Male-Killing Wolbachia and Mitochondrial DNA: Selective Sweeps, Hybrid Introgression and Parasite Population Dynamics

Francis M. Jiggins[1]

Department of Genetics, University of Cambridge, Cambridge, CB2 3EH, United Kingdom

ABSTRACT

Mitochondrial DNA (mtDNA) sequences are widely used as neutral genetic markers in insects. However, patterns of mtDNA variability are confounded by the spread of maternally transmitted parasites, which are genetically linked to the mitochondria. We have investigated these effects in the butterflies *Acraea encedon* (which is host to two strains of male-killing Wolbachia bacteria) and *A. encedana* (which is host to one strain). Within a population, the mitochondria are in linkage disequilibrium with the different male-killers. Furthermore, there has been a recent selective sweep of the mtDNA, which has led to the loss of mitochondrial variation within populations and erased any geographical structure. We also found that one of the male-killers, together with the associated mtDNA, has introgressed from *A. encedana* into *A. encedon* within the last 16,000 years. Interestingly, because butterflies are female heterogametic, this will presumably have also led to the introgression of genes on the W sex chromosome. Finally, in *A. encedon* the mitochondria in uninfected females are unaltered by the spread of the male-killer and have diverse, geographically structured mtDNA. This means we can reject the hypothesis that the male-killer is at a stable equilibrium maintained by imperfect transmission of the bacterium. Instead, some other form of balancing selection may be maintaining uninfected females in the population and preventing the species from going extinct due to a shortage of males.

MITOCHONDRIAL DNA is widely used both in phylogenetics and as a neutral marker in population genetics. However, these studies may be confounded because many arthropods are infected by parasites that are transmitted from mother to offspring. Maternal transmission means that these parasites are genetically linked to the mitochondrial genome, and, in female heterogametic taxa like butterflies, the W sex chromosome. Therefore, patterns of mtDNA variation may often reflect the evolutionary history of these parasites rather than their hosts.

The effect of such parasites on mitochondrial evolution has been best studied in *Drosophila simulans*, which is host to several strains of Wolbachia bacteria that cause cytoplasmic incompatibility. It has been found that different populations are host to different strains of Wolbachia, each of which is associated with a distinct mitotype (MONTCHAMP-MOREAU *et al.* 1991; ROUSSET and SOLIGNAC 1995). More directly, Wolbachia was observed as it spread through Californian *D. simulans*, and, as it invaded, so did the associated mtDNA (TURELLI *et al.* 1992; TURELLI and HOFFMANN 1995). The spread of Wolbachia is therefore expected to cause a selective sweep of mtDNA, and, as predicted, the diversity of the *D. simulans* mtDNA is significantly reduced compared to the predictions of neutral theory (BALLARD *et al.* 1996; BALLARD 2000a). In a final twist to the story, the *D. simulans* mitochondria appear to have introgressed into *D. mauritiana*, probably due to the introgression of Wolbachia (BALLARD 2000b).

In this study we investigated the association between mtDNA and male-killing Wolbachia. These bacteria kill male hosts early in their development and will spread through the population if this act increases the lifetime reproductive success of the sibling female hosts, who transmit the infection to the next generation (HURST 1991). Theoretical studies have shown that male-killers are expected to significantly reduce mitochondrial diversity for three reasons (JOHNSTONE and HURST 1996). First, the initial invasion of the bacterium will result in a selective sweep of the mitotype associated with the infection. Second, at equilibrium the effective population size of mtDNA is reduced because mutations in uninfected females are rapidly lost. This occurs as imperfect vertical transmission of the bacteria results in gene flow from the infected gene pool into the uninfected population, but not in the opposite direction. Finally, there is likely to be strong selection on the parasites themselves, possibly exerted by host defenses. This will result in additional selective sweeps of the male-killer and associated mitochondria as beneficial bacterial mutations spread through the parasite population.

The only empirical study of male-killers and mito-

chondria comes from the ladybird beetle *Adalia bipunctata*, which is host to four different male-killing bacteria. Each bacterium is associated with a different set of mitochondrial sequences, and the mitochondrial phylogeny was used to infer the order in which the different male-killers invaded the population (SCHULENBURG *et al.* 2002).

In this study we have examined the association between mitochondria and male-killers in two closely related species of African butterfly, *Acraea encedon* and *A. encedana*. In both species, from 50% to >95% of females are infected with male-killing Wolbachia bacteria (JIGGINS *et al.* 1998, 2000a,b; HURST *et al.* 1999). Only a single strain of Wolbachia has been identified in *A. encedana*, while two distantly related strains are found in *A. encedon* (JIGGINS *et al.* 2001b). One of the *A. encedon* strains (Ug) is identical in the sequence of three genes (*ftsZ*, *wsp*, and the *groE* operon) to the *A. encedana* male-killer, suggesting that these bacteria are extremely closely related (JIGGINS *et al.* 2002a). This strain occurs throughout Africa, while the second strain (Tz) is known only from Tanzania, where it is found within the same population as strain Ug (JIGGINS *et al.* 2001b). The Wolbachia strains all belong to the B supergroup.

MATERIALS AND METHODS

Butterfly strains: The collection of the butterflies and the identification of Wolbachia-infected specimens have been described elsewhere (JIGGINS *et al.* 1998, 2000a, 2001b; HURST *et al.* 1999). *A. encedon* is host to two different male-killing Wolbachia strains called Ug and Tz, both of which were included in the study. Strain Tz is known only from Dar es Salaam in Tanzania, while strain Ug has been found across sub-Saharan Africa in Ghana, Uganda, Zimbabwe, and Tanzania (JIGGINS *et al.* 2001b). We also used uninfected specimens from Ghana, Uganda, and Tanzania. We have collected *A. encedana* from Uganda only, where it is infected with the strain Ug Wolbachia (JIGGINS *et al.* 2000a).

