PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2006) **273**, 1603–1611
doi:10.1098/rspb.2006.3493
Published online 3 April 2006

Evolutionary history of a mosquito endosymbiont revealed through mitochondrial hitchhiking

Jason L. Rasgon[1,2,*], Anthony J. Cornel[3] and Thomas W. Scott[4]

[1]*The W. Harry Feinstone Department of Molecular Microbiology and Immunology, Bloomberg School of Public Health, Johns Hopkins University, and* [2]*Johns Hopkins Malaria Research Institute, Baltimore, MD 21205, USA,*
[3]*Department of Entomology, University of California at Davis, Mosquito Control Research Laboratory, Parlier, CA 93648, USA*
[4]*Department of Entomology, University of California at Davis, Davis, CA 95616, USA*

Due to cytoplasmic inheritance, spread of maternally inherited *Wolbachia* symbionts can result in reduction of mitochondrial variation in populations. We examined sequence diversity of the mitochondrial NADH dehydrogenase subunit 4 (*ND4*) gene in *Wolbachia*-infected (South Africa (SA), California and Thailand) and uninfected (SA) *Culex pipiens* complex populations. In total, we identified 12 haplotypes (A–L). In infected populations, 99% of individuals had haplotype K. In the uninfected SA population, 11 haplotypes were present, including K. Nuclear allozyme diversity was similar between infected and uninfected SA populations. Analysis of nuclear DNA sequences suggested that haplotype K presence in uninfected SA *Cx. pipiens* was probably due to a shared ancestral polymorphism rather than hybrid introgression. These data indicate that *Wolbachia* spread has resulted in drastic reduction of mitochondrial variability in widely separated *Cx. pipiens* complex populations. In contrast, the uninfected SA population is probably a cryptic species where *Wolbachia* introgression has been prevented by reproductive isolation, maintaining ancestral levels of mitochondrial diversity. Molecular clock analyses suggest that the *Wolbachia* sweep occurred within the last 47 000 years. The effect of *Wolbachia* on mitochondrial dynamics can provide insight on the potential for *Wolbachia* to spread transgenes into mosquito populations to control vector-borne diseases.

Keywords: *Culex pipiens*; species complex; *Wolbachia*; mitochondria; genetic hitchhiking

1. INTRODUCTION

Understanding the structure of mosquito populations is critical for addressing public health issues such as evolution and spread of insecticide resistance alleles, epidemiology of mosquito-borne pathogens and developing and testing transgenic strategies for control of vector-borne diseases (Tripet *et al.* 2001; Bennett *et al.* 2002; Fanello *et al.* 2003). Mitochondrial DNA sequences have been shown to be useful markers for studying the structure of medically important insect populations such as mosquitoes, tsetse flies (*Glossina morsitans*) and sand flies (*Phlebotomus papatasi*; Besansky *et al.* 1997; Gorrochotegui-Escalante *et al.* 2000; Krafsur *et al.* 2000; Donnell *et al.* 2001; Krafsur *et al.* 2001; Gorrochotegui-Escalante *et al.* 2002; Parvizi *et al.* 2003; Marquez *et al.* 2004). Patterns of mitochondrial variability can be confounded, however, by the spread of maternally inherited symbionts such as *Wolbachia* (Turelli & Hoffmann 1999; Hurst & Jiggins 2005). *Wolbachia* is associated with reproductive alterations such as cytoplasmic incompatibility (CI); i.e. reduced egg hatch when uninfected females mate with infected males. Matings between infected females and infected or uninfected males are fertile. Consequently, infected females have a reproductive advantage in a mixed population, allowing infection to spread (Turelli & Hoffmann 1999). As *Wolbachia* spreads, there can be a linked sweep of the associated mitochondria, resulting in reduction of mitochondrial diversity in the infected populations (Turelli *et al.* 1992, 1999; Hurst & Jiggins 2005). Selective mitochondrial sweeps have been observed in concert with *Wolbachia* invasions in a number of species (Turelli *et al.* 1992; Ballard *et al.* 1996; Baurdy *et al.* 2003; Jiggins 2003; Parvizi *et al.* 2003).

Among mosquitoes, the population biology of *Wolbachia* has been most thoroughly studied in the *Culex pipiens* species complex (Cornel *et al.* 2003; Rasgon & Scott 2003; 2004). The two most common and widespread members of the complex are the subspecies *Culex pipiens pipiens* (L.) and *Culex pipiens quinquefasciatus* Say (Barr 1982). Both subspecies have a global distribution and complicated population structure. For example, in North America and Asia extensive gene flow occurs between subspecies (Tabachnick & Powell 1983; Urbanelli *et al.* 1997; Fonseca *et al.* 2004), whereas in some parts of Africa gene flow between subspecies is partial (Urbanelli *et al.* 1985, 1995) or restricted (Jupp 1978; Cornel *et al.* 2003). Depending on location, members of this complex can be important vectors of nematodes that cause lymphatic filariasis and arboviruses such as St Louis encephalitis, West Nile and Rift Valley fever viruses (Hoogstraal *et al.* 1979; Krida *et al.* 1998; Day 2001; Nasci *et al.* 2001; Fonseca *et al.* 2004).

Mosquitoes in *Cx. pipiens* complex are commonly infected with *Wolbachia* (Rasgon & Scott 2003, 2004), and have been previously shown to exhibit reduced levels

*Author and address for correspondence: Johns Hopkins Malaria Research Institute, Bloomberg School of Public Health, Johns Hopkins University, 615 North Wolfe Street Room E4626 Baltimore, MD 21205, USA (jrasgon@jhsph.edu).

Received 22 December 2005
Accepted 23 January 2006

of mitochondrial variability among colonized strains (Guillemaud *et al.* 1997) and in naturally bottlenecked populations (Fonseca *et al.* 2000). With few exceptions all populations are believed to be infected with *Wolbachia* (Hoffmann & Turelli 1997). The published records of uninfected populations occur in *Culex pipiens australicus* Dobrotwortsky and Drummond from Australia and *Cx. p. pipiens* from Rhodesia (Zimbabwe; Irving-Bell 1974). We recently described a *Wolbachia*-uninfected *Cx. p. pipiens* population in South Africa (SA) near Johannesburg that was reproductively isolated from sympatric *Cx. p. quinquefasciatus* infected populations (Cornel *et al.* 2003). Reproductive isolation between these two populations was inferred through morphological assessment of male genitalia DV/D ratios (Sundararaman 1949), fixed allozyme differences and *Wolbachia* infection status. If reproductive isolation of this uninfected population from other infected populations predates the sweep of *Wolbachia* through *Cx. pipiens sensu latu*, then it can be used as a baseline to directly test the influence of *Wolbachia* spread on mitochondrial variability in the *Cx. pipiens* species complex and perhaps to predict the ramification of this process in insects in general. To examine this issue, we (i) compared mitochondrial sequence diversity from the *Wolbachia*-uninfected SA population to infected populations from SA, California and Thailand, (ii) assessed *Wolbachia* frequency and *Wolbachia* surface protein (*wsp*) gene diversity in each infected population and (iii) assessed levels of nuclear diversity in infected and uninfected SA populations by sequencing nuclear DNA and by allozymes. The data suggest that *Wolbachia* has recently swept through *Cx. pipiens* complex populations and has dramatically affected patterns of mitochondrial variability in this mosquito. The observed patterns of mitochondrial variation in infected and uninfected populations provide important insights for understanding drive mechanisms for transgenic mosquito disease prevention strategies.

