

## DIVERGENCE OF MITOCHONDRIAL DNA IS NOT CORROBORATED BY NUCLEAR DNA, MORPHOLOGY, OR BEHAVIOR IN *DROSOPHILA SIMULANS*

J. WILLIAM O. BALLARD,<sup>1,2</sup> BARRY CHERNOFF,<sup>3</sup> AND AVIS C. JAMES<sup>1</sup>

<sup>1</sup>Department of Biological Sciences, University of Iowa, Iowa City, Iowa 52242

<sup>2</sup>E-mail: bill-ballard@uiowa.edu

<sup>3</sup>Field Museum, 1400 South Lake Shore Drive, Chicago, Illinois 60605-2496

**Abstract.**—We ask whether the observed mitochondrial DNA (mtDNA) population subdivision of *Drosophila simulans* is indicative of organismal structure or of specific processes acting on the mitochondrial genome. Factors either intrinsic or extrinsic to the host genome may influence the evolutionary dynamics of mtDNA. Potential intrinsic factors include adaptation of the mitochondrial genome and of nucleomitochondrial gene complexes specific to the local environment. An extrinsic force that has been shown to influence mtDNA evolution in invertebrates is the bacterial endosymbiont *Wolbachia*. Evidence presented in this study suggests that mtDNA is not a good indicator of organismal subdivision in *D. simulans*. Furthermore, there is no evidence to suggest that *Wolbachia* causes any reduction in nuclear gene flow in this species. The observed differentiation in mtDNA is not corroborated by data from *NADH: ubiquinone reductase 75kD subunit precursor* or the *Alcohol dehydrogenase-related* loci, from the shape or size of the male genital arch, or from assortative premating behavior. We discuss these results in relation to a mitochondrial genetic species concept and the potential for *Wolbachia*-induced incompatibility to be a mechanism of speciation in insects. We conclude with an iterated appeal to include phylogenetic and statistical tests of neutrality as a supplement to phylogenetic and population genetic analyses when using mtDNA as an evolutionary marker.

**Key words.**—Behavior, *Drosophila simulans*, mitochondrial variability, morphometrics, species concepts, *Wolbachia*.

Received July 6, 2001. Accepted November 12, 2001.

Investigation of selective forces that influence the dynamics of distinct biological processes has been a major research focus of evolutionary biologists for more than a century. The focus of research shifted to mitochondrial DNA (mtDNA) in the early 1970s because the maternal, generally nonrecombining mode of inheritance often provides multiple alleles or haplotypes that can be ordered into gene genealogies interpretable as the matriarchal component of the organisms' history. MtDNA haplotypes and clades within many species have proven to be geographically localized (Avise et al. 1987). Factors either intrinsic or extrinsic to the host genome may cause the evolutionary dynamics of mtDNA to deviate from neutral expectations. Potential intrinsic factors include adaptation of the mitochondrial genome and adaptation of specific nucleomitochondrial gene complexes to the local environment (Ballard and Dean 2001). *Wolbachia* is an example of an extrinsic force that has been shown to influence mtDNA evolution in invertebrates in theory (Caspari and Watson 1959), population cages (Nigro and Prout 1990), and natural populations (Hale and Hoffmann 1990; Turelli and Hoffmann 1991, 1995). *Wolbachia* is a group of maternally inherited intracellular alpha-proteobacteria that have been found in a number of invertebrate species and can have dramatic affects on reproduction in their hosts (Werren 1997; Bourtzis and O'Neill 1998). Most insects, including *Drosophila*, can be cured of *Wolbachia* infection with tetracycline (O'Neill and Karr 1990).

A primary goal of the study is to test whether mtDNA divergence can be employed to define species genetically. Moore (1997, p. 627) stated "there are very few circumstances where the mt-haplotype tree would not be the best bet for the species tree." Bradley and Baker (2001) examined cytochrome *b* sequences from four genera of rodents and seven genera of bats to investigate the robustness of a mi-

tochondrial species concept. The authors suggested that in rodents and bats genetic distances of less than 2% were indicative of intraspecific variation, values between 2% and 11% had a high probability of being conspecific populations of genetical species, and values greater than 11% were indicative of separate species. This thinking is reflected dramatically in the conclusions concerning the species status of the Indonesian coelacanth, *Latimeria menadoensis*. Pouyaud et al. (1999) based the original description of *L. menadoensis* on "significant differences" in morphology and sequence divergence from its Comoros Island relative, *L. chalumnae*. Holder et al. (1999), however, found that significant morphological differences were based on inadequate comparisons of morphology by Pouyaud et al. (1999). Nonetheless, Holder et al. (1999) concluded that a 4.1% sequence divergence found within 4823 bp of mitochondrial DNA constitutes substantial divergence and, therefore, is indicative of separate species. However, it is clear that mitochondrial divergence is a poor equivalency of taxonomic rank. Johns and Avise (1998) investigated the sequence divergence of the cytochrome *b* gene across the vertebrates and noted that avian species have been placed into separate genera at a lower level of genetic divergence than is often true for other vertebrates. One explanation for this result is that birds have experienced an elevated rate of morphological evolution.

It is difficult to empirically investigate the processes that result in species formation (but see Perez et al. 1993; Ting et al. 2000). Instead, we compare the mtDNA divergence among populations of a single species. Specifically, we contrast the genetic structure in each of the three well-defined, reciprocally monophyletic mtDNA haplotype groups of *D. simulans* (*siI*, *siII*, and *siIII*; Solignac et al. 1986; Satta et al. 1987; Baba-Aissa et al. 1988; Ballard 2000a) with data from two nuclear loci, morphology, and behavior. The genetic di-

vergence of these three haplotypes is 2.0–2.6%, but it should be noted that these three haplotypes are not monophyletic relative to *D. sechellia* and *D. mauritiana mal* (Solignac and Monnerot 1986; Ballard 2000c). If mtDNA divergence is a general process that is diagnostic of species differentiation, then it may be expected to be diagnostic of population differentiation. Ballard (2000c) estimated the divergence time of the *D. simulans* haplotypes to be around 1.75 million years. Despite this age, few mutations have accumulated within any of the mtDNA lineages. The incongruity of high divergence among haplotypes and low diversity within each of them remains unexplained. One hypothesis is that *Wolbachia*-induced selective sweeps have reduced the variation within but not among haplotypes (Ballard 2000a) by a process of genetic hitchhiking (Maynard Smith and Haigh 1974).