Mitochondrial sequences: Initially, a short fragment of the mitochondrial *COI* gene was sequenced from a large number of individuals to assess the distribution of the different mitotypes. This was amplified using the PCR primers C1-J-1751 and C1-N-2191 (SIMON *et al.* 1994). The mitochondrial primers are all numbered according to where the 3′ base occurs in the *D. yakuba* sequence (SIMON *et al.* 1994). The temperature cycle was 95° for 2 min; then 35 cycles of 95° for 30 sec, 55° for 1 min, and 70° for 1 min; and finally a single cycle at 70° for 10 min. This fragment was then sequenced directly using the forward primer.

To investigate the phylogenetic relationships of the mitotypes identified from the short sequences, a longer region was sequenced, which encompassed most of *COI* and all of *COII*. Two overlapping fragments were amplified using the primer pairs C1-J-1751/C2-N-3389 and C1-J-2183/TK-N-3795 (SIMON *et al.* 1994), using the above cycle but with a 2-min extension time. The purified PCR product was then sequenced using the primers C1-J-1751, C1-N-2191, C1-J-2630 (5′-CTTTCTATAG GAGCTGTATTTGC), C2-N-3389, C1-J-2183, and TK-N-3795 (5′-GGGCTATAATATGGTTTAAGAGA). Primer sequences not detailed can be found in SIMON *et al.* (1994). To root the phylogeny, the same region was sequenced from single specimens of *A. aurivilli* and *A. jodutta*, both of which were collected from southern Uganda. The sequences were submitted to EMBL under accession nos. AJ414617–AJ414628.

Longer sequences were used to investigate the level of polymorphism in *A. encedana*. In addition to the regions described above, we sequenced a 1256-bp fragment from five Ugandan *A. encedana* and one specimen of *A. jodutta* consisting of tRNA-Ser, tRNA-Glu, tRNA-Phe, and part of the ND5 gene. This region was amplified and sequenced using the primers TN-J-6172 (5′-GAGGTAAATCACTGTTAATGA) and N5-N-7493, using the thermal cycle described above. The sequence accession numbers are AJ439626–AJ439627.

To date the split of *A. encedon* and *A. encedana*, the rest of the ND5 gene was sequenced from a single Ug-infected individual of both species. This was amplified using the reverse primers TN-J-6172 and TH-N-8176 (5′-TGATTTGTGGTAT CAATGATA). The sequence accession numbers are AJ-508686–AJ508687.

Nuclear sequences: We also sequenced three nuclear regions. All three were used to provide phylogenies for comparison with the mitochondrial sequences, and the two single-copy nuclear genes were used to compare with the diversity of the mitochondria.

The first region sequenced was the first internally transcribed spacer of the ribosomal DNA transcription unit (ITS1). This is present in multiple copies that are subject to concerted evolution, which homogenizes the sequences of different copies within individuals and populations. For this reason, ITS1 was sequenced only from a single individual from each population and a single specimen of the outgroup *A. jodutta*. The ITS1 region was amplified under identical conditions to the short mtDNA region, but using the primers BD1 and 4S (HILLIS and DIXON 1991). The PCR product was sequenced directly using the PCR primers. The sequences obtained included the entire ITS1 region and short flanking sequences of the 5.8S and 18S rDNA. There was little evidence of "double peaks" in the sequence, suggesting that a single sequence predominates within an individual. The sequences were submitted to EMBL under accession nos. AJ414611–AJ414616.

We also sequenced individual alleles of the two single-copy nuclear genes by first cloning the PCR product and then sequencing individual clones. PCR copying errors cause both nucleotide substitutions and recombination between different target sequences (ZYLSTRA *et al.* 1998). These were minimized by using only 25 thermal cycles and *Taq* polymerase, which causes a low rate of *in vitro* recombination (although a high rate of substitution errors; ZYLSTRA *et al.* 1998). To correct any remaining errors, three to five separate PCR reactions were set up from each individual. Each PCR product was ligated into a T-tailed vector and transformed into competent *Escherichia coli* (strain JM109), and the inserts were sequenced using primers designed from the flanking vector sequence and, where necessary, internal primers. This procedure means that each clone was derived from a separate PCR reaction, which avoids the potential duplication of errors either by replication of the bacteria before they are plated out or during the PCR reaction. A second source of error when estimating allele frequencies arises from the difficulty of identifying homozygotes—if several identical clones are sequenced they may come from a homozygote or the same allele of a heterozygote. This was avoided by retaining only one allele from each individual.

The first single-copy nuclear gene sequenced was glucose 6-phosphate dehydrogenase (*g6pd*). Initially short fragments were amplified and sequenced using degenerate primers 56F and 209R, which amplify a fragment at the 5′ end of the gene (SOTOADAMES *et al.* 1994). This was used to design a forward primer at the 3′ end of this sequence. An alignment of human

and Drosophila protein sequences (FOUTS *et al.* 1988) was used to design a new degenerate reverse primer at the 3′ end of the gene (g6pd-488r, 5′-RTCNGCYTCNKTNGGNCCNCK NSWNCCRTA). The fragment amplified by these primers was sequenced. These two sequences were then used to design two primers specific to the Acraea butterflies that amplify most of the gene (g6pd-Acraea-5′F, 5′-GGCTATGGTTCTTGTAT CGTGATGAT; and g6pd-Acraea-3r, 5′-SWNCCRTATGTAT AAGGCACAGGAG). These primers were then used to amplify this region, which was then cloned and sequenced as described above. Unlike all previously published *g6pd* sequences, the gene in Acraea does not contain any introns. The sequences were submitted under accession nos. AJ439639–AJ439648.