2. MATERIAL AND METHODS

(a) *Mosquito collections and identification*

(i) *South Africa*

Wild gravid and recently blood-fed females were collected resting inside geese and chicken coops on the outskirt of Johannesburg (26°06′ S 27°50′ E) in March 2000. Females were allowed to oviposit and each egg raft was reared separately, allowing results for mitochondrial and nuclear sequence variation and *Wolbachia*-infection status to be traced to individual isofemale lines. For *a priori* identification of each family as either *Cx. p. pipiens* or *Cx. p. quinquefasciatus* the male genitalia from four males (older than 24 h) were dissected and slide mounted for DV/D ratio measurements (the distance between the dorsal and ventral arms of the male phallosome divided by the distance between the two dorsal arms) to determine what proportion of the sample were represented by *Cx. p. pipiens*, *Cx. p. quinquefasciatus* and hybrids between the two (Sundararaman 1949). No hybrids were detected. Study populations are denoted as: SAP, South Africa *Cx. p. pipiens*; SAQ, South Africa *Cx. p. quinquefasciatus*.

(ii) *California and Thailand*

Mosquitoes were collected as larvae, reared to adults in the laboratory, killed by freezing and stored at −80 °C or in 95% ethanol until processed for DNA extraction. The male genitalia from at least 40 males from each location were dissected and slide mounted for DV/D ratio measurements to determine what proportion of each sample was represented by *Cx. p. pipiens*, *Cx. p. quinquefasciatus* and hybrids. No hybrids were detected. Study populations are denoted as: CAP, California *Cx. p. pipiens* (Shasta Co.); CAQ, California *Cx. p. quinquefasciatus* (Riverside Co.); THQ, Thailand *Cx. p. quinquefasciatus* (Mesot, Thailand).

(b) *DNA extraction*

Ethanol stored specimens were re-hydrated in phosphate-buffered saline before DNA extraction. DNA from individual mosquitoes was extracted by salt extraction/ethanol precipitation as previously described (Rasgon & Scott 2003), re-constituted in deionized water and stored at −20 °C until used for PCR.

(c) Wolbachia *infection*

(i) *Polymerase chain reaction*

PCR was conducted using primers 99F and 994R (O'Neill *et al.* 1992), which amplify an approximately 900 bp fragment from *Wolbachia* 16S rDNA and are designed to be specific to *Wolbachia* of all strains. PCR conditions were as stated by Rasgon & Scott (2003). Known infected (LIN) and uninfected (LINT) colony mosquitoes (Rasgon & Scott 2003) were included in every reaction as positive and negative controls, respectively. Template DNA quality was assessed by successful amplification of a 400 bp fragment from insect 12S mtDNA using primers 12SA1 and 12SB1 (Simon *et al.* 1991) as previously described (Rasgon & Scott 2003). Amplified fragments were separated by agarose gel electrophoresis, stained with ethidium bromide (1 μl ml^{-1}) and visualized with ultraviolet light.

(ii) Wolbachia *surface protein gene sequencing*

From each infected population, 3–5 infected individuals were randomly chosen and subjected to PCR amplification of the *Wolbachia* Surface Protein (*wsp*) gene using primers 81F and 691R as previously described (Zhou *et al.* 1998). Amplified fragments were separated by agarose gel electrophoresis, purified from the gel using Qiaquick columns (QIAGEN, Valencia, CA) and directly sequenced in both directions using an ABI Prism 377 DNA sequencer with Big Dye chemistry (Perkin-Elmer Applied Biosystems, Foster City, CA). SEQUENCHER, DNA Sequence Analysis Software v. 4.0.5 (Gene Codes Corporation, Ann Arbor, MI) was used to align sequences.

(d) *Mosquito mitochondrial variability*

(i) *PCR and sequencing*

The primers ND4+ and ND4− were used to amplify a 389 bp fragment from the NADH dehydrogenase subunit 4 (*ND4*) gene from all specimens examined. Sequences were deposited in the GenBank database under accession numbers AY793688–AY793703. *ND4* sequences from *Culex restuans* (Baltimore Co., MD) and *Culex tarsalis* (Kern Co., CA) were obtained as out group taxa and deposited in GenBank under accession numbers AY788866–AY788867. Primer sequences and PCR conditions were as stated by

Gorrochotegui-Escalante *et al.* (2000). Amplified fragments were separated, purified and sequenced as described above.

(ii) *Phylogenetic analysis of* ND4 *haplotypes*

After removing the primer sequences, sequences were aligned with manual correction using CLUSTAL X (Thompson *et al.* 1997). Haplotype networks were constructed using statistical parsimony criteria with a 95% cut-off (Templeton *et al.* 1992) using TCS v. 1.13 (Clement *et al.* 2000). The statistical parsimony algorithm calculates a cut-off number (parsimony limit) of mutational steps below which the haplotypes can be connected with 95% confidence. We also conducted maximum-likelihood phylogenetic analysis using PAUP* v. 4.01b 10 (Swofford 1998). For analyses, the GTR+I+G model was selected as the most appropriate evolutionary model of DNA substitution using MODELTEST v. 3.06 (Posada & Crandall 1998). Tree robustness was evaluated by bootstrapping (500 replicates).

(iii) *Statistical analysis of mitochondrial variability*

The null hypothesis that all mutations were neutral was statistically ascertained by four tests. DnaSP (Rozas & Rozas 1999) was used to implement both Tajima's *D*-test (Tajima 1989) and Fu & Li's *D*-test; ALLELIX software (obtained from S. Mousset) was used to implement both Depaulis' *H*-and *K*-tests (Depaulis & Veuille 1998), where *p* values were calculated using 10 000 independent replicates. Partitioning of mitochondrial haplotype variation was analysed by Analysis of Molecular Variance (AMOVA; Excoffier *et al.* 1992) using ARLEQUIN software v. 2.0 (http://lgb.unige.ch/arlequin), which uses a non-parametric permutation test to calculate the significance of variance components associated with each level of genetic structure. Significance levels for population differentiation pairwise comparisons were computed using a Markov chain exact-test procedure and Bonferroni correction for multiple tests. For each population, DnaSP was used to estimate haplotype diversity (*Hd*), the number of polymorphic sites, the average number of nucleotide differences (*k*) and the nucleotide diversity with a Jukes–Cantor correction (π_2).

