Sequential Evolution of a Symbiont Inferred From the Host: *Wolbachia* and *Drosophila simulans*

J. William O. Ballard

The Roy J. Carver Center for Comparative Genomics, Department of Biological Science, University of Iowa

This study aims to unravel the biogeography of a model symbiont/host system by exploiting the prediction that a symbiont will leave a signature of infection on the host. Specifically, a global sample of 1,442 *Drosophila simulans* from 33 countries and 64 sampling localities was employed to infer the phylogeography of the maternally inherited alpha-proteobacteria *Wolbachia*. Phylogenetic analyses, from three symbiont genes and 24 mtDNA genomes (excluding the A + T–rich region), showed that each of four *Wolbachia* strains infected *D. simulans* once. The global distribution and abundance of the *Wolbachia* strains and the three mtDNA haplogroups (*D. simulans si*I, *si*II and *si*III) was then determined. Finally, network analyses of variable regions within *si*I (584 bp from seven additional lines) and *si*II (1,701 bp from 383 lines) facilitated a detailed biogeographic discussion. There is little variation in *si*III and the haplogroup is restricted in its distribution. These data show how the history of an infection can be mapped by combining data from the symbiont and the host. They say little about the organismal history of the host because the mtDNA genome is a biased representation of the whole genome.

Introduction

This study aims to unravel the biogeography of a model symbiont/host system by exploiting the prediction that a symbiont will leave a signature of infection on the host. Specifically, I examine the hypothesis that *Wolbachia* infection will leave a statistical signature of infection on host mtDNA because both are maternally inherited. *Wolbachia* are closely related to *Anaplasma marginale*, *Ehrlichia risticii*, and the *Rickettsia* spp., all arthropod-borne pathogens of mammals. They are obligate mutualists in nematodes (Bandi et al. 1998), and in one study they infected 76% of all insects tested (Jeyaprakash and Hoy 2000), making this group of bacteria the most prevalent symbiont (other than organelles) on the planet. *Wolbachia* cause a number of phenotypic effects in insects and terrestrial crustaceans, including thelytokous parthenogenesis (Breeuwer and Werren 1990), lethality to male embryos (Hurst, Hurst, and Majerus 1993; Jiggins, Hurst, and Majerus 2000), feminization of genetic males (Martin, Juchault, and Legrand 1973), and incompatibility in a variety of organisms including mosquitoes and *Drosophila* (Yen and Barr 1973; Hoffmann, Turelli, and Simmons 1986; James and Ballard 2000). It is intriguing that, while these phenotypes are diverse, they all subvert the host's reproductive system for the benefit of the bacterium. This drive is a significant force that explains the high prevalence of *Wolbachia* in the invertebrate world.

A model *Wolbachia* system occurs in *D. simulans*, a human commensal with a cosmopolitan distribution. There is little autosomal subdivision in this species (Begun and Aquadro 1993; Eanes et al. 1996; Hamblin and Veuille 1999; Kliman et al. 2000; Andolfatto 2001); however, it is well known to have three distinct mitochondrial DNA haplogroups (*si*I, *si*II, and *si*III). These haplogroups were first identified by Solignac and Monnerot (1986), who divided 13 isofemale lines of *D. simulans* into three mtDNA cleavage morphs based on 12 restriction enzymes. Baba-Aïssa et al. (1988) extended the survey and found that the three mitochondrial types differed by 10 to 15 restriction sites, and that variability was absent or was restricted to a single site within a type. Evidence for reduced levels of sequence variation within the *si*II haplotype group was subsequently observed at the sequence level at both the NADH dehydrogenase subunit 5 (Rand, Dorfsman, and Kann 1994) and cytochrome *b* (Ballard and Kreitman 1994) loci. Ballard (2000*a*) extended previous studies and compared the complete mtDNA sequence, excluding the A + T–rich region, from 22 *D. simulans* isofemale lines with that observed at intron 1 of the *alcohol dehydrogenase-repeated* locus. In that study, patterns of variation suggested that distinct forces are influencing the evolution of mtDNA and autosomal DNA in *D. simulans*.

One explanation for the high interhaplogroup divergence and low intrahaplogroup diversity in the mtDNA of *D. simulans* is adaptation of the mtDNA genome, or of specific nuclear-mitochondrial gene complexes, to the local environment. A given mutation may confer a selective advantage directly on the mitochondrial genome or by epistatic interactions with proteins imported from the nucleus (Ballard and Dean 2001). We have compared the distinct population genetic structure shown in the mtDNA with three nuclear loci (Ballard et al. 1996; Ballard 2000*a*; Ballard, Chernoff, and James 2002; Dean et al. 2003). In each case there was no correlation. However, only one was a nuclear locus that produces a protein that is imported into the mitochondria and essential for oxidative phosphorylation (Ballard, Chernoff, and James 2002).

An alternate, or perhaps additional, explanation for the observed population subdivision is that maternally inherited *Wolbachia*-induced cytoplasmic incompatibility, or *Wolbachia*-induced fitness increase, has significantly influenced the evolution of *D. simulans* mtDNA. If this is true, it is predicted that *Wolbachia* infection will leave a signature of infection on the mtDNA genome. In the simplest case, incompatibility occurs when an uninfected female mates with an infected male, causing a reduction in the egg hatch rate (see Hoffman and Turelli [1997] for

Key words: *Wolbachia*, mitochondria, *Drosophila simulans*, co-evolution, symbiosis.

E-mail: bill-ballard@uiowa.edu.

Mol. Biol. Evol. 21(3):428–442. 2004
DOI: 10.1093/molbev/msh028
Advanced Access publication December 5, 2003
Molecular Biology and Evolution vol. 21 no. 3

review). *Wolbachia*-induced incompatibility will cause the symbiont and the linked maternally inherited mitochondrial genotype to rise in frequency, in theory (Caspari and Watson 1959), in population cages (Nigro and Prout 1990; Kambhampati, Rai, and Verleye 1992), and in nature (Turelli and Hoffmann 1991). Compelling evidence that *Wolbachia* cause incompatibility in Drosophila came from treating infected lines with antibiotics, to cure the fly line of the bacteria (Hoffmann and Turelli 1988), and microinjection to introduce *Wolbachia* to uninfected lines (Boyle et al. 1993). In the first case incompatibility was abated, and in the second case incompatibility was induced. The physiological mechanism of incompatibility is not known. Sperm enter the egg normally (Lassy and Karr 1996), but paternal chromosomes fail to participate in first mitosis, leaving a haploid embryo (Stouthamer and Kazmer 1994; Callaini et al. 1996; Tram and Sullivan 2002). One intriguing hypothesis is that the density of bacteria in the host will correlate with the intensity of incompatibility expression within some strains (Breeuwer and Werren 1993; Clancy and Hoffmann 1998). Clearly, however, both symbiont and host affect the expression of incompatibility.

In *D. simulans*, six *Wolbachia* have been named. The names follow the location or country where the infection was first collected. The first *Wolbachia* to be identified was classified by the incompatibility phenotype of infected *D. simulans* Riverside (DSR) males (Hoffmann, Turelli, and Simmons 1986). This *Wolbachia* subsequently has been designated *w*Ri (*Wolbachia* from Riverside). A fly line from Hawaii was found to harbor a second *Wolbachia*, *w*Ha (from Hawaii), and is bi-directionally incompatible with the Riverside infected line (O'Neill and Karr 1990). *Wolbachia w*Ma from northern Madagascar (Mont Ambre, called $w\text{Ma}_{\text{Ma}}$ here), and *w*No from Nouméa (called $w\text{Ma}_{\text{No}}$ here) were described by a unique 16S rDNA sequence (Rousset, Vautrin, and Solignac 1992). Flies from New Caledonia and the Seychelles were found to be doubly infected with *w*Ha and $w\text{Ma}_{\text{No}}$ *Wolbachia* (Rousset and Solignac 1995) and were shown to be bidirectionally incompatible with *w*Ri (Merçot et al. 1995). Hoffmann, Clancy, and Duncan (1996) then described a *Wolbachia* from Australia that does not induce high incompatibility in the host (*w*Au from Australia). The sixth *Wolbachia* to be named came from flies collected near Mount Kilimanjaro in Tanzania and was termed *w*Ki (called $w\text{Ma}_{\text{Ki}}$ in this study) by Merçot and Poinsot (1998*a*). Charlat, Le Chat, and Merçot (2003) previously noted the similarity among the *w*Ma variants. I follow James and Ballard (2000) and designate uninfected fly lines as *w*-.

Data presented in this study show that combining data from the symbiont and the host can unravel the history of an infection. Linked phylogenetic studies and network analyses suggest that the *w*Ma strain is the oldest infection in the species infecting *si*I and *si*III flies. The *w*Ha infection likely occurred in the Seychelles Islands before the divergence of *D. simulans si*I and *D. sechellia*. Doubly infected (*w*Ma + *w*Ha) *si*I flies then dispersed to New Caledonia. The *w*Ma strain was then lost in *si*I flies prior to, or during, colonization of Tahiti and Hawaii. It is hypothesized that *Wolbachia*-uninfected *si*II dispersed out of Africa and were subsequently infected with the *w*Au or *w*Ri strains in Ecuador (no double infections have been found). Both infections have independently spread back to Africa, opening up the potential for admixture in African populations. The *w*Ma-infected *si*III flies occur in Kenya, Tanzania, Madagascar, and Reunion Island, and two uninfected populations have been found in coastal Kenya and Tanzania.

Materials and Methods

This study employed a multifaceted approach to investigate the association of *Wolbachia* with *D. simulans* mtDNA. As the nuances of the subdivision have unfolded, more specific assays have been developed, and multiple references are made to particular techniques that have been employed over time to increase efficiency. First, *D. simulans* flies were collected in the field or they were obtained from colleagues. Second, the number of *Wolbachia* strains infecting the host was determined. In this study three regions of *Wolbachia* DNA were sequenced. Third, the distribution and abundance of the *Wolbachia* strains in *D. simulans* were determined. Fourth, five additional complete mtDNA genomes, excluding the A + T–rich region were sequenced and added to a preexisting data set (Ballard 2000*a*) so the *Wolbachia* strains could be mapped onto the mtDNA genealogy. The observation that the three *D. simulans* mtDNA haplogroups are not monophyletic relative to *D. sechellia* and *D. mauritiana ma*I is not novel and is not discussed here (Solignac and Monnerot 1986; Ballard 2000*c*). Fifth, the distribution and abundance of haplogroups were determined. Finally, network analyses of variable regions within mtDNA haplogroups enabled biogeographic analyses. When combined with the phylogenetic studies the network analyses enable a temporal component to be included in the biogeographic discussion. However, I do not attempt to date the events in this study because of the high error associated with the low numbers of changes and because *Wolbachia* clearly influences the evolution of mtDNA in Drosophila.

