MOLECULAR ECOLOGY

Molecular Ecology (2011) **20**, 2805–2817 doi: 10.1111/j.1365-294X.2011.05128.x

Wolbachia-mediated persistence of mtDNA from a potentially extinct species

KELLY A. DYER,* CRISTA BURKE†‡ and JOHN JAENIKE†

*Department of Genetics, University of Georgia, Athens, GA 30602, USA, †Department of Biology, University of Rochester, Rochester, NY 14627, USA

Abstract

***Drosophila quinaria* is polymorphic for infection with *Wolbachia*, a maternally transmitted endosymbiont. *Wolbachia*-infected individuals carry mtDNA that is only distantly related to the mtDNA of uninfected individuals, and the clade encompassing all mtDNA haplotypes within *D. quinaria* also includes the mtDNA of several other species of *Drosophila*. Nuclear gene variation reveals no difference between the *Wolbachia*-infected and uninfected individuals of *D. quinaria*, indicating that they all belong to the same interbreeding biological species. We suggest that the *Wolbachia* and the mtDNA with which it is associated were derived via interspecific hybridization and introgression. The sequences in the *Wolbachia* and the associated mtDNA are ≥6% divergent from those of any known *Drosophila* species. Thus, in spite of nearly complete species sampling, the sequences from which these mitochondria were derived remain unknown, raising the possibility that the donor species is extinct. The association between *Wolbachia* infection and mtDNA type within *D. quinaria* suggests that *Wolbachia* may be required for the continued persistence of the mtDNA from an otherwise extinct *Drosophila* species. We hypothesize that pathogen-protective effects conferred by *Wolbachia* operate in a negative frequency-dependent manner, thus bringing about a stable polymorphism for *Wolbachia* infection.**

Keywords: *Drosophila quinaria*, introgression, mtDNA, MLST, phylogenetics

Received 27 December 2010; revision revised 7 April 2011; accepted 18 April 2011

Introduction

Wolbachia may be the most widespread endosymbiont in terrestrial ecosystems, infecting perhaps two-thirds of present-day insect species, as well as a substantial fraction of species in other arthropod groups (Stouthamer *et al.* 1999; Hilgenboecker *et al.* 2008). In many host species, *Wolbachia* spread and persist by manipulating host reproduction in ways such as cytoplasmic incompatibility and male-killing, which enhance the relative fitness of infected cytoplasmic lineages, even though such tactics can exact a severe demographic and genetic toll on the host population (Engelstädter & Hurst 2009).

Correspondence: Kelly A. Dyer, E-mail: kdyer@uga.edu; John Jaenike, E-mail: john.jaenike@rochester.edu

‡Present address: Crista Burke, Department of Biology, Tufts University, Medford, MA 02155, USA. E-mail: crista.burke@tufts.edu

Wolbachia, however, are not invariably parasitic: several recent reports indicate that some strains are beneficial to their hosts, boosting fertility (Dedeine *et al.* 2001), fecundity (Weeks *et al.* 2007; Brownlie *et al.* 2009), and resistance to pathogens and parasites (Hedges *et al.* 2008; Teixeira *et al.* 2008; Kambris *et al.* 2009).

Regardless of their fitness effects, the sojourn times of *Wolbachia* within host lineages are typically short on an evolutionary time scale, as phylogenetic analyses reveal very few examples of codivergence of insect host species and their associated *Wolbachia* (Werren *et al.* 1995; but see Raychoudhury *et al.* 2009; Stahlhut *et al.* 2010). Such patterns indicate that there must be relatively high rates of *Wolbachia* colonization of new host species, as well as high rates of extinction of *Wolbachia* from infected species, although the mechanisms by which these processes occur are largely unknown.

Inferences about the evolutionary history of *Wolbachia* within a species are often made using the host's mtDNA, which is maternally transmitted together with *Wolbachia* through the egg cytoplasm. Because the mtDNA can have a higher mutation rate than *Wolbachia*, it may more accurately reflect the recent history of an infection within a host species and between closely related host species. Such analyses of host mtDNA have revealed, for instance, the recent spread of *Wolbachia* and its associated mtDNA haplotype (Turelli & Hoffmann 1991), the species-wide sweep of a single mtDNA lineage associated with *Wolbachia* (Dyer & Jaenike 2004), and multiple, independent colonizations of a host species by *Wolbachia* (James & Ballard 2000).

It has recently been found that mtDNA and *Wolbachia* can move from one species of insect to another via interspecific hybridization (Ballard 2000; Jiggins 2003; Bachtrog *et al.* 2007; Raychoudhury *et al.* 2009). Here, we report another likely example of *Wolbachia* colonization by hybridization, but in this case the source species has not been identified and may be extinct: the only genetic remnant of this putatively extinct species is its mtDNA, which hitchhiked along with *Wolbachia* during the introgression and is probably dependent on *Wolbachia* for its continued persistence. We suggest that the long-term polymorphism of two mitochondrial clades associated with *Wolbachia* infection status may be a consequence of negative frequency-dependent selection on *Wolbachia* infection.

Methods

Fly collections and culturing

Drosophila quinaria, a member of the quinaria species group of the subgenus Drosophila, is native to eastern North America, where it breeds in rotting skunk cabbage (*Symplocarpus foetidus*; Jaenike 1978). For this study, we collected *D. quinaria* in Linesville, PA, USA and Rochester, NY, USA in 1995 and 2004, and in Deer Isle, ME, USA in 1995 (Table 1). Flies were collected by sweep-netting over baits of decaying cucumbers and bananas. Individuals of *D. quinaria* were identified by their distinctive wing and abdominal spotting patterns (for a recent species identification key see Markow & O'Grady 2006). Wild-caught females from Linesville were placed individually in culture at 22°C on Formula 4–24 Drosophila medium (Carolina Biological Supply, Burlington, NC) supplemented with *Agaricus bisporus* commercial mushroom. The F1 sex ratio was determined for females collected in 1995.

Molecular methods

Genomic DNA was extracted from single flies using Qiagen Puregene DNA Purification Kit (Valencia, CA), and PCR was conducted using standard protocols. For all DNA sequencing, amplicons were purified using Exosap (USB), and DNA was sequenced in both directions using Big Dye v3.1 chemistry (Applied Biosystems) and run out on an ABI 3730 at core facilities at the University of Georgia, University of Rochester, or Cornell University. Chromatograms were analysed using Sequencher (Gene Codes Corporation, Ann Arbor, MI, USA) and sequences were aligned manually in Se-Al (Rambaut 1996). For some loci, heterozygous sites were called based on double peaks in both sequencing directions in the chromatograms, and alleles were inferred using the program PHASE (Scheet & Stephens 2006). All sequences from this study have been deposited in Genbank (JF63872–JF64081).

Flies were screened for *Wolbachia* infection using primers that amplify a portion of the rapidly evolving *Wolbachia* gene *wsp* (*Wolbachia* surface protein), using the primers wsp81F and wsp691R (Braig *et al.* 1998). For all samples infected with *Wolbachia*, we sequenced the *Wolbachia* genes *wsp* (using the same primers as above) and five MLST loci (*gatB, coxA, hcpA, ftsZ,* and *fbpA*), using primers and protocols described in Baldo *et al.* (2006).

