected) and Seq B (Uninfected) when compared with a common outgroup sequence. Since Seq A is always Infected and Seq B is always Uninfected, $d \geq 0$. For comparison, t-tests were also separately conducted within infected species (*A. albopictus* and *C. pipiens*) and uninfected (*A. vexans* and *C erraticus*) species assuming, H0: $\mu_d \geq 10$, Alternative Hypothesis H1: $\mu_d < 10$. The μ_d values represent higher and lower thresholds for determining differences between infected and uninfected species and similarities within infected and uninfected species, respectively.

Post-hoc analysis of phylogenetic discordance between mitochondrial (*CO1*) gene regions for select *Aedes* and *Culex* species was conducted using a Shimodaira-Hasegawa test (SH test) of tree topology with constraints for *Aedes* and *Culex* monophyly using IQ-TREE (Nguyen et al. 2015).

RESULTS

CO1 sequences were examined for 75 individual mosquitoes of 14 separate species and 7 genera. *Cx. pipiens* and *Culex quinquefaciatus* Say were considered a single species for phylogenetic analyses, hereafter referred to as *Cx. pipiens*, given across species monophyly and within species paraphyly of the taxa (Fig. 1). Among these sequences, 30 distinct mitochondrial haplotypes were identified as a result of 203 polymorphic sites considered parsimony informative. *Wolbachia* infection was detected via PCR assay using *Wolbachia* specific primers for the 16S *Wolbachia*-specific gene in 40 individuals (53.33%) representative of 6 species, with no discrepancy between these results for both primers (see Supp. Table 2). Infected individuals were comprised of only 12 mitochondrial haplotypes, with 1 single haplotype representing 23 *Aedes albopictus* individuals. All individuals from species in which Wolbachia was detected tested positive for Wolbachia, i.e. there were no infection frequencies less than one (Supp. Table 2).

Wolbachia positive individuals underwent additional PCR assay using strain-specific primers for the *Wolbachia coxA* and *gatB* genes. PCR assay for these genes was successful for a subset of infected individuals and concatenation of these sequences revealed 93 parsimony informative sites with 11 haplotypes. Maximum likelihood trees were created for all mitochondrial, *Wolbachia* concatenation, and *ITS2* sequences utilizing the General Time Reversible, Hasegawa-Kishino-Yano, and Jukes-Cantor models, respectively, with 1000 bootstrap pseudoreplications. Each prefix to species name in the phylogenies represents a unique haplotype, with the exception of the *A. albopictus* taxa in the *CO1* tree. The *A. albopictus* prefixes (Xn) are designated to represent number of individuals (n) with that unique haplotype.

The phylogenetic tree for all mitochondrial sequences was composed of five discernable clades (Fig. 1). Mosquitoes of the *Anopheles* genus formed a monophyletic clade (except *Anopheles rangeli* Galbodon, Cova, Garcia, Lopez) basal to all other sequences, followed by a larger monophyletic group containing the following defined clades: *Culex erraticus*, *Psorophora,* Infected *Aedes/Culex*, Uninfected *Aedes vexans,* and *Aedes/Ochlerotatus* (Figure 1). Four of the five clades correspond to recognized genera designations, with bifurcated *Aedes/Ochlerotatus* clades distinguished by *Aedes japonicus* Theobald/*Ochlerotatus triseriatus* Say sub-clades. It should be noted that *Ochlerotatus* is considered an *Aedes* sub-genus in some revisions of the Culicidae tribe Aedini (Reinert et al. 2004). The other clade, *Aedes/Culex* infected, grouped *Wolbachia*-infected species in the *Aedes* and *Culex* genera in a distinct clade bifurcated according to genus. The topological structure of the maximum likelihood tree was confirmed with identical clade designations in maximum parsimony and neighbor-joining phylogenetic analysis of the sequence data.