At this time we are aware of a single study that specifically includes nuclear data from each of the three mtDNA haplotypes of *D. simulans*. Ballard (2000a) compared the nucleotide variation among 22 complete mitochondrial genomes of the three distinct *D. simulans* haplotypes with intron 1 of the *alcohol dehydrogenase-repeated* (*Adhr*) locus. This autosomal region was sequenced because it has considerable variation and the pattern is consistent with a neutrally evolving locus (Sumner 1991). Patterns of variation suggest that distinct forces are influencing the evolution of mtDNA and *Adhr*. However, it is possible that the gene trees are not concordant because of selection, introgression, poor resolution of the data, or stochastic lineage sorting (Avise 1987). Moore (1995) showed that when the effective population size of mtDNA is one-quarter that of nuclear-autosomal loci, the mtDNA genealogy has a higher probability of tracking the species tree in fewer generations. In a reply to Moore (1995), Hoelzer (1997) provided multiple examples in which nuclear genes will coalesce more quickly than mitochondrial genes when the operational sex ratio is seven females to one male. It is known that X-linked drivers are present in *D. simulans* (Atlan et al. 1997; Cazemajor et al. 1997), but autosomal and Y-linked suppressors of drive are also abundant and thus may not affect the sex ratio in natural populations (Montchamp-Moreau et al. 2001). There are no good estimates of the variance in the reproductive success of *D. simulans* males in nature, although in *D. melanogaster* larger males have greater mating success (Partridge et al. 1987).

Nuclear variation in *D. simulans* was further investigated by sequencing a portion of the *NADH: ubiquinone reductase 75kD subunit precursor* (*ND75*). This nuclear gene is located on the X chromosome and encodes a mitochondrial polypeptide that forms part of the NADH complex in the inner mitochondrial membrane. This locus was selected to bias the result in favor of finding an association between the mitochondrial and autosomal genealogies. If no association is found, we suggest the result is conservative. The divergence between the specific *D. simulans* mitochondrial genes with which *ND75* interacts is in the range 2.6–2.9%.

Morphological variation was investigated by examining the size and shape of the male genital arch. The consistent morphological differences between members of the *D. melanogaster* subgroup include the shape of the posterior process of the male genital arch and the number of bristles on the sex comb (Coyne 1983; True et al. 1997). Furthermore,

differences in male genitalia are often considered premating isolation factors, the so-called lock-and-key concept. Coyne (1993) suggested that the short duration of mating between *D. simulans* females and *D. mauritiana* males, probably caused by a mismatch of the genitalia, reduced gene flow among these species.

Behavioral variation was also investigated by testing whether there is random mating among isofemale lines of each *D. simulans* mtDNA haplotype. *Drosophila simulans* is reproductively isolated from closely related species but produces some fertile female hybrids when mated with either *D. mauritiana* or *D. sechellia* (Coyne and Charlesworth 1989). Although the genetic basis of reproductive isolation within this clade has been studied more extensively than in any other group of species (Coyne et al. 1998), the extreme genetic differentiation observed in the mtDNA genome leads us to question whether the *D. simulans* haplotypes were randomly mating. Wu et al. (1995) observed strong sexual isolation among highly differentiated populations of *D. melanogaster* from Zimbabwe and the United States (Begin and Aquadro 1993).

The second primary aim of this study was to investigate whether there is any evidence to support the hypothesis that *Wolbachia*-induced incompatibility can induce pangenomic divergence. Bordenstein et al. (2001) suggested that *Wolbachia*-induced cytoplasmic incompatibility between the parasitic wasps *Nasonia giraulti* and *N. longicornis* occurred before the evolution of other postmating isolating mechanisms, including hybrid inviability and sterility. This led the authors to speculate that *Wolbachia*-induced incompatibility may be an important mechanism of speciation in invertebrates, although they note that allopatric divergence may be important in the case of *N. giraulti* and *N. longicornis*. In *D. simulans*, cytoplasmic incompatibility may be expressed when a male harboring a strain of *Wolbachia* mates with a female that does not carry that same bacterial strain or is uninfected. Sperm enters the egg normally, but defects in fertilization cause a reduction in egg hatchability (Hoffmann et al. 1986; Lassy and Karr 1996).

*Drosophila simulans* is known to be infected by multiple strains of *Wolbachia* (*wAu*, *wHa*, *wMa*, *wNo*, and *wRi*) that may cause bidirectional, unidirectional, or no significant incompatibility (O'Neill and Karr 1990; Hoffmann et al. 1996; Merçot and Poinsot 1998; James and Ballard 2000; James et al. 2002). Each of the *Wolbachia* strains appears to be associated with a specific mtDNA haplotype in *D. simulans*. The *siL* haplotype, found on Indian and Pacific Ocean islands, may be infected with *wHa* and/or *wNo*. The *siII* haplotype has a global distribution, but is allopatric with the *siL* haplotype. Flies with the *siII* haplotype may be infected with either the *wAu* or the *wRi* *Wolbachia* strains. The *siIII* haplotype has been collected in Reunion, Madagascar, and East Africa and may be infected with the *wMa* strain of *Wolbachia*. In this study, we characterize *Wolbachia* genotypes using 16S rDNA and a *Wolbachia* surface protein (*wsp*) coding gene. Flies not infected with *Wolbachia* have been found in most populations studied. Ancestors of these uninfected flies may never have been infected or the infection may have been lost (Turelli and Hoffmann 1995; Ballard et al. 1996).