(e) *Mosquito nuclear variability*

(i) *Nuclear gene phylogenetics*

We conducted sequence analysis of *Cx. pipiens* spp. using internal transcribed spacer (ITS) and ribosomal DNA sequences because there are a large number of complementary sequences available for comparison from a wide geographical sample in the GenBank database (Miller *et al.* 1996). The entire ITS 2 region, the entire 5.8S gene and a portion of the ITS1 region were amplified from five SA *pipiens* and five SA *quinquefasciatus* specimens using primers PQ10 and CP16. Primer sequences and PCR conditions were as stated by Miller *et al.* (1996). Specimens were randomly chosen with the exception of specimen SAP3.54a, which was selected due to its unique haplotype/*Wolbachia* infection status (*Wolbachia*-negative, haplotype K—see §3). PCR products were purified using the MinElute clean-up kit (Qiagen, Valencia, VA), cloned into the pCR 4-TOPO vector and transformed into competent *Escherichia coli* cells (Invitrogen, Carlsbad, CA). One clone per individual was sequenced using the M13F primer. Sequences were deposited in the GenBank database under accession numbers DQ341106–DQ341115. Sequences were aligned with sequences from other geographical isolates (Miller *et al.* 1996, fig. 3) using BIOEDIT, for a total alignment of 748 bp including indels. The HKY+G model was selected as the most appropriate evolutionary model of DNA substitution using MODELTEST v. 3.06 (Posada & Crandall 1998). These settings were used with a neighbour-joining algorithm to conduct phylogenetic analyses using PAUP* v. 4.01b 10 (Swofford 1998). Tree robustness was evaluated by bootstrapping (1000 replicates).

(ii) *Estimates of nuclear variation*

Allozyme analysis from previously published data (Cornel *et al.* 2003) was undertaken to estimate differences in nuclear diversity between infected and uninfected SA *Cx. pipiens* complex populations. We used GENEPOP v. 1.2 (Raymond & Rousset 1995) to calculate (i) average heterozygosity and proportion of polymorphic loci in each population and (ii) probability of differentiation between SA populations by a Markov-chain exact test.

3. RESULTS

In total, we identified 12 mitochondrial haplotypes (A–L) in the *Cx. pipiens s.l.* member populations sampled, consisting of 27 variable sites (table 1) Phylogenetic analysis of mitochondrial sequences identified two main clades; A–J and K–L. These two clades form two unconnected networks as calculated by statistical parsimony (95% cut-off: seven steps; figure 1), but group together with strong bootstrap support (99%) by phylogenetic analysis (figure 2).

All individuals tested from populations CAP, CAQ, THQ and SAQ were found to be infected with *Wolbachia*. There were no differences in *wsp* sequences within or between populations; sequences were identical to those obtained from *Wolbachia* infecting North American and Asian *Cx. pipiens s.l.* (e.g. AF301010, AF216859, AF216860). In addition, recently colonized individuals from populations SAQ and CAQ were reproductively compatible when crossed in either direction (79% versus 93%, ANOVA, NS). Populations SAQ ($n=50$), THQ ($n=10$) and CAP ($n=10$) were fixed for mitochondrial haplotype K, which was the majority haplotype in population CAQ as well. One individual tested in population CAQ ($n=10$) possessed haplotype L, which differs from haplotype K by a single nucleotide substitution.

In contrast, none of the individuals assayed from population SAP contained detectable *Wolbachia* infections. In SAP mosquitoes, 11 haplotypes (A–K) were identified from the 21 families we examined. Population SAP exhibited an over 35 fold increase in *Hd* and an over 170 fold increase in nucleotide diversity (with a conservative Jukes–Cantor correction for multiple hits) compared to the combined infected populations (table 2). Neutrality of mitochondrial variation in population SAP was assessed by four statistical tests. Some tests indicated deviation from neutrality due to the large genetic distance of haplotype K from the other haplotypes present in population SAP (Tajima's $D=-1.824$, $p<0.05$; Fu and Li's $D=-2.796$, $p<0.05$). Depaulis' *H*- and *K*-tests did not suggest a deviation from neutrality and indicated that the number of haplotypes present and the *Hd* in population SAP did not differ from what was expected by chance (observed number of haplotypes = 12,

Table 1. Alignment of 27 variable sites from *Cx. pipiens* spp. *ND4* sequences.

haplotype	site	34	46	64	67	73	82	88	97	106	109	123	142	143	151	157	179	181	190	208	223	232	247	251	262	265	301	322
A		G	A	G	T	T	A	T	G	G	A	A	T	T	A	T	C	T	C	A	T	G	T	T	A	C	C	T
B		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
C		A	.	.	.	.	.	.	A	A	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.
D		.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
E		.	.	.	.	.	.	.	A	.	G	.	.	C	G	.	.	.	.	.	.	.	.	.	.	.	T	.
F		.	.	.	.	.	.	.	A	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	C	.	.	.
G		.	.	.	.	.	.	.	A	.	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
H		.	.	.	.	.	.	.	A	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
I		.	.	.	.	.	.	.	A	.	.	.	C	C	.	.	.	.	.	.	.	.	.	.	.	.	T	.
J		.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	T	.
K		A	T	A	A	A	G	C	A	A	T	T	.	.	G	A	.	A	T	T	.	A	G	.	T	T	T	G
L		A	T	A	A	A	G	C	A	A	T	T	.	.	G	A	T	A	T	T	.	A	G	.	T	T	T	G

calculated 95% confidence intervals: 6–15, $p=0.65$; observed $Hd=0.895$, calculated 95% confidence intervals: 0.694–0.916, $p=0.85$).

Analysis of mt haplotype variation (AMOVA) was performed (i) among populations with variance partitioned within geographical regions (South Africa, SAP and SAQ; California, CAP and CAQ; Asia, Thailand) and (ii) among populations with variance partitioned within *Wolbachia* infection status. In the first AMOVA, there was a large amount of negative among-region variance, indicating a lack of among-region population structure with individuals more related between rather than within geographical regions; a result consistent with a selective mt sweep. The majority of the significant variation indicative of population structure was found among populations within geographical regions, highlighting the substantial difference in *Hd* between populations SAP and SAQ. When haplotype variation was partitioned between *Wolbachia* infection status, the greatest amount of variation (more than 95%) was due to variation between infection types, suggesting differences in *Hd* that are correlated with presence or absence of *Wolbachia* (table 3). Markov-chain exact pairwise comparisons between all populations indicated that population SAP had significantly greater mitochondrial variation compared to every infected population ($p<0.0001$), but the infected populations did not differ significantly among themselves ($p>0.17$).