We also sequenced several loci in all *Wolbachia*-infected *D. quinaria* and a sample of uninfected individ-

Table 1 Sampled populations and sites, with abbreviations used in figures

Population	Population abbreviation	Site	Site abbreviation	Year sampled	Number of samples
Linesville, PA	PA	Pymatuning	PY	1995	39
		Pymatuning	PY04	2004	57
Rochester, NY	NY	Pittsford Fairport	LM, CHP TH, QLX	1995	39
		Pittsford	ROC04	2004	12
Deer Isle, ME	ME	Dunham's Point	DPT	1995	20
		Water Pump	WP		

uals from all three populations. Loci included mitochondrial *COI* (using primers LCO1490 and HCO2198 from Folmer *et al.* 1994); *plexA*, located on the dot chromosome in *D. melanogaster* and other species (primers in Dyer *et al.* 2011); *tpi*, located on Müller's element E in *D. melanogaster* and other species (primers in Shoemaker *et al.* 2004); and the X-linked genes *mof* (primers in Dyer *et al.* 2007), *skpA* (primers in Dyer *et al.* 2011), and *R1B* (primers in Gentile *et al.* 2001). For *COI*, phylogenetic analysis was carried out on a dataset of 373 bp sequences, while the population genetic analyses were completed on a dataset of 650 bp sequences. Because *Wolbachia* and the mtDNA are both maternally transmitted through the egg cytoplasm, we use both the *Wolbachia* and mtDNA to infer the history of the *Wolbachia* infection in *D. quinaria* (Ballard & Rand 2005; Hurst & Jiggins 2005).

Phylogenetic and population genetic analysis

Phylogenetic analyses were used to infer the gene tree of each sequenced locus. For *COI*, we incorporated sequences from all but one of the known Nearctic species and all but four of the Palearctic species in the quinaria group (Perlman *et al.* 2003; Sidorenko 2009). The only Nearctic species not included is *D. suffusca*, which has apparently been collected only twice, once in Hart Prairie AZ (Patterson 1943) and once in Silver City NM (collected by M. Wheeler; D. Grimaldi, pers. comm.). The male genitalia of the *D. suffusca* specimens are extremely similar to that of *D. occidentalis*, raising the question of whether *D. suffusca* is really a separate species (D. A. Grimaldi, personal communication). The Palearctic species we have been unable to obtain include *D. schachti* (which has been found only in Turkey), *D. sannio* (found only in Kazakhstan), *D. parakuntzei* (found only in Mongolia and central Siberia), and *D. angularis* (a rare species found only in Japan and the nearby Primorskii Territory on the Asian mainland).

For all loci, sequences for *D. virilis*, *D. tripunctata*, and/or *D. immigrans* were used as outgroups to the quinaria group. Sequences incorporated from previous studies include *COI* (from Perlman *et al.* 2003), *R1B* (from Gentile *et al.* 2001) *tpi*, *mof*, *skpA*, *and plexA* (from Dyer *et al.* 2011). We also included the previously unpublished *COI* sequences of *D. tenebrosa*, *D. magnaquinaria*, and *D. rellima*. Phylogenetic analyses were conducted for each locus using MrBayes v3.1 (Ronquist & Huelsenbeck 2003). Based on the results of Akaike Information Criterion tests in jModelTest (Posada 2008; Guindon & Gascuel 2003), we used the most appropriate model of DNA evolution in the phylogenetic analysis (Lanave *et al.* 1984; Tavaré 1986; Yang 1994). This included GTR + I + Γ for *COI*, *mof*, *plexA*, *skpA*, and *tpi*, and GTR + Γ for *R1B* (Lanave *et al.* 1984; Tavaré 1986; Yang 1994). We ran the program for 2,000,000 iterations, sampling every 1,000 iterations. For each locus, we conducted four independent runs to ensure convergence of the chains. The first 500,000 samples of each run were excluded as burn-in, and the most probable tree was constructed from the final 6,000,000 samples, combined from the four runs. Similar results were found using other phylogenetic methods (data not shown).

We inferred the time to the most recent common mtDNA ancestor (TMRCA) using BEAST v 1.6.1 (Drummond & Rambaut 2007). We used a Hasagawa–Kishino–Yano (HKY) substitution model of sequence evolution, assumed a constant population size, and enforced a strict molecular clock. The mutation rate of the *Drosophila* mitochondrial DNA has been estimated as 6.2×10^{-8} substitutions/site/generation (Haag-Liautard *et al.* 2008). As this was estimated from mutation accumulation lines, it includes deleterious mutations and is thus an overestimate of the substitution rate in natural populations. Therefore, the use of this mutation rate provides a minimum estimate of TMRCA. We further assume that *D. quinaria* has three generations per year, and thus $\mu = 18.6 \times 10^{-8}$ substitutions/site/year. For each run we ran the chain for 10^7 iterations, sampling every 1000 steps and excluding the first 1000 samples as burn-in. We combined the results from four independent runs to estimate the posterior density of the TMRCA and effective population size (N_e).

For each locus, we estimated various population genetic statistics, including intraspecific polymorphism, among-group differentiation, and deviation of the site frequency spectrum (Tajima's *D*), using the program DnaSP v5.1 (Librado & Rozas 2009). We compared the levels of polymorphism and divergence in the mtDNA of *D. quinaria* to other loci using an HKA test (Hudson *et al.* 1987), as implemented in MLHKA (Wright & Charlesworth 2004). The MLHKA tests used only silent sites and only samples from the Linesville, PA, USA population, because this was the only site where *D. quinaria* was found to be polymorphic for *Wolbachia* infection, and there is some evidence for geographic differentiation among populations (Table S1, Supporting information). Divergence was calculated from *D. innubila*, a member of the quinaria group that is an outgroup to all of the mtDNA haplotypes found in *D. quinaria*. Simulations accounted for the inheritance pattern of the mtDNA versus the autosomes, a 10-fold higher mutation rate in the mtDNA versus the autosomes (as found by Haag-Liautard *et al.* 2008), and the relative abundance of each mtDNA clade in the Linesville, PA, USA population (4/96 = 0.04 for the *F* clade and 93/ 6 = 0.96 for the *Q* clade), each of which is expected to affect the equilibrium levels of polymorphism and divergence.

The HKA tests included all nuclear loci except *R1B*, whose mode of concerted evolution effectively eliminates polymorphism within species.

Results

In the 1995 collections, *Wolbachia*-infected *D. quinaria* were found only in Linesville, PA, USA, where 3 out of 39 (8%) females carried this endosymbiont. None of the females collected from Deer Isle ME ($n = 20$) nor Rochester NY ($n = 39$) were infected with *Wolbachia*. Two F1 individuals of each infected female were PCR-screened for the *Wolbachia* infection, and all were found to be infected, indicating that the infection is maternally transmitted. We surveyed the same Linesville, PA, USA and Rochester, NY, USA populations again in 2004, but found no individuals infected with *Wolbachia* ($n = 57$ in Linesville and $n = 12$ in Rochester).

We surveyed the offspring sex ratio of 18 wild-caught females from Linesville, PA, USA, of which three were infected with *Wolbachia*. One of the three *Wolbachia*-infected females produced exclusively female offspring, whereas the other two produced broods with an even sex ratio (Fig. 1). Among the 15 females not infected with *Wolbachia*, 6 produced strongly female-biased progeny, and the remaining nine produced ~1:1 offspring sex ratios. There is not a clear distinction among progeny sex ratios between *Wolbachia*-infected and uninfected females (two-tailed Fisher's exact test, $P = 0.7$). It is likely that the female-biased sex ratios resulted from matings with males that carried a driving X chromosome; such X-chromosomes result in the production of all-female progeny and have previously been found in *D. quinaria* (Jaenike 1996). Therefore, the *Wolbachia* in *D. quinaria* is probably not a male-killer.