In both *D. simulans* and *Nasonia*, there is *Wolbachia*-in-

duced postmatting incompatibility, but it is not clear if it significantly inhibits gene flow in either case. The prediction is that, if all else is equal, taxa with the highest bidirectional incompatibility may be expected to show the least gene flow. Overall, there is higher incompatibility between *D. simulans* *siI wHa-* and *siII wRi*-infected individuals than between *N. giraulti* and *N. longicornis* (James and Ballard 2000; Bordenstein et al. 2001). Here, we investigate the potential for a reduction in nuclear gene flow among *siI wHa-* and *siII wRi*-infected individuals.

This study exemplifies the strategies that can be employed to investigate hypotheses of genetic subdivision inferred from a single locus (Faith and Trueman 2001). Specifically, the genetic subdivision observed in *D. simulans* mtDNA is not corroborated by nuclear DNA, morphology of the genital arch, or mating behavior. These data suggest that distinct evolutionary forces are acting on the mitochondrial and nuclear genomes and that it is invalid to infer organismal divergence from mtDNA subdivision. We also show that *Wolbachia*-induced bidirectional cytoplasmic incompatibility is not an important mechanism of pangenomic divergence in *D. simulans*.

## MATERIALS AND METHODS

### *Fly Lines*

Twenty-two isofemale lines of *D. simulans* are included (Ballard 2000a). Six isofemale lines of *siI*, seven of *siII*, and nine of *siIII* were included to maximize geographic and mitochondrial diversity; these lines are not a random sample of *D. simulans*.

For the sequencing studies we employ two outgroup isofemale lines of *D. melanogaster*. The Oregon R line was collected in Roseburg, Oregon, by D. E. Lancefield. The Zimbabwe 53 line was collected in Sengwa, Zimbabwe, by R. R. Ramey and L. Brown.

### *Wolbachia Strains*

It is important to know which lines are infected with which strains of *Wolbachia* so that specific tests can be conducted concerning the potential for *Wolbachia*-induced reduction in nuclear gene flow. Total genomic DNA was isolated from three whole adult flies from each line using the fixed tissue protocol from Gentra's (Minneapolis, MN) PureGene Kit. The 16S rDNA was amplified following O'Neill et al. (1992) and *wsp* following Zhou et al. (1998). Both strands were sequenced, and each *D. simulans* infected line was characterized as being singly infected with *wAu*, *wHa*, *wMa*, *wNo*, or *wRi* or doubly infected with *wHa* + *wNo* (James and Ballard 2000). If the line appeared to be uninfected by *Wolbachia*, the procedure was repeated twice. Positive controls were employed to test whether the extracted DNA was amplifiable.

### *Drosophila Sequence Data*

We sequenced a portion of *ND75*. This nuclear gene, which encodes a mitochondrial polypeptide, is located on the X chromosome at map position 7D20–22. To conduct specific statistical tests of neutrality we compare the *ND75* data with

TABLE 1. Primers employed to sequence *ND75*. Primer numbers refer to the 3'-most base from the *Drosophila melanogaster* cDNA (Caggese et al. 1999).

		5' → 3' sequence
Amplification primers	X45+	CAG GAA TAT TTT CCA GCA TTT CG
	X1968–	TCG TCC AGG TTG TCG TAC GGC AG
Sequencing primers	X317+	GGC AAC TGC AGG ATG TGC CTC
	X612+	CGC TGC ATC CAC TGC ACC CG
	X1118+	ACT TGC CGA CCT GGA GGC TC
	X1520+	CCA TCA TTA TTG GAG CCG AT
	X297–	GAG GCA CAT CCT GCA GTT GCC
	X635–	TCT CGG AGG CGA AAC GCA CA
	X1150–	CAG ACG GTT CAG CAG ATC CT
	X1501–	ATC GGC TCC AAT AAT GAT GG

that previously collected from intron 1 of *Adh*. To contrast the genealogies and to map the *Wolbachia* strains onto each topology, we compare the *ND75* data with that previously collected from the mtDNA. We do not include the mtDNA data in the tests of neutrality because Ballard (2000a) has previously shown a significant difference among *Adhr* and mtDNA.

DNA was isolated from single whole male flies as described above. Polymerase chain reaction (PCR) amplification primers (Table 1) were designed from the *D. melanogaster* messenger RNA sequence in GenBank (Caggese et al. 1999). Internal sequencing primers were designed from single-stranded *D. simulans* sequences by primer walking from each end. Amplifications were performed using the following thermal profile: 95°C for 2 min, followed by 35 cycles of 95°C for 30 sec, 60°C for 20 sec, 72°C for 150 sec. Both strands were sequenced using Taq-Big Dye Deoxy Terminator Cycle sequencing (Applied Biosystems, Foster City, CA) using 75–100 ng of template DNA and 20 ng of primer. Sequences were imported into Sequencher (Gene Codes, Ann Arbor, MI), where they were checked and the contig for each line constructed.

### *Statistical analyses*

Nucleotide diversity ( $\pi$ ) and the neutral parameter ( $\theta$ ) based on the number of segregating sites were calculated for *D. melanogaster* and *D. simulans*. Nucleotide diversity (within each species or haplotype) and divergence (between species) were calculated for the total sequence and for silent sites using DnaSP 3.5 (Rozas and Rozas 1997).

Tajima's *D* (Tajima 1989), Fu and Li's *D\** (Fu and Li 1993), and Fu's *Fs* (Fu 1997) test whether the observed substitution patterns in the data are consistent with a neutral model of molecular evolution. Tajima's *D* tests whether there is a significant difference in estimates of  $\theta$  derived from the number of segregating sites. The average heterozygosity is measured by  $\pi$ . Fu and Li's *D\** test statistic is based on differences between the number of singletons and the total number of mutations. Fu's *Fs* tends to be positive when there is an excess of old mutations; it is the most powerful of the three tests for detecting hitchhiking when there is no recombination (Fu 1997). These tests assume that each sample is

taken from a single randomly mating population. This assumption is violated in this study, because approximately equal numbers of each haplotype were included and because of geographical biases. However, data presented here suggest that there is no association between the mtDNA genotypes and those derived from nuclear DNA. As a consequence, these lines may be considered to be an effectively random sample of *D. simulans* nuclear DNA.