These data indicate a significant reduction in mitochondrial variability in infected populations compared to the uninfected population SAP. An alternative explanation to a *Wolbachia* sweep for this result is that the infected populations experienced a bottleneck prior to a population range expansion. If this is true, we would expect that the infected populations would exhibit reduction in the diversity of nuclear loci as well as mitochondrial loci. Although there are fixed allozyme differences indicating that the two SA populations are highly differentiated from one another (Markov-chain exact test, $p<0.0001$), levels of diversity as calculated from 13 allozyme loci are remarkably similar between both populations (SAP: average heterozygosity=0.236, proportion polymorphic loci=0.85; SAQ: average heterozygosity=0.214, proportion polymorphic loci=0.85). Our data are inconsistent with the hypothesis that reduced mt variation is attributable to a bottleneck followed by population expansion.

Phylogenetic analysis of mosquito ribosomal and ITS DNA sequences indicates that SAP and SAQ individuals do not cluster together. SAP individuals were in the same clade and clustered with other *pipiens* isolates from around the world (e.g. Virginia, Colorado and Sweden). This was true even for individual SAP 3.54a, which had *ND4* haplotype K. In contrast, SAQ individuals did not cluster and were more scattered on the tree, but in no case did they group closely with SAP individuals (figure 3). We suggest that the most parsimonious explanation for the presence of haplotype K in population SAP is shared ancestral polymorphism rather than hybrid introgression and that it is likely that population SAP predates the sweep of *Wolbachia* through *Cx. pipiens s.l.*

Mitochondria sequence data can be used to calculate an approximate time-interval of the *Wolbachia* sweep through *Cx. pipiens* populations. Clade K–L contains no

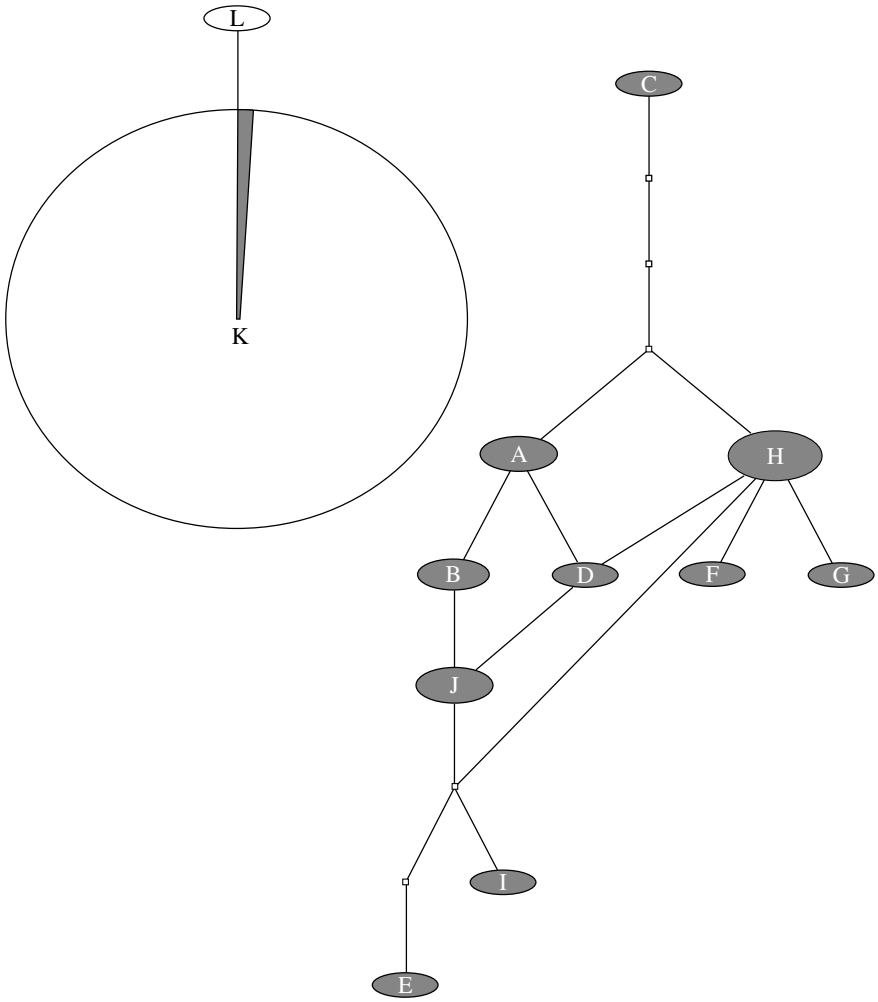

Figure 1. Statistical parsimony haplotype network of *Culex pipiens s.l. ND4* sequences. Unlabelled nodes represent inferred haplotypes. Oval size is proportional to haplotype frequency. Shading represents lack of *Wolbachia* infections (population SAP), no shading represents infection with *Wolbachia* (all other populations). 95% cut-off = 7 mutational steps.

synonymous substitutions. The genera *Culex* and *Aedes* diverged approximately 38 000 000 years ago (Besansky & Fahey 1997). We calculated the nucleotide divergence between haplotype K and *Aedes aegypti* (GenBank accession number AF334848) and estimated the mutation rate for this gene (substitutions/site/year) at twofold (2.74×10^{-9}) and fourfold (1.19×10^{-8}) degenerate sites with a conservative Jukes–Cantor correction for multiple hits. Assuming a molecular clock (Likelihood ratio test, $p=0.21$), we calculated the maximum time-interval in which we would expect to observe at least one synonymous substitution (Rich *et al.* 1998). The 95% confidence interval for the time of the initial *Wolbachia* sweep through *Cx. pipiens s.l.* is 0–47 000 years ago. This date should be considered a preliminary estimate because the neutrality assumption was not supported by all tests.

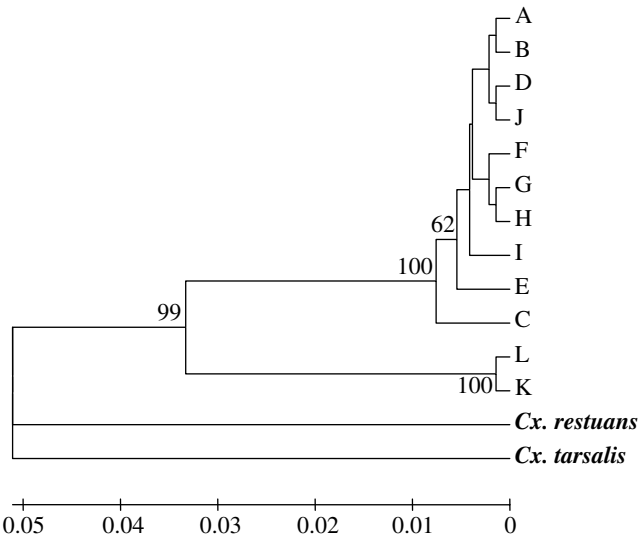

Figure 2. Phylogenetic analysis of *Culex pipiens s.l.* mitochondrial *ND4* haplotypes. Phylogenetic tree was generated using maximum likelihood. Numbers at tree nodes represent bootstrap support values (500 replicates).