Drosophila Lines and Wild-Caught Males

Females within the *D. melanogaster* subgroup were sorted in the field and placed individually into vials. Genital arch morphology of male offspring confirmed species identification. When it was not possible to maintain live lines, or when the density of *D. simulans* was low, males were placed into 2-ml screw-top vials containing 100% ethanol and included in subsequent analyses. The remaining *D. simulans* were obtained from colleagues either as isofemale lines or as wild-caught flies. A total of 1,442 *D. simulans* are included in this study. The two *D. sechellia* included in the study were collected in the Seychelles Islands by the author.

Genomic DNA extraction, polymerase chain reaction (PCR) amplification and sequencing followed Ballard (2000*a*) and Dean et al. (2003). Unless otherwise stated, both strands were sequenced using Taq-Dye Deoxy Terminator Cycle sequencing. Sequences were imported

into the Sequencher software program, the chromatograms investigated, and contigs constructed.

Wolbachia Strains and Isolates

It is not simple to define a strain of *Wolbachia*. To avoid confusion in this study, a strain is defined on the basis of ''common ancestry'' (Lincoln, Boxshall, and Clark 1998). Specifically, strains must be monophyletic as determined by DNA sequence data. Here, a sequence isolate is defined as having a unique DNA sequence. Thus, multiple sequence isolates may occur within a strain, just as multiple mtDNA haplotypes may occur within a monophyletic haplogroup. Isolates will be shown with a subscript following the strain designation. In this study ''common physiological traits'' and ''characteristic properties'' are not employed to help define strains (Lincoln, Boxshall, and Clark 1998), because they have not been determined in a common host.

Wolbachia sequence data were obtained from 13 isofemale lines of *D. simulans* (table 1). A total of 2,532 bp was obtained from three loci: 16S rDNA (848 bp), *Wolbachia* surface binding protein (*wsp*) (627 bp), and the rapidly evolving cell-cycle gene *ftsz* (1,057bp). The 16S rDNA was amplified following O'Neill et al. (1992), *wsp* following Zhou, Rousset, and O'Neill (1998) and James and Ballard (2000), and *ftsz* following Werren, Zhang, and Guo (1995). When a fly line was doubly infected each PCR amplicon was cloned and a minimum of two copies of each *Wolbachia* were sequenced. The outgroup *Wolbachia* infects the nematode *Onchocerca gibsoni* (Bandi et al. 1998).

Sequences were aligned in ClustalX using a gap-opening penalty of 50 and gap extension penalty of 5 (Thompson et al. 1997). Sequences were exported into PAUP* (Swofford 1998) and a single data matrix constructed. Gaps were treated as missing and 10 additional characters were coded as insertion or deletion events (indels).

The data were analyzed by likelihood. To establish the most appropriate likelihood model for analyzing the *Wolbachia* data, a Neighbor-Joining search was conducted using PAUP* (Swofford 1998). The likelihood ratio test obtained the most appropriate model for the analysis (Swofford et al. 1996). The general time reversible (GTR) model, with the proportion of invariable sites and the gamma distribution estimated from the data, was selected as the model (GTR + I + Γ). Parsimony was then employed to map the number of changes onto each branch. There was no evidence that the genes generate a different phylogenetic signal.

Wolbachia Distribution and Abundance

To determine infection status of flies, the conserved 16S rDNA primers (O'Neill et al. 1992) were employed. There are at least three possible explanations for a failed 16S rDNA amplification. First, the line or male may be *Wolbachia* uninfected. All presumptive *Wolbachia*-uninfected DNA samples were checked with *wsp* primers. Second, the DNA in the extraction may not be amplifiable.

Table 1
Lines of *Drosophila simulans* and the *Wolbachia* Used for Strain and Isolate Identification

Fly Line	*Wolbachia* Strain	Isolate	Haplogroup	Collection Site
NC48	*w*Ha & *w*Ma	$w\text{Ma}_{\text{No}}$	*si*I	Nouméa, New Caledonia[1]
NC102	*w*Ha & *w*Ma	$w\text{Ma}_{\text{No}}$	*si*I	Nouméa, New Caledonia[2]
HW00	*w*Ha		*si*I	Honolulu, Hawaii, USA[3]
TT01	*w*Ha		*si*I	Papeete, Tahiti[2]
N7No	*w*Ma	$w\text{Ma}_{\text{No}}$	*si*I	Nouméa, New Caledonia[4]
NC117	*w*Ma	$w\text{Ma}_{\text{No}}$	*si*I	Nouméa, New Caledonia[2]
Coffs	*w*Au		*si*II	Coffs Harbour, Australia[5]
MD225	*w*Au		*si*II	Joffreville, Madagascar[2]
DSR	*w*Ri		*si*II	Riverside, CA, USA[5]
C167	*w*Ri		*si*II	Nanyuki, Kenya[6]
MD199	*w*Ma	$w\text{Ma}_{\text{Ma}}$	*si*III	Joffreville, Madagascar[2]
RU07	*w*Ma	$w\text{Ma}_{\text{Ki}}$	*si*III	Salazie, Reunion[2]
KC9	*w*Ma	$w\text{Ma}_{\text{Ki}}$	*si*III	Mt. Kilimanjaro, Tanzania[7]

NOTE.—Lines collected or provided by: 1 M. Solignac, 2 J.W.O. Ballard, 3 K. Kaneshiro, 4 H. Merçot, 5 A. A. Hoffmann, 6 A.V. Olembo, 7 D. Lachaise.

In all cases, amplification using conserved mtDNA primers tested whether the DNA was amplifiable. This was particularly important in the case of ethanol-preserved wild-caught males where the DNA may have been degraded (Dean and Ballard 2000). Third, individual offspring from an infected female may have lost the infection in the laboratory. In all these cases, three independent fly extractions with multiple primer pairs confirmed that the line was uninfected. This was an important step as some *Wolbachia* strains (e.g., *w*Ma) have less than 100% transmission fidelity in the laboratory.

Restriction fragment length polymorphism (RFLP) analysis, primer-specific amplifications, or sequence data from the *wsp* locus identified each infecting strain (James and Ballard 2000; James et al. 2002; Dean et al. 2003). Phylogenetic analysis presented in the Results section shows that the *wsp* locus accurately identifies four strains of *Wolbachia* infecting *D. simulans*.

MtDNA Genealogy

Phylogenetic analysis was employed to test the hypothesis that each *Wolbachia* strain invaded *D. simulans* once. Twenty-four mtDNA genomes, excluding the A + T–rich region, were included. Five genomes were sequenced for this study. Four lines (KY07, KY45, KY201, and KY215) were collected in Kenya. They were selected because they show high mtDNA diversity (Dean et al. 2003) and had the potential to break the long inter-haplogroup branches observed by Ballard (2000*a*). The AU23 line from Australia was included after completion of preliminary intrahaplogroup network analyses because of its key position in the network. The remaining 19 mtDNA genomes were from Ballard (2000*a*; 2000*b*) (16 *D.*

simulans, one *D. sechellia*, one *D. mauritiana ma*I and one *D. mauritiana ma*II). Here *D. mauritiana ma*II is employed as the outgroup. Six sequences from Ballard (2000*a*) were not included in the analysis. The excluded lines are homosequential with lines included in the analyses and carry the same strain of *Wolbachia*, or are uninfected and so carry little additional information.

In all cases, the DNA was extracted from individuals less than 14 days of age. The 15,034-bp mitochondrial genome was PCR amplified in 11 overlapping fragments (available at www.myweb.uiowa.edu/ballard). To minimize the possibility of contamination, each genome was completed before the next was commenced. Negative controls confirmed that there was no contamination. To sequence the mitochondrial molecule, 63–68 cycle sequencing reactions were employed. No inconsistencies between the sequences derived from independent PCR products were detected.

The alignment of the five additional mitochondrial genomes with those previously published was straightforward for the majority of the 15,034 bp. Ballard (2000*a*, 2000*b*) deleted 76 bp from the analysis because it is difficult to unequivocally determine the alignment between 5,535–5,584 and 6,022–6,047. Ambiguous alignment among haplogroups has the potential to increase homoplasy among haplogroups; however, inclusion of these characters has the potential to increase intrahaplogroup resolution. Preliminary analyses showed that the relationships among haplogroups are robust to the inclusion or exclusion of these regions, and they are included in this study. Fifty-six indel characters were included at the end of the matrix. Gaps were then scored as missing.

The genealogical relationships of the mtDNA data were analyzed with the HKY + I + Γ maximum likelihood model using PAUP* 4 (Swofford 1998). Steinbachs et al. (2001) investigated the efficiency of genes and the accuracy of 83 tree-building methods (27 distance, 4 parsimony, 50 maximum likelihood, and 2 Bayesian) in recovering a well-supported Drosophila mitochondrial genealogy. Here the HKY + I + Γ likelihood model is employed as it was shown to be robust with mtDNA sequence data obtained from the *D. melanogaster* subgroup.

A backbone constraint was employed to test the hypothesis that each strain of *Wolbachia*-infected *D. simulans* once. This is an appropriate constraint. In nature, it is hypothesized that parasitoid or mite-mediated horizontal transfer mediates the interspecific movement of *Wolbachia*; however, this infection mechanism has not been found stable in any species tested (Heath et al. 1999). Loss of infection, on the other hand, is an important character defining the frequency of *Wolbachia* infections (Hoffmann and Turelli 1988; Turelli and Hoffmann 1991; 1995).

MtDNA Distribution and Abundance

Determination of the mtDNA genomes facilitated the development of techniques for the rapid screening of mtDNA type. The *D. simulans* mtDNA haplogroup of isofemale lines/males was determined by PCR/RFLP (James and Ballard 2000), multiplex PCR (Dean et al. 2003), or direct sequencing (Ballard 2000*a*). The frequency of the three haplogroups was then plotted for sites where more than 20 individuals were sampled.

Biogeographic Analyses

For *si*I, 584 bp of mtDNA from five *D. simulans* lines (two from the Seychelles and three from Tahiti) were amplified, sequenced, and added to the data set of James et al. (2002). The amplified region spanned an intervening spacer between ND3 and the alanine tRNA, where a variable number of AT repeats has been observed (Ballard 2000*b*). *D. sechellia* is the outgroup to *si*I and two additional *D. sechellia* isofemale lines were sequenced. There is low variability in these flies, and both DNA strands were not sequenced for each line. Rather, any ambiguous or potentially informative site was confirmed by double-stranded sequencing.

For *si*II, three regions of mtDNA totaling 1,701 bp were amplified and sequenced from 383 isofemale lines/males following Ballard (2000*a*). The three regions were sampled because mtDNA positions 1558, 3441, 8175, and 8202 were variable in this haplogroup. The three primer pairs (1128+ and 1815–, 3182A+ and 3929–, and 7780+ and 8475– (www.myweb.uiowa.edu/ballard) amplified regions of five protein-coding genes (ND2, COI, COII, ND5, and ND4), six transfer RNAs ($tRNA_{trp}$, $tRNA_{cys}$, $tRNA_{tyr}$, $tRNA_{lys}$, $tRNA_{asp}$, and $tRNA_{his}$) and four intervening spacer regions. Again, sequencing was single-stranded, but ambiguous or potentially informative sites were always confirmed by double-stranded sequencing.