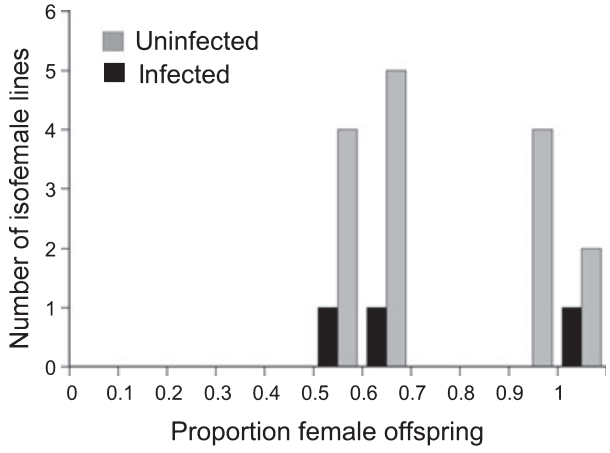


Fig. 1 Lack of association between *Wolbachia* infection and sex-ratio distortion. F1 sex ratios obtained from females collected in Linesville PA in 1995 that were either infected (*black*) or uninfected (*gray*) with *Wolbachia*. There were an average of 101 ± 11 offspring per family.

Inspection of the chromatograms for all sequenced *Wolbachia* genes revealed no double peaks, indicating that each of the three *Wolbachia*-infected females carried a single strain of *Wolbachia*. Two flies carried *Wolbachia* with identical *wsp* sequences, which differed from the third *wsp* sequence at one synonymous and three non-synonymous sites (total divergence = 0.7%). Neither of these *wsp* sequences match any published sequences in GenBank or the *Wolbachia* MLST Database (http://pubmlst.org/wolbachia/), and they are more closely related to each other than to any other known *wsp* sequence. The most similar *wsp* sequences identified by a BLAST search are 4% divergent across 561 bp, and group with infections in the *Wolbachia* supergroup A (Zhou *et al.* 1998). Several of these most similar sequences are found in tropical *Drosophila* (including *D. paulistorum*, *D. willistoni*, *D. yakuba*, and *D. santomea*) and *Asobara tabida*, a common generalist parasitoid of *Drosophila* (Fleury *et al.* 2009). The most similar *Wolbachia* strain that infects a quinaria group species is from *Drosophila recens*, and is 5% divergent at the nucleotide level (GenBank AY154399 from Shoemaker *et al.* 2004).

The *Wolbachia* from all three infected *D. quinaria* samples were identical at all five MLST loci. When compared to the *Wolbachia* MLST Database, this strain is an exact match to allele 1 at *coxA*, *fbpA*, *gatB*, and *hcpA*, and to allele 3 at *ftsZ*. The MLST data confirm that the *Wolbachia* in *D. quinaria* belongs to supergroup A, and that, furthermore, it falls within the ST-13 complex, as defined by Baldo *et al.* (2006). This complex contains many highly similar *Wolbachia* strains, including those from *D. recens*, *D. innubila*, and *D. munda*, all of which belong to the quinaria group of *Drosophila*. The ST-13 complex also includes the *Wolbachia* strains wMel from *D. melanogaster* and wAu from *D. simulans* (Ishmael *et al.* 2009), as well as the *Wolbachia* that infects *Asobara tabida* (Baldo *et al.* 2006).

The *COI* sequences indicate that the mtDNA haplotypes within *D. quinaria* belong to two distinct clades (Fig. 2). One clade is found in the majority of *D. quinaria* flies, including those sequenced previously (Perlman *et al.* 2003). The other mtDNA clade is phylogenetically distant and was found in all three *Wolbachia*-infected flies and one uninfected individual. These two clades are denoted *Q* (*quinaria*) and *F* (foreign), respectively. There is a highly significant association between *Wolbachia* infection status and mtDNA clade ($P < 0.0002$; Fisher's exact test). The two unique *COI* haplotypes in clade *F* are very similar, differing at 3 out of 650 bp (0.3% divergent; Table S2, Supporting information). The *COI* haplotypes in the *Q* clade are also quite similar, differing on average by 0.9% (Table S2). However, the *Q*

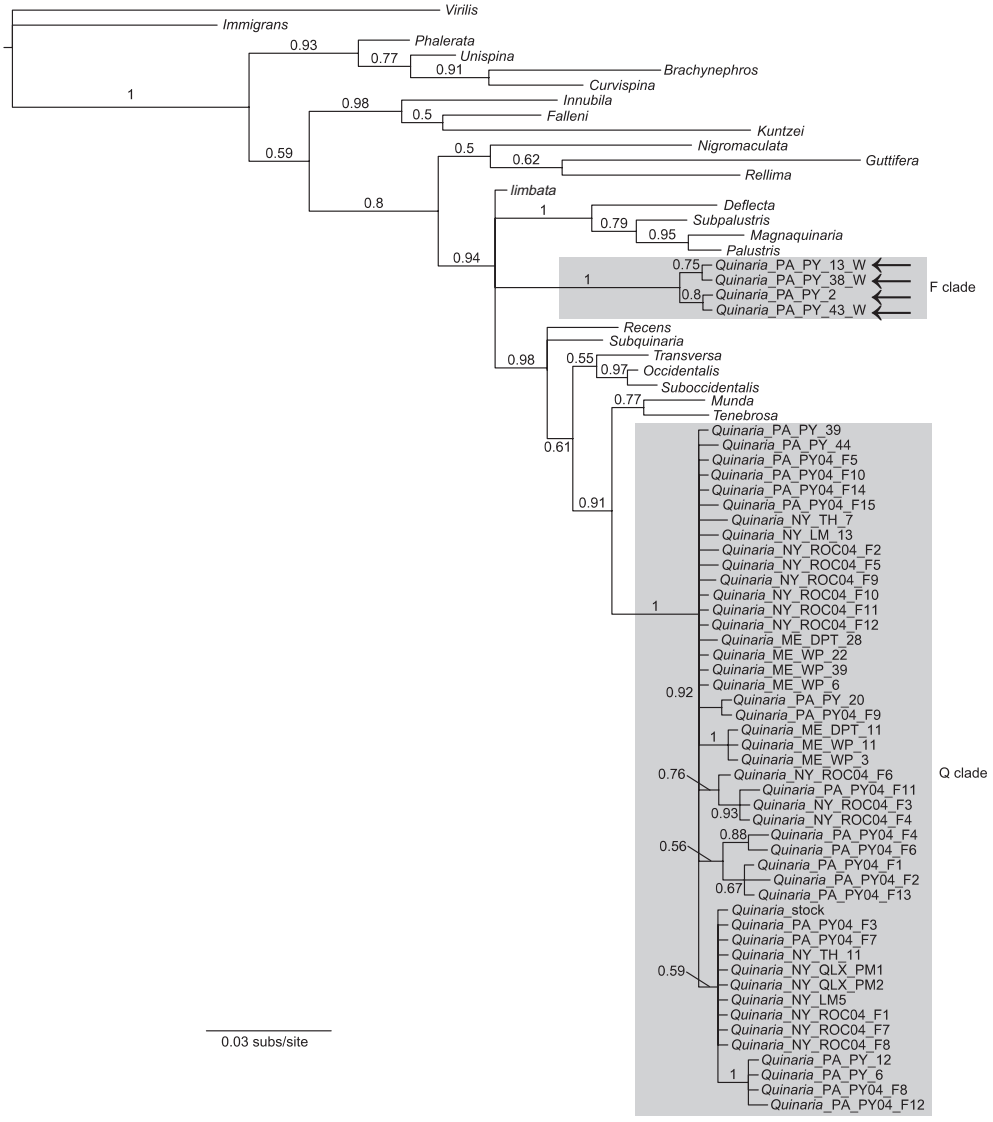
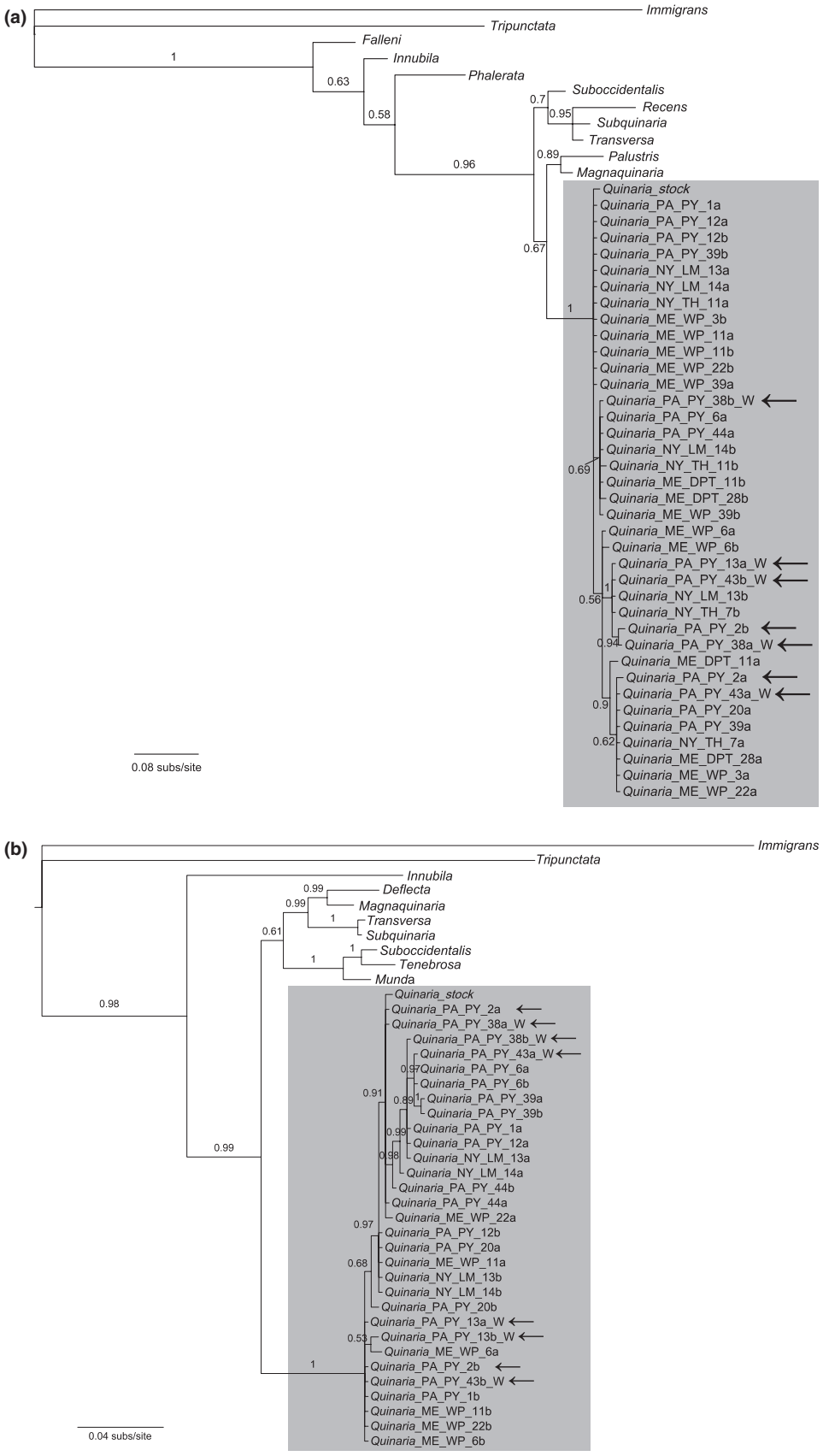