The second gene was the long wavelength opsin *ops1* (HSU *et al.* 2001). Initially, a short fragment at the 3′ end of the gene was amplified with the degenerate primers 80 and OPSRD (HSU *et al.* 2001). This sequence was used to design a reverse primer at the 3′ end of the gene (ops-Acraea-2r, 5′-GGCTGATCGTCATGGTTTCGAA). Long wavelength opsin amino acid sequences from nymphalid butterflies were aligned and a degenerate forward primer was designed at the 5′ end of the gene (ops-universal-2F, 5′-RTTYCCNCCNATG AAYCCNYTNTGGCA). These two primers were then used to amplify the gene. The resulting sequences contained 720 bp of exon and 399–402 bp of intron. The sequences were submitted under accession nos. AJ439628–AJ439638.

Phylogenetics: The sequences, which contained few insertions or deletions, were aligned by eye. Their phylogeny was reconstructed by maximum likelihood with the program PAUP* v.4.0b8 (SWOFFORD 1998). The phylogenies were reconstructed using the HKY85 model, which accounts for the base frequency and transition:transversion ratio (HASEGAWA *et al.* 1985). These parameters were estimated from a maximum-parsimony tree. The maximum-likelihood topology was then reconstructed by a heuristic search using nearest-neighbor interchanges, and its robustness assessed by 10,000 nonparametric bootstrap replicates.

We then tested the hypothesis that the mtDNA sequences from *A. encedon* and *A. encedana* formed two monophyletic groups. First, the maximum-likelihood tree was reconstructed under the constraint of monophyly of the two species. We compared the log likelihood of this tree to that of the unconstrained maximum-likelihood tree by the Shimodaira-Hasegawa (SH) test (SHIMODAIRA and HASEGAWA 1999; GOLDMAN *et al.* 2000), using the program SH tests v. 1.0 (RAMBAUT 2001).

RESULTS

Mitochondrial diversity in *A. encedana*: A single mitochondrial sequence was found in both infected and uninfected *A. encedana.* No polymorphisms were detected in 430 bp from seven infected and seven uninfected female butterflies (Table 1). Similarly, there was no variation in a longer sequence of 3225 bp from four infected and one uninfected female.

It is possible to test whether this lack of variation is significantly lower than that expected were the gene evolving neutrally by comparing this data to the level of polymorphism at an autosomal locus. Neutral theory predicts that the level of polymorphism at different loci will be positively correlated with the degree to which those loci have diverged from an outgroup. Deviations from this prediction may be tested using the Hudson-Kreitman-Aguadé (HKA) test (HUDSON *et al.* 1987), cor-

recting for the fact that mitochondria have one-quarter the effective population size of autosomal genes because they are both haploid and uniparentally inherited.

The level of nucleotide polymorphism and the extent to which these loci have diverged from *A. jodutta* are shown in Table 2. There was no variation in the *A. encedana* mitochondrial sequences, but there were 274 interspecific differences. This level of polymorphism was significantly less than that observed in *g6pd* ($\chi^2 = 7.32$, d.f. $= 1$, $P = 0.0068$). The mitochondrial polymorphism was also less than that of *ops1* ($\chi^2 = 6.66$, d.f. $= 1$, $P = 0.0099$).

The reduced polymorphism of the mitochondrial sequences could be explained by balancing selection acting on both the nuclear loci. However, this is unlikely. First, we were unable to detect any deviation from neutrality in either of the two loci using Tajima's *D*, Fu and Li's *D*, or Fu and Li's *F*-statistics (TAJIMA 1989; FU and LI 1993). Second, balancing selection is unlikely as all the nucleotide variation in *g6pd* was at silent sites, meaning that there was no variation in amino acid sequence. Therefore, we can conclude that the diversity of mtDNA is significantly less than that expected under neutrality.

The next question is, Why has the diversity of mitochondria been reduced? Two factors result in male-killers reducing mitochondrial diversity: selective sweeps and a reduced effective population size of mitochondria at equilibrium (JOHNSTONE and HURST 1996). The latter effect occurs as imperfect vertical transmission of the bacteria results in gene flow from the infected gene pool into the uninfected population, but not in the opposite direction. Therefore, alleles found only in uninfected females are rapidly lost. The reduction in the effective population size of mitochondria is greatest when the proportion of infected females infected is low.

Male-killers will also reduce the effective population size of autosomal genes (N_e) because of the biased population sex ratio. This effect is greatest when the proportion of females infected is high. In the HKA test above we assumed the population size of mitochondria ($N_{\text{mitochondria}}$) to be $0.25N_e$. In Uganda, 96% of females are infected with the male-killer. Therefore, if at this prevalence $N_{\text{mitochondria}} > 0.25N_e$, we can reject the hypothesis that the system is at equilibrium and conclude that there has been a recent selective sweep.