Flies with *si*III mtDNA were collected in Madagascar, Reunion Island, and continental eastern Africa. Ballard (2000*a*) sequenced nine mtDNA genomes from Madagascar and Reunion Island and observed just three singleton segregating sites. Dean et al. (2003) observed no segregating sites in 37 lines from Tanzania and Kenya. The exceptionally low variability in *si*III flies coupled with the apparent lack of diagnostic single-nucleotide polymorphic (SNP) sites precluded further analysis of this mtDNA type.

Networks were built using statistical parsimony (Templeton, Crandall, and Sing 1992) implemented in TCS version 1.13 (Clement, Posada, and Crandall 2000). Network analyses take into account the persistence of ancestral sequences and recombination by allowing multifurcations (Posada and Crandall 2001). TCS collapses identical haplotypes and calculates the number of mutational steps, below which, sequences can be joined with 95% confidence (Templeton, Crandall, and Sing 1992). The sampling strategy employed in this study is not random and only unique haplotypes are included in the analyses. To facilitate the *si*I analysis, each AT repeat was treated as a single character in the network analysis. However, it is not clear that this is the appropriate coding strategy as the biological mechanism causing the repeats is not known. An alternate approach is to examine the frequency of repeats (James et al. 2002). This alternative was rejected here because it does not facilitate inclusion of the outgroup and biogeographic discussion.

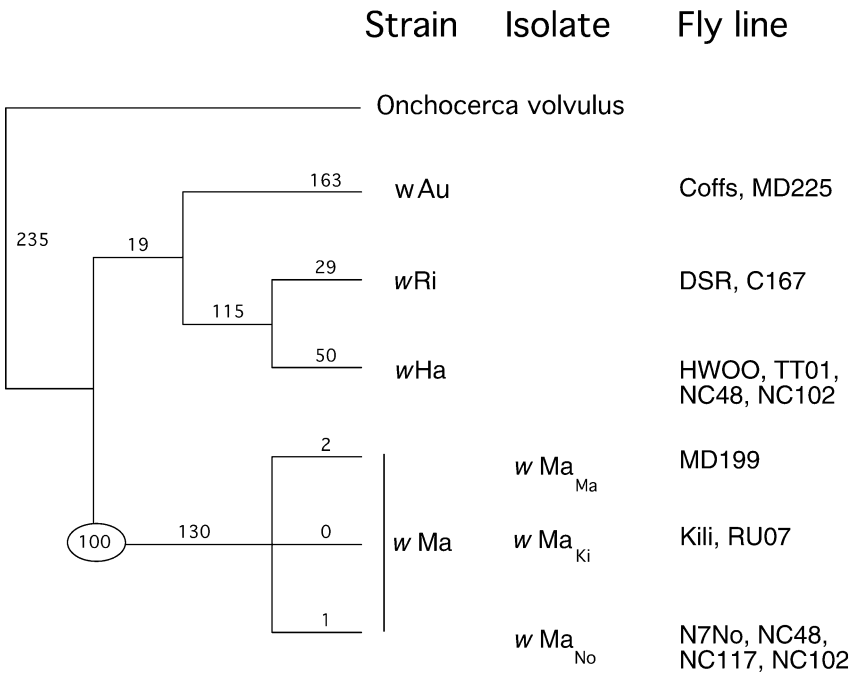

FIG. 1.—Phylogenetic hypotheses considering the relationships of the *Wolbachia* infecting *D. simulans*. The strains are *w*Au, *w*Ri, *w*Ha, and *w*Ma. The *w*Ma strain has three sequence isolates shown with a subscript. A maximum likelihood GTR + I + Γ exhaustive search (proportion of invariable sites = 0.50 and the gamma shape = 0.88) generates one tree with score of –ln L 6175.877. The criterion was changed to parsimony and the changes mapped onto the genealogy. Parsimony analysis recovered the identical tree topology. Bootstrap values above 95% are shown in circles.

Results

Wolbachia Strains and Isolates

When homosequential sequences were pooled, phylogenetic analysis of the three loci showed that four distinct strains and three isolates of *Wolbachia* infect *D. simulans* (fig. 1; table 1). Here, I follow precedence and employ the four strain names *w*Au, *w*Ha, *w*Ri, and *w*Ma. The strain *w*Au infects Coffs and MD225; *w*Ha infects HW00, TT01, NC48, and NC102; and *w*Ri infects DSR and C167. The three sequence variants of *w*Ma are considered isolates because of the low sequence variability and the lack of significant monophyly. The nomenclature of the isolates follows existing published designations: $w\text{Ma}_{\text{Ma}}$ infects MD199; $w\text{Ma}_{\text{Ki}}$ infects Kili and RU07; and $w\text{Ma}_{\text{No}}$ infects NC48, N7No, NC117, and NC102. Phylogenetic analysis considering all sequences showed the same pattern. Each strain is significantly monophyletic, with a bootstrap value over 95% (Felsenstein and Kishino 1993).

Following the procedure of Conditional Data Combination (Bull et al. 1993), the data were divided into four partitions: the 16S rDNA, *wsp*, *Ftsz*, and indels. The incongruence length difference (ILD) test (Farris et al. 1995) was employed to test the null hypothesis that the partitions are evolving under homogeneous processes. This test supported the hypothesis that the signal in four data sets does not conflict (611 steps $P = 0.75$). However, Barker and Lutzoni (2002) warned that the ILD test is not a good method for testing the combinability of data sets, so the topology of each locus was also analyzed independently.

The 16S rDNA data set is 848 characters in length (GenBank X61769, X61770, AF390865, X64266, AF312372). Thirteen characters are parsimony informative. A GTR + I + Γ maximum likelihood model generates a tree with two major clades (–ln L 1363.31 with the likelihood parameters estimated from the data). One clade includes *w*Ma; the second includes *w*Au, *w*Ha, and *w*Ri. Within the *w*Ma clade, a T → C substitution in a loop region identifies $w\text{Ma}_{\text{No}}$ (NC48, N7No, NC117, NC102), while $w\text{Ma}_{\text{Ma}}$ and $w\text{Ma}_{\text{Ki}}$ are homosequential (MD199, Kili, RU07). Within the second clade, a C → T substitution identifies *w*Ha (TT01, NC48, NC102, HW09). There are no differences between *w*Ri (DSR, C167) and *w*Au (Coffs, MD225).

The *wsp* data set reliably distinguishes the four *Wolbachia* strains (GenBank AF020068, AF020070, AF020067, AF020074). It is 627 bases in length and contains 127 parsimony-informative characters. A GTR + I + Γ maximum likelihood model generates a tree with four clades consisting of the four *Wolbachia* strains with no intrastrain variation: –ln L 2138.57.

The *ftsz* data set is 1,057 bases in length and contains 129 parsimony-informative characters (GenBank no. AY508998-901). A likelihood model generates a tree with two clades (–ln L 2325). One clade clusters *w*Ma. Within this clade, one nonsynonymous and one synonymous substitution differentiate $w\text{Ma}_{\text{Ma}}$ (ATA → ATG at position 255 and TTT → TTC at position 699 of *ftsz*) in MD199. The *Wolbachia* $w\text{Ma}_{\text{No}}$ (NC48, N7No, NC117, NC102) and $w\text{Ma}_{\text{ki}}$ are homosequential (Kili, RU07). Within a second clade, a single substitution 25 bp upstream of the *ftsz* initiation codon distinguishes *w*Au from *w*Ri and *w*Ha. The sequence obtained from DSR is

identical to that previously published (GenBank U28178), whereas the sequence from HW00 differs by a single nucleotide (ACA → ACG at position 924) from the published sequence (GenBank U28185). The latter difference supports the hypothesis that variation exists within a *Wolbachia* strain.

The indel data set is 10 characters in length and contains four parsimony-informative characters. This data set consists of zeros and ones, and these data were analyzed by parsimony. Parsimony analysis (seven equally parsimonious trees of length 10 steps) distinguishes four distinct strains with no intrastrain variation. All strains differ from *w*Ri. The *w*Au strain has two indel events, and the *w*Ha and *w*Ma strains both have four independent indel events.

Wolbachia Distribution and Abundance

Table 2 collates all *Wolbachia* infections and figure 2 shows infection frequencies, where more than 20 flies were collected from a specific site. Flies infected with *w*Au were identified from Australia, Cameroon, Ecuador, Japan, and the continental USA, whereas flies singly infected with the strain *w*Ma were identified from Kenya, Madagascar, New Caledonia, Reunion, and Tanzania. Flies singly infected with *w*Ha were collected from the Pacific Islands of Hawaii, New Caledonia, and Tahiti. Flies doubly infected with *w*Ma and *w*Ha were only collected from New Caledonia and the Seychelles. Flies infected with *w*Ri were collected from Bolivia, China, Congo, Cook Islands, Ecuador, France, Gabon, Greece, Israel, Japan, Kenya, Malawi, Mexico, Morocco, Seychelles, South Africa, Spain, Tanzania, Tunisia, Ukraine, and the continental USA.

MtDNA Genealogy

Five mtDNA genomes were sequenced (GenBank no. AY518670-4) and added to the preexisting dataset (Ballard 2000*b*; 2000*c*). Of the 15,091 characters (1–15,034 mtDNA sequence and 15,035–15,091 indels), 551 were parsimony informative. A heuristic HKY + I + Γ likelihood search with 10 random starting trees recovered one tree of length –ln L 24250.22 (fig. 3*A*). This data set was not divided into process partitions because there is no evidence for recombination in the mtDNA genome of Drosophila.

The hypothesis that each strain of *Wolbachia* invaded *D. simulans* once is not rejected by these data. The backbone constrained search yielded one tree that differed only in the placement of the DSR fly (fig. 3*B*). Analysis of the unconstrained and constrained trees using the Shimodaira-Hasegawa test (Shimodaira and Hasegawa 2001) showed no significant difference in tree lengths ($P = 0.357$).

MtDNA Distribution and Abundance

Distributions of haplogroups are presented in table 2. Figure 4 plots the abundances of each haplogroup where more than 20 flies were sampled in a population. Flies with *si*I mtDNA were collected in the Seychelles and the Pacific Islands of Hawaii, New Caledonia, and Tahiti. Flies with *si*II mtDNA were collected in Australia, South America (Bolivia and Ecuador), North America (Mexico and the continental USA), Asia (China, Japan, Cook Islands, and India), Africa (Cameroon, Congo, Egypt, Ethiopia, Gabon, Kenya, Madagascar, Malawi, Morocco, South Africa, and Tanzania), Europe (France, Greece, Spain, Israel, and Ukraine), and the islands of Jamaica, Reunion, and the Seychelles. The *si*III haplotype was collected in continental eastern Africa (Kenya and Tanzania) and in the islands of Madagascar and Reunion. No *D. simulans* were collected on the Pacific islands of Upolu, independent Samoa, or Viti Levu, Fiji, in 2001.