4. DISCUSSION

Our results indicate that a sweep of a microbial symbiont, *Wolbachia* had profound effects on the worldwide population structure of a complex of closely related insects. We observed almost no mitochondrial variation in infected *Cx. pipiens* populations collected at geographically distinct locations, compared to high levels of variation in a single uninfected population (SAP) that was sympatric with an infected population (SAQ). Haplotype K was shared by virtually all (98.75%) infected mosquitoes regardless of location. The only other haplotype found among infected mosquitoes (L, 1.025%) differed from K by a single substitution. Levels of nuclear diversity between

Table 2. Mitochondrial *ND4* haplotype variability.

population	no. of haplotypes	*N*	*Hd*	*k*	π_2	no. of polymorphic sites
SAP	11	21	0.895	4.105	0.0121	26
SAQ	1	50	0	0	0	0
CAP	1	10	0	0	0	0
CAQ	2	10	0.2	0.2	0.00057	1
THQ	1	10	0	0	0	0
infected	2	80	0.025	0.025	0.00007	1
uninfected	11	21	0.895	4.105	0.0121	26

Table 3. AMOVA results for mitochondrial haplotype variation.

grouping	source of variation	d.f.	SS %	variation
region	between regions	2	34.7	−99.94
	between populations within regions	2	278.2	188.2
	within populations	96	41.9	11.74
infection status	between infection status	1	312.8	95.83
	between populations within infection status	3	0.087	−0.28
	within populations	96	41.95	4.45

population SAP and the sympatric infected population SAQ were similar, indicating that the reduced variation was due to a selective sweep of the K haplotype, not to a genetic bottleneck followed by population expansion. Haplotype K was observed at low levels (4.8%) in the uninfected population SAP. Phylogenetic analysis of nuclear DNA sequences supports the hypothesis that the presence of haplotype K in population SAP reflects an ancestral polymorphism rather than hybrid introgression. However, is should be noted that the hypothesis of introgression cannot be formally rejected since the nuclear region we sequenced exists in high copy number (Miller *et al.* 1996) and undergoes concerted evolution, making direct comparisons with mtDNA data somewhat problematic. Nevertheless, the observed patterns of nuclear and mitochondrial variation between infected and uninfected populations are so striking that we feel the most likely explanation is genetic hitchhiking of haplotype K with the initial *Wolbachia* invasion into the *Cx. pipiens* complex.

Our current and previous data (Cornel *et al.* 2003) suggest that population SAP may represent a new cryptic sibling species within the *Cx. pipiens* species complex. Reproductive isolation of this population from other sympatric infected populations has been confirmed by multiple nuclear and cytoplasmic markers. The role of *Wolbachia* in speciation events is controversial, but theoretical and empirical support for the idea is growing (Werren 1997; Rokas 2000; Bordenstein *et al.* 2003; Telschow *et al.* 2005). It is possible that *Wolbachia*-induced unidirectional CI can contribute to reproductive isolation between infected and uninfected populations. For instance, *Drosophila recens* is infected with *Wolbachia*, while its sister taxa *Drosophila subaquinaria* is uninfected. In laboratory experiments, gene flow via matings between *D. subaquinaria* males and *D. recens* females is hindered by behavioural isolation. *D. recens* males will readily mate with *D. subaquinaria* females, but gene flow in this cross is prevented by strong CI-induced sterility (Shoemaker *et al.* 1999). We hypothesize *Wolbachia* may be acting in a similar manner to prevent gene flow between sympatric infected and uninfected *Cx. pipiens* complex populations in SA, i.e. prezygotic isolation between *Cx. p. pipiens* males and *Cx. p. quinquefasciatus* females and CI-induced sterility between *Cx. p. quinquefasciatus* males and *Cx. p. pipiens* females. Before the initial horizontal transfer of *Wolbachia* into *Cx. pipiens s.l.*, non-reciprocal gene flow might have existed between these two populations. A *Wolbachia* sweep would explain the pattern observed today of two sympatric but genetically isolated *Culex* populations, both with equal levels of nucleotide diversity, high levels of mt variability in the uninfected population and fixation of a single mt haplotype in the infected population that is shared with the uninfected population. Future studies of population mating structure are warranted to confirm or refute this hypothesis.

Worldwide, populations of *Cx. pipiens* exhibit what is perhaps the greatest variation in CI crossing patterns observed in any insect (Laven 1967). Our analysis of *wsp* sequences showed no variation, indicating that worldwide *Culex Wolbachia* strains are closely related. Considering that the *Wolbachia* sweep of the *Cx. pipiens* complex appears to have been recent; how could this multitude of crossing types evolve in such a short evolutionary time period? It is possible that genetic differences in mosquito host factors are partially responsible; such factors have recently been implicated in modifying CI phenotype in *Culex* (Sinkins *et al.* 2005) and in other insects (Bordenstein *et al.* 2003; Mercot & Charlat 2004). Alternatively or in concert with host related factors, mobile genetic entities in *Wolbachia*, such as phages (Sinkins *et al.* 2005; Duron *et al.* 2006) or transposable elements (Bordenstein & Wernegreen 2004; Sanogo & Dobson 2004; Wu *et al.* 2004; Duron *et al.* 2005) could facilitate rapid evolution by causing mutations from insertion/transposition and/or by acting as a mechanism for horizontal transfer of genetic material between *Wolbachia* strains (Sinkins *et al.* 2005).

In addition to an improved understanding of the mechanisms that contributed to the complicated population structure of an insect complex, spread of *Wolbachia*