Conservatively, we can take the effective population size of mitochondria at equilibrium to be equal to the number of infected females (the true value will be greater than this). The number of individuals in a population that are infected females ($N_{\text{f infected}}$) can be derived from the proportion of females that are infected, *p*, and the total population size, *N*:

$$N_{\text{f infected}} = Np/(2 - p).$$

It is also conservative to assume that the effective population size of autosomal genes, N_e, is equivalent to that of an uninfected population with a sex ratio equal to that

TABLE 1

The mitotypes of infected and uninfected *A. encedon* and *A. encedana*

| Sequence | Infected with strain Ug | | | | | Infected strain Tz: | Uninfected | | | |
| | *encedana*: | *encedon* | | | | *encedon*: | *encedana*: | *encedon* | | |
	Uganda	Uganda	Tanzania	Ghana	Zimbabwe	Tanzania	Uganda	Uganda	Tanzania	Ghana
Ug	7	2	4	2	2		7	1		
Tz						10				
1								2	5	
2									1	
3									2	
4								3		
5										2
6										1
7										2

Sequences with the same number or letter (Ug, Tz, and 1–7) are identical at all 430 nucleotide sites.

of the *A. encedana* population (the true value of N_e may be less than this),

$$1/N_e = 1/4N_f + 1/4N_m,$$

where N_f and N_m are the numbers of females and males in the population, respectively. The sex ratio (proportion male) of a population infected with a male-killer is

$$N_m/(N_f + N_m) = (1 - p)/(2 - p).$$

Therefore, under these conservative assumptions, when >59% of females are infected, then $N_{f \text{ infected}} > 0.25N_e$. This leads us to conclude that the reduction in mitochondrial diversity that we observed is not expected at equilibrium and the *A. encedana* cytoplasm has been subject to a selective sweep.

Mitochondrial diversity in *A. encedon*: The two strains of Wolbachia in *A. encedon* were each associated with their own mitotype (Table 1). In samples from across sub-Saharan Africa, the strain Ug Wolbachia is associated with a single mitochondrial type (Table 1). In the Tanzanian population, which contains both bacterial

TABLE 2

Mitochondrial and autosomal polymorphism within *A. encedana* and divergence between *A. encedana* and *A. jodutta*

	Mitochondria	*g6pd*	*ops1*
Total no. sites	3225	1236	1119
No. segregating sites in *A. encedana*	0	4	10
Sample size in *A. encedana*	5	5	9
Mean no. interspecific differences	274	23.60	35.33

strains, the strain Tz is associated with a second mitotype. There was no sequence variation within either of these mitotypes in the 430 bp sequenced.

In *A. encedon,* the strain Ug male-killer is indistinguishable from the male-killer in *A. encedana.* Longer (1970 bp) sequences were obtained from two *A. encedon* strain Ug females. One of these was identical to the *A. encedana* sequences, and the other differed by a single substitution. The two 1970-bp *A. encedon* Tz sequences were identical.

The uninfected specimens of *A. encedon* contrast with the infected specimens in that they had different, diverse, and geographically structured mitotypes (Table 1). First, the mitotypes associated with the male-killers are rare in the uninfected population; in 19 specimens only 1 shared the same sequence as the infected population. Second, the uninfected population contained far more diverse mitotypes. The 19 uninfected specimens had eight different mitotypes, in contrast to the total absence of diversity within the 10 strain Ug and 10 strain Tz specimens. Third, the uninfected females had geographically structured mitochondrial types, with the different mitotypes being nonrandomly distributed among the populations (Monte Carlo simulation, 50,000 replicates, $P < 0.0001$).

Mitochondrial and nuclear phylogenies: The observation that the strain Ug male-killer was associated with identical mitotypes in both host species suggests that it may have moved between the two species by hybrid introgression. Alternatively, the infection could have been present before the two species diverged, and this mitotype has been inherited from a common ancestor. These two hypotheses were distinguished by reconstructing the phylogenies of the different mitotypes and three nuclear genes.

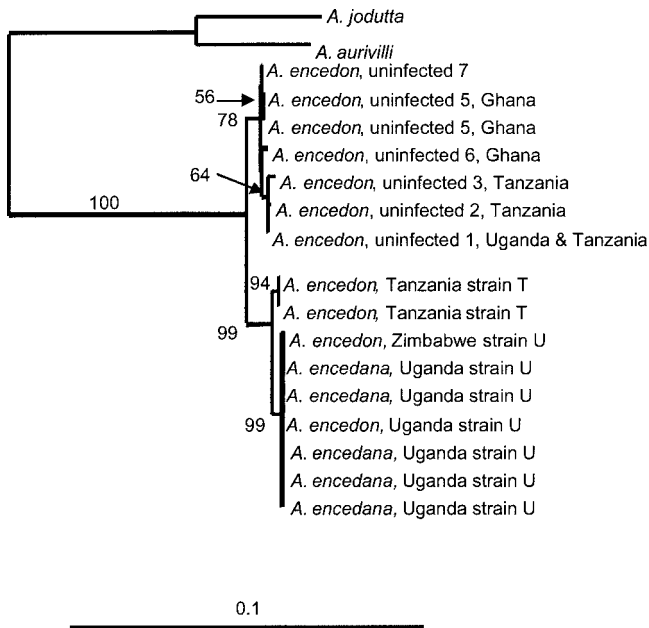The mitochondrial phylogeny contains two monophyletic groups (Figure 1). The first contains all the unin-

FIGURE 1.—Maximum-likelihood phylogeny of Acraea mitochondria. The percentage of 10,000 bootstrap replicates supporting each node is shown. The scale represents the mean number of substitutions per site.