Biogeography

In *si*I flies, network analysis of AT repeat number in the intervening sequence between ND3 and the alanine tRNA is a linear array (fig. 5). There are six AT repeats in all *si*I lines collected in the Seychelles. James et al. (2002) reported that the number of AT repeats ranged from 5 to 11 in 54 flies collected from New Caledonia. In this study, 10 AT repeats were observed in all *w*Ha infected lines from Tahiti. Ballard (2000*a*) sequenced two *w*Ha-infected lines from Tahiti: TT01 line exhibits 10 AT repeats and TT00 exhibits 9 AT repeats. In 18 *w*Ha-infected flies from Hawaii, the number of repeats ranged from 6 to 10. Phylogenetic analysis of complete mtDNA data showed that *D. sechellia* is the outgroup to *si*I mtDNA (fig. 3). In all lines of *D. sechellia* sequenced, the intervening sequence is ATACACATATAT. However, phylogenetic analysis of this region suggests that both T → C substitutions occurred in the branch to *D. sechellia*. These data suggest that six AT repeats is the ancestral repeat number in *si*I (fig. 5), and *si*I flies spread from the Seychelles Islands.

In *si*II, 389 sequences of 1,701 bp identified 28 distinct haplotypes (table 3 and fig. 6). Phylogenetic analysis clearly showed that the basal *si*II lineages collected from Kenya, Tanzania, and Madagascar are not infected with *Wolbachia* (fig. 2). It is hypothesized that uninfected flies with the KY45 haplotype migrated to Ecuador, where the 1,626 T → C mutation generated the AU23 haplotype. The *w*- AU23 haplotype was then infected with the *w*Au strain of *Wolbachia* (in Australia, 28 of 33 flies with this haplotype were uninfected). In parallel, AU23 had the common synonymous 8,201 G → A mutation to the DSR fly haplotype, which subsequently became *w*Ri infected. The *w*Ri infection then spread worldwide. All key haplotypes were collected in the Sangoqui Markets, Ecuador, suggesting this as a possible region for both infection events (KY45, AU23, DSR, DSW, LA28, and Coffs; fig. 6; table 3).

Discussion

Defining a ''strain'' of *Wolbachia* is challenging, and the approach used in this study is one of many possible. Clearly, raising each unique sequence type to the level of a strain is not biologically informative as each ''strain'' will map to a terminal branch in a phylogeny and little can be inferred biogeographically or evolutionarily. Data indicated that four *Wolbachia* strains infect the three *D. simulans* mtDNA haplogroups. It is possible, however, that additional sampling, sequence data, or data from

Table 2
Fly Lines Assayed in this Study

Site	Locality	*Wolbachia* Strain	mtDNA	*N*	Year
Australia	Brisbane	*w*Au	*si*II	17	1999[1]
Australia	Brisbane	*w*-	*si*II	21	1999[1]
Australia	Coffs Harbour	*w*Au	*si*II	1	1995[6]
Australia	Coffs Harbour	*w*Au	*si*II	32	1999[1]
Australia	Coffs Harbour	*w*-	*si*II	16	1999[1]
Australia	Richmond	*w*Au	*si*II	3	1999[1]
Australia	Richmond	w-	*si*II	3	1999[1]
Bolivia	?	*w*Ri	*si*II	1	1993–7[7]
Cameroon	Yaounde	*w*Au	*si*II	2	1998[8]
Cameroon	Yaounde	*w*-	*si*II	2	1999[1]
China	Li Jiang	*w*Ri	*si*II	25	2002[2]
China	Li Jiang	*w*	*si*II	5	2002[2]
Congo	Brazzaville?	*w*Ri	*si*II	2	?[9]
Cook Islands	Rarotonga	*w*Ri	*si*II	42	2001[3]
Cook Islands	Rarotonga	*w*-	*si*II	1	?[10]
Cook Islands	Rarotonga	*w*-	*si*II	1	2001[3]
Ecuador	Rocafuerta	*w*Au	*si*II	1	2000[11]
Ecuador	Rocafuerta	*w*Ri	*si*II	29	2000[11]
Ecuador	Rocafuerta	*w*-	*si*II	7	2000[11]
Ecuador	Sangoqui	*w*Au	*si*II	22	2000[1,11]
Ecuador	Sangoqui	*w*Ri	*si*II	7	2000[1,11]
Ecuador	Sangoqui	*w*-	*si*II	9	2000[1,11]
Egypt	Cairo?	*w*-	*si*II	1	1992[7]
Ethiopia	Welo, Ataye Ri.	*w*-	*si*II	1	1990[12]
France	Villeurbanne	*w*Ri	*si*II	1	1992?[7]
France	Valence	*w*Ri	*si*II	1	1993?[7]
Gabon	Franceville	*w*Ri	*si*II	29	2002[1,13]
Gabon	Franceville	*w*-	*si*II	1	2002[1,13]
Greece	Athens	*w*Ri	*si*II	54	2000[1,2,3]
Greece	Chania	*w*Ri	*si*II	31	2000[1,2,3]
Greece	Crete	*w*Ri	*si*II	39	2000[1]
Greece	Crete	*w*-	*si*II	1	2000[1]
India	Dehli	*w*-	*si*II	1	1998[14]
Israel	''Evolution'' Canyon	*w*Ri	*si*II	1	1993[7]
Jamaica	Runaway Bay	*w*Au	*si*II	2	2000[15]
Japan	Chiba	*w*Ri	*si*II	48	2000[9]
Japan	Chiba	*w*-	*si*II	4	2000[9]
Japan	Chichi-jima Is.	*w*Au	*si*II	6	1994[9]
Japan	Chichi-jima Is.	*w*-	*si*II	28	1994[9]
Japan	Miyuki-no-hana,	*w*Au	*si*II	1	1997[9]
Japan	Haha-jima Is.	*w*-	*si*II	5	1997[9]
Japan	Ogasawara	*w*Au	*si*II	1	1996[9]
Japan	Ogasawara	*w*Ri	*si*II	1	1996[9]
Japan	Ogasawara	*w*-	*si*II	5	1996[9]
Japan	Tajima	*w*Au	*si*II	4	1997[9]
Japan	Haha-jima Is.	*w*-	*si*II	6	1997[9]
Kenya	Malindi	*w*Ri	*si*II	1	2001[1]
Kenya	Malindi	*w*-	*si*II	35	2001[1]
Kenya	Malindi	*w*-	*si*III	24	2001[1]
Kenya	Nairobi	*w*Ri	*si*II	1	2001[1,16]
Kenya	Nairobi	*w*-	*si*II	5	1979[9]
Kenya	Nairobi	*w*Ma	*si*II	27	2001[1,16]
Kenya	Nairobi	*w*-	*si*III	14	2001[1,16]
Kenya	Nanuki	*w*Ri	*si*II	1	1973[17]
Kenya	?	*w*-	*si*II	1	1988[7]
Madagascar	Ambositra	*w*Au	*si*II	7	1998[1]
Madagascar	Ambositra	wMa	*si*III	3	1998[1]
Madagascar	Ambositra	*w*-	*si*III	1	1998[1]
Madagascar	Antananarivo	*w*Au	*si*II	19	1998[1]
Madagascar	Antananarivo	*w*-	*si*II	1	?[18]
Madagascar	Antananarivo	*w*-	*si*II	4	~1987[19]
Madagascar	Antananarivo	*w*-	*si*II	5	1993[9]
Madagascar	Antananarivo	*w*-	*si*II	21	1998[1]
Madagascar	Antananarivo	*w*-	*si*III	14	1993[9]
Madagascar	Antananarivo	*w*-	*si*III	1	~1987[19]
Madagascar	Antananarivo	*w*-	*si*III	29	1998[1]
Madagascar	Antsirabe	*w*Au	*si*II	21	1998[1]

Table 2
Continued

Site	Locality	*Wolbachia* Strain	mtDNA	*N*	Year
Madagascar	Antsirabe	*w*-	*si*II	19	1998[1]
Madagascar	Antsirabe	*w*Ma	*si*III	5	1998[1]
Madagascar	Antsirabe	*w*-	*si*III	15	1998[1]
Madagascar	Joffreville	*w*Au	*si*II	11	1998[1]
Madagascar	Joffreville	*w*-	*si*II	4	1998[1]
Madagascar	Joffreville	*w*Ma	*si*III	6	1998[1]
Madagascar	Joffreville	*w*-	*si*III	12	1998[1]
Madagascar	Ranomafana	*w*Au	*si*II	12	1998[1]
Madagascar	Ranomafana	*w*-	*si*II	9	1998[1]
Madagascar	Ranomafana	*w*Ma	*si*III	2	1998[1]
Madagascar	Ranomafana	*w*-	*si*III	7	1998[1]
Malawi	Mwanza	*w*Ri	*si*II	28	2001[1]
Mexico	Taxco	*w*Ri	*si*II	44	2000[1,20]
Morocco	?	*w*Ri	*si*II	1	?[7]
New Caledonia	Nouméa	*w*Ma	*si*I	11	1989[8]
New Caledonia	Nouméa	*w*Ma	*si*I	3	1999[1]
New Caledonia	Nouméa	*w*Ha	*si*I	4	1999[1]
New Caledonia	Nouméa	*w*Ha & *w*Ma	*si*I	3	1991[21]
New Caledonia	Nouméa	*w*Ha & *w*Ma	*si*I	47	1999[1]
New Caledonia	Nouméa	*w*-	*si*I	1	1999[1]
Reunion	St.-Pierre	*w*-	*si*II	1	1998[1]
Reunion	St.-Pierre	*w*Ma	*si*III	4	1998[1]
Reunion	St.-Denis	*w*-	*si*III	1	1979[22]
Reunion	St.-Denis	*w*-	*si*III	1	1987[23]
Reunion	Bois des Nefles	*w*-	*si*III	2	1993[19]
Seychelles	Mahe?	*w*Ha & *w*Ma	*si*I	1	?[8]
Seychelles	Mahe?	*w*Ha & *w*Ma	*si*I	1	?[24]
Seychelles	Mahe?	*w*Ha & *w*Ma	*si*I	5	1987[9]
Seychelles	Mahe?	*w*Ri	*si*II	1	?[25]
Spain	Canary Is.	*w*Au	*si*II	1	1994[7]
Spain	?	*w*-	*si*II	1	1992[7]
South Africa	Cape Town	*w*Ri	*si*II	4	2000[26]
South Africa	Cape Town	*w*-	*si*II	1	2000[26]
South Africa	Pretoria	*w*Ri	*si*II	24	2001[27]
South Africa	Pretoria	*w*-	*si*II	1	2001[27]
South Africa	?	*w*Ri	*si*II	2	?[9]
South Africa	?	*w*-	*si*II	1	?[9]
Tahiti	Morea	*w*Ha	*si*I	1	?[25]
Tahiti	Papeete	*w*Ha	*si*I	4	1998[1]
Tanzania	Mt. Kilamanjaro	*w*Ma	*si*III	1	1997[8]
Tanzania	Dar es Salaam	*w*Ri	*si*II	1	2001[1,28]
Tanzania	Dar es Salaam	*w*-	*si*II	17	2001[1,28]
Tunisia	?	*w*Ri	*si*II	11	?[9]
Tunisia	?	*w*-	*si*II	1	?[9]
Ukraine	Kiev	*w*Ri	*si*II	38	2002[4]
Ukraine	Kiev	*w*-	*si*II	4	2002[4]
Ukraine	Yalta	*w*Ri	*si*II	26	2002[4]
Ukraine	Yalta	*w*-	*si*II	2	2002[4]
USA (Hawaii)	Oahu	*w*Ha	*si*I	14	1998[29]
USA (Hawaii)	Oahu	*w*Ha	*si*I	4	1998[30]
USA (Hawaii)	Oahu	*w*Ha	*si*I	36	1998[1]
USA (Hawaii)	Kauai	*w*Ha	*si*I	62	1998[1]
USA (Continental)	Riverside, Calif.	*w*Ri	*si*II	1	1987[6]
USA (Continental)	Watsonville, Calif.	*w*-	*si*II	1	1985[6]
USA (Continental)	Gainesville, Fla.	*w*Ri	*si*II	1	2002[5]
USA (Continental)	Gainesville, Fla.	*w*-	*si*II	20	2002[5]
USA (Continental)	Lantana, Fla.	*w*Au	*si*II	1	1994[31]
USA (Continental)	Lantana, Fla.	*w*-	*si*II	1	1994[31]
USA (Continental)	Tallahassee, Fla.	*w*Ri	*si*II	131	1999[2]
USA (Continental)	Tallahassee, Fla.	*w*-	*si*II	1	1999[2]
USA (Continental)	Brooksville, Fla.	*w*Au	*si*II	2	2002[1]
USA (Continental)	Brooksville, Fla.	*w*Ri	*si*II	23	2002[1]
USA (Continental)	Brooksville, Fla.	*w*-	*si*II	1	2002[1]
USA (Continental)	Iowa City, Iowa	*w*Ri	*si*II	15	2001[2]
USA (Continental)	Iowa City, Iowa	*w*Ri	*si*II	15	2002[2]
USA (Continental)	Iowa City, Iowa	*w*-	*si*II	1	2002[2]