Fig. 2 *COI* gene tree of quinaria group species of *Drosophila*. *D. immigrans* and *D. virilis* are outgroups. Individuals of *D. quinaria* are blocked in gray, with the *F* clade flies here and in subsequent figures indicated with arrows next to sample names. The W after the sample name here and in all subsequent figures denotes *Wolbachia* infection; all *F* clade flies carry *Wolbachia* except PA_PY_2. See Table 1 for population and site abbreviations. The support value for each node is shown if that partition was found in >50% of the sampled trees. See text for details about the analyses.

and *F* clades are 6.7% divergent, a difference typically characteristic of different species within a genus (Hebert *et al*. 2003a,b). Indeed, the *Q* and *F* clades of *D. quinaria* do not form a monophyletic group within the quinaria group, as the *Q* clade sequences are, with strong support, more similar to those of several other species than they are to the *F* clade sequences (Fig. 2). None of the sampled *COI* trees in the phylogenetic analysis place all of the *D. quinaria* mtDNA haplotypes as a monophyletic group. Furthermore, the *COI* sequences most similar to the *F* clade are from *Drosophila recens*, and these are 6.3% divergent. This suggests that the source of these mitochondrial haplotypes is either an extinct or an as yet undiscovered or unsampled species within the quinaria group.

Despite the considerable sequence divergence between the *F* and *Q* clades of mtDNA within *D. quinaria*, the flies carrying mtDNA of the two clades are identical or very similar for every nuclear gene we sequenced, including two autosomal and three X-linked genes. For example, the X-linked *R1B* locus is a retrotransposable element that is located within the rDNA complex and thus undergoes concerted evolution within species, but diverges rapidly among species (Eickbush *et al*. 1995, 1997; Gentile *et al*. 2001). It is therefore expected to be an especially good marker for species identification. We



find no *R1B* polymorphism within *D. quinaria*, but high divergence from closely related species (10% divergence to *D. palustris*). For each of the nuclear genes we sequenced, the *D. quinaria* samples formed a monophyletic group with Bayesian support values above 0.99 in the phylogenetic analyses (Fig. 3). Within *D. quinaria*, the nuclear gene sequences from the flies carrying mtDNA haplotypes from the two distinct clades were interspersed, indicating that all of these flies belong to a single interbreeding biological species. Statistically, there is little or no evidence of nuclear genetic differentiation between flies carrying the *Q* and *F* clades of mtDNA (Table 2).

Assuming a single *Wolbachia* introgression event, we estimate a minimum age of this event using an inferred the genealogy of *COI* sequences within the *F* clade. Because mtDNA is co-transmitted with *Wolbachia*, the genealogy of this mtDNA clade should reflect the evolutionary history of *Wolbachia* in *D. quinaria* back to the time of the most recent common ancestor (TMRCA). Assuming $\mu = 18.6 \times 10^{-8}$ substitutions/site/year and the inferred mean N_e of 11,000 for the *F* mtDNA clade only, the median TMRCA for the *F COI* clade was 9,000 year, with a 95% highest posterior density from 1,200–23,500 year (Fig. 4).

There is no evidence that either the *F* or the *Q* mtDNA clade has experienced a recent selective sweep or bottleneck in population size, in contrast to what is often found in species with *Wolbachia* infections. First, we find segregating polymorphisms in both the *Q* and *F* mtDNA clades, with no deviation from the neutral frequency spectrum, using Tajima's *D*, for either mtDNA clade (Table S2). This suggests that there has not been a recent bottleneck in N_e or selective sweep in either mtDNA clade, although we caution that Tajima's *D* for the *F* mtDNA clade is based on only three segregating sites. Second, an HKA test including only the *F* or *Q* clade of *COI* sequences does not reveal a significant departure from neutrality for either mtDNA clade (*F* clade: $k_{COI\text{-}F} = 4.53$; $\Lambda = 3.44$, $P = 0.10$; *Q* clade: $k_{COI\text{-}Q} = 0.87$; $\Lambda = 0.19$, $P = 0.67$). The lack of a deficiency of segregating polymorphism suggests that there has not been a recent bottleneck within either clade, or a selective spread of *Wolbachia*-associated mtDNA within the *F* clade. Thus, the population genetic analyses indicate that *Wolbachia* has infected *D. quinaria* for a considerable evolutionary period.

Discussion

Drosophila quinaria, like many other insects, is polymorphic for infection with endosymbiotic *Wolbachia*. However, we have found that the infected and uninfected flies carry distinct—and quite distantly related—lineages of mtDNA. Most strikingly, the clade that includes both of these mtDNA lineages is paraphyletic, encompassing the mtDNA of several other species of *Drosophila*.