FIGURE 2.—Maximum-likelihood phylogeny of the nuclear ITS1 region.

fected *A. encedon*. The second consists of both *A. encedana* and the infected *A. encedon* mitotypes. Therefore, *A. encedon* is paraphyletic, with the mitotypes clustering according to their infection status rather than the species they were sequenced from. Furthermore, this pattern is statistically significant, as we are able to reject the hypothesis that the two species were monophyletic (Shimodaira-Hasegawa test, $P = 0.02$).

In contrast to the mitochondrial phylogenies, the three nuclear genes all support the hypothesis that the two species are monophyletic (Figures 2–4). Therefore, the pattern seen in the mitochondrial phylogeny, where *A. encedon* was paraphyletic, is not reflected in autosomal loci.

Dating the spread of Wolbachia: Using these data we can estimate the date at which the strain Ug male-killer invaded both species of butterflies. The first event is the spread of the male-killer through *A. encedana*. The second event is the introgression of this male-killer from *A. encedana* into *A. encedon*.

We can date the introgression of Wolbachia into the *A. encedon* population by examining the number of substitutions at fourfold degenerate (silent) sites that have occurred between the two species in the mtDNA associated with the strain Ug Wolbachia. For this purpose, I used the longest sequences, which were obtained from a single individual of each species. These sequences were identical at all 3886 bp, which included 456 fourfold degenerate (silent) sites and 1052 twofold degenerate sites. The 95% binomial confidence interval of four-
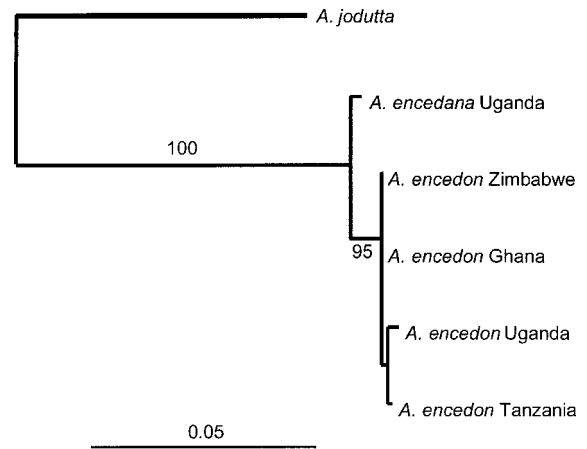
fold degenerate sites differing between the two species is 0–0.65%. Assuming a substitution rate of 57×10^{-9} substitutions/silent site/year (TAMURA 1992), this gives a date of introgression of 0–57,000 years ago (95% confidence interval). This approach is conservative, as it ignores the lack of variation at the other 3430 sites.

We have a far larger data set of mtDNA sequences from the *A. encedana* population, but it is difficult to date the selective sweep because we do not know the effective population size of the mitochondria. This can be overcome if we assume that the selective sweep was recent and the population size is large, so that the mitochondrial geneology is star shaped. The coding sequences contain 380 fourfold degenerate sites and 850 twofold degenerate sites, none of which are polymorphic in the five alleles sequenced. Using the method of RICH *et al.* (1998), and the mutation rate given above for fourfold degenerate sites, the selective sweep in *A. encedana* occurred 0–28,000 years before present (95% confidence limit).

A more accurate estimate can be obtained by including the twofold degenerate sites. The alignment of the *A. encedana* Ug sequence with the outgroup *A. jodutta* contains 343 fourfold degenerate sites and 805 twofold degenerate sites. There are 114 differences between the species at the fourfold sites and 99 differences at the twofold sites. These values can be corrected for multiple substitutions, using the Jukes-Cantor model (JUKES and CANTOR 1969; for the two-state model for the twofold sites see RICH *et al.* 1998), to give 150.6 substitutions at fourfold sites and 113.6 substitutions at twofold sites. Therefore, the rate of synonymous substitutions per twofold degenerate site is 0.321 times that at fourfold degenerate sites, or 18×10^{-9} substitutions/silent site/year (TAMURA 1992).

Therefore, including all 380 fourfold degenerate sites and 850 twofold degenerate sites in the polymorphism data, the method of RICH *et al.* (1998) gives a date of
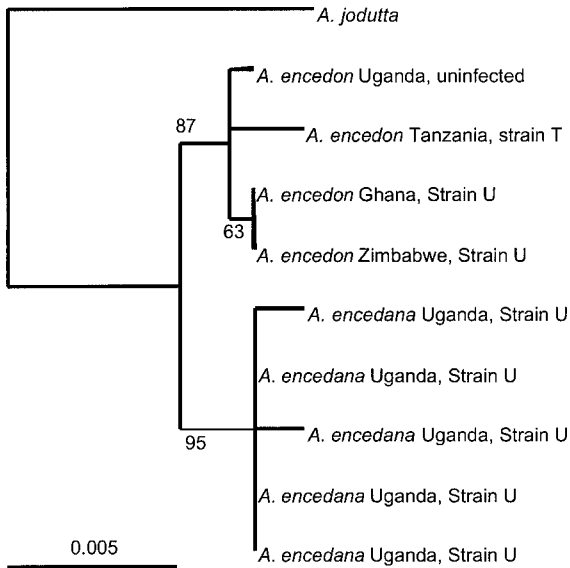
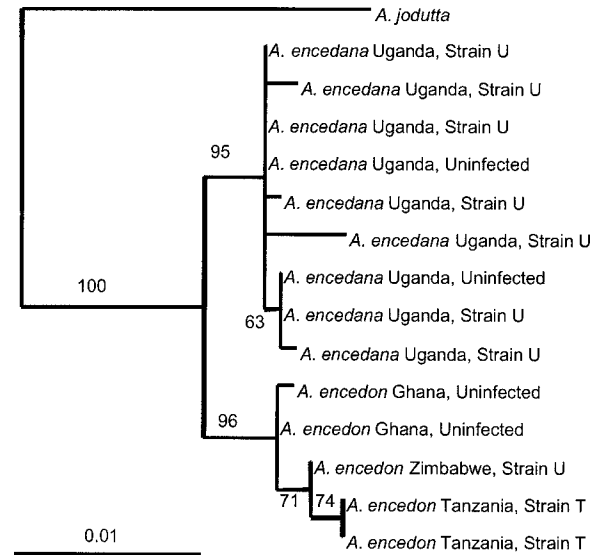FIGURE 3.—Maximum-likelihood phylogeny of the nuclear gene *g6pd*.