Table 2
Continued

Site	Locality	*Wolbachia* Strain	mtDNA	*N*	Year
Zimbabwe	Harare	*w*Ri	*si*II	2	1994[32]
Zimbabwe	Harare	*w*-	*si*II	8	1994[32]
Zimbabwe	Victoria Falls	*w*Ri	*si*II	48	2001[1,33]

NOTE.—Lines collected or provided by: 1 J. W. O. Ballard, 2 A. C. James, 3 M. D. Dean, 4 T. Nosenko, 5 M. Zickell, 6 A. A. Hoffmann, 7 C. Vieira-Heddi, 8 H. Merçot. 9 R. Kondo, 10 Species stock center, 11 D. L. Vela, L. Lopez, and T. Moran, 12 F. Lemeunier, 13 S. Charlat, 14 M. Habibula, 15 J. Bond, 16 K. Maes, 17 A. Olembo, 18 J. Roote, 19 R. Russell, 20 P. Grace, 21 F. Baba-Aissa, 22 O. Kitagawa, 23 S. I. Chigusa, 24 D. Presgraves, 25 M. Solignac, 26 R. Bowie, 27 A. Potts, 28 C. Meena and G. Mtoka, 29 K. Kaneshiro, 30. D. Baer, 31 M. Kreitman, 32 T. Mutangadura, 33 E. Zaranyika.

phenotypic expression of transinfected strains will demonstrate that additional strains exist.

The *w*Ma strain appears to have been associated with *D. simulans* for the longest period and is the only strain observed to have multiple sequence isolates. A possible site for *w*Ma infection is Madagascar, as this is probably the region of endemicity for *D. simulans* (Lachaise et al. 1988). Data from mtDNA support this hypothesis. The *w*Ma strain may infect *si*I and *si*III haplogroup flies, but it has been lost in the *si*II haplogroup. The wMa_{No} isolate may infect flies with the *si*I haplotype (N7No, NC48, NC117, NC102). The wMa_{Ma} isolate infects the MD199 *si*III line that was collected in Madagascar. The wMa_{Ki} isolate infects *si*III flies in Reunion (RU07) and eastern Africa (Tanzania). It has been suggested that *Wolbachia* from Tanzanian *si*III flies ''rescued'' the incompatibility of some singly infected *si*I flies (Bourtzis et al. 1998; Merçot and Poinsot 1998*a*). An alternate explanation for these intriguing data is not one of rescue but rather that the *Wolbachia* were isolates of the *w*Ma strain and *w*Ma exhibits high variance in inducing incompatibility (James and Ballard 2000), possibly a result of host effects or variation in the experimental protocol.

I will now consider each mtDNA haplogroup and their *Wolbachia* infections separately.

*D. simulans si*I mtDNA

Wolbachia data, complete mtDNA analyses, and network analyses suggest that *si*I migrated from the Seychelles to New Caledonia and then moved independently to Hawaii and Tahiti. *D. simulans si*I collected from the Seychelles and from New Caledonia may be doubly infected with *w*Ha and wMa_{No}, whereas flies from Hawaii and Tahiti are singly infected with *w*Ha. Consistent with a more recent infection, *w*Ha-induced incompatibility expression levels are higher in Hawaii and Tahiti than in New Caledonia (James and Ballard 2000; James et al. 2002). Also compatible with this hypothesis is the documentation of apparent *w*Ha and wMa_{No} double infections in *D. sechellia* (Charlat et al. 2002; Charlat, Bonnavion, and Merçot 2003).

Wolbachia may be mechanistically involved in

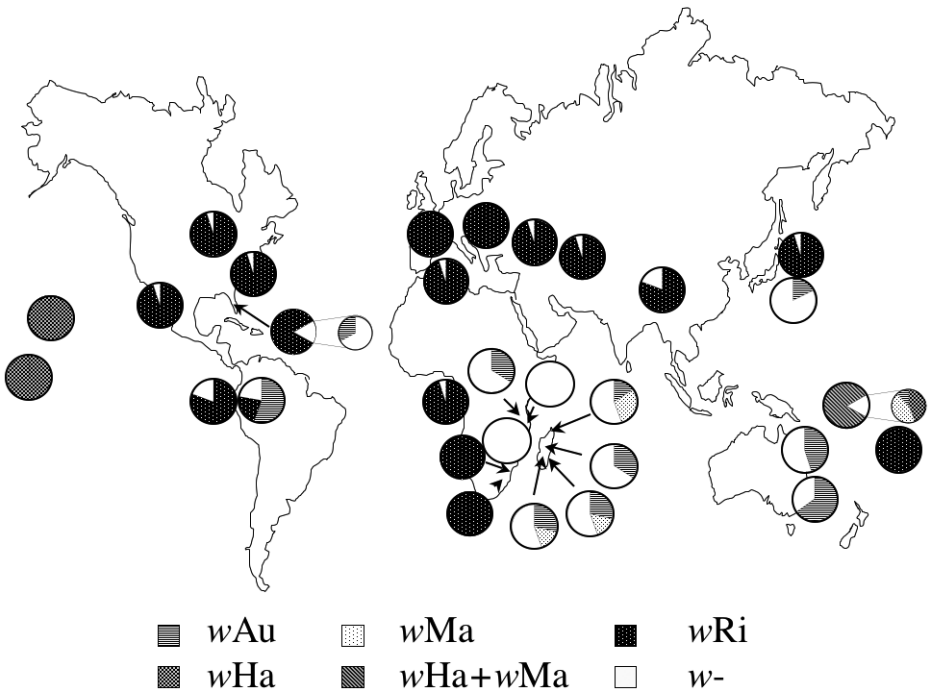

FIG. 2.—Frequencies of *Wolbachia*-infected *D. simulans* (when more than 20 individuals were collected from a single site during a single collection period).

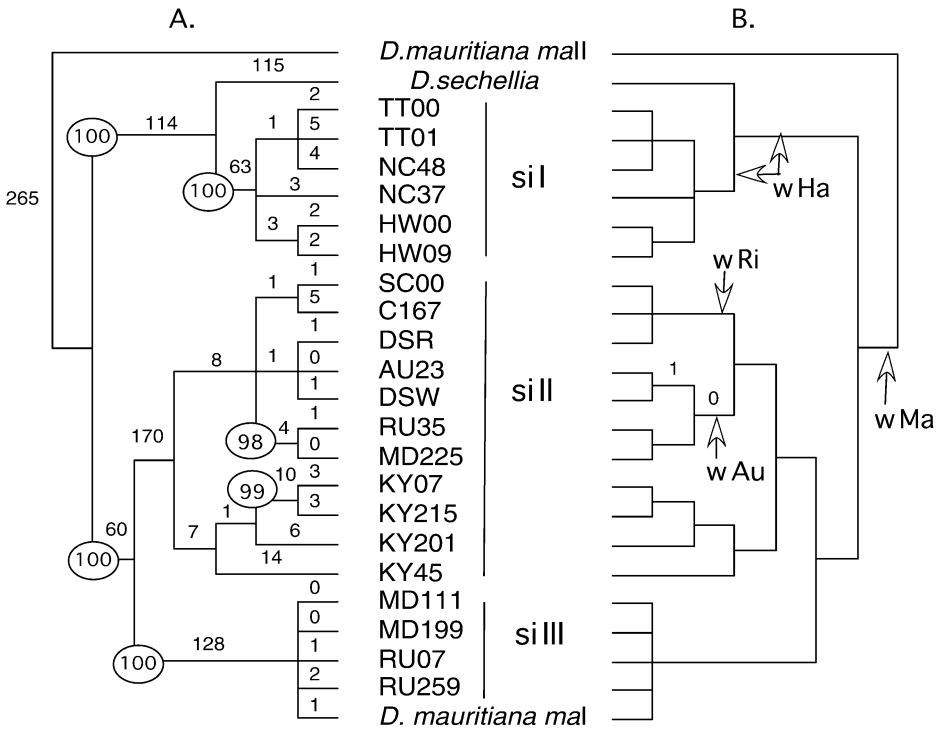

FIG. 3.—Genealogical relationships of complete *Drosophila mauritiana*, *D. sechellia*, and *D. simulans* mitochondrial DNA genomes. **A.** A Neighbor-Joining search estimated the transition/transversion ratio to be 6.318, the proportion of invariant sites to be 0.56, and the gamma shape to be 0.019. **A**. A heuristic HKY + I + Γ likelihood search with 10 random starting trees recovered one tree of length –ln L 24250.22. Bootstrap values above 95% are shown in circles. **B**. A heuristic search under the backbone constraint that each *Wolbachia* strain infected *D. simulans* once finds a single tree of length –ln L 24253.13. The potential sites of *Wolbachia* infection of *D. simulans* are indicated. In the case of *w*Ha, a double-headed arrow is indicated. *Drosophila sechellia* harbors two *Wolbachia* strains that appear to be *w*Ha and $w\text{Ma}_{\text{No}}$; however, *Wolbachia* from *D. sechellia* were not sequenced in this study so their identity is uncertain. The unconstrained and constrained topologies differ only in the placement of the DSR line. Parsimony branch lengths are shown above each line. Branch lengths are the same, unless otherwise noted.

maintaining interhaplogroup diversity and reducing intrahaplogroup variation in host mtDNA. The *w*Ha strain causes the strongest incompatibility phenotype and infects almost 100% of *si*I flies (O'Neill and Karr 1990; Turelli and Hoffmann 1995; Merçot et al., 1995; Merçot and Poinsot, 1998*b*; James and Ballard 2000; James et al. 2002). It may enhance global mitochondrial diversity by ''protecting'' what may be the less fit *si*I mtDNA haplogroup from extinction. In an elegant paper, de Stordeur (1997) conducted microinjection studies between eggs carrying the three mtDNA types and assayed the frequencies of the foreign injected mtDNA. He demonstrated that flies with *si*I mtDNA have lowest fitness following microinjection into a fly line harboring a different mtDNA type. James and Ballard (2003) found that *si*I flies had the shortest development time and the shortest longevity, and that males had the lowest activity.