We can envision several scenarios by which this unusual mtDNA–*Wolbachia* association arose in *D. quinaria*. First, *Wolbachia* may have colonized the ancestor of *D. quinaria* in the distant past and been maintained in polymorphic condition ever since. This could happen if a species ancestrally polymorphic for *Wolbachia* infection underwent several rounds of speciation, with loss of *Wolbachia* and its associated mtDNA in all descendant species except *D. quinaria*, thus yielding the pattern shown in Fig. 2. Alternatively, suppose only one geographic population of an ancestral species became infected with *Wolbachia* and that subsequently the uninfected populations experienced several rounds of speciation. If the *Wolbachia*-infected flies remained genetically coherent with only one of these uninfected species, specifically *D. quinaria*, this could also yield the observed *Wolbachia*–mtDNA association.

However, given the high level of sequence divergence (6.7%) between the *Q* and *F* mtDNA clades within *D. quinaria*, it is more likely in our view that *D. quinaria* acquired both *Wolbachia* and its associated mtDNA via hybridization and introgression from another *Drosophila* species, as is known to have occurred in other *Drosophila* (*e.g.*, Ballard 2000; Jaenike *et al.* 2006; Bachtrog *et al.* 2007). Such introgression requires that female hybrids between species be fertile, which in turn requires either that pre-mating isolation between species evolves more quickly than post-mating isolation, or that post-mating isolating mechanisms follow Haldane's rule, both of which are common patterns in *Drosophila* (Coyne & Orr 1997). Based on the genetic distance among mtDNA sequences, it appears that the quinaria group, especially the section encompassing *D. quinaria*, speciates very rapidly (Perlman *et al.* 2003). Among the loci sequenced here we find no consistent phylogenetic placement of *D. quinaria*, which suggest that there has been rapid radiation—and thus incomplete lineage sorting—of

Fig. 3 Gene trees of nuclear genes in *D. quinaria* and other members of the quinaria group. Outgroup species include *D. immigrans* and/or *D. tripunctata*. Genes include (a) *mof*, (b) *plexA*, (c) *R1B*, (d) *skpA*, and (e) *tpi*. *Arrows* identify sequences from flies carrying *F* clade of mtDNA, and the *grey box* indicates the sequences from individuals identified phenotypically as *D. quinaria*. See Table 1 for population and site abbreviations. Because of heterozygosity at all loci except *R1B*, there can be 2 different nuclear gene haplotypes for individual flies. The support value for each node is shown if that partition was found in >50% of the sampled trees. See text for details about the analyses.

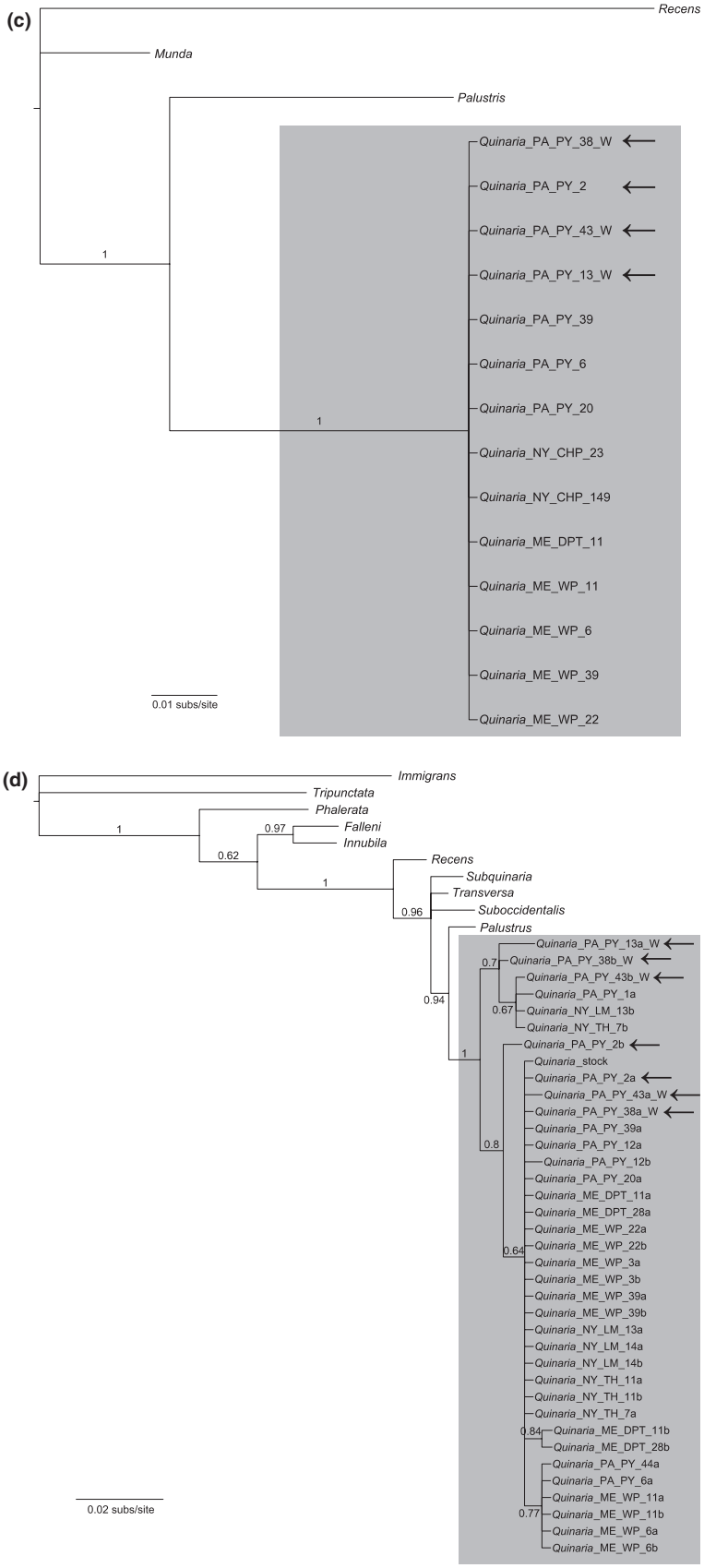


Fig. 3 (*Continued*)