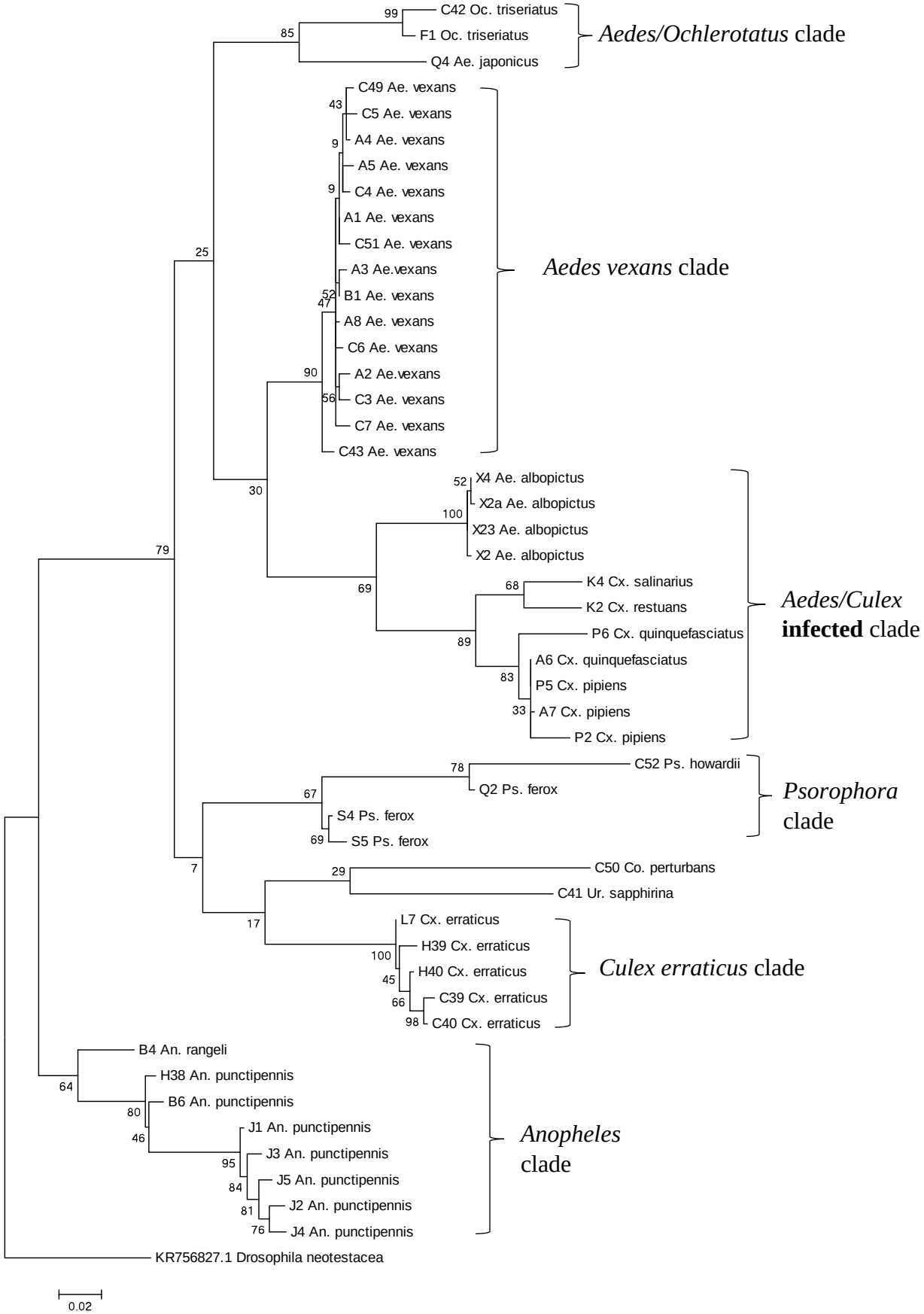
C42 Oc. triseriatus
F1 Oc. triseriatus
Q4 Ae. japonicus
Aedes/Ochlerotatus clade
C49 Ae. vexans
C5 Ae. vexans
A4 Ae. vexans
A5 Ae. vexans
C4 Ae. vexans
A1 Ae. vexans
C51 Ae. vexans
A3 Ae.vexans
B1 Ae. vexans
A8 Ae. vexans
C6 Ae. vexans
A2 Ae.vexans
C3 Ae. vexans
C7 Ae. vexans
C43 Ae. vexans
Aedes vexans clade
X4 Ae. albopictus
X2a Ae. albopictus
X23 Ae. albopictus
X2 Ae. albopictus
K4 Cx. salinarius
K2 Cx. restuans
P6 Cx. quinquefasciatus
A6 Cx. quinquefasciatus
P5 Cx. pipiens
A7 Cx. pipiens
P2 Cx. pipiens
Aedes/Culex infected clade
C52 Ps. howardii
Q2 Ps. ferox
S4 Ps. ferox
S5 Ps. ferox
Psorophora clade
C50 Co. perturbans
C41 Ur. sapphirina
L7 Cx. erraticus
H39 Cx. erraticus
H40 Cx. erraticus
C39 Cx. erraticus
C40 Cx. erraticus
Culex erraticus clade
B4 An. rangeli
H38 An. punctipennis
B6 An. punctipennis
J1 An. punctipennis
J3 An. punctipennis
J5 An. punctipennis
J2 An. punctipennis
J4 An. punctipennis
Anopheles clade
KR756827.1 Drosophila neotestacea
0.02

Figure 1. The evolutionary history for CO1 gene sequences of sampled mosquito species was inferred by using the Maximum Likelihood method based on the General Time Reversible model. The tree with the highest log likelihood (-5465.58) is shown. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 49 nucleotide sequences. There was a total of 609 positions in the final dataset.

Phylogenetic analysis of concatenated *Wolbachia coxA-gatB* gene sequences, using *Brugia malayi Wolbachia* as an outgroup recovered two distinct clades corresponding to Type A and B *Wolbachia* (Fig. 2). Within the *Wolbachia* B clade a monophyletic *Culex* clade was recovered. Within the *Wolbachia* A clade *Aedes* sequences form a basal paraphyletic root giving way to a monophyletic *D. melanogaster*, *M. uniraptor* and *Uranotaenia sapphirina* Osten Sacken clade, with *M. uniraptor* and *Ur. sapphirina* as sister taxa. Recombinant *Wolbachia* A and B strain alleles among *A. albopictus*, in particular *coxA*, have been previously reported (Shaikevich et al. 2019) and may be responsible for the paraphyletic structure of the *Ae. albopictus Wolbachia* gene sequences.

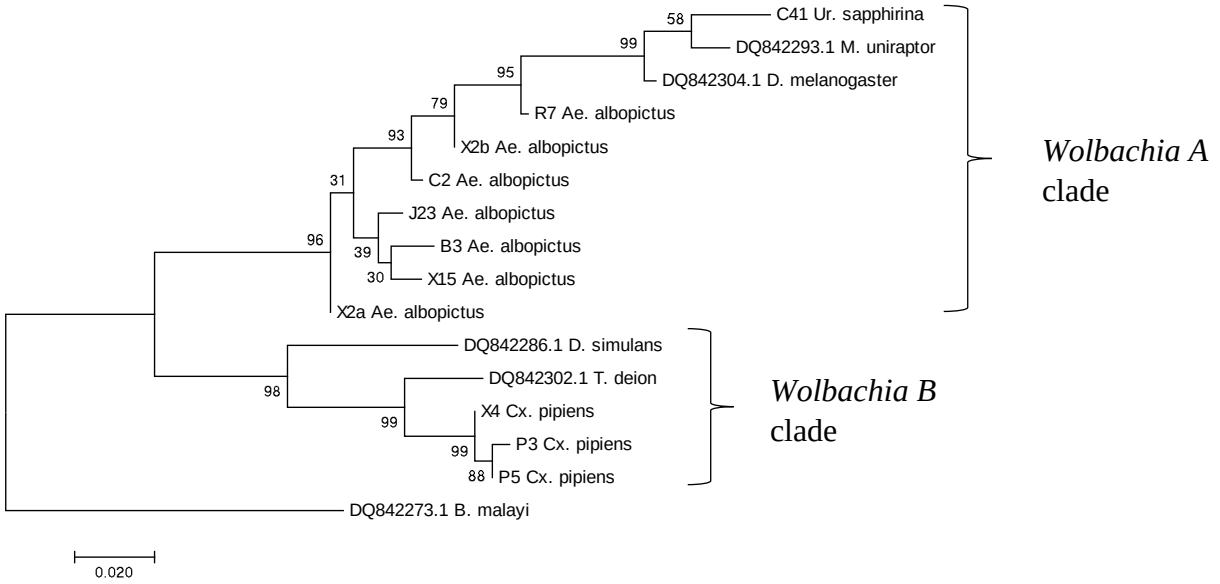

Figure 2. The evolutionary history for concatenation of the *Wolbachia* coxA-gatB gene sequences from sampled mosquito species was inferred by using the Maximum Likelihood method based on the Hasegawa-Kishino-Yano model. The tree with the highest log likelihood (-2227.02) is shown. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. The analysis involved 16 nucleotide sequences. There were a total of 698 positions in the final dataset. Accession numbers signify samples drawn from Genbank to root the tree and identify *Wolbachia* strain types.