*D. simulans siI*I mtDNA

The *si*II haplogroup is globally the most common. The basal lineages within this haplogroup have uninfected flies, implying a loss of the ancestral *w*Ma infection. In Tanzania and in Kenya, the mtDNA variation in these flies is consistent with a neutral equilibrium model of evolution (Dean et al. 2003). Within the *si*II haplogroup five distinct mtDNA haplotypes were associated with *w*Ri and four with *w*Au. The *w*Ri strain is possibly the most studied strain of *Wolbachia* in *D. simulans* and causes high levels of incompatibility (Hoffmann and Turelli 1988; Boyle et al. 1993; Turelli and Hoffmann 1995; Lassy and Karr 1996; James and Ballard 2000; Snook et al. 2000; Dean et al. 2003). The *w*Au strain causes no incompatibility in flies collected from Australia, Madagascar, and the Cameroon (Hoffmann, Clancy, and Merton 1994; James and Ballard 2000; Charlat, Le Chat and Merçot 2003) but intermediate incompatibility in flies from Florida (Ballard et al. 1996; James and Ballard 2000). This apparent conflict may be caused by variation in the *w*Au strain, the host genotype, or by the experimental design (Reynolds and Hoffmann 2002).

I propose that *si*II *w*- females migrated out of east Africa to Ecuador. In Ecuador, a female harboring the AU23 mitochondrial haplotype mutated to the DSR haplotype before being infected with *w*Ri. The *w*Ri-infected DSR haplotype then spread. Also in Ecuador, a female fly with the AU23 haplotype was infected with *w*Au and then spread worldwide. It is of particular interest that the common Malagasy MD225 mtDNA haplotype, which may be *w*Au infected, is derived from the Coffs genotype that was collected in high numbers from Australia and in low numbers from Ecuador. In Australia, 44 of 57 flies with the Coffs haplotype were *w*Au infected. These data suggest that Madagascar was reinvaded by derived *si*II *w*Au-infected *D. simulans*. Certainly, many

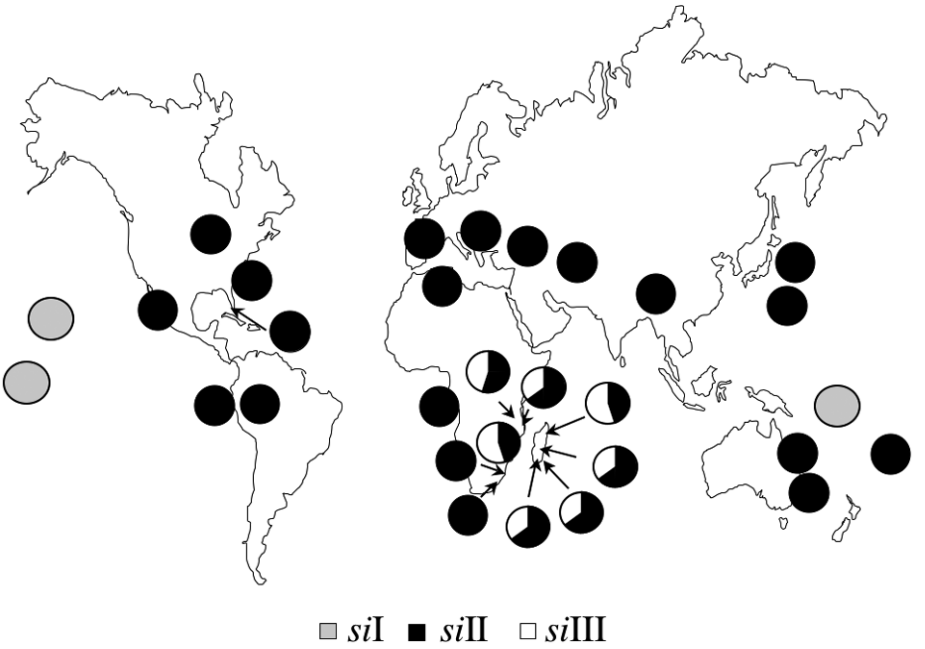

FIG. 4.—Frequencies of the mtDNA haplogroups of *D. simulans* (when more than 20 individuals were collected from a single site during a single collection period).

trees and shrubs have been transported from Australia to Madagascar to help combat deforestation. The uninfected DSW and AU117 haplotypes are also derived from AU23. DSW has been collected extensively in North America (Hoffmann and Turelli 1988; Hoffmann, Turelli, and Harshman 1990; Turelli and Hoffmann 1995) while AU117 is a singleton line from Australia.

Several alternate biogeographic hypotheses exist. First, it is possible that one or more of the mtDNA haplotypes migrated into, but did not arise in, Ecuador. The LA28, DSR, and AU23 haplotypes have been found in Florida, whereas the AU23, Coffs, and AU117 haplotypes have been found in Australia. Consequently, one or more infections may have occurred in Florida or Australia. Second, the DSR haplotype may have arisen from a *w*Au-infected AU23 haplotype fly. This alternative is less parsimonious and requires an additional step (two gains and a loss compared to two infection gains). Third, it is possible that AU23 arose by mutation from the DSR mtDNA type prior to DSR's infection with *w*Ri. This alternative is considered less likely as it implies that DSR arose from a hypothetical ancestor (fig. 6). It is possible, however, that the hypothetical haplotype or, indeed, the site of infection has been lost in a *w*Ri-induced or mtDNA-induced population genetic sweep.

The preferred *Wolbachia* infection hypothesis is more resolved than that of Ballard (2000*a*). This difference occurs because (1) five additional genomes within *si*II were included in this study, (2) constrained phylogenetic analyses did not reject the hypothesis that *Wolbachia* infection arose once in *D. simulans*, (3) a likelihood, and not parsimony, analysis of complete genome sequence was conducted, and (4) the hypothesis was tested with 1,701 bp of data from 383 *si*II lines.

*D. simulans si*III mtDNA

Within the *si*III haplogroup, there is a significant deficiency of mtDNA variation, and it is not possible to infer any biogeographic patterns. Indeed, only singleton SNP sites were detected in nine mtDNA genomes. The *si*III haplogroup is infected with *w*Ma, but it is not clear that *Wolbachia* in and of itself can cause this reduction in mtDNA diversity. Sequencing three genes has identified three distinct sequence isolates within *w*Ma, and it is possible that these isolates differentially express incompatibility. Laboratory incompatibility assays show that *w*Ma incompatibility is variable (Rousset and Solignac 1995;

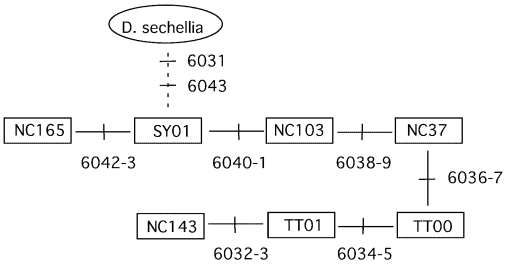

FIG. 5.—Intrahaplogroup network analysis of *D. simulans si*I using 584 bp from the intervening sequence between ND3 and the alanine tRNA. In this analysis, each AT repeat is coded as a single character. The *D. sechellia* root was then added (drawn with a dotted line to designate some uncertainty).

Table 3
Polymorphism in 1,701 bp of mtDNA in 383 Lines of *D. simulans si*II

Fly Line[c]	Position[a] **II**NSNSSSSNSSNSI**SSSNSSNNS**I 1111111133333337777888888 1445555623344678999000122 7585569201414193458778708 5041800683321678674586517	Number[b]	Distribution
DSR	TTGTGCCCCTCCTATTCGCGGCGA-	118	China, Ecuador, Gabon, Greece, Japan, Kenya, Malawi, North America, Seychelles, Tanzania, Ukraine, Zimbabwe
MD225	....A.......A.........AG.	115	Madagascar
Coffs	....A.......A.........G.	57	Ecuador, Australia (56)
AU23	........................G.	41	Australia, Cook Islands, Ecuador
KY201	......TT.......C...A...G.	10	Ethiopia, Kenya, India,
MD07	.......T...........A...G.	7	Madagascar
KY215	.......TT....T.....A.....	5	Kenya, Tanzania
LA28	....A..................G.	3	Ecuador, Lantana
KY07	.......TT.....C....A...G.	3	Kenya, Tanzania
KY09	.......T........T......G.	3	Kenya, Tanzania
KY45	.......T..............G.	3	Ecuador, Kenya, Madagascar
SA01	...G.....................	2	South Africa
UR25	.........A..............	2	Ukraine
DSW	.C.....................G.	1	North America
AU117	................A.....G.	1	Australia
KY17	.....A.T..............G.	1	Kenya
KY19	.......TT.....C....A...GT	1	Kenya
KY22	.......TT.........A...G.	1	Kenya
KY28	C.....TT.......C...A...G.	1	Kenya
KY216	......TT..T....C...A...G.	1	Kenya
KY259	.......T...........A.....	1	Kenya
MD85	.......T...T.............	1	Madagascar
MW30	................T......	1	Malawi
RU35	....A.......A.........AGT	1	Reunion Island
TZ09	..A....TT.....C....A...G.	1	Tanzania
TZ33	.......TT.........A...GT	1	Tanzania
TZ41	.......TT....T.....A.T...	1	Tanzania
UR8	..................A....	1	Ukraine

[a] Position in the mtDNA genome (relative to Ballard 2000a): I = silent, S = synonymous, and N = nonsynonymous substitutions. Boldface indicates minor strand coding regions. – indicates an indel event.
[b] Number of sequences for each type. This is not a random sample of *si*II sequences.
[c] Fly line representing each haplotype ranked by frequency of collection.