FIGURE 4.—Maximum-likelihood phylogeny of the nuclear gene *ops1*.

the selective sweep as 0–16,000 years before present (95% confidence limits).

DISCUSSION

Linkage disequilibrium: In *A. encedon* there is perfect linkage disequilibrium within a population between mitochondrial type and two different male-killers. Therefore, we expect selection acting on these parasites to also affect the pattern of mitochondrial diversity. This association also suggests that horizontal transmission of the infection is rare. Therefore, we can reject the hypothesis that horizontal transmission plays an important role in driving these infections to extremely high prevalence (JIGGINS *et al.* 2000a).

Selective sweeps: The first hypothesis we tested was that male-killing bacteria reduce the diversity of their host's mtDNA. By comparing the levels of intraspecific polymorphism and interspecific divergence of both mitochondrial and nuclear genes, we found that the level of mitochondrial variation is significantly less than expected were it evolving neutrally. A likely explanation of this is that the spread of the male-killer has resulted in a selective sweep of the host's mitochondria. Alternatively, this pattern may result from the spread of a beneficial mutation on either the mitochondria or the W sex chromosome. Also, selection against deleterious mutations may cause the loss of variation by background selection (CHARLESWORTH *et al.* 1993). The effects of background selection are likely to be greater on the combination of Wolbachia, mitochondria, and the W chromosome than on autosomal genes because of the lack of recombination in the former.

Fortunately, data from the second species, *A. encedon*,

allowed us to differentiate between these hypotheses and identify the male-killer as the cause of the reduction in mitochondrial diversity. In *A. encedon*, the mitotype associated with the male-killer has not replaced the ancestral mitotypes in the uninfected females. This allows us to compare the evolution of mitochondria with and without a male-killer. The uninfected population has relatively diverse and geographically differentiated mitochondria. However, infected females lack both inter- and intrapopulation diversity. Therefore, male-killing bacteria are the likely cause of the loss of mtDNA diversity in these butterflies.

There are two reasons why male-killers reduce mtDNA diversity (JOHNSTONE and HURST 1996). First, the initial invasion of Wolbachia, and the subsequent spread of beneficial parasite genes, will result in selective sweeps of mtDNA. Second, at equilibrium the effective population size of mtDNA will be reduced, because mutations in uninfected females will tend to be lost. In *A. encedana* the second effect will be small, as virtually all the females are infected (96%). This allowed us to show that the effects of the biased population sex ratio on the effective population size of nuclear genes will tend to be far greater than the reduction in the effective population size of mitochondria. Therefore, the reduction in mtDNA diversity relative to nuclear loci that we observed must be the result of a selective sweep.

Hybrid introgression: Phylogenetic analysis showed that *A. encedon* mitochondria are paraphyletic. There are two monophyletic groups in the mtDNA phylogeny; the first contains *A. encedana* and infected *A. encedon*, and the second contains uninfected *A. encedon*. This pattern contrasts with that found in three nuclear loci, all of which supported the two species as monophyletic groups. Balancing selection could maintain ancestral

polymorphisms in mtDNA after they have been lost from nuclear genes. However, this could not account for the total lack of variation between the two species in their mtDNA, while much slower evolving nuclear genes have numerous interspecific differences. Therefore, we can conclude that the male-killer has spread from *A. encedana* into *A. encedon* by hybrid introgression and that this has resulted in introgression of the *A. encedana* mitochondria into *A. encedon*.

This is particularly interesting as butterflies are female heterogametic and therefore contain a maternally transmitted W sex chromosome. Therefore, the male-killer will presumably also have introgressed any nuclear genes on this chromosome between the two species. This could be tested using genetic markers on the W chromosome.

The date of the selective sweep: The polymorphism data suggest that the last selective sweep in *A. encedana* occurred within the last 16,000 years. This could be either the date when the male-killer invaded the population or a more recent selective sweep due to positive selection acting on the bacterial genome. Either way, it must have occurred before the introgression of the male-killer into *A. encedon*, which has therefore spread 6000 km across sub-Saharan Africa (Sierra Leone to Tanzania) in this time.

The population dynamics of male-killers: Most of the male-killing bacteria that have been described are imperfectly transmitted from mother to offspring. This means that each generation infected females produce uninfected daughters. This prevents the male-killer spreading to fixation and maintains a stable polymorphism of infected and uninfected females (Hurst 1991).

A clear prediction of this hypothesis is that, at equilibrium, all of the uninfected females in the population will be descended through the maternal line from uninfected females. Therefore, the hypothesis predicts that uninfected females will have mtDNA descended from infected females. In the case of *A. encedana* this prediction is met. However, in *A. encedon* we can reject the hypothesis that the population is at a stable equilibrium maintained by imperfect transmission.