Phylogenetic analysis of nuclear *ITS2* gene region (a limited subset of the mitochondrial data due to amplification failures) returns two distinct clades corresponding to the genera *Aedes* and *Culex*, with *Coquillettidia perturbans* Walker designated as an outgroup (Fig. 3). Within the *Aedes* clade strong support for two monophyletic groups are recovered for the two species, *Ae. albopictus* and *Ae. vexans*.

Likewise, within the *Culex* clade, strong support for the monophyly of the two *Culex* species, *Cx. pipiens* and *Cx. erraticus* was observed.

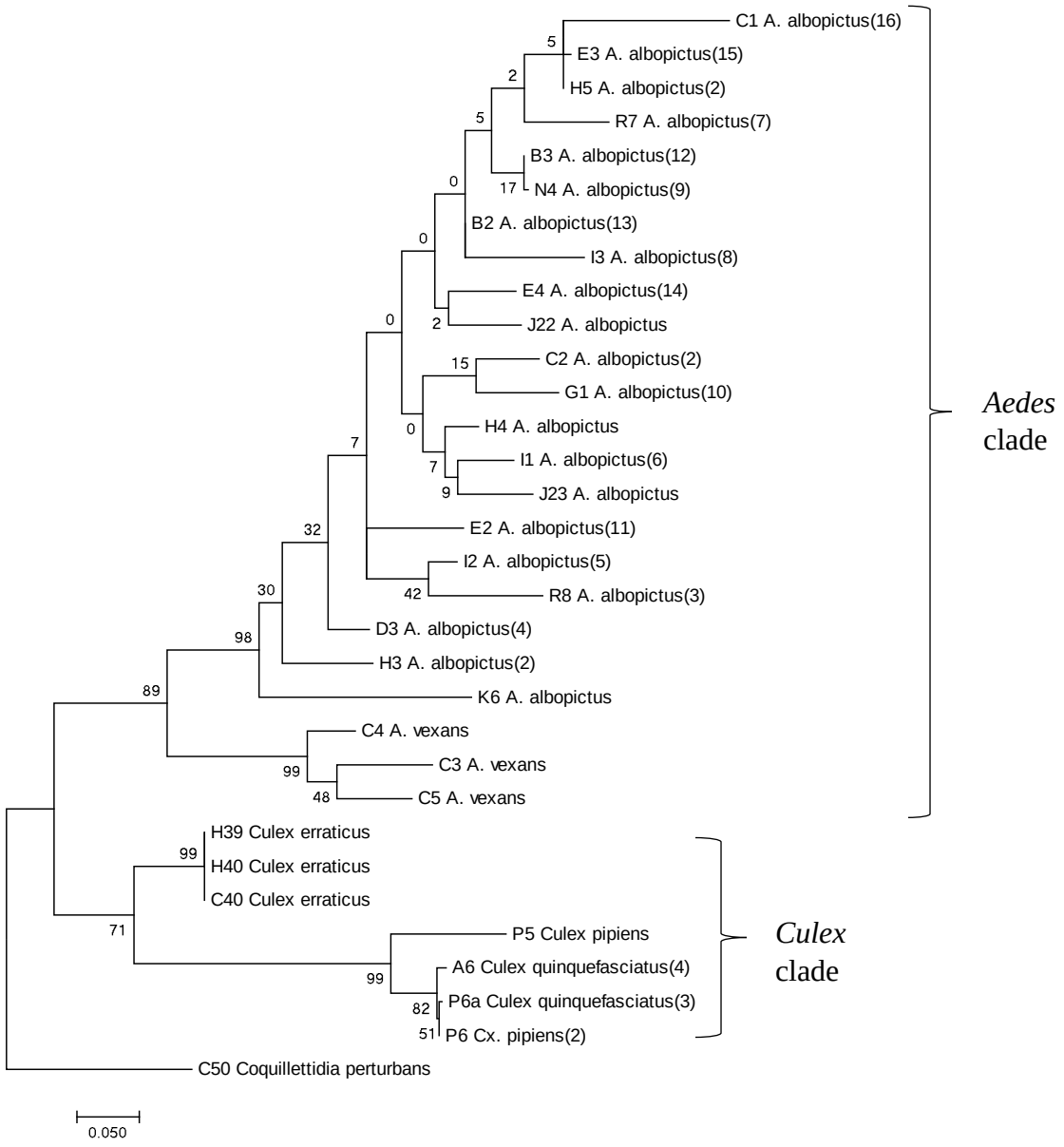

Figure 3. The evolutionary history for ITS2 gene sequences from sampled mosquito species was inferred by using the Maximum Likelihood method based on the Jukes-Cantor model. The tree with the highest log likelihood (-5618.92) is shown. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. The analysis involved 32 nucleotide sequences. There were a total of 606 positions in the final dataset. Major clades are shown on the right.

Phylogenetic discordance of *A. albopictus* and *Cx. pipiens*, relative to their uninfected congeners *A. vexans* and *Cx. erraticus*, prompted investigation into population genetic measures associated with these respective species. Sample size constraints limited analyses to these particular species. Tajima's D, a relative measure of population variance, values for the *CO1* gene region of the four species were insignificant for all species; *ITS2* gene region values were significant for *A. albopictus*

(Tajima's D -2.27, p < 0.01) and *Cx. pipiens* (Tajima's D -0.93, p < 0.001), the only two species with samples sizes sufficient for analyses (see Supp Table 3). The significance of these results is questionable, though, given the non-coding nature of the *ITS2* gene region and the larger sample sizes needed to estimate within population genetic variance and departure from neutral evolution (Simonsen et al. 1995).