To investigate protein evolution, the MK test of neutrality (McDonald and Kreitman 1991) was used to compare the number of nonsynonymous to synonymous substitutions in *ND75*. The MK test predicts that the ratio of nonsynonymous to synonymous differences should be equal within and among species. The index of neutrality (NI) summarizes the substitution pattern (Rand and Kann 1996) and reflects the extent to which the levels of amino acid variation within species depart from the strictly neutral model. Strict neutrality has an index of 1.0. Values greater than 1.0 typically indicate an excess of amino acid variation within species, whereas values less than 1.0 indicate an excess of nonsynonymous substitutions among species, relative to the number of synonymous substitutions.

The HKA test (Hudson et al. 1987) was used to compare the levels of silent and synonymous variation in *ND75* and *Adhr*. This is a conservative test of an equilibrium neutral model's prediction that polymorphism within species and divergence between species will be positively correlated. It assumes that each sample is taken from a single randomly mating population. For each test, the effective population size of *ND75* (X chromosome) was considered to be three-quarters that of *Adhr* (autosomal). The first test considers the polymorphism in the 22 lines of *D. simulans* with the divergence from the two lines of *D. melanogaster*. These results should be treated with some caution because the total sample is clearly not taken from a panmictic population. The three subsequent tests consider the polymorphism within each haplotype (*siI*, *siII*, and *siIII*, respectively) with the divergence from *D. melanogaster*.

#### Trees from sequence data

It may be predicted that the *ND75*, *Adhr*, and mtDNA genealogies will not be correlated if they are independently segregating in a randomly mating population or if they are evolving under distinct evolutionary processes. As an alternative, we consider that distinct mtDNA haplotypes exist in populations that have distinct evolutionary trajectories. In this latter case, we expect that the autosomal and mitochondrial data would be correlated. To investigate whether the different datasets are significantly associated, the incongruence length difference (ILD) test (Farris et al. 1995) was employed to test the null hypothesis that each biological partition is evolving under homogeneous biological processes. In each case, 1000 randomized partitions of the same size were employed to test significance. This random-partitioning test is an extension of a measure originally reported by Mickovich and Farris (1981) and is based on the null hypothesis of congruence. Cunningham (1997) suggested that the ILD was the most useful test to examine the relationship between incongruence and phylogenetic accuracy, defined as the de-

gree to which each data partition supported the expected tree calculated using the percent clade correct index (Hillis et al. 1994). The ILD test distinguished between cases in which combining the data generally improved phylogenetic accuracy ( $P > 0.01$ ) and cases in which accuracy of the combined data suffered relative to the individual partitions ( $P < 0.001$ ).

The genealogical relationship inferred from *ND75* and those previously published from the mitochondrial genome of *D. simulans* (Ballard 2000a) were analyzed by maximum parsimony using PAUP\* 4.0b8 (Swofford 1998). The conclusions drawn from these analyses were corroborated by maximum-likelihood and neighbor-joining analyses. Bootstrapping was used to test monophyly (Efron 1982; Felsenstein 1985). For this study, 1000 pseudosamples were generated to estimate the bootstrap proportions.

#### Morphology

The male genital arch was examined because it is known to differ among members of the *D. melanogaster* subgroup. In *Drosophila*, the size and shape of the male genitalia appear to be controlled by many genes scattered throughout the genome (Coyne 1983, 1985; Coyne and Kreitman 1986; Liu et al. 1996).

Flies were raised at constant temperature (25°C) and density (30 individuals per vial, conditions in which there is low larval crowding). One to two days after eclosion, males were collected and preserved in 70% alcohol. The genital arch of six males from each line was dissected in saline and mounted in Permount (Fisher Scientific, Pittsburgh, PA). Each slide was numbered and given blindly to B. Chernoff such that the numbering reflected replicates but he was unaware of the genealogical relationships.

Images were captured at identical magnifications (140.8 times), thus preserving relative size (Fig. 1A). The images were modified in Photoshop 5.5 (Adobe, San Jose, CA) to eliminate cilia. A diagonal standard line was drawn to signify the bottom of the genital arch (Fig. 1B). In the digitization program TPSDIG 1.20 (Rohlf 1998), images were first smoothed. One hundred evenly spaced coordinates beginning at the tip of the hook were generated with the outline tool (Fig. 1C). No bias in the direction or position of error or in digitization was discovered following repeated measurements of multiple genital arches.

#### Size

The size of each genital arch was estimated as their centroid size (Bookstein 1991), which is the sum of the squared distances from each of the outline points to the center of the form (i.e., the mean  $x$  and mean  $y$ ). To remove the artifacts of orientation and position on the monitor as well as the effect of size, all arches were aligned to their consensus and scaled to a centroid size of one using the Procrustes module of NTSYSPC 2.10a (Rohlf 2000).