This difference between the two species can be explained by measurements of the Ug male-killer's transmission efficiency in wild females. In *A. encedon*, transmission is perfect, and all the offspring produced by infected females are themselves infected (Jiggins *et al.* 2002b). In contrast, the transmission efficiency is significantly lower in *A. encedana*, where 4% of the offspring of infected females are uninfected (Jiggins *et al.* 2000a). Another factor that may be important in maintaining the ancestral mtDNA in uninfected females is the observation that the mtDNA and W chromosome associated with the Ug male-killer in *A. encedon* have a hybrid origin. Therefore, it may carry some fitness cost and be re-moved from the uninfected population by natural selection.

These factors do not provide a mechanism to maintain a stable polymorphism of infected and uninfected females in *A. encedon*, and we are left with the question as to why the male-killer has not spread to fixation. This is clearly an important question, as, if all the uninfected females had been lost, then the species would have gone extinct. The first hypothesis that can explain this paradox is if the system is not yet at equilibrium, and in the future either the species will go extinct or an equilibrium will be reached that is maintained by very low levels of imperfect transmission (too small to be detected in our sample). This seems unlikely as throughout sub-Saharan Africa the infection is at an extremely high prevalence and uninfected females have diverse mtDNA. Intuitively, it appears improbable that in all these populations the male-killer is just about to reach fixation, especially as biased sex ratios were first recorded >500 generations ago (Poulton 1914).

Alternatively, balancing selection may be maintaining the uninfected females in the population. Two potential sources of balancing selection have been tested, and both have been rejected. First, costly suppressors of the male-killer's transmission or phenotype can maintain a stable polymorphism of infected and uninfected females (Randerson *et al.* 2000b). However, an extensive search in natural populations failed to provide any evidence of either type of suppressor (Jiggins *et al.* 2002b). Second, if males choose to mate with uninfected females rather than infected ones, this also results in a stable polymorphism (Randerson *et al.* 2000a). However, there is no evidence of mate choice in natural populations (Jiggins *et al.* 2002b). I suggest two alternative hypotheses. First, uninfected females occur in areas containing more males and therefore have a higher reproductive success when males are scarce. This could occur if siblings mate or at the population level. Various forms of such group selection have been modeled for cytoplasmic male sterility in plants (McCauley and Taylor 1997) and sex chromosome meiotic drive in animals (Heuch 1978; Heuch and Chanter 1982). Alternatively, environmental or ecological differences may mean the selection favors infected cytoplasm in some populations and uninfected cytoplasm in others. Migration between these areas may result in the observed polymorphism.

In conclusion maternally transmitted parasites can reduce intraspecific polymorphism and cause interspecific introgression of mtDNA. Because maternally transmitted parasites are extremely common in arthropods (Werren *et al.* 1995; Jiggins *et al.* 2001a), this may confound many studies that use mtDNA to make inferences about the evolutionary history of populations and the relationships of species. Mitochondria can be a useful tool for understanding the population genetics of these parasites.

LITERATURE CITED

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

BALLARD, J. W. O., 2000b When one is not enough: introgression of mitochondrial DNA in *Drosophila*. Mol. Biol. Evol. **17:** 1126–1130.

BALLARD, J. W. O., J. HATZIDAKIS, T. L. KARR and M. KRIETMAN, 1996 Reduced variation in *Drosophila simulans* mitochondrial DNA. Genetics **144:** 1519–1528.

CHARLESWORTH, B., M. T. MORGAN and D. CHARLESWORTH, 1993 The effect of deleterious mutations on neutral genetic variation. Genetics **134:** 1289–1303.

FOUTS, D., R. GANGULY, A. G. GUTIERREZ, J. C. LUCCHESI and J. E. MANNING, 1988 Nucleotide-sequence of the Drosophila glucose-6-phosphate-dehydrogenase gene and comparison with the homologous human-gene. Gene **63:** 261–275.

FU, Y.-X., and W.-H. LI, 1993 Statistical tests of neutrality of mutations. Genetics **133:** 693–709.

GOLDMAN, N., J. P. ANDERSON and A. G. RODRIGO, 2000 Likelihood-based tests of topologies in phylogenetics. Syst. Biol. **49:** 652–670.

HASEGAWA, M., H. KISHINO and T. YANO, 1985 Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. **21:** 160–174.

HEUCH, I., 1978 Maintenance of butterfly populations with all-female broods under recurrent extinction and recolonization. J. Theor. Biol. **75:** 115–122.

HEUCH, I., and D. O. CHANTER, 1982 The persistence of abnormal sex ratios in the African butterfly *Acraea encedon*. Oikos **38:** 228–233.

HILLIS, D. M., and M. T. DIXON, 1991 Ribosomal DNA: molecular evolution and phylogenetic inference. Q. Rev. Biol. **66:** 411–453.

HSU, R., A. D. BRISCOE, B. S. W. CHANG and N. E. PIERCE, 2001 Molecular evolution of a long wavelength-sensitive opsin in mimetic Heliconius butterflies (Lepidoptera: Nymphalidae). Biol. J. Linn. Soc. **72:** 435–449.

HUDSON, R. R., M. KREITMAN and M. AGUADÉ, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

HURST, G. D. D., F. M. JIGGINS, J. H. G. SCHULENBERG, D. BERTRAND, S. A. WEST *et al.*, 1999 Male-killing *Wolbachia* in two species of insect. Proc. R. Soc. Lond. Ser. B **266:** 735–740.

HURST, L. D., 1991 The incidences and evolution of cytoplasmic male killers. Proc. R. Soc. Lond. Ser. B **244:** 91–99.