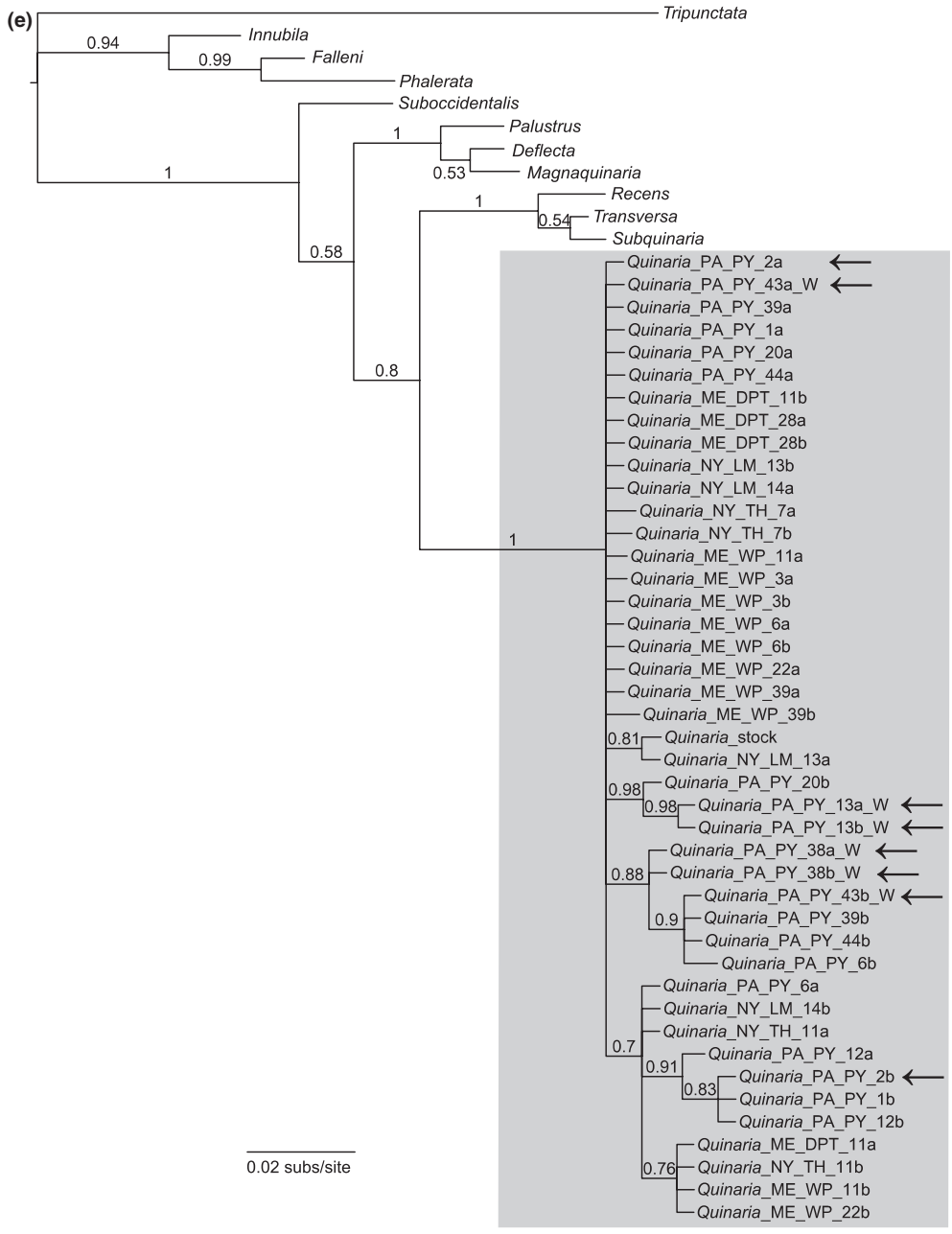


Fig. 3 (*Continued*)

quinaria group species. Furthermore, numerous pairs of closely related species within this group can produce fertile hybrid females (Sears 1947; Blumel 1949; K. Dyer, unpublished), although hybrid males are usually sterile (Shoemaker *et al*. 1999). Although we have not measured the fertility of these hybrid females (other than *D. recens–D. subquinaria* hybrids, which are highly fertile; Shoemaker *et al*. 1999), in our experience most hybrid females types can produce many offspring (K. Dyer and J. Jaenike, personal observations). As a result of such hybrid female fertility, many species pairs in the quinaria group could potentially exchange maternally transmitted elements, including mtDNA and *Wolbachia*.

If the *Wolbachia* and its associated mtDNA in *D. quinaria* were derived via hybridization with another quinaria group species, we cannot identify the source species. The *F*-clade *COI* sequences in *D. quinaria* are more than 6% divergent from any known *Drosophila COI* sequence (Fig. 2), even though we have attempted to collect and sequence *COI* from as many species in the quinaria

Table 2 Genetic differentiation between flies carrying *F* and *Q* clades of mtDNA. Only Linesville, PA, samples are included. *R1B* is not included because it contained no polymorphic sites within *D. quinaria*

Locus	K_{ST}*	S_{nn}	Fixed differences	Shared polymorphisms	D_{xy}	D_a
COI	0.276 ($P < 0.0001$)	1.0 ($P < 0.0001$)	36	2	0.06669	0.06089
mof	0.128 ($P = 0.028$)	0.689 ($P = 0.069$)	0	5	0.00625	0.00148
tpi	−0.012 ($P = 0.570$)	0.592 ($P = 0.208$)	0	7	0.00731	0.00038
skpA	0.025 ($P = 0.204$)	0.571 ($P = 0.197$)	0	4	0.00652	0.00039
plexA	0.013 ($P = 0.26$)	0.529 ($P = 0.34$)	0	7	0.00491	0.00029

* K_{ST} and S_{nn} are measures of genetic differentiation, D_{xy} is the average number of nucleotide substitutions per site between *F* and *Q* clades, and D_a is net number of nucleotide substitutions per site between *F* and *Q* clades.

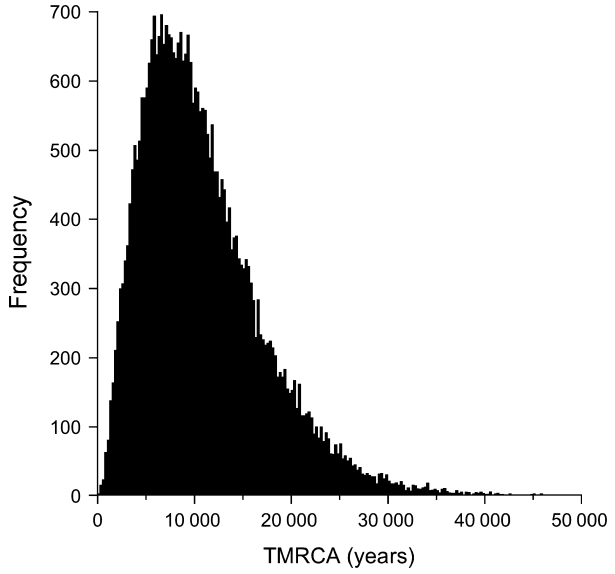


Fig. 4 Estimate of the time to the most recent common ancestor (TMRCA) of the *Wolbachia*-associated *F* clade of mtDNA. Shown is the posterior density; the mean is 11,000 years and the median is 10,000 years.

group as possible, including most of the Palearctic species and all but one Nearctic species. Similarly, the *Wolbachia wsp* sequence in these flies is not closely similar to any published *Wolbachia* sequence. One explanation of these findings is that the *Drosophila* species from which *Wolbachia* and mtDNA were introduced to *D. quinaria* went extinct some time after—or even as a consequence of—the introgression event. Thus, the mtDNA carried by some individuals of *D. quinaria* may be the only genetic remnant of an extinct *Drosophila* species. It is also possible that the donor species is an extant, but unsampled or undiscovered species.

Our analysis suggests a coalescent time of 1,000–24,000 years for the *F* clade mtDNA within *D. quinaria*. Thus, the species has probably been polymorphic two major cytoplasmic clades—one with 'foreign' (*F*) mtDNA and harboring *Wolbachia*, and the other with native, 'quinaria' (*Q*) mtDNA but lacking *Wolbachia*—for at least that long. The absence of *Wolbachia* in flies carrying *Q* clade mtDNA indicates that there has been little or no horizontal transmission of *Wolbachia* following its introduction to *D. quinaria*. Furthermore, the very strong association between the *F* clade mtDNA and *Wolbachia* infection, despite *Wolbachia*'s imperfect maternal transmission, suggests that *Wolbachia* has been essential for the long-term persistence of the *F* clade cytoplasmic lineages. This in turn suggests that, in the absence of *Wolbachia*, the fitness of cytoplasmic lineages with foreign mtDNA is less than those with native mtDNA, indicative of a cytonuclear interaction. A similar long-term association of different *Wolbachia* strains with divergent host mtDNA lineages has also been found in the two-spotted ladybird *Adalia bipuntata*, although in this case the mtDNA lineages are monophyletic within the host species and are apparently not a case of introgression (Jiggins & Tinsley 2005).