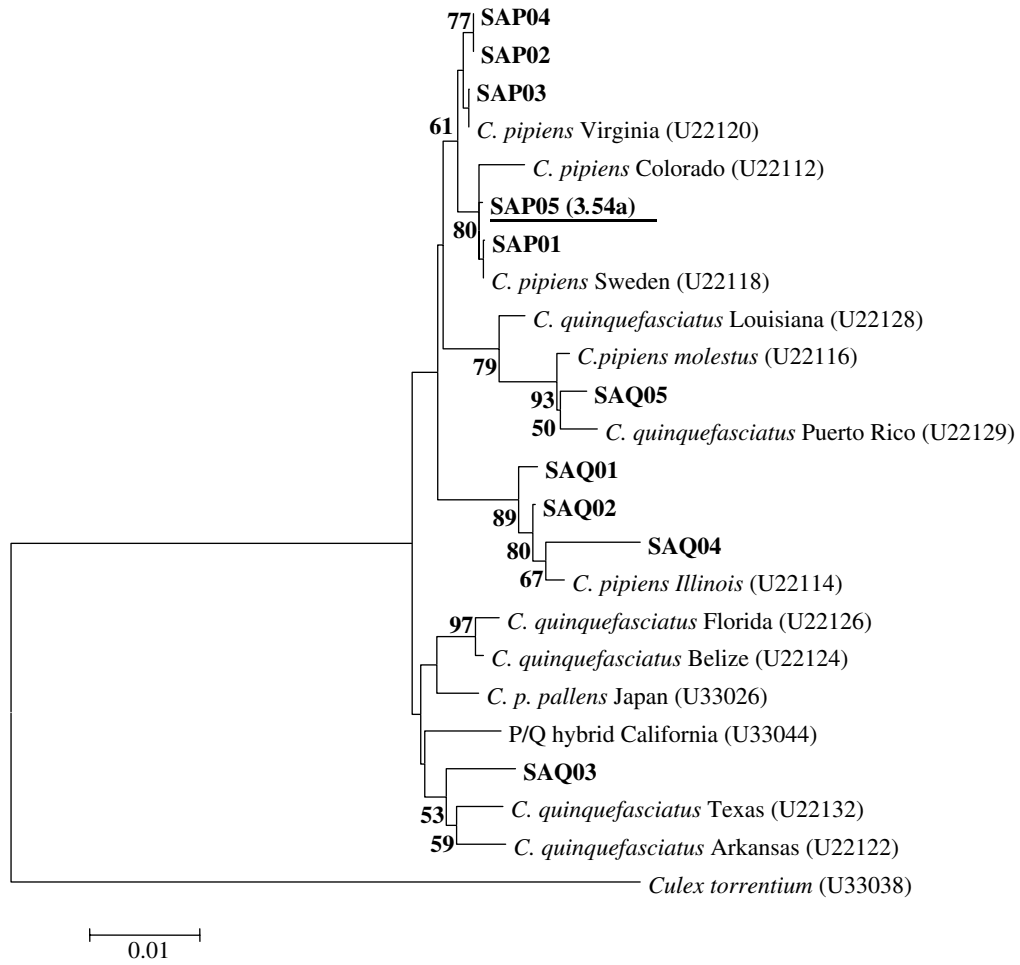

Figure 3. Phylogenetic analysis of *Cx. pipiens s.l.* ITS and ribosomal DNA sequences. Phylogenetic tree was generated using neighbour-joining with maximum-likelihood settings. Numbers at tree nodes represent bootstrap support values (1000 replicates). Taxa in bold represent clones from South African individuals from this study. Clone SAP05 (individual 3.54a, underlined) possessed *ND4* haplotype 'K'. Alphanumeric codes represent GenBank accession numbers from Miller *et al.* (1996).

in mosquito populations is of applied interest for the control of vector-borne diseases (Turelli & Hoffmann 1999; Rasgon & Scott 2003; Sinkins 2004). For more than a decade a high-profile effort has been underway to genetically modify mosquitoes so that they no longer transmit pathogens (Beaty 2000). Transgenes that block pathogen transmission have been introduced into and expressed in mosquitoes (Ito *et al.* 2002). However, no empirically demonstrated method to spread or 'drive' these engineered genetic traits into wild mosquito populations currently exists. Strategies exploiting *Wolbachia*-induced CI to drive introduced transgenic traits into vector populations are being considered (Turelli & Hoffmann 1999; Rasgon & Scott 2003; Sinkins 2004). In one strategy, separate cytoplasmically inherited transgenes can be driven into a population along with *Wolbachia* as long as the transgene construct is transmitted to 100% of the offspring (Turelli & Hoffmann 1999).

Because mitochondria are cytoplasmically inherited and transmitted with near 100% fidelity, their dynamics can be used to predict how introduced transgenes might spread under the influence of a *Wolbachia* driver (Turelli & Hoffmann 1999). The essentially single *Wolbachia* mitochondrial haplotype and lack of variation in geographically separated *wsp* sequences indicate that *Wolbachia* invasion of the *Cx. pipiens* complex was recent, rapid and initiated by one or very few infected females. This implies that under proper circumstances, the number of transgenic mosquitoes that must be released may be low and that the gene of interest could spread in a relatively short period of time across an extensive geographical range.

We thank Julie Christiansen (UC Davis Mosquito Control Research Laboratory) for conducting crossing experiments between California and SA *Cx. p. quinquefasciatus*, Michael Petridis (Johns Hopkins University) for obtaining nuclear DNA sequences and Sylvain Mousset for helpful advice regarding Allelix software. We thank three anonymous reviewers for comments and suggestions that significantly improved the manuscript. This research was supported by the Johns Hopkins Malaria Research Institute and by NIH grant GM020092 to J.L.R., funds from the state of California to T.W.S. and from the University of California Mosquito Research Laboratory to A.J.C.

REFERENCES

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

Barr, A. R. 1982 Symbiont control of reproduction in *Culex pipiens*. In *Recent developments in the genetics of insect disease vectors* (ed. W. W. M. Steiner, W. J. Tabachnick, K. S. Rai & S. Narang), pp. 153–158. Champaign, IL: Stipes.

Baurdy, E., Bartos, J., Whitworth, T. & Werren, J. H. 2003 *Wolbachia* and genetic variability in the birdnest blowfly *Protocalliphora sialia*. *Mol. Ecol.* **12**, 1843–1854. (doi:10.1046/j.1365-294X.2003.01855.x)

Beaty, B. J. 2000 Genetic manipulation of vectors: a potential novel approach for control of vector-borne diseases. *Proc. Natl Acad. Sci. USA* **97**, 10 295–10 297. (doi:10.1073/pnas.97.19.10295)

Bennett, K. E., Olson, K. E., Munoz, M. L., Fernandez-Salas, I., Farfan-Ale, J. A., Higgs, S., Black 4th, W. C. & Beaty, B. J. 2002 Variation in vector competence for dengue 2 virus among 24 collections of *Aedes aegypti* from Mexico and the United States. *Am. J. Trop. Med. Hyg.* **67**, 85–92.

Besansky, N. J. & Fahey, G. T. 1997 Utility of the white gene in estimating phylogenetic relationships among mosquitoes (Diptera: Culicidae). *Mol. Biol. Evol.* **14**, 442–454.

Besansky, N. J., Lehmann, T., Fahey, G. T., Fontenille, D., Braack, L. E., Hawley, W. A. & Collins, F. H. 1997 Patterns of mitochondrial variation within and between African malaria vectors, *Anopheles gambiae* and *An. arabiensis*, suggest extensive gene flow. *Genetics* **147**, 1817–1828.

Bordenstein, S. R. & Wernegreen, J. J. 2004 Bacteriophage flux in endosymbionts (*Wolbachia*): infection frequency, lateral transfer, and recombination rates. *Mol. Biol. Evol.* **21**, 1981–1991. (doi:10.1093/molbev/msh211)

Bordenstein, S. R., Uy, J. J. & Werren, J. H. 2003 Host genotype determines cytoplasmic incompatibility type in the haplodiploid genus *Nasonia*. *Genetics* **164**, 223–233.