#### Shape

The strategy for the analysis of shape variation was based on the fact that, with the possible exception of the tip of the hook and the sides of the base of the arch, there are no

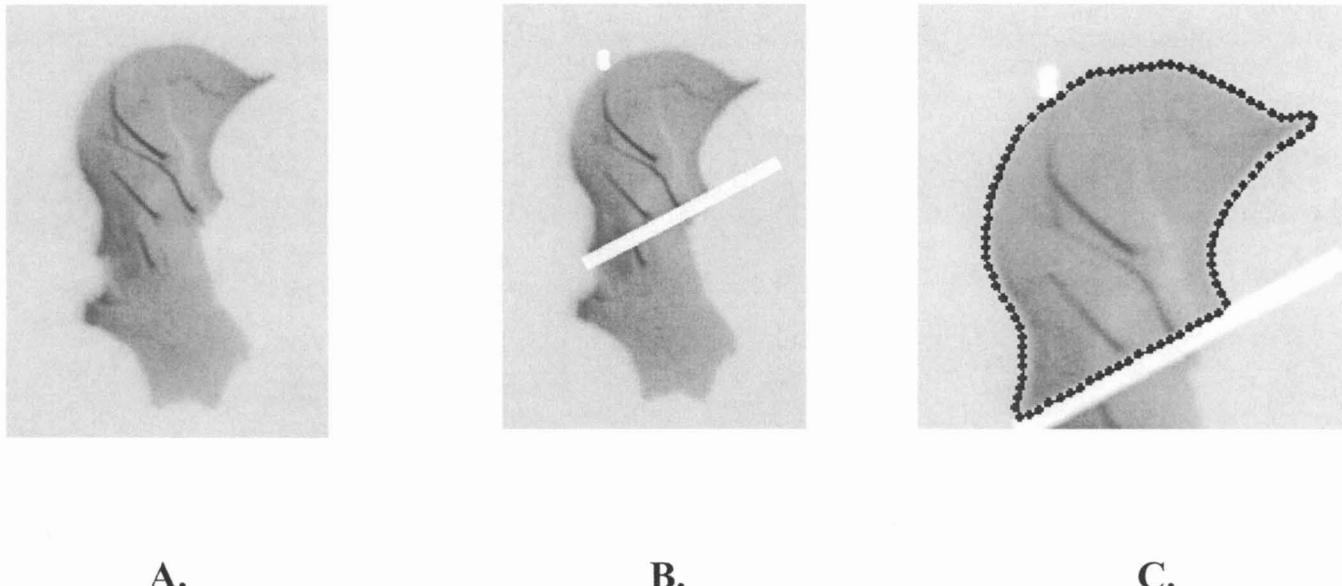


FIG. 1. Genital arch of *Drosophila simulans*. (A) Images of the specimens were captured through the Zeiss AxioCam digital camera at identical magnifications and  $x$ - and  $y$ -coordinate of dissected genital arches were generated for each specimen. (B) Cilia were eliminated in Adobe Photoshop 5.5 (note the two white circles that eliminate a cilium) and a diagonal standard line was drawn to signify the bottom of the genital helmet. (C) One hundred evenly spaced coordinates beginning at the tip of the hook, were generated with the outline tool. All damaged arches were discarded.

topological points that can unequivocally be construed as homologous. Our strategy involved a two-step process. First, the shape of each specimen was described by elliptical Fourier analysis (Rohlf and Archie 1984; Ferson et al. 1985). Second, a shape space was created from the principal components (PCs) computed from the correlation matrix of 78 Fourier coefficients. Eighteen PCs captured 91.2% of the variation in shape. The specimens represented by their Fourier descriptors (coefficients) were projected onto the eigenvectors to generate scores. The PC scores were plotted on combinations of eigenvectors as well as used for the statistical analyses described below. Elliptical Fourier and principal component analyses were carried out in NTSYSPC (Rohlf 2000).

#### Statistical analyses of morphological data

The analyses of size and shape data proceeded in a hierarchical manner. We began with the null hypothesis that the sizes (centroid size) and shapes (PC scores) of the genital arches of the isofemale lines comprise a single statistical population. Eigenvectors are orthogonal and independent of all others, so they were subjected to analysis of variance (ANOVA), analysis of covariance (ANCOVA), or multivariate analysis of variance (MANOVA).

#### Classification

We tested group structure for the a posteriori classifications of the morphological data against the group structures inherent in mtDNA and nuclear DNA data using mixture models (McLachlan and Basford 1988; McLachlan and Jones 1988; Ballard et al. 1991; McLachlan and Krishnan 1997; McLachlan and Peel 1998; Hunt and Chapman 2001). Mix-

ture models evaluate univariately or multivariately in a series of hierarchical hypotheses about the likelihood of more than a single statistical group being present in the data. A log-likelihood-ratio test is computed for each of the competing hypotheses, and bootstrapping, with 1000 iterations, tests the  $\chi^2$  statistic (McLachlan and Basford 1988). These analyses were carried out with EMMIX software (available via <http://www.maths.uq.edu.au/~gjm/>). Once the maximum-likelihood solution for the number of statistical populations in the data was known, the probability of assigning each individual into each of the size or shape populations was calculated. We then used  $G$ -tests to test the independence.

#### Trees from morphology

Unrooted minimum evolution networks among the isofemale lines were calculated from generalized distance matrices. Generalized distance, also known as Mahalanobis distance, is the Euclidean distance among the centroids of the isofemale lines in the full space of the canonical variates of the PC scores. NTSYSPC 2.10a (Rohlf 2000) and PAUP\* 4.0b8 (Swofford 1998) were used to generate the generalized distances matrices and the minimum evolution networks, respectively. A  $t$ -test was used to test the null hypothesis that branches connecting to terminal taxa (i.e., isofemale lines) have the same mean length as branches connecting internal nodes. Because of the lack of independence within a network, the usual statistical tables for judging significance cannot be used. Therefore, we randomized the elements in the rows and columns of the generalized distance matrix 1000 times and recalculated a  $t$ -statistic for each iteration. The  $t$ -statistic calculated from the actual data was then compared to the distribution of  $t$ -statistics from the randomized data.

### Mating Behavior

The genetic differentiation observed in the mtDNA genome leads us to question whether the haplotypes were randomly mating. Multiple nuclear genes control premating behavior of *D. simulans*. Uenoyama and Inoue (1995) observed that the high compatibility of the S2 line of *D. simulans* with *D. melanogaster* is caused by at least two genes, one on the second and the other on the third chromosome.