Merçot and Poinsot 1998*a*; James and Ballard 2000; Charlat, Le Chat and Merçot 2003). Little is known about the *w*Ma strain in nature, but it is possible that it confers a fitness advantage to infected flies. Dobson et al. (2002) examined a *Wolbachia* superinfection in the mosquito *Aedes albopictus* and found the infection to be associated with both cytoplasmic incompatibility and increased host fecundity. Relative to uninfected females, infected females lived longer, produced more eggs, and had higher hatching rates in compatible crosses. One obvious alternate possibility for the low *si*III variation is that an advantageous mutation (in the mtDNA or a nuclear gene interacting with a specific gene product) has caused flies with the observed haplotype to have an increased fitness. Fitness of flies could be tested directly in population cages or indirectly by monitoring the frequency of *si*III and *si*II flies where they occur in sympatry.

Conclusions

Populations of nearly all species exhibit at least some degree of differentiation among geographic locales (Ehrlich and Raven 1969). A continuing challenge is to describe population genetic architectures within species and to identify, and order, the evolutionary forces responsible for the observed subdivision. This study gives insight into the biogeography of *Wolbachia* infections and genetic subdivision in the mtDNA genome of *D. simulans*. These data say little about the organismal history of *D. simulans* because the mtDNA genome is a small piece of

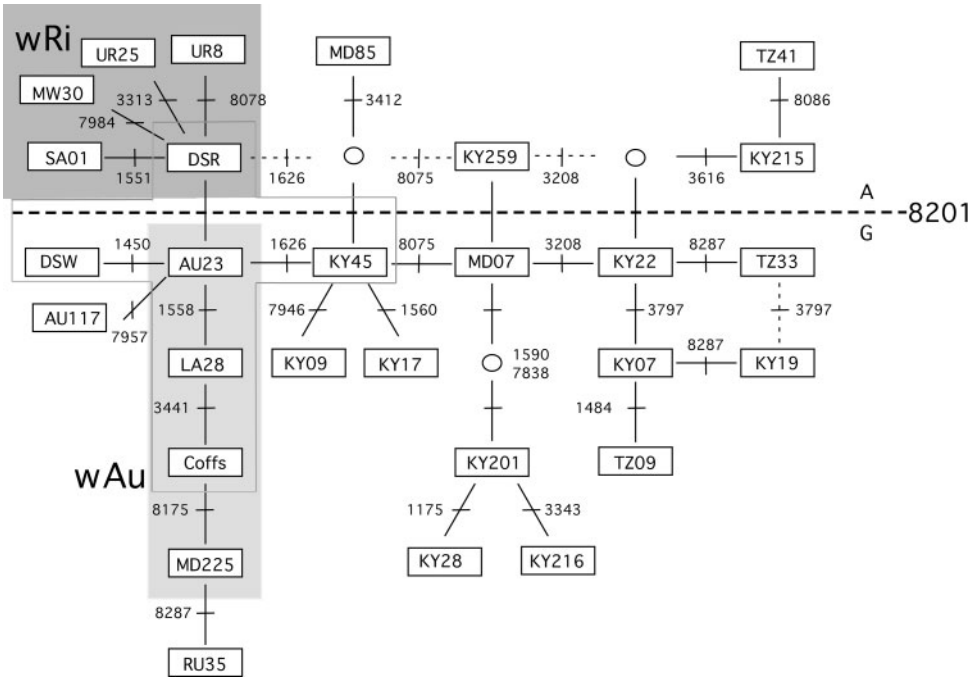

FIG. 6.—Intrahaplogroup network analysis of *D. simulans si*II. Analysis from 1,701 bp identified 28 distinct haplotypes and three missing genotypes. There were as many as four independent synonymous A → G changes at position 8201, and potentially multiple substitutions at positions 1626, 3208, 8075, and 3797. Also, it seems likely that 8287 has changed multiple times and that 3797 has changed once. Based on the frequency of the ancestral genotype, dotted lines are placed on the less likely mutation. Possibility of *Wolbachia* infection is overlain on the plot (fly lines may also be uninfected). Haplotypes inside the gray cross were found in Sangoqui Market in Ecuador.

the whole genome. Additional studies with multiple nuclear loci are required to give a more complete picture of the host's history.

This study is the accumulation of 14 international collectors from 33 countries and 64 sampling localities specifically investigating the biogeography of *Wolbachia* in *D. simulans*. In this study, I present new data and integrate information published from my laboratory (Ballard 2000*a*; James and Ballard 2000; Dean et al. 2003) to present a comprehensive summary. I do not include studies completed in other laboratories (Solignac and Monnerot 1986; Solignac, Monnerot, and Mounolou 1986; Baba-Aïssa et al. 1988; Montchamp-Moreau, Ferveur, and Jacques 1991; Rousset and Solignac 1995; Turelli and Hoffmann 1995; Merçot et al. 1995; Charlat, Le Chat, and Merçot (2003), but note that they have made substantial contributions to our understanding of this system.

Acknowledgments

Thanks to M. Dean, A. James, M. Whitlock, and two anonymous reviewers for constructively commenting on the manuscript. I acknowledge the following people and institutions for help with collecting: A. Berry (Madagascar and Reunion), C. Treese (Hawaii), ICIPIE (Madagascar), J. Featherston (Australia), K. Ballard (Australia, Crete, east Africa, Florida, Greece, Hawaii, New Caledonia, Tahiti), K. Maas (Kenya), P. Grace (Mexico), S. Charlat (Gabon), C. Meena and G. Mtoka (Tanzania), Greenwell and Kingsley Wallani (Malawi), V. Rafael, D. Vela, L. Lopez, and T. Moran (Ecuador), E. Zaranyika (Zimbabwe), and A. James and M. Dean (Crete and Greece). In addition, specific collection trips for this study were completed by A. James (China, Iowa, Talahassee), M. and M. Dean (Rarotonga), T. Nosenko (Ukraine), M. Zickell (Florida), and K. Kaneshiro and D. Baer (Hawaii). Additional flies were obtained from R. Kondo, M. Solignac, H. Merçot, A. Hoffmann, M. Habibula, F. Lemeunier, D. Lachaise, J. Bond, A. Olembo, J. Roote, R. Russell, F. Baba-Aïssa, O. Kitagawa, S. Chigusa, D. Presgraves, R. Bowie, A. Potts, M. Kreitman, and T. Mutangadura. A. James, M. Dean, A. Glass, P. Mena, Z. Preckwinkle helped with maintaining fly stocks and a variety of laboratory chores. The mitochondrial sequence data were generated in the Roy J. Carver Center for Comparative Genomics at the University of Iowa. Helen Kurtz gathered the sequence data for the *Wolbachia tsz* locus in the Pritzker Laboratory at the Field Museum. National Science Foundation grants DEB-9702824 and DEB-0296086 and the Field Museum Marshall Field Fund provided funds.

Literature Cited

Andolfatto, P. 2001. Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogaster* and *Drosophila simulans*. Mol. Biol. Evol. **18**:279–290.

Baba-Aïssa, F., M. Solignac, N. Dennebouy, and J. R. David. 1988. Mitochondrial DNA variability in *Drosophila simulans*: quasi absence of polymorphism within each of the three cytoplasmic races. Heredity **61**:419–426.

Ballard, J. W. O. 2000*a*. Comparative genomics of mitochondrial DNA in *Drosophila simulans*. J. Mol. Evol. **51**:64–75.

———. 2000*b*. Comparative genomics of mitochondrial DNA in members of the *Drosophila melanogaster* subgroup. J. Mol. Evol. **51**:48–63.

———. 2000c. When one is not enough: introgression of mitochondrial DNA in *Drosophila*. Mol. Biol. Evol. **17**:1126–1130.

Ballard, J. W. O., B. Chernoff, and A. C. James. 2002. Divergence of mitochondrial DNA is not corroborated by nuclear DNA, morphology, or behavior in *Drosophila simulans*. Evolution **56**:527–545.

Ballard, J. W. O., and M. D. Dean. 2001. The mitochondrial genome: mutation, selection and recombination. Curr. Opin. Genet. Dev. **11**:667–672.

Ballard, J. W. O., J. Hatzidakis, T. L. Karr, and M. Kreitman. 1996. Reduced variation in *Drosophila simulans* mitochondrial DNA. Genetics **144**:1519–1528.

Ballard, J. W. O., and M. Kreitman. 1994. Unraveling selection in the mitochondrial genome of *Drosophila*. Genetics **138**:757–772.

Bandi, C., T. J. C. Anderson, C. Genchi, and M. L. Blaxter. 1998. Phylogeny of *Wolbachia* in filarial nematodes. Proc. R. Soc. Lond. Ser. B **265**:2407–2413.

Barker, F. K., and F. M. Lutzoni. 2002. The utility of the incongruence length difference test. Syst. Biol. **51**:625–637.

Begun, D. J., and C. F. Aquadro. 1993. African and North American populations of *Drosophila melanogaster* are very different at the DNA level. Nature **365**:548–550.

Bourtzis, K., S. L. Dobson, H. R. Braig, and S. L. O'Neill. 1998. Rescuing *Wolbachia* have been overlooked. Nature **391**:852–853.

Boyle, L., S. L. O'Neill, H. M. Robertson, and T. L. Karr. 1993. Interspecific and intraspecific horizontal transfer of *Wolbachia* in *Drosophila*. Science **260**:1796–1799.

Breeuwer, J. A., and J. H. Werren. 1990. Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. Nature **346**:558–560.

———. 1993. Cytoplasmic incompatibility and bacterial density in *Nasonia vitripennis*. Genetics **135**:565–574.

Bull, J. J., J. P. Huelsenbeck, C. W. Cunningham, D. L. Swofford, and P. J. Waddell. 1993. Partitioning and combining data in phylogenetic analysis. Syst. Biol. **42**:384–397.

Callaini, G., M. G. Riparbelli, R. Giordano, and R. Dallai. 1996. Mitotic defects associated with cytoplasmic incompatibility in *Drosophila simulans*. J. Invert. Pathol. **67**:55–64.

Caspari, E., and G. S. Watson. 1959. On the evolutionary importance of cytoplasmic sterility in mosquitoes. Evolution **13**:568–570.

Charlat, S., P. Bonnavion, and H. Merçot. 2003. Wolbachia segregation dynamics and levels of cytoplasmic incompatibility in *Drosophila sechellia*. Heredity **90**:157–161.

Charlat, S., L. Le Chat, and H. Merçot. 2003. Characterization of non-cytoplasmic incompatibility inducing Wolbachia in two continental African populations of *Drosophila simulans*. Heredity **90**:49–55.

Charlat, S., A. Nirgianaki, K. Bourtzis, and H. Merçot. 2002. Evolution of *Wolbachia*-induced cytoplasmic incompatibility in *Drosophila simulans* and *D. sechellia*. Evolution **56**:1735–1742.

Clancy, D. J., and A. A. Hoffmann. 1998. Environmental effects on cytoplasmic incompatibility and bacterial load in *Wolbachia* infected *Drosophila simulans*. Entomol. Exp. Appl. **86**:13–24.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. **9**:1657–1659.

de Stordeur, E. 1997. Nonrandom partition of mitochondria in heteroplasmic *Drosophila*. Heredity **79**:615–623.