Clement, M., Posada, D. & Crandall, K. A. 2000 TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**, 1657–1659. (doi:10.1046/j.1365-294x.2000.01020.x)

Cornel, A. J., McAbee, R. D., Rasgon, J., Stanich, M. A., Scott, T. W. & Coetzee, M. 2003 Differences in extent of genetic introgression between sympatric *Culex pipiens* and *Culex quinquefasciatus* (Diptera: Culicidae) in California and South Africa. *J. Med. Entomol.* **40**, 36–51.

Day, J. F. 2001 Predicting St. Louis encephalitis virus epidemics: lessons from recent, and not so recent, outbreaks. *Annu. Rev. Entomol.* **46**, 111–138. (doi:10.1146/annurev.ento.46.1.111)

Depaulis, F. & Veuille, M. 1998 Neutrality tests based on the distribution of haplotypes under an infinite-site model. *Mol. Biol. Evol.* **15**, 1788–1790.

Donnell, M. J., Licht, M. C. & Lehmann, T. 2001 Evidence for recent population expansion in the evolutionary history of the malaria vectors *Anopheles arabiensis* and *Anopheles gambiae*. *Mol. Biol. Evol.* **18**, 1353–1364.

Duron, O., Lagnel, J., Raymond, M., Bourtzis, K., Fort, P. & Weill, M. 2005 Transposable element polymorphism of *Wolbachia* in the mosquito *Culex pipiens*: evidence of genetic diversity, superinfection and recombination. *Mol. Ecol.* **14**, 1561–1573. (doi:10.1111/j.1365-294X.2005.02495.x)

Duron, O., Fort, P. & Weill, M. 2006 Hypervariable prophage WO sequences describe an unexpected high number of *Wolbachia* variants in the mosquito *Culex pipiens*. *Proc. R. Soc. B* **273**, 495–502.

Excoffier, L., Smouse, P. E. & Quattro, J. M. 1992 Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491.

Fanello, C. *et al.* 2003 The pyrethroid knock-down resistance gene in the *Anopheles gambiae* complex in Mali and further indication of incipient speciation within *An. gambiae s.s.* *Insect Mol. Biol.* **12**, 241–245. (doi:10.1046/j.1365-2583.2003.00407.x)

Fonseca, D. M., LaPointe, D. A. & Fleischer, R. C. 2000 Bottlenecks and multiple introductions: population genetics of the vector of avian malaria in Hawaii. *Mol. Ecol.* **9**, 1803–1814. (doi:10.1046/j.1365-294x.2000.01070.x)

Fonseca, D. M., Keyghobadi, N., Malcolm, C. A., Mehmet, C., Schaffner, F., Mogi, M., Fleischer, R. C. & Wilkerson, R. C. 2004 Emerging vectors in the *Culex pipiens* complex. *Science* **303**, 1535–1538. (doi:10.1126/science.1094247)

Gorrochotegui-Escalante, N., Munoz, M. L., Fernandez-Salas, I., Beaty, B. J. & Black 4th, W. C. 2000 Genetic isolation by distance among *Aedes aegypti* populations along the northeastern coast of Mexico. *Am. J. Trop. Med. Hyg.* **62**, 200–209.

Gorrochotegui-Escalante, N., Gomez-Machorro, C., Lozano-Fuentes, S., Fernandez-Salas, L., De Lourdes Munoz, M., Farfan-Ale, J. A., Garcia-Rejon, J., Beaty, B. J. & Black 4th, W. C. 2002 Breeding structure of *Aedes aegypti* populations in Mexico varies by region. *Am. J. Trop. Med. Hyg.* **66**, 213–222.

Guillemaud, T., Pasteur, N. & Rousset, F. 1997 Contrasting levels of variability between cytoplasmic genomes and incompatibility types in the mosquito *Culex pipiens*. *Proc. R. Soc. B* **264**, 245–251. (doi:10.1098/rspb.1997.0035)

Hoffmann, A. A. & Turelli, M. 1997 Cytoplasmic incompatibility in insects. In *Influential passengers* (ed. S. L. O'Neill, A. A. Hoffmann & J. H. Werren), pp. 42–80. Oxford, UK: Oxford University Press.

Hoogstraal, H., Meegan, J. M., Khalil, G. M. & Adham, F. K. 1979 The Rift Valley fever epizootic in Egypt 1977–78. 2. Ecological and entomological studies. *Trans. R. Soc. Trop. Med. Hyg.* **73**, 624–629. (doi:10.1016/0035-9203(79)90005-1)

Hurst, G. D. & Jiggins, F. M. 2005 Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc. R. Soc. B* **272**, 1525–1534. (doi:10.1098/rspb.2004.3004)

Irving-Bell, R. J. 1974 Cytoplasmic factors in the gonads of *Culex pipiens* complex mosquitoes. *Life Sci.* **14**, 1149–1151. (doi:10.1016/0024-3205(74)90239-2)

Ito, J., Ghosh, A., Moreira, L. A., Wimmer, E. A. & Jacobs-Lorena, M. 2002 Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. *Nature* **417**, 452–455. (doi:10.1038/417452a)

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

Jupp, P. G. 1978 *Culex* (*Culex*) *pipiens pipiens* Linnaeus and *Culex* (*Culex*) *pipiens quinquefasciatus* Say in South Africa: morphological and reproductive evidence in favour of their status as two species. *Mosq. Syst.* **10**, 461–473.

Krafsur, E. S., Madsen, M., Wohlford, D. L., Mihok, S. & Griffiths, N. T. 2000 Population genetics of *Glossina morsitans submorsitans* (Diptera: Glossinidae). *Bull. Entomol. Res.* **90**, 329–335.

Krafsur, E. S., Endsley, M. A., Wohlford, D. L., Griffiths, N. T. & Allsopp, R. 2001 Genetic differentiation of *Glossina morsitans centralis* populations. *Insect Mol. Biol.* **10**, 387–395. (doi:10.1046/j.0962-1075.2001.00277.x)

Krida, G., Bouattour, A., Rodhain, F. & Failloux, A. B. 1998 Variability among Tunisian populations of *Culex pipiens*: genetic structure and susceptibility to a filarial parasite, *Brugia pahangi*. *Parasitol. Res.* **84**, 139–142. (doi:10.1007/s004360050371)

Laven, H. 1967 A possible model for speciation by cytoplasmic isolation in the *Culex pipiens* complex. *Bull. WHO* **37**, 263–266.

Marquez, J. G., Vreysen, M. J., Robinson, A. S., Bado, S. & Krafsur, E. S. 2004 Mitochondrial diversity analysis of

Glossina palpalis gambiensis from Mali and Senegal. *Med. Vet. Entomol.* **18**, 288–295. (doi:10.1111/j.0269-283X.2004.00508.x)

Mercot, H. & Charlat, S. 2004 *Wolbachia* infections in *Drosophila melanogaster* and *D. simulans*: polymorphism and levels of cytoplasmic incompatibility. *Genetica* **120**, 51–59. (doi:10.1023/B:GENE.0000017629.31383.8f)

Miller, B. R., Crabtree, M. B. & Savage, H. M. 1996 Phylogeny of fourteen *Culex* mosquito species, including the *Culex pipien* complex, inferred from the internal transcribed spacers of ribosomal DNA. *Insect Mol. Biol.* **5**, 93–107.