Two sets of lines were used in replicate  $3 \times 3$  mating cage studies. Set A contained HW09, MD225, and MD199 and set B contained TT01, MD106, and RU07. These lines were chosen because they represent each haplotype, they were collected within two years of the study, and they were all infected with *Wolbachia* (see Results). For the *siII* haplotype, we employed flies infected with the *wAu* strain rather than *wRi* strain of *Wolbachia*. Our *wRi*-infected strains had been in laboratory culture for more than five years, and maintenance in captivity may affect life-history traits (Latter and Mulley 1995).

Flies were raised at constant temperature ( $25^{\circ}\text{C}$ ) and density (30 larvae per vial). Adults were collected as virgins and aged for four to six days. Fifty virgin males and females from each line (four to six days old) were placed on *Drosophila* food that had been dyed green, red, or violet. These colors were visible through their abdomen and thus marked each haplotype. To control the effect of food color, three cages from each set were run simultaneously with all line  $\times$  food color combinations. No effect of color was detected (results not presented). Two days after the virgins were placed on the colored medium, 300 virgins (50 of each sex by each of the three lines) were released into each  $22 \times 21 \times 36$  cm cage. For two hours, copulating pairs were aspirated into vials and then frozen. Each pair was then scored using the colored medium in their alimentary canal. Data were analyzed by *G*-tests.

There are two limitations to this study. First, with two sets of three lines there is limited power to distinguish haplotype from line effects. Second, four- to six-day-old virgin flies may not be particularly discriminatory. An alternative approach is to study remating behavior because the intensity of sexual selection of a nonvirgin female may be higher than a virgin female.

### Trees from mating behavior

The data from the mating studies was used to generate unrooted networks to examine the hypothesis that mating was not structured hierarchically by mitochondrial haplotype. Each of the  $3 \times 3$  mating experiments resulted in  $3 \times 3$  tables of mating occurrences. The latter was turned into a  $6 \times 6$  table by assigning males and females of each haplotype to both rows and columns. The total number of matings was used as the denominator from which to re-express mating occurrences as percentages. In the cells of male  $\times$  male or female  $\times$  female interactions a value of 0.0 was assigned. The diagonal was assigned a value of 1.0. Each value was subtracted from 1.0 to produce a dissimilarity (distance) matrix. A minimum evolution network was then calculated parsimoniously from the dissimilarity matrix in PAUP\* 4.0b (Swofford 1998) as unrooted, with all terminal taxa con-

strained to be monophyletic, using the heuristic search algorithm. In each case, only a single network was found.

## RESULTS

### *Wolbachia* Strains

Of the 22 *D. simulans* isofemale lines included in this study, 14 were infected with *Wolbachia*. Isofemale lines HW00, HW09, TT00, and TT01 were infected with *wHa* (GenBank X61769 16S rDNA and AF020068 *wsp*), and NC48 and NC37 were doubly infected with *wHa* + *wNo* (AF312372 16S rDNA and AF020074 *wsp*). C167, SC00, and DSR were infected with *wRi* (X61770 16S rDNA and AF020070 *wsp*), and MD106 and MD225 were infected with *wAu* (AF390865 16S rDNA and AF020067 *wsp*). RU07, MD112, and MD199 were infected with the *wMa* strain of *Wolbachia* (AF390864 16S rDNA and AF020069 *wsp*). The *D. melanogaster* lines were infected with the *wMel* strain of *Wolbachia* (Z28983 16S rDNA and AF020063 *wsp*).

### *Drosophila* Sequence Data

A total of 2353 bp from *ND75* was obtained for all 24 isofemale lines (GenBank AY074994–AY075017). The sequenced region can be divided into an upstream region (positions 1–213), four exons (214–396, 468–799, 868–1894, and 1958–2353), and three introns (397–467, 800–867, and 1895–1953). Three indel events were observed. A complex indel event occurred between nucleotide positions 1940 and 1944 and five bases were deleted prior to phylogenetic analysis. Single-base interspecific indels at positions 844 and 1929 were each scored as a single additional character.

Sequence data from the same 24 isofemale lines are reported for intron 1 of *Adhr* (GenBank 2AF201423–201424, AF201428–201449) and from the mtDNA genome (GenBank AF200828–200829, AF200833–200854; Ballard 2000a). All 501 sites of *Adhr* were readily aligned. In the mtDNA, 76 bp of 15,034 bp were deleted from all analyses because it was not possible to unequivocally determine the alignment (Ballard 2000a).

### Statistical analyses

Two standardized estimators of *ND75* nucleotide polymorphism—nucleotide diversity and the neutral parameter—indicate similar levels of nucleotide polymorphism for *D. melanogaster* and *D. simulans* (Table 2). Somewhat surprisingly, however, the *siIII* mtDNA haplotype has higher nucleotide diversity than either the *siII* or the *siI* mtDNA haplotypes (Table 2). Given the geographic distribution of the haplotypes, it is difficult to explain this result and it may be an artifact of the sampling strategy employed. Five of the nine *siIII* lines are from Madagascar. In contrast, just two of the *siII* lines are from Madagascar. The *siI* haplotype has not been collected on the island.

Within each haplotype, Tajima's *D* and Fu and Li's *D\** are negative, whereas Fu's *Fs* is positive (Table 2). None of the values suggest that there is a significant deviation from a neutral equilibrium model of evolution.

According to the MK test, there is no evidence of selection on nonsynonymous sites at *ND75*. Statistical tests of protein

TABLE 2. Variation at *ND75* of *Drosophila melanogaster* (*mel*) and *D. simulans* (*sim*). T, total sequence; S, silent sites. The 22 lines of *D. simulans* included in this study were not chosen at random.