We model the coexistence problem by assuming that there are three types of cytoplasmic lineages: (i) 'quinaria' mtDNA, uninfected with *Wolbachia* (designated *Q*-), (ii) 'foreign' mtDNA, infected with *Wolbachia* (*F*+); and (iii) 'foreign' mtDNA, uninfected with *Wolbachia* (*F*-). First, consider the population of lineages carrying foreign mtDNA (*F*+ and *F*-). The equilibrium prevalence of *Wolbachia* infection among the *F* lineages ($\hat{P}_F$) is expected to be $\hat{P}_F = (s + \beta - 1)/s$, where s is the selective advantage of *Wolbachia* infection among *F*-type flies and β is the fidelity of maternal transmission of *Wolbachia* by *F*+ females (Dyer & Jaenike 2004). The equilibrium mean fitness of the *F* lineages is then expected to be $\bar{W}_F = \hat{P}_F W_{F+} + (1 - \hat{P}_F) W_{F-}$. With frequency-independent fitnesses, the *F* clade should sweep to fixation if $\bar{W}_F > W_{Q-}$, and be lost from the population if $\bar{W}_F < W_{Q-}$. Cytoplasmic incompatibility (CI) results in positive frequency-dependent selection (Caspari & Watson 1959; Turelli 1994), and should lead to the loss or fixation of the *F* clade, depending on whether the

prevalence of *Wolbachia* infection falls below or above a certain threshold frequency. Thus, CI cannot bring about a stable persistence of the *F* and *Q* lineages. Therefore, in the absence of negative frequency-dependent selection, a stable polymorphism of the two mitochondrial types is not expected.

There have been several recent reports that *Wolbachia* confers resistance to various types of pathogens and parasites (Hedges *et al.* 2008; Teixeira *et al.* 2008; Kambris *et al.* 2009). If *Wolbachia* reduces the transmission rate of a pathogen, then the force of pathogen infection (*i.e.*, the probability per unit time that an individual becomes infected) should be negatively correlated with the prevalence of *Wolbachia* in a population. Consequently, the selective advantage of carrying *Wolbachia* would be greatest when the prevalence of *Wolbachia* infection is lowest. The mitochondrial type associated and co-transmitted with *Wolbachia* will also experience such negative frequency-dependent dynamics. In this manner, a stable polymorphism of *Q* and *F* mitochondrial types, as well as *Wolbachia* infection, could be maintained. The population genetic analyses presented above indicate that there has been a long-term polymorphism of the *F* and *Q* mtDNA clades within *D. quinaria*, consistent with this scenario.

It is interesting to note that the wMel strain of *Wolbachia* within *D. melanogaster* that confers protection from viral infection belongs to the ST-13 complex. Furthermore, *D. simulans* is infected with several strains of *Wolbachia*, but only those closely related to wMel provide strong antiviral protection (Osborne *et al.* 2009). The fact that the *Wolbachia* found within *D. quinaria* belong to the ST-13 complex supports our hypothesis these *Wolbachia* may provide protection from pathogenic infections in this species and thus be maintained by negative frequency-dependent selection. Such a frequency-dependent mechanism would be most effective for pathogens transmitted primarily within rather than between host species. Currently, little is known about rates of viral transmission within and among *Drosophila* species in nature, although recent studies indicate that viral transmission can occur both vertically and laterally within species, and that they do have the potential to move from one species to another, at least on evolutionary time scales (Jousset & Plus 1975; Habayeb *et al.* 2009; Longdon *et al.* 2011a,b). Although we do not yet know the selective benefit of *Wolbachia* in *D. quinaria*, our findings suggest that these maternally transmitted endosymbionts can drive introgression of mtDNA among newly diverging species during the interval between the initial evolution of pre-mating isolation and the final closing of the post-mating window. Our findings also suggest that some organisms may carry genomic vestiges of species now extinct.

Acknowledgments

We thank the Pymatuning Laboratory of Ecology for their hospitality; M. Minhas, L. Boelio, and B. White for technical assistance; and D. Grimaldi for discussion of *D. suffusca*. This research was supported by NSF grants DEB 0922977 and an REU Supplement (NSF 1025344) to JJ and funds from the University of Georgia Research Foundation and Ellison Medical Foundation to KAD.

References

Bachtrog D, Thornton K, Clark A, Andolfatto P (2007) Extensive introgression of mitochondrial DNA relative to nuclear genes in the *Drosophila yakuba* species group. *Evolution*, **60**, 292–302.

Baldo L, Dunning Hotopp JC, Jolley KA *et al.* (2006) Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Applied and Environmental Microbiology*, **72**, 7098–110.

Ballard JWO (2000) When one is not enough: introgression of mitochondrial DNA in *Drosophila*. *Molecular Biology and Evolution*, **17**, 1126–1130.

Ballard JWO, Rand DM (2005) The population biology of mitochondrial DNA and its phylogenetic implications. *Annual Review of Ecology, Evolution, and Systematics*, **36**, 621–642.

Blumel J (1949) Additional tests within the *quinaria* species-group of Drosophila. Studies in the genetics of Drosophila VI. *University of Texas Publication*, **4920**, 31–38.

Braig HR, Zhou W, Dobson S, O'Neill SL (1998) Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. *Journal of Bacteriology*, **180**, 2373–2378.

Brownlie JC, Cass BN, Riegler M *et al.* (2009) Evidence for metabolic provisioning by a common invertebrate endosymbiont, *Wolbachia pipientis*, during periods of nutritional stress. *PloS Pathogens*, **5**(10), e1000630.

Caspari E, Watson GS (1959) On the evolutionary importance of cytoplasmic sterility in mosquitoes. *Evolution*, **13**, 568–570.

Coyne JA, Orr HA (1997) 'Patterns of Speciation in Drosophila' revisited. *Evolution*, **51**, 295–30.

Dedeine F, Vavre F, Fleury F, Loppin B, Hochberg ME, Bouletreau M (2001) Removing symbiotic *Wolbachia* bacteria specifically inhibits oogenesis in a parasitic wasp. *Proceedings of the National Academy of Sciences USA*, **98**, 6247–6252.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Dyer KA, Jaenike J (2004) Evolutionarily stable infection by a male-killing endosymbiont in *Drosophila innubila*: molecular evidence from the host and parasite genomes. *Genetics*, **168**, 1443–1455.

Dyer KA, Charlesworth B, Jaenike J (2007) Chromosome-wide linkage disequilibrium as a consequence of meiotic drive. *Proceedings of the National Academy of Sciences, USA*, **104**, 1587–1592.

Dyer KA, White BE, Bray MJ, Piqué DG, Betancourt AJ (2011) Molecular evolution of a Y chromosome to autosome gene duplication in *Drosophila*. *Molecular Biology and Evolution*, **28**, 1293–1306.

Eickbush DG, Lathe III WC, Francino MP, Eickbush TH (1995) R1 and R2 Retrotransposable elements of *Drosophila* evolve at rates similar to those of nuclear genes. *Genetics*, **139**, 685–695.

Eickbush TH, Burke WD, Eickbush DG, Lathe WC (1997) Evolution of R1 and R2 in the rDNA units of the genus *Drosophila*. *Genetica*, **100**, 49–61.

Engelstädter J, Hurst GDD (2009) The ecology and evolution of microbes that manipulate host reproduction. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 127–149.

Fleury F, Gibert P, Ris N, Allemand R (2009) Ecology and life history evolution of frugivorous *Drosophila* parasitoids. In:*Parasitology: Parasitoids of Drosophil* (ed Prevost G.). pp. 3–44, Elsevier, London.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.

Gentile KL, Burke WD, Eickbush TH (2001) Multiple lineages of R1 retrotransposable elements can coexist in the rDNA loci of *Drosophila*. *Molecular Biology and Evolution*, **18**, 235–245.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.

Haag-Liautard C, Crown N, Houle D, Lynch M, Charlesworth B, Keightley PD (2008) Direct estimation of the mitochondrial DNA mutation rate in *D. melanogaster*. *PLoS Biology*, **6**, 1706–1714.