Measures of mitochondrial (*CO1*) nucleotide diversity and divergence among infected and uninfected *Aedes* and *Culex* species were conducted. An order of magnitude lower nucleotide diversity (Pi) was observed for *A. albopictus*, relative to other sampled species (Table I). Such results may be indicative of a mitochondrial selective sweep, possibly associated with a more recent *Wolbachia* strain invasion and spread in the sampled population, supporting conclusions found among *Ae. albopictus* populations in other studies (Armbruster et al. 2003). Nucleotide diversity measures differ from previously reported results (Atyame et al. 2011) for *Ae. vexans* and *Ae. albopictus* (lower in present study) and *Cx. pipiens* (higher in present study). Between species population divergence (Dxy), a measure of average between-population mitochondrial nucleotide substitutions, was lowest between the two infected species populations, *Ae. albopictus* and *Cx. pipiens* (Table I).

Table I. Cytochrome oxidase subunit one (*CO1*) gene region nucleotide diversity (Pi) within species is shown on the shaded diagonal below for select species populations. The upper off-diagonal values represent the average number of nucleotide substitutions per site between species, or between species population divergence (Dxy).

	***Ae. vexans* (U)**	***Ae. albopictus* (I)**	***Cx. pipiens* (I)**	***Cx. erraticus* (U)**
***Ae. vexans* (U)**	0.010	0.113	0.125	0.141
***Ae. albopictus* (I)**		0.001	0.107	0.149
***Cx. pipiens* (I)**			0.022	0.151
***Cx. erraticus* (U)**				0.013

One sample hypothesis testing on uninfected and *Wolbachia*-infected mosquitoes showed significant support for differences in mean *CO1* nucleotide substitutions based on infection status when both are compared to outgroup *D. neotestacea* using Tajima's relative rate test ($H_A = \mu_d > 15$; $t = 3.5177$, $df = 179$, $p < 0.001$, Table II). Difference values of 10 and 15 were used to test the hypothesis that infected species' elevated rate of evolution (substitution) can be accounted for by mean differences from uninfected species by greater than 15 substitutions. The elevated rate of evolution was associated with *Wolbachia* infection since similar rate differences were not observed when within infected and uninfected species are compared (mean differences between infected species was significantly less than 10 substitutions (H_A: $\mu_d < 10$; $t = 27.362$, $df = 19$, $p < 0.001$), as it was for mean differences between uninfected species ($t = 20.431$, $df = 74$, $p < 0.001$). Given upper and lower thresholds for determining differences and similarities, respectively, statistical results support greater nucleotide substitution rates for infected *A. albopictus* and *Cx. pipiens*, when compared with uninfected *A. vexans* and *Cx. erraticus*.

Table II. Contingency table of counts (upper right off-diagonal) used to determine cytochrome oxidase subunit one (*CO1*) gene region unique mean nucleotide differences (lower left off-diagonal), obtained from intermediate calculations used in Tajima's Relative Rate test, species A (top row) – species B (left column), from outgroup (*D. neotestacea*) among infected (I) and uninfected (U) *Aedes* and *Culex* species is shown below.

	***Ae. vexans* (U)**	***Ae. albopictus* (I)**	***Cx. pipiens* (I)**	***Cx. erraticus* (U)**
***Ae. vexans* (U)**		60	75	75
***Ae. albopictus* (I)**	-14.6		20	20
***Cx. pipiens* (I)**	-15.4	-2.6		25
***Cx. erraticus* (U)**	4.68	17.4	20.0	

Elevated substitution rates for the *CO1* sequences among infected *Aedes* and *Culex* species suggested a potential long-branch attraction effect that may be responsible for the observed discordance among the two genera (Sanderson & Shaffer 2002). Phylogenetic analysis of translated (amino acid) sequences was subsequently conducted to minimize the effect of long-branch attraction (Bergsten 2005). The discordant infected *Aedes-Culex* clade was partially resolved in the amino acid phylogenetic tree with a monophyletic *Aedes* clade composed of uninfected *A. vexans* and infected *A. albopictus*, and an unresolved *Culex* grouping (Fig. 4). Post-hoc analyses of the CO1 tree topologies, with monophyletic constraints placed on *Aedes* and *Culex* taxa, failed to reject monophyletic constraints by the Shimodaira-Hasegawa test ($p > 0.328$).