Nasci, R. S. *et al*. 2001 West Nile virus in overwintering *Culex* mosquitoes, New York City, 2000. *Emerg. Infect. Dis.* **7**, 742–744.

O'Neill, S. L., Giordano, R., Colbert, A. M., Karr, T. L. & Robertson, H. M. 1992 16S rRNA phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic incompatibility in insects. *Proc. Natl Acad. Sci. USA* **89**, 2699–2702.

Parvizi, P., Benlarbi, M. & Ready, P. D. 2003 Mitochondrial and *Wolbachia* markers for the sandfly *Phlebotomus papatasi*: little population differentiation between peridomestic sites and gerbil burrows in Isfahan province, Iran. *Med. Vet. Entomol.* **17**, 351–362. (doi:10.1111/j.1365-2915.2003.00451.x)

Posada, D. & Crandall, K. A. 1998 MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818. (doi:10.1093/bioinformatics/14.9.817)

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

Rasgon, J. L. & Scott, T. W. 2004 An initial survey for *Wolbachia* (Rickettsiales: Rickettsiaceae) infections in selected California mosquitoes (Diptera: Culicidae). *J. Med. Entomol.* **41**, 255–257.

Raymond, M. & Rousset, F. 1995 GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* **86**, 248–249.

Rich, S. M., Licht, M. C., Hudson, R. R. & Ayala, F. J. 1998 Malaria's Eve: evidence of a recent population bottleneck throughout the world populations of *Plasmodium falciparum*. *Proc. Natl Acad. Sci. USA* **95**, 4425–4430. (doi:10.1073/pnas.95.8.4425)

Rokas, I. I. 2000 *Wolbachia* as a speciation agent. *Trends Ecol. Evol.* **15**, 44–45. (doi:10.1016/S0169-5347(99)01783-8)

Rozas, J. & Rozas, R. 1999 DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15**, 174–175. (doi:10.1093/bioinformatics/15.2.174)

Sanogo, Y. O. & Dobson, S. L. 2004 Molecular discrimination of *Wolbachia* in the *Culex pipiens* complex: evidence for variable bacteriophage hyperparasitism. *Insect Mol. Biol.* **13**, 365–369. (doi:10.1111/j.0962-1075.2004.00498.x)

Shoemaker, D. D., Katju, V. & Jaenike, J. 1999 The role of *Wolbachia* in the reproductive isolation between *Drosophila recens* and *Drosophila subquinaria*. *Evolution* **53**, 1157–1164.

Simon, C., Franke, A. & Martin, A. 1991 Polymerase chain reaction: DNA extraction and amplification. In *Molecular techniques in taxonomy* (ed. G. M. Hewitt, A. W. B. Johnston & J. W. P. Young), pp. 329–355. Berlin, Germany: Springer.

Sinkins, S. P. 2004 *Wolbachia* and cytoplasmic incompatibility in mosquitoes. *Insect Biochem. Mol. Biol.* **34**, 723–729. (doi:10.1016/j.ibmb.2004.03.025)

Sinkins, S. P., Walker, T., Lynd, A. R., Steven, A. R., Makepeace, B. L., Godfray, H. C. & Parkhill, J. 2005 *Wolbachia* variability and host effects on crossing type in *Culex* mosquitoes. *Nature* **436**, 257–260. (doi:10.1038/nature03629)

Sundararaman, S. 1949 Biometrical studies on intergradation in the genitalia of certain populations of *Culex pipiens* and *Culex quinquefasciatus* in the United States. *Am. J. Hyg.* **50**, 307–314.

Swofford, D. L. 1998. *PAUP* 4.0—phylogenetic analysis using parsimony, v. 4.01b*, 10. Sunderland, MA: Sinauer Associates.

Tabachnick, W. J. & Powell, J. R. 1983 Genetic analysis of *Culex pipiens* populations in the central valley of California. *Ann. Entomol. Soc. Am.* **76**, 715–720.

Tajima, F. 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.

Telschow, A., Hammerstein, P. & Werren, J. H. 2005 The effect of *Wolbachia* versus genetic incompatibilities on reinforcement and speciation. *Evol. Int. J. Org. Evol.* **59**, 1607–1619.

Templeton, A. R., Crandall, K. A. & Sing, C. F. 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., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. 1997 The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882. (doi:10.1093/nar/25.24.4876)

Tripet, F., Toure, Y. T., Taylor, C. E., Norris, D. E., Dolo, G. & Lanzaro, G. C. 2001 DNA analysis of transferred sperm reveals significant levels of gene flow between molecular forms of *Anopheles gambiae*. *Mol. Ecol.* **10**, 1725–1732. (doi:10.1046/j.0962-1083.2001.01301.x)

Turelli, M. & Hoffmann, A. A. 1999 Microbe-induced cytoplasmic incompatibility as a mechanism for introducing transgenes into arthropod populations. *Insect Mol. Biol.* **8**, 243–255. (doi:10.1046/j.1365-2583.1999.820243.x)

Turelli, M., Hoffmann, A. A. & McKechnie, S. W. 1992 Dynamics of cytoplasmic incompatibility and mtDNA variation in natural *Drosophila simulans* populations. *Genetics* **132**, 713–723.

Urbanelli, S., Bullini, L. & Villani, F. 1985 Electrophoretic studies on *Culex quinquefasciatus* Say from Africa: genetic variability and divergence from *Culex pipiens* L. (Diptera: Culicidae). *Bull. Entomol. Res.* **75**, 291–304.

Urbanelli, S., Silvestrini, F., Sabatinelli, G., Raveloarifera, F., Petrarca, V. & Bullini, L. 1995 Characterization of the *Culex pipiens* complex (Diptera: Culicidae) in Madagascar. *J. Med. Entomol.* **32**, 778–786.

Urbanelli, S., Silvestrini, F., Reisen, W. K., De Vito, E. & Bullini, L. 1997 Californian hybrid zone between *Culex pipiens pipiens* and *Cx. p. quinquefasciatus* revisited (Diptera: Culicidae). *J. Med. Entomol.* **34**, 116–127.

Werren, J. H. 1997 Biology of *Wolbachia*. *Annu. Rev. Entomol.* **42**, 587–609. (doi:10.1146/annurev.ento.42.1.587)

Wu, M., Sun, L. V. *et al*. 2004 Phylogenomics of the reproductive parasite *Wolbachia pipientis wMel*: a streamlined genome overrun by mobile genetic elements. *PLoS Biol.* **2**, 327–341. (doi:10.1371/journal.pbio.0020327)

Zhou, W., Rousset, F. & O'Neil, S. 1998 Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. *Proc. R. Soc. B* **265**, 509–515. (doi:10.1098/rspb.1998.0324)