	All	<i>mel</i>	<i>sim</i>	<i>siI</i>	<i>siII</i>	<i>siIII</i>
<i>N</i>	24	2	22	6	7	9
Segregating sites	181	33	121	38	54	100
Variants	18	2	16	5	5	8
Synonymous	114	22	77	24	34	63
Nonsynonymous	4	0	3	1	1	1
$\pi$ (T) <sup>1</sup>	—	.0141	.0120	.00656	.00829	.0143
$\pi$ (S)	—	.0368	.0310	.0168	.0214	.0373
$\theta$ (T) <sup>2</sup>	—	.0140	.0147	.00727	.00973	.0165
$\theta$ (S)	—	.0368	.0376	.00186	.0251	.0427
Tajima's <i>D</i>	—	—	-0.76	-0.63	-0.84	-0.67
Fu and Li's <i>D</i> *	—	—	-0.74	-0.55	-0.93	-0.59
Fu's <i>Fs</i>	—	—	0.84	1.53	3.12	1.20

<sup>1</sup> Nucleotide diversity (per site).

<sup>2</sup> Neutral parameter (per site).

evolution suggest that the null hypothesis of an equal ratio of nonsynonymous to synonymous substitutions within and among species cannot be rejected. There are 18 fixed and 96 polymorphic synonymous substitutions and one fixed and three polymorphic nonsynonymous changes. The neutrality index is 0.53 and the *G*-value with William's correction is 0.17 ( $P = 0.68$ ).

The HKA tests support the hypothesis that *ND75* and *Adhr* are evolving in a manner consistent with a neutral equilibrium model of evolution. HKA tests show there is no significant difference between the levels of polymorphism in *ND75* and intron 1 of *Adhr* in *D. simulans* or *D. simulans* haplotypes relative to their divergence from *D. melanogaster* (Table 3). These data support the tenet that there is a significant reduction in mtDNA diversity in each *D. simulans* haplotype relative to its divergence with *D. melanogaster* (Ballard 2000a).

#### Trees from sequence data

Phylogenetic analyses of two nuclear genes do not support the mtDNA genetic subdivision. As such, we suggest there is no evidence of organismal subdivision in *D. simulans*. Sequence data from *ND75*, *Adhr*, and mtDNA were analyzed

independently because the ILD test rejected the null hypothesis that the three regions are evolving under a single homogeneous biological process (Bull et al. 1993). The sum of lengths for the original partition for the three datasets was 1418 steps; for *ND75* and the mtDNA data 1327 steps; for *ND75* and *Adhr* 368 steps; and for mtDNA and *Adhr* 1141 steps. In all cases this is significantly shorter than the sum of lengths of 1000 randomized partitions of the same size ( $P < 0.001$ ). There is no evidence of organismal subdivision in *D. simulans* lines infected with strains of *Wolbachia* that induce bidirectional incompatibility, and the results are not presented ( $P < 0.001$ ). *ND75* was analyzed as a single region because the ILD test did not reject the null hypothesis that introns and exons are evolving under a single homogeneous process (267 steps,  $P = 0.15$ ). One 501-bp intron of *Adhr* was considered. The mitochondrial genome was also considered as a single region because there is no evidence of recombination (Ballard 2000b). The increase in the number of steps in each partition compared with the number of steps determined from a constrained analysis is presented in Table 4.

There is no evidence of an association of *Wolbachia* with *ND75* (Fig. 2A) or *Adhr* (Ballard 2000a). In contrast, there is a clear association of symbiont strains with mtDNA hap-

TABLE 3. HKA test comparing the intraspecific silent and synonymous polymorphism in *Drosophila simulans* to the interspecific divergence from *D. melanogaster*. The neutral parameter ( $\theta$ ) is defined as  $4Nm$  for autosomal loci of diploid organisms, where  $N$  is the effective population size (diploid individuals) and  $m$  is the neutral mutation rate (per gene or per base pair) per generation. Assuming equal population sizes of males and females,  $\theta$  is assumed to be  $3Nm$  for X-linked loci of diploid organisms.

	<i>D. simulans</i>		<i>siI</i>		<i>siII</i>		<i>siIII</i>	
	<i>ND75</i>	<i>Adhr</i>	<i>ND75</i>	<i>Adhr</i>	<i>ND75</i>	<i>Adhr</i>	<i>ND75</i>	<i>Adhr</i>
<b>Polymorphism data</b>								
No. segregating sites (o)	123	39	38	17	55	27	104	9
No. segregating sites (e)	122.23	39.77	39.03	15.97	57.64	24.36	95.88	17.12
Total number of sites†	896.4	459	897.39	477	896.31	477	896.18	461
Sample size	22	22	6	6	7	7	9	9
<b>Divergence data</b>								
Mean no. differences (o)	74.91	21.23	75	21.67	74.93	21.85	75.50	20.44
Mean no. differences (e)	75.68	20.46	73.97	22.70	72.29	24.49	83.62	12.32
Total no. of sites†	896.54	443	897.39	460	896.31	460	896.18	445
$\chi^2\ddagger$		0.08		0.03		0.12		2.27

† Excludes sites with alignment gaps.

‡ No value is significant.

TABLE 4. Increase in the number of steps in each partition compared with the number of steps determined from a constrained analysis. In this case, the constraint was the 70% bootstrap topology in that partition.

Partition <sup>1</sup>	70% bootstrap constraint			
	ND75	Adhr	mtDNA	All data
ND75	0	+77	+96	+94
Adhr	+35	0	+24	+26
mtDNA	+1358	+1078	0	+2
All data	+1346	+1059	+1	0

<sup>1</sup> The number of parsimony-informative positions in ND75 was 112, in Adhr 42, in the mtDNA genome 884, and in the combined data 1038.

lotypes (Fig. 2B). The *siI* haplotype may be infected by *wHa* and/or *wNo*, the *siII* haplotype by *wRi* or *wAu*, and the *siIII* mtDNA haplotype may be infected with *wMa*. Flies not infected with *Wolbachia* have been collected in most populations and in all haplotypes.