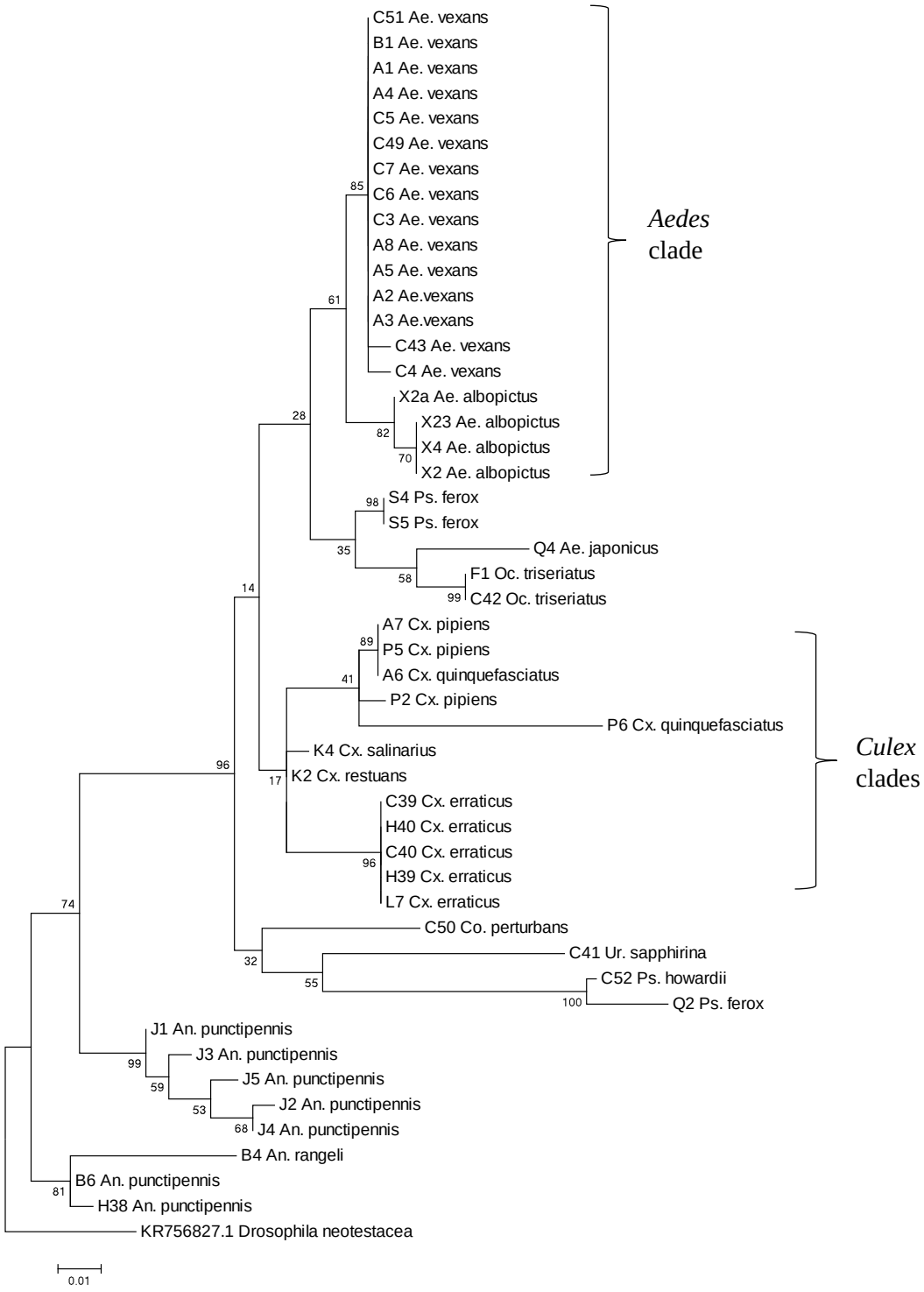

Figure 4. The evolutionary history for *CO1* amino acid sequences was inferred by using the Maximum Likelihood method based on the Whelan And Goldman + Freq. model [1]. The tree with the highest log likelihood (-1315.75) is shown. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. The analysis involved 49 amino acid sequences. There were a total of 203 positions in the final dataset. Relevant clades are shown on the right.

DISCUSSION

Initial phylogenetic patterns observed for mosquito species sampled in this study appeared to show a discordance associated with *Wolbachia* infection status for mitochondrial nucleotide sequences among *Aedes* and *Culex* species (Fig. 1). This finding led to statistical analyses and support for elevated rates of molecular evolution (substitution) among *Wolbachia*-infected species (Table II, one sample *t*-test values in Results). What are the proximal explanations for *Wolbachia*-induced evolutionary rate increases, and how could elevated substitution rates explain the observed phylogenetic discordance? *A. albopictus* and *Cx. pipiens* are known to be infected with cytoplasmic incompatibility (CI) *Wolbachia* (Dobson et al. 2001, Yen and Barr 1973). CI-*Wolbachia* have the well-documented capacity to induce mitochondrial selective sweeps via molecular hitchhiking (Galtier et al. 2009; Hurst and Jiggins 2005, Jiggins et al. 2002), leading to reductions in mitochondrial effective population sizes (Cariou et al. 2017). Under this scenario, while rates of adaptive evolution would appear to decrease, genetic drift in smaller populations would elevate the rate of neutral or slightly deleterious substitutions (Lanfear et al. 2014; Shoemaker et al. 2004). One consequence of drift-induced increase in substitution rates is a compromised capacity for adaptive mitochondrial evolution in populations subject to driving cytoplasmic elements like *Wolbachia* (Cariou et al. 2017). We were unable to discern the phenotypic effect of the substitutions responsible for the elevated rates observed in *A. albopictus* and *C. pipiens*. However, our observation of lower nucleotide diversity for *A. albopictus* (Table 1) is consistent with expectations derived from a mitochondrial selective sweep and an associated increase in substitution rate.

Regardless of proximal explanations, elevated substitution rates in infected *Aedes* and *Culex* species appear to result in an artifactual phylogenetic grouping of the *Aedes/Culex* infected clade (Fig. 1) by long-branch attraction (LBA). Long-branch attraction is a known phylogenetic consequence of elevated substitution rates (Sanderson & Shaffer 2002) that can result in discordant nuclear-mitochondrial phylogenies within genera (Omilian and Taylor 2001), as well as phylogenetic misplacement across genera (Hassanin 2006). One recommended solution to the problems imposed by long-branch attraction is to limit the effect of potentially fast-evolving third codon positions (Bergsten 2005). The translated *CO1* (amino acid) sequence phylogeny for the mosquito species accomplished this goal with expected resolution of the *Aedes CO1* clade (Fig. 4). However, the loss of character states associated with this methodology is also known to result in some loss of phylogenetic resolution (Bergsten 2005), which was observed for the *Culex* genus in this study (see low bootstrap support values for *Culex* clade, Fig. 4). Limited genomic coverage from the use of a single gene region (*CO1*, 609 bp) also contributed to low bootstrap values in phylogenetic analyses.

Discordant mitochondrial phylogeny associated with *Wolbachia* infection within the *Drosophila quinaria* Loew species complex has been observed for *CO1*, in which *Wolbachia* infected *D. quinaria* are sister to a clade encompassing uninfected *D. quinarian* and seven other *Drosophila* species (Dyer et al. 2011), with a species-level measure of divergence between infected and uninfected *D. quinaria CO1* sequences (6.7%). In this study the authors suggested the observed pattern could be explained

through congeneric hybridization and cytoplasmic introgression. The widespread occurrences of cyto-nuclear discordance associated with *Wolbachia* infection is frequently attributed to cytoplasmic introgression of this type (Graham and Wilson 2012; Jiggins 2003; Narita et al. 2006; Sahoo et al. 2018; Whitworth et al. 2007).