JIGGINS, F. M., G. D. D. HURST and M. E. N. MAJERUS, 1998 Sex ratio distortion in *Acraea encedon* (Lepidoptera: Nymphalidae) is caused by a male-killing bacterium. Heredity **81:** 87–91.

JIGGINS, F. M., G. D. D. HURST, C. E. DOLMAN and M. E. N. MAJERUS, 2000a High prevalence male-killing *Wolbachia* in the butterfly *Acraea encedana*. J. Evol. Biol. **13:** 495–501.

JIGGINS, F. M., G. D. D. HURST and M. E. N. MAJERUS, 2000b Sex ratio distorting *Wolbachia* cause sex role reversal in their butterfly hosts. Proc. R. Soc. Lond. Ser. B **267:** 69–73.

JIGGINS, F. M., J. K. BENTLEY, M. E. N. MAJERUS and G. D. D. HURST, 2001a How many species are infected with Wolbachia?: cryptic sex ratio distorters revealed to be common by intensive sampling. Proc. R. Soc. Lond. Ser. B **268:** 1123–1126.

JIGGINS, F. M., G. D. D. HURST, J. H. G. VON DER SCHULENBURG and M. E. N. MAJERUS, 2001b Two male-killing *Wolbachia* strains coexist within a population of the butterfly *Acraea encedon*. Heredity **86:** 161–166.

JIGGINS, F. M., J. K. BENTLEY, M. E. N. MAJERUS and G. D. D. HURST, 2002a Recent changes in phenotype and patterns of host specialisation in Wolbachia bacteria. Mol. Ecol. **11:** 1275–1283.

JIGGINS, F. M., J. P. RANDERSON, G. D. D. HURST and M. E. N. MAJERUS, 2002b How can sex ratio distorters reach extreme prevalences? A natural population in which male-killing *Wolbachia* have near-perfect vertical transmission efficiency and are not suppressed. Evolution **56:** 2290–2295.

JOHNSTONE, R. A., and G. D. D. HURST, 1996 Maternally inherited male-killing microorganisms may confound interpretation of mtDNA variation in insects. Biol. J. Linn. Soc. **53:** 453–470.

JUKES, T. H., and C. R. CANTOR, 1969 Evolution of protein molecules, pp. 21–132 in *Mammalian Protein Metabolism*, edited by H. N. MUNRO. Academic Press, New York.

McCAULEY, D. E., and D. R. TAYLOR, 1997 Local population structure and sex ratio: evolution in gynodioecious plants. Am. Nat. **150:** 406–419.

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.

POULTON, E. B., 1914 W. A. Lamborn's breeding experiments upon Acraea encedon (Linn.) in the Lagos district of West Africa, 1910–1912. J. Linn. Soc. Lond. **32:** 391–416.

RAMBAUT, A., 2001 *SHTests*. Department of Zoology, University of Oxford, Oxford.

RANDERSON, J. P., F. M. JIGGINS and L. D. HURST, 2000a Male-killing can select for male mate choice: a novel solution to the paradox of the lek. Proc. R. Soc. Lond. Ser. B **267:** 867–874.

RANDERSON, J. P., N. G. C. SMITH and L. D. HURST, 2000b The evolutionary dynamics of male-killers and their hosts. Heredity **84:** 152–160.

RICH, S. M., M. C. LIGHT, R. R. HUDSON and F. J. AYALA, 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.

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

SHIMODAIRA, H., and M. HASEGAWA, 1999 Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. **16:** 1114–1116.

SIMON, C., F. FRATI, A. BECKENBACH, B. CRESPI, H. LIU *et al.*, 1994 Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. **87:** 651–701.

SOTOADAMES, F. N., H. M. ROBERTSON and S. H. BERLOCHER, 1994 Phylogenetic utility of partial DNA-sequences of G6pdh at different taxonomic levels in Hexapoda with emphasis on Diptera. Ann. Entomol. Soc. Am. **87:** 723–736.

SWOFFORD, D. L., 1998 *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sinauer Associates, Sunderland, MA.

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

TAMURA, K., 1992 The rate and pattern of nucleotide substitution in Drosophila mitochondrial DNA. Mol. Biol. Evol. **9:** 814–825.

TURELLI, M., and A. A. HOFFMANN, 1995 Cytoplasmic incompatability in *Drosophila simulans*: dynamics and parameter estimates from natural populations. Genetics **140:** 1319–1338.

TURELLI, M., A. A. HOFFMANN and S. W. McKECHNIE, 1992 Dynamics of cytoplasmic incompatibility and mtDNA variation in natural *Drosophila simulans* populations. Genetics **132:** 713–723.

SCHULENBURG, J. H. G. V. D., G. D. D. HURST, D. TETZLAFF, G. E. BOOTH, I. A. ZAKHAROV *et al.*, 2002 History of infection with different male-killing bacteria in the two-spot ladybird beetle *Adalia bipunctata* revealed through mitochondrial DNA sequence analysis. Genetics **160:** 1075–1086.

WERREN, J. H., D. WINDSOR and L. R. GUO, 1995 Distribution of *Wolbachia* among neotropical arthropods. Proc. R. Soc. Lond. Ser. B **262:** 197–204.

ZYLSTRA, P., H. S. ROTHENFLUH, G. F. WEILLER, R. V. BLANDEN and E. J. STEELE, 1998 PCR amplification of murine immunoglobulin germline V genes: strategies for minimization of recombination artefacts. Immunol. Cell Biol. **76:** 395–405.

Communicating editor: D. CHARLESWORTH