Dean, M. D., and J. W. O. Ballard. 2000. Factors affecting mitochondrial DNA quality from museum preserved *Drosophila simulans*. Entomol. Exp. Appl. **98**:279–280.

Dean, M. D., K. J. Ballard, A. Glass, and J. W. O. Ballard. 2003. Influence of two *Wolbachia* strains on population structure of east African *Drosophila simulans*. Genetics **165**:1959–1969.

Dobson, S. L., E. J. Marsland, Z. Veneti, K. Bourtzis, and S. L. O'Neill. 2002. Characterization of *Wolbachia* host cell range via the in vitro establishment of infections. Appl. Environ. Microbiol. **68**:656–660.

Eanes, W. F., M. Kirchner, J. Yoon, C. H. Biermann, I. N. Wang, M. A. McCartney, and B. C. Verrelli. 1996. Historical selection, amino acid polymorphism and lineage-specific divergence at the G6pd locus in *Drosophila melanogaster* and *D. simulans*. Genetics **144**:1027–1041.

Ehrlich, P. R., and P. H. Raven. 1969. Differentiation of populations. Science **165**:1228–1232.

Farris, J. S., M. Kallersjo, A. Kluge, and C. Bult. 1995. Testing significance of incongruence. Cladistics **10**:315–319.

Felsenstein, J., and H. Kishino. 1993. Is there something wrong with the bootstrap on phylogenies? A reply to Hillis and Bull. Syst. Biol. **42**:193–200.

Hamblin, M. T., and M. Veuille. 1999. Population structure among African and derived populations of *Drosophila simulans*: evidence for ancient subdivision and recent admixture. Genetics **153**:305–317.

Heath, B. D., R. D. Butcher, W. G. Whitfield, and S. F. Hubbard. 1999. Horizontal transfer of *Wolbachia* between phylogenetically distant insect species by a naturally occurring mechanism. Curr. Biol. **9**:313–316.

Hoffmann, A. A., D. Clancy, and J. Duncan. 1996. Naturally-occurring *Wolbachia* infection in *Drosophila simulans* that does not cause cytoplasmic incompatibility. Heredity **76**:1–8.

Hoffmann, A. A., D. J. Clancy, and E. Merton. 1994. Cytoplasmic incompatibility in Australian populations of *Drosophila melanogaster*. Genetics **136**:993–999.

Hoffmann, A. A., and M. Turelli. 1988. Unidirectional incompatibility in *Drosophila simulans*: inheritance, geographic variation and fitness effects. Genetics **119**:435–444.

———. 1997. Cytoplasmic incompatibility in insects. Pp. 42–80 *in* S. L. O'Neill, A. A. Hoffmann, and J. H. Werren, eds. Influential passengers: inherited microorganisms and invertebrate reproduction. Oxford University Press, Oxford, U.K.

Hoffmann, A. A., M. Turelli, and L. G. Harshman. 1990. Factors affecting the distribution of cytoplasmic incompatibility in *Drosophila simulans*. Genetics **126**:933–948.

Hoffmann, A. A., M. Turelli, and G. M. Simmons. 1986. Unidirectional incompatibility between populations of *Drosophila simulans*. Evolution **40**:692–701.

Hurst, G. D. D., L. D. Hurst, and M. E. N. Majerus. 1993. Altering sex ratios: the games microbes play. BioEssays **15**:695–697.

James, A. C., and J. W. O. Ballard. 2000. Expression of cytoplasmic incompatibility in *Drosophila simulans* and its impact on infection frequencies and distribution of *Wolbachia pipientis*. Evolution **54**:1661–1672.

———. 2003. Mitochondrial genotype affects fitness in *Drosophila simulans*. Genetics **164**:187–194.

James, A. C., M. D. Dean, M. E. McMahon, and J. W. O. Ballard. 2002. Dynamics of double and single *Wolbachia* infections in *Drosophila simulans* from New Caledonia. Heredity **88**:182–189.

Jeyaprakash, A., and M. A. Hoy. 2000. Long PCR improves *Wolbachia* DNA amplification: wsp sequences found in 76% of 63 arthropod species. Insect Mol. Biol. **9**:393–405.

Jiggins, F. M., G. D. D. Hurst, and M. E. N. Majerus. 2000. Sex-ratio-distorting *Wolbachia* causes sex-role reversal in its butterfly host. Proc. R. Soc. Lond. Ser. B **267**:69–73.

Kambhampati, S., K. S. Rai, and D. M. Verleye. 1992. Frequencies of mitochondrial DNA haplotypes in laboratory cage populations of the mosquito, *Aedes albopictus*. Genetics **132**:205–209.

Kliman, R. M., P. Andolfatto, J. A. Coyne, F. Depaulis, M. Kreitman, A. J. Berry, J. McCarter, J. Wakeley, and J. Hey. 2000. The population genetics of the origin and divergence of the *Drosophila simulans* complex species. Genetics **156**: 1913–1931.

Lachaise, D., C. ML, D. JR, L. F, T. L, and A. M.(??) 1988. Historical biography of the *Drosophila melanogaster* species subgroup. Evol. Biol. **22**:159–225.

Lassy, C. W., and T. L. Karr. 1996. Cytological analysis of fertilization and early embryonic development in incompatible crosses of *Drosophila simulans*. Mech. Dev. **57**:47–58.

Lincoln, R., G. Boxshall, and P. Clark. 1998. A dictionary of ecology, evolution and systematics. Cambridge University Press, Cambridge, U.K.

Martin, G., P. Juchault, and J. J. Legrand. 1973. Mise en évidence d'un micro-organism intracytoplasmique symbiote de l'Oniscoïde *Armadillidium vulgare* L., dont la présence accompagne l'intersexualité ou la féminisation totale des mâles génétiques de la lignée thélygene. C. R. Acad. Sci. Paris III **276**:2313–2316.

Merçot, H., B. Llorente, M. Jacques, A. Atlan, and C. Montchamp-Moreau. 1995. Variability within the Seychelles cytoplasmic incompatibility system in *Drosophila simulans*. Genetics **141**:1015–1023.

Merçot, H., and D. Poinsot. 1998*a*. . . .and discovered on Mount Kilimanjaro. Nature **391**:853.

Merçot, H., and D. Poinsot. 1998*b*. *Wolbachia* transmission in a naturally bi-infected *Drosophila simulans* strain from New Caledonia. Entomol. Exp. Applicata **86**:97–103.

Montchamp-Moreau, C., J. F. Ferveur, and M. Jacques. 1991. Geographic distribution and inheritance of three cytoplasmic incompatibility types in *Drosophila simulans*. Genetics **129**:399–407.

Nigro, L., and T. Prout. 1990. Is there selection on RFLP differences in mitochondrial DNA? Genetics **125**:551–555.

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. Proc. Natl. Acad. Sci. USA **89**:2699–2702.

O'Neill, S. L., and T. L. Karr. 1990. Bidirectional incompatibility between conspecific populations of *Drosophila simulans*. Nature **348**:178–180.

Posada, D., and K. A. Crandall. 2001. Intraspecific gene genealogies: trees grafting into networks. Trends Ecol. Evol. **16**:37–45.

Rand, D. M., M. Dorfsman, and L. M. Kann. 1994. Neutral and non-neutral evolution of *Drosophila* mitochondrial DNA. Genetics **138**:741–756.

Reynolds, K. T., and A. A. Hoffmann. 2002. Male age, host effects and the weak expression or non-expression of cytoplasmic incompatibility in *Drosophila* strains infected by maternally transmitted *Wolbachia*. Genet. Res. **80**:79–87.

Rousset, F., and M. Solignac. 1995. Evolution of single and double *Wolbachia* symbioses during speciation in the *Drosophila simulans* complex. Proc. Natl. Acad. Sci. USA **92**:6389–6393.

Rousset, F., D. Vautrin, and M. Solignac. 1992. Molecular identification of *Wolbachia*, the agent of cytoplasmic incompatibility in *Drosophila simulans*, and variability in relation with host mitochondrial types. Proc. R. Soc. Lond. Ser. B **247**:163–168.

Shimodaira, H., and M. Hasegawa. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics **17**:1246–1247.

Snook, R. R., S. Y. Cleland, M. F. Wolfner, and T. L. Karr. 2000. Offsetting effects of *Wolbachia* infection and heat shock on sperm production in *Drosophila simulans*: analyses of fecundity, fertility and accessory gland proteins. Genetics **155**:167–178.

Solignac, M., and M. Monnerot. 1986. Race formation and introgression within *Drosophila simulans*, *D. mauritiana* and *D. sechellia* inferred from mitochondrial DNA analysis. Evolution **40**:531–539.

Solignac, M., M. Monnerot, and J. C. Mounolou. 1986. Mitochondrial DNA evolution in the melanogaster species subgroup of *Drosophila*. J. Mol. Evol. **23**:31–40.

Steinbachs, J. C., N. V. Schizas, and J. W. O. Ballard. 2001. Efficiency of different genes and accuracy of different methods in recovering a known *Drosophila* genealogy. Pp. 606–617 *in* R. B. Altman, A. K. Dunker, L. A. Hunter, K. Lauderdale, and T. E. Klein, eds. Proceedings of the Pacific Symposium on Biocomputing. World Scientific, Singapore.

Stouthamer, R., and D. J. Kazmer. 1994. Cytogenetics of microbe-associated parthenogenesis and its consequence for gene flow in *Trichogramma* wasps. Heredity **73**:317–327.

Swofford, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, Mass.

Swofford, D. L., G. J. Olsen, P. J. Wadell, and D. M. Hillis. 1996. Phylogenetic Inference. Pp. 407–514 *in* D. M. Hillis, C. Moritz, and K. Mable, eds. Molecular systematics. Sinauer Associates, Sunderland, Mass.

Templeton, A. R., K. A. Crandall, and C. F. Sing. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics **132**:619–633.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24**:4876–4882.

Tram, U., and W. Sullivan. 2002. Role of delayed nuclear envelope breakdown and mitosis in *Wolbachia*-induced cytoplasmic incompatibility. Science **296**:1124–1126.

Turelli, M., and A. A. Hoffmann. 1991. Rapid spread of an inherited incompatibility factor in California *Drosophila*. Nature **353**:440–442.

———. 1995. Cytoplasmic incompatibility in *Drosophila simulans*: dynamics and parameter estimates from natural populations. Genetics **140**:1319–1338.

Werren, J. H., W. Zhang, and L. R. Guo. 1995. Evolution and phylogeny of *Wolbachia*: reproductive parasites of arthropods. Proc. R. Soc. Lond. Ser. B **261**:55–63.

Yen, J. H., and A. R. Barr. 1973. The etiological agent of cytoplasmic incompatibility in *Culex pipiens*. J. Invert. Pathol. **22**:242–250.

Zhou, W., F. Rousset, and S. L. O'Neill. 1998. Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. Proc. R. Soc. Lond. Ser. B **265**:509–515.

Pierre Capy, Associate Editor

Accepted September 29, 2003