The unreliability of mitochondrial phylogenetics has been well documented (Hurst and Jiggins 2005), with particular emphases on the hitchhiking effect that comes with co-inherited cytoplasmic symbionts, like *Wolbachia*, capable of inducing strong selective drive forces. Discordant mitochondrial phylogenies associated with driving cytoplasmic symbionts can be explained by population genetic consequences that come with a reduced effective population size for mtDNA having undergone selective sweeps (Cariou et al. 2017), and by hybrid introgression, as mentioned above. The support found for elevated rates of molecular evolution among infected *Aedes* and *Culex* species, relative to uninfected congeneric species supports previous findings of recurrent selective sweeps and increased rates of molecular substitution among infected species (Cariou et al. 2017; Shoemaker et al. 2004). These population genetic effects, as well limited sample sizes for both genomic coverage and taxa, may be responsible for the observed artifactual cladogenesis of infected *Aedes* and *Culex* species in the present study.

In conclusion, cyto-nuclear phylogenetic discordance among *Aedes* and *Culex* mosquito species in our limited sample set appears to be due to elevated substitution rates associated with *Wolbachia* infection, and a resultant long-branch attraction effect. Evidence of elevated rates of molecular evolution for infected mitochondrial sequences may be a result of molecular hitch-hiking with driving *Wolbachia* variants in *Aedes* and *Culex* species. Disjunct *CO1* topologies, and evidence of elevated substitution rates, and phylogenetic resolution of long-branch effects associated with *Wolbachia*-infected *CO1* gene regions supports previous assertions regarding the difficulty of phylogenetic inference using cytoplasmic molecular markers with *Wolbachia*-infected species. The limited taxon (four species) and genetic (two loci) dataset used to analyze mitochondrial-nuclear discordance herein was a function of collection constraints. More extensive sampling of congeneric *Wolbachia*-infected and uninfected species and populations would help resolve potential phylogenetic discordance associated with sampling bias, and the long-branch attraction artifacts associated with elevated substitution rates. Larger sample sizes would also make population variance measures, such as Tajima's D, and hypotheses regarding departures from neutral evolution more tractable.

ACKNOWLEDGEMENTS

We thank professors Wendy Dustman, Rebekah Ward, Fengjie Sun, and the Lawrenceville Science Tavern for useful discussion on research results.

REFERENCES

Armbruster P., W.E. Damsky, R. Giordano, J. Birungi, L.E. Munstermann, and J.E. Conn. 2003. Infection of New- and Old-World Aedes Albopictus (Diptera: Culicidae) by the Intracellular Parasite Wolbachia: Implications for Host Mitochondrial DNA Evolution. J. Med. Entomol., 40(3): 356-360.

Atyame C.M., F. Delsuc, N. Pasteur, M. Weill, and O. Duron. 2011. Diversification of Wolbachia Endosymbiont in the Culex pipiens Mosquito. Mol. Biol. Evol., 28(10): 2761-2772.

Baldo, L., J.C.D. Hotopp, K.A. Jolley, S.R. Bordenstein, S.A. Biber, R.R. Choudhury, … J.H. Werren. 2006. Multilocus Sequence Typing System for the Endosymbiont Wolbachia pipientis. Appl. Environ. Microb., 72(11):7098–7110. doi: 10.1128/aem.00731-06

Bergsten, J. 2005. A review of long-branch attraction. Cladistics, 21(2):163-193. doi.org/10.1111/j.1096-0031.2005.00059.x

Binnington, K. C. and A.A. Hoffmann. 1989. *Wolbachia*-like organisms and cytoplasmic incompatibility in Drosophila simulans. J. Invertebr. Pathol., 54(3):344-352.

Burkett-Cadena, N. D. 2013. Mosquitoes of the Southeastern United States. Tuscaloosa, AL: University of Alabama Press.

Cariou, M., L. Duret, and S. Charlat. 2017. The global impact of *Wolbachia* on mitochondrial diversity and evolution. J. Evolution. Biol., 30(12):2204–2210. doi: 10.1111/jeb.13186

Dobson, S. L., E. J. Marsland, W. Rattanadechakul. 2001. *Wolbachia*-induced cytoplasmic incompatibility in single- and superinfected *Aedes albopictus* (Diptera: Culicidae). J. Med. Entomol., 38(3): 382-387.

Dyer, K. A., C. Burke, and J. Jaenike. 2011. Wolbachia-mediated persistence of mtDNA from a potentially extinct species. Mol. Ecol., 20(13):2805–2817. doi: 10.1111/j.1365-294x.2011.05128.x

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotech., 3:294–299.

Galtier, N., B. Nabholtz, S. Glemin, and G.D.D. Hurst. 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Molecular Ecology, 18:4541-4550. doi: 10.1111/j.1365-294X.2009.04380.x

Graham, R. I., and K. Wilson. 2012. Male-killing Wolbachia and mitochondrial selective sweep in a migratory African insect. BMC Evol. Biol., 12(1):204. doi: 10.1186/1471-2148-12-204

Hassanin, A. 2006. Phylogeny of Arthropoda inferred from mitochondrial sequences: Strategies for limiting the misleading effects of multiple changes in pattern and rates of substitution. Mol. Phylogenet. Evol., 38(1): 100-116. doi.org/10.1016/j.ympev.2005.09.012.

Hoffmann, A. A., B.L. Montgomery, J. Popovici, I. Iturbe-Ormaetxe, P.H. Johnson, F. Muzzi, … S.L. O'Neill. 2011. Successful establishment of Wolbachia in Aedes populations to suppress dengue transmission. Nature, 476(7361): 454–457. doi: 10.1038/nature10356

Hurst, G. D., and F.M. Jiggins. 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. P. Roy. Soc. B-Biol. Sci., 272(1572): 1525–1534. doi: 10.1098/rspb.2005.3056

Jiggins, F.M. 2003. Male-Killing Wolbachia and Mitochondrial DNA: Selective Sweeps, Hybrid Introgression and Parasite Population Dynamics. Genetics, 164: 5-12.