Habayeb MS, Cantera R, Casanova G *et al.* (2009) The *Drosophila* Nora virus is an enteric virus, transmitted via feces. *Journal of Invertebrate Pathology*, **101**, 29–33.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003a) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, 313–321.

Hebert PDN, Ratnasingham S, de Waard JR (2003b) Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, S96–S99.

Hedges LM, Brownlie JC, O'Neill SL, Johnson KN (2008) *Wolbachia* and virus protection in insects. *Science*, **322**, 702.

Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH (2008) How many species are infected with *Wolbachia* ?—a statistical analysis of current data. *FEMS Microbiology Letters*, **281**, 215–220.

Hudson R, Kreitman M, Aguadé M (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics*, **116**, 153–159.

Hurst GDD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society of London B: Biological Sciences*, **272**, 1525–1534.

Ishmael N, Hotopp JCD, Ioannidis P *et al.* (2009) Extensive genomic diversity of closely related *Wolbachia* strains. *Microbiology*, **155**, 2211–2222.

Jaenike J (1978) Resource predictability and niche breadth in the *Drosophila quinaria* species group. *Evolution*, **32**, 676–678.

Jaenike J (1996) Sex-ratio meiotic drive in the *Drosophila quinaria* group. *American Naturalist*, **148**, 237–254.

Jaenike J, Dyer KA, Cornish C, Minhas MS (2006) Asymmetrical reinforcement and *Wolbachia* infection in *Drosophila*. *PLoS Biology*, **4**, 1852–1862.

James AC, Ballard JWO (2000) Expression of cytoplasmic incompatibility in Drosophila simulans and its impact on infection frequencies and distribution of Wolbachia pipientis. *Evolution*, **54**, 1661–1672.

Jiggins FM (2003) Male-killing *Wolbachia* and mitochondrial DNA: selective sweeps, hybrid introgression and parasite population dynamics. *Genetics*, **164**, 5–12.

Jiggins FM, Tinsley MC (2005) An ancient mitochondrial polymorphism in *Adalis bipunctata* linked to a sex-ratio distorting bacterium. *Genetics*, **171**, 1115–1124.

Jousset FX, Plus N (1975) Étude de la transmission horizontale et de la transmission verticale des picornavirus de Drosophila melanogaster et de Drosophila immigrans [Study of the vertical transmission and horizontal transmission of Drosophila melanogaster and Drosophila immigrans picornavirus]. *Annals of Microbiology (Paris)*, **126**, 231–249.

Kambris Z, Cook PE, Phuc HK, Sinkins SP (2009) Immune activation by life-shortening Wolbachia and reduced filarial competence in mosquitoes. *Science*, **326**, 134–136.

Lanave C, Preparata G, Sacone C, Serio G (1984) A new method for calculating evolutionary substitution rates. *Journal of Molecular Evolution*, **20**, 86–93.

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Longdon B, Wilfert L, Obbard DJ, Jiggins FM (2011a) Rhabdoviruses in two species of *Drosophila*: vertical transmission and a recent selective sweep. *Genetics*, **188**, 141–150.

Longdon B, Wilfert L, Osei-Poku J, Cagney H, Obbard DJ, Jiggins FM (2011b) Host-switching by a vertically transmitted rhabdovirus in *Drosophila*. *Biology Letters*, doi:10.1098/rsbl.2011.0160.

Markow TA, O'Grady PM (2006) *Drosophila: A Guide to Species Identification and Use*. Elsevier, London.

Osborne SE, Leong YS, O'Neill SL, Johnson KN (2009) Variation in antiviral protection mediated by different *Wolbachia* strains in *Drosophila simulans*. *PLoS Pathogens*, **5**, e1000656. doi:10.1371/journal.ppat.1000656.

Patterson JT (1943) Studies in the genetics of Drosophila III. The Drosophilidae of the Southwest. *University of Texas Publication*, **4313**, 7–216.

Perlman SJ, Spicer GS, Shoemaker DD, Jaenike J (2003) Associations between mycophagous *Drosophila* and their *Howardula* nematode parasites: a worldwide phylogenetic shuffle. *Molecular Ecology*, **12**, 237–249.

Posada D (2008) jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Rambaut A (1996) Se-Al: sequence alignment editor. Available: http://evolve.zoo.ox.ac.uk/.

Raychoudhury R, Baldo L, Oliveira DCSG, Werren JH (2009) Modes of acquisition of *Wolbachia*: horizontal transfer, hybrid introgression and co-divergence in the *Nasonia* species complex. *Evolution*, **63**, 165–183.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Scheet P, Stephens M (2006) A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *American Journal of Human Genetics*, **78**, 629–644.

Sears JW (1947) Relationships within the *quinaria* species group of *Drosophila*. *University of Texas Publication*, **4720**, 137–156.

Shoemaker DD, Katju V, Jaenike J (1999) *Wolbachia* and the evolution of reproductive isolation between *Drosophila recens* and *Drosophila subquinaria*. *Evolution*, **53**, 1157–1164.

Shoemaker DD, Dyer KA, Ahrens M, McAbee K, Jaenike J (2004) Decreased diversity but increased substitution rate in host mtDNA as a consequence of *Wolbachia* endosymbiont infection. *Genetics*, **168**, 2049–2058.

Sidorenko VS (2009) A review of species of the *Drosophila quinaria* Loew species-group (Diptera Drosophilidae) of the Asian part of Russia and adjacent countries. *Entomological Review*, **89**, 85–90.

Stahlhut J, Desjardins C, Clark M *et al.* (2010) The mushroom habitat as an ecological arena for global exchange of *Wolbachia*. *Molecular Ecology*, **19**, 1940–1952.

Stouthamer R, Breeuwer JAJ, Hurst GDD (1999) *Wolbachia pipientis*: microbial manipulators of arthropod reproduction. *Annual Review of Microbiology*, **53**, 71–102.

Tavaré S (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures in Mathematics in the Life Sciences*, **17**, 57–86.

Teixeira L, Ferreira A, Ashburner M (2008) The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. *PloS Biology*, **6**, 2753–2763.

Turelli M (1994) Evolution of incompatibility-inducing microbes and their hosts. *Evolution*, **48**, 1500–1513.

Turelli M, Hoffmann AA (1991) Rapid spread of an inherited incompatibility factor in California *Drosophila*. *Nature*, **353**, 440–442.

Weeks AR, Turelli M, Harcombe WR, Reynolds KT, Hoffmann AA (2007) From parasite to mutualist: rapid evolution of *Wolbachia* in natural populations of *Drosophila*. *PloS Biology*, **5**, 997–1005.

Werren JH, Zhang W, Guo LR (1995) Evolution and phylogeny of *Wolbachia* - reproductive parasites of arthropods. *Proceedings of the Royal Society of London B: Biological Sciences*, **261**, 55–63.

Wright S, Charlesworth B (2004) The HKA test revisited: a maximum-likelihood-ratio test of the standard neutral model. *Genetics*, **168**, 1071–1076.

Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution*, **39**, 306–314.

Zhou W, Rousset F, O'Neil S (1998) Phylogeny and PCR-based classification of *Wolbachia* strains using wsp gene sequences. *Proceedings of the Royal Society of London B: Biological Sciences*, **265**, 509–515.

K.D. studies the population genetics and the genetics of adaptation and speciation in *Drosophila*. C.B. is currently a graduate student at Tufts University researching temporal isolation as a reproductive barrier in a recently introduced invasive agricultural pest species. Most of J.J.'s current research focuses on the evolutionary ecology of host–parasite and host–symbiont interactions.

Data Accessibility

DNA sequences: Genbank accessions JF63872–JF64081

Treebase accession URL for this study: http://purl.org/phylo/treebase/phylows/study/TB2:S11452

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Geographic differentiation among populations of *D. quinaria* from PA, NY, and ME.

Table S2 Polymorphism in *D. quinaria*.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.