Jiggins, F. M., G.D.D. Hurst, and Z. Yang. 2002. Host-Symbiont Conflicts: Positive Selection on an Outer Membrane Protein of Parasitic but not Mutualistic Rickettsiaceae. Mol. Biol. Evol., 19(8): 1341–1349. doi: 10.1093/oxfordjournals.molbev.a004195

Kikuchi, Y., & T. Fukatsu. 2003. Diversity of Wolbachia Endosymbionts in Heteropteran Bugs. Appl. Environ. Microb., 69(10):6082–6090. doi: 10.1128/aem.69.10.6082-6090.2003

Kodandaramaiah, U., T.J. Simonsen, S. Bromilow, N. Wahlberg, and F. Sperling. 2013. Deceptive single-locus taxonomy and phylogeography:Wolbachia-associated divergence in mitochondrial DNA is not reflected in morphology and nuclear markers in a butterfly species. Ecol. Evol., 3(16): 5167–5176. doi: 10.1002/ece3.886

Kriesner, P., A.A. Hoffmann, S.F. Lee, M. Turelli, and A.R. Weeks. 2013. Rapid Sequential Spread of Two Wolbachia Variants in Drosophila simulans. PLoS Pathog., 9(9). doi: 10.1371/journal.ppat.1003607

Lanfear, R., H. Kokko, and A. Eyre-Walker. 2014. Population size and the rate of evolution. Trends Ecol. Evol., 29(1): 33–41. doi: 10.1016/j.tree.2013.09.009

Manonmani, A., H. Townson, T. Adeniran, P. Jambulingam, S. Sahu, and T. Vijayakumar. 2001. rDNA-ITS2 polymerase chain reaction assay for the sibling species of Anopheles fluviatilis. Acta Trop., 78(1): 3–9. doi: 10.1016/s0001-706x(00)00154-6

Narita, S., M. Nomura, Y. Kato, and T. Fukatsu. 2006. Genetic structure of sibling butterfly species affected by Wolbachia infection sweep: evolutionary and biogeographical implications. Mol. Ecol., 15(4): 1095–1108. doi: 10.1111/j.1365-294x.2006.02857.x

Nguyen, L. T., H.A. Schmidt, A. von Haeseler, and B.Q. Minh. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Mol. Biol. Evol., 32: 268-274.

Omilian, A. R., and D.J. Taylor. 2001. Rate Acceleration and Long-branch Attraction in a Conserved Gene of Cryptic Daphniid (Crustacea) Species, Mol. Biol. Evol., 18(12): 2201-2212. doi.org/10.1093/oxfordjournals.molbev.a003767

O'Neill, S. L., R. Giordano, A.M. Colbert, T.L. Karr, and H.M. Robertson. 1992. 16S rRNA phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic incompatibility in insects. P. Natl. Acad. Sci. USA, 89(7): 2699–2702. doi: 10.1073/pnas.89.7.2699

Rasgon, J. L. and T.W. Scott. 2003. Wolbachia and cytoplasmic incompatibility in the California Culex pipiens mosquito species complex: parameter estimates and infection dynamics in natural populations. Genetics, 165: 2029-2038.

Rasgon, J.L., and T.W. Scott. 2004. An initial survey for Wolbachia (Rickettsiales: Rickettsiaceae) infections in selected California mosquitoes (Diptera: Culicidae). J. Med. Entomol., 41: 255–257. doi:10.1603/0022-2585-41.2.255.

Reinert, J.F., R.E. Harbach, and I.J. Kitching. 2004. Phylogeny and classification of Aedini (Diptera:Culicidae), based on morphological characters of all life stages. Zoological Journal of the Linnean Society, 142:289-368

Richardson, M.F., L.A. Weinert, J.J. Welch, R.S. Linheiro, M.M. Magwire, F.M. Jiggins, and C.M. Bergman. 2012. Population genomics of the Wolbachia endosymbiont in *Drosophila melanogaster*. PLoS Genetics, 8(12): e1003129. doi:10.1371/journal.pgen.1003129

Rozas, J., A. Ferrer-Mata, J.C. Sánchez-Delbarrio, S. Guirao-Rico, P. Librado, S.E. Ramos-Onsins, and A. Sánchez-Gracia. 2017. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Mol. Biol. Evol., 34(12): 3299–3302. doi: 10.1093/molbev/msx248

Ruang-Areerate, T., and P. Kittayapong. 2006. Wolbachia transinfection in Aedes aegypti: A potential gene driver of dengue vectors. P. Natl. Acad. Sci. USA, 103(33): 12534–12539. doi: 10.1073/pnas.0508879103

Sahoo, R. K., D.J. Lohman, N. Wahlberg, C.J. Müller, O. Brattström, S.C. Collins, … and U. Kodandaramaiah. 2018. Evolution of Hypolimnas butterflies (Nymphalidae): Out-of-Africa origin and Wolbachia-mediated introgression. Mol. Phylogenet. Evol., 123: 50–58. doi: 10.1016/j.ympev.2018.02.001

Sanderson, M. J., and H.B. Shaffer. 2002. Troubleshooting molecular phylogenetic analysis. Annu. Rev. Ecol. Syst., 33: 49-72. doi.org/10.1146/annurev.ecolsys.33.010802.15059.

Schilthuizen, M. O., and R. Stouthamer. 1997. Horizontal transmission of parthenogenesis–inducing microbes in Trichogramma wasps. P. Roy. Soc. B-Biol. Sci., 264(1380): 361–366. doi: 10.1098/rspb.1997.0052

Schmidt, T. L., N.H. Barton, G. Rašić, A.P. Turley, B.L. Montgomery, I. Iturbe-Ormaetxe, …and M. Turelli. 2017. Local introduction and heterogeneous spatial spread of dengue-suppressing Wolbachia through an urban population of Aedes aegypti. PLOS Biol., 15(5). doi: 10.1371/journal.pbio.2001894

Shaikevich, E., A. Bogacheva, V. Rakova, L. Ganushkina, and Y. Ilinsky. 2019. Wolbachia symbionts in mosquitoes: Intra- and intersupergroup recombinations, horizontal transmission and evolution. Mol. Phylogenet. Evol., 134: 24–34. doi: 10.1016/j.ympev.2019.01.020

Shoemaker, D. D., K.A. Dyer, M. Ahrens, K. Mcabee, and J. Jaenike. 2004. Decreased Diversity but Increased Substitution Rate in Host mtDNA as a Consequence of Wolbachia Endosymbiont Infection. Genetics, 168(4): 2049–2058. doi: 10.1534/genetics.104.030890

Simonsen, K. L., G. A. Churchill, C. F. Aquadro. 1995. Properties of statistical tests of neutrality for DNA polymorphism data. Genetics, 141: 413-429.

Sinkins, S. P., H.R. Braig, and S.L. O'Neill. 1995. Wolbachia superinfections and the expression of cytoplasmic incompatibility. P. Roy. Soc. B-Biol. Sci., 261(1362): 325–330. doi: 10.1098/rspb.1995.0154

Smith M.A., C. Bertrand, K. Crosby, E.S. Eveleigh, J. Fernandez-Triana, …and X. Zhou. 2012. Wolbachia and DNA Barcoding Insects: Patterns, Potential, and Problems. PLoS ONE, 7(5): e36514. doi:10.1371/journal.pone.0036514

Stouthamer, R., J.A.J. Breeuwer, and G.D.D. Hurst. 1999. Wolbachia Pipientis: Microbial Manipulator of Arthropod Reproduction. Annu. Rev. Microbiol., 53(1): 71–102. doi: 10.1146/annurev.micro.53.1.71

Sun, X. J., J.H. Xiao, J.M. Cook, G. Feng, and D.W. Huang. 2011. Comparisons of host mitochondrial, nuclear and endosymbiont bacterial genes reveal cryptic fig wasp species and the effects of Wolbachia on host mtDNA evolution and diversity. BMC Evol. Biol., 11(1). doi: 10:1186/1471-2148-11-86.

Turelli, M., and A.A. Hoffmann. 1991. Rapid spread of an inherited incompatibility factor in California Drosophila. Nature, 353(6343): 440–442. doi: 10.1038/353440a0

Turelli, M., and A.A. Hoffmann. 1995. Cytoplasmic Incompatibility in Drosophila simulans: Dynamics and Parameter Estimates from Natural Populations. Genetics, 140: 1319-1338.

Turelli, M., A.A. Hoffmann, and S.W. McKechnie. 1992. Dynamics of Cytoplasmic Incompatibility and mtDNA Variation in Natural *Drosophila simulans* populations. Genetics, 132: 713-723.

Turelli, M., B.S. Cooper, K.M. Richardson, P.S. Ginsberg, B. Peckenpaugh…and A.A. Hoffmann. 2018. Rapid global spread of wRI-like Wolbachia across multiple Drosophila. Current Biology, 28: 963-971. https://doi.org/10.1016/j.cub.2018.02.015

Werren, J. H., W. Zhang, and L.R. Guo. 1995. Evolution and phylogeny of Wolbachia: reproductive parasites of arthropods. P. Roy. Soc. B-Biol. Sci., 261(1360): 55–63. doi: 10.1098/rspb.1995.0117

Werren, J. H., L. Baldo, and M.E. Clark. 2008. Wolbachia: master manipulators of invertebrate biology. Nat. Rev. Microbiol., 6(10): 741–751. doi: 10.1038/nrmicro1969

Whitworth, T., R. Dawson, H. Magalon, and E. Baudry. 2007. DNA barcoding cannot reliably identify species of the blowfly genus Protocalliphora (Diptera: Calliphoridae). P. Roy. Soc. B-Biol. Sci., 274(1619): 1731–1739. doi: 10.1098/rspb.2007.0062

Xi, Z., C.C. Khoo, and S.L. Dobson. 2005. Wolbachia Establishment and Invasion in an Aedes aegypti Laboratory Population. Science, 310(5746): 326–328. doi: 10.1126/science.1117607

Yen, J. H., A. R. Barr. 1973. The etiological agent of cytoplasmic incompatibility in *Culex pipiens*. J. Invertebr. Pathol., 22(3): 242-250.

Zhou, W., F. Rousset, and S. O'Neill. 1998. Phylogeny and PCR–based classification of Wolbachia strains using wsp gene sequences. P. Roy. Soc. B-Biol. Sci., 265(1395): 509–515. doi: 10.1098/rspb.1998.0324

Supplemental information

Discordant cytoplasmic-nuclear phylogenies associated with *Wolbachia* infection

Russell, J.E., M. Saum, R. Williams

Supp. Table 1. Primer sequences for gene regions used for analyses are shown below.

Gene		Primer sequence
CO1	HCO	5'TAA ACT TCA GGG TGA CCA AAA AAT CA-3'
	LCO	5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3'
16S	99F	5'-TTG TAG CCT GCT ATG GTA TAA CT-3'
	994R	5'-GAA TAG GTA TGA TTT TCA TGT-3'
coxA	*coxA*F	5'-TTG GRG CRA TYA ACT TTA TAG-3'
	*coxA*R	5'-CTA AAG ACT TTK ACR CCA GT -3'
gatB	*gatB*F	5'-GAK TTA AAY CGY GCA GGB GTT -3'
	*gatB*R	5'-TGG YAA YTC RGG YAA AGA TGA -3'
ITS2	5.8S	5'-ATC ACT CGG CTC GTG GAT CG -3'
	28S	5'-ATG CTT AAA TTT AGG GGG TAG TC -3'

Supp Table 2. Mosquito species samples collected for gene sequence analyses, including number of individuals tested and presence/absence (+/-) of detected Wolbachia infection. **All collected individuals from *Wolbachia* positive species tested positive for infection.**

Genus / Species	No. tested	Infection
Aedes vexans	15	-
Aedes albopictus	31	+
Aedes japonicus	1	-
Coquillettidia perturbans	1	+
Culex erraticus	5	-
Culex pipiens/quinquefaciatus	5	+
Culex restuans	1	+
Culex salinarius	1	+
Ochleratatus triseriatus	2	-
Psorophora howardii	1	-
Psorophora ferox	3	-
Anopheles punctipennis	7	-
Anopheles rangeli	1	-
Uranotaenia sapphirina	1	+

Supp Table 3. Tajima's D values for the CO1 and ITS2 gene region of selected Aedes and Culex species is shown below.

Genus / Species	*CO1*	*ITS2*
Aedes vexans	-1.85	NA
Aedes albopictus	0.17	-2.27
Culex erraticus	-1.25	NA
Culex pipiens/quinquefaciatus	0.04	-0.93