### Morphology

The size and shape of the genital arch differs as much within isofemale lines as it does among lines, and there are

no significant differences in the size or shape of the genital arch among the three mtDNA haplotypes. However, isofemale lines that are *siIII* and infected with *Wolbachia* have significantly larger genital arches than do *siIII* *Wolbachia*-uninfected lines. There is no evidence to suggest that bidirectionally incompatible *D. simulans* lines (*siI*, *wHa/wNo* and *siII*, *wRi*) have significantly different size or shape of the genital arch, and the results for these lines are not presented separately.

### Size

Two populations of sizes exist in the data that have little fidelity with respect to isofemale lines but are a function of mitochondrial haplotype and *Wolbachia* infection status. The mean values of centroid sizes vary considerably among the isofemale lines from 573.5 units<sup>2</sup> to 672.0 units<sup>2</sup> (Fig. 3). The differences among isofemale lines are significant (ANOVA,  $F_{1,21} = 4.28$ ,  $P < 0.001$ ). The isofemale lines are almost bimodal in genital arch sizes with 12 lines in the large group, seven in the small, and three somewhat intermediate (Fig. 3A). Isofemale lines DSW and MD199 are the most variable, with standard deviations  $\geq 8.0\%$  of their mean values.

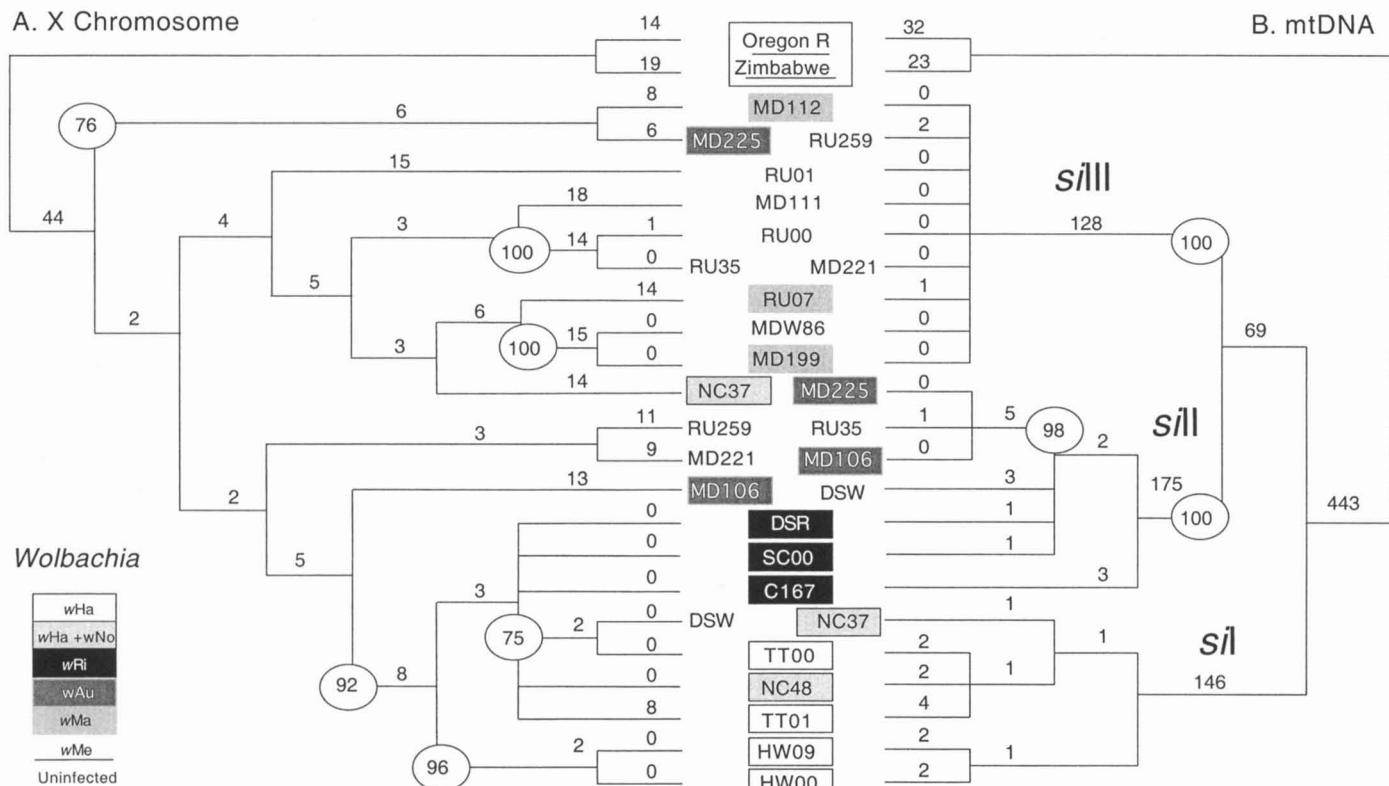


FIG. 2. Strict consensus trees generated from the mitochondrial and autosomal data. Taxa are placed in the center aisle. The two taxa in the box are lines of *Drosophila melanogaster* and are the designated outgroup. If a single taxon is present in a given row, it applies to both the ND75 and mitochondrial datasets. Where two taxa are in a single row, the left taxon refers to the autosomal dataset and the right taxon to the mitochondrial dataset. Both datasets were bootstrapped 1000 times and the resulting proportions ( $>70\%$ ) are shown in circles. The number of substitutions is shown above each branch. (A) Strict consensus trees generated from ND75. A total of 2348 bp are included: 2167 characters are constant, 69 variable characters are parsimony uninformative, and 112 are parsimony informative. Three equally parsimonious trees of length 277 steps were found. These trees have a consistency index of 0.68. (B) Mitochondrial data. A total of the 14,958 bp are included: 14,023 characters are constant, 51 variable characters are parsimony uninformative, and 884 parsimony informative. Five equally parsimonious trees of length 1050 steps were found. These trees have a consistency index of 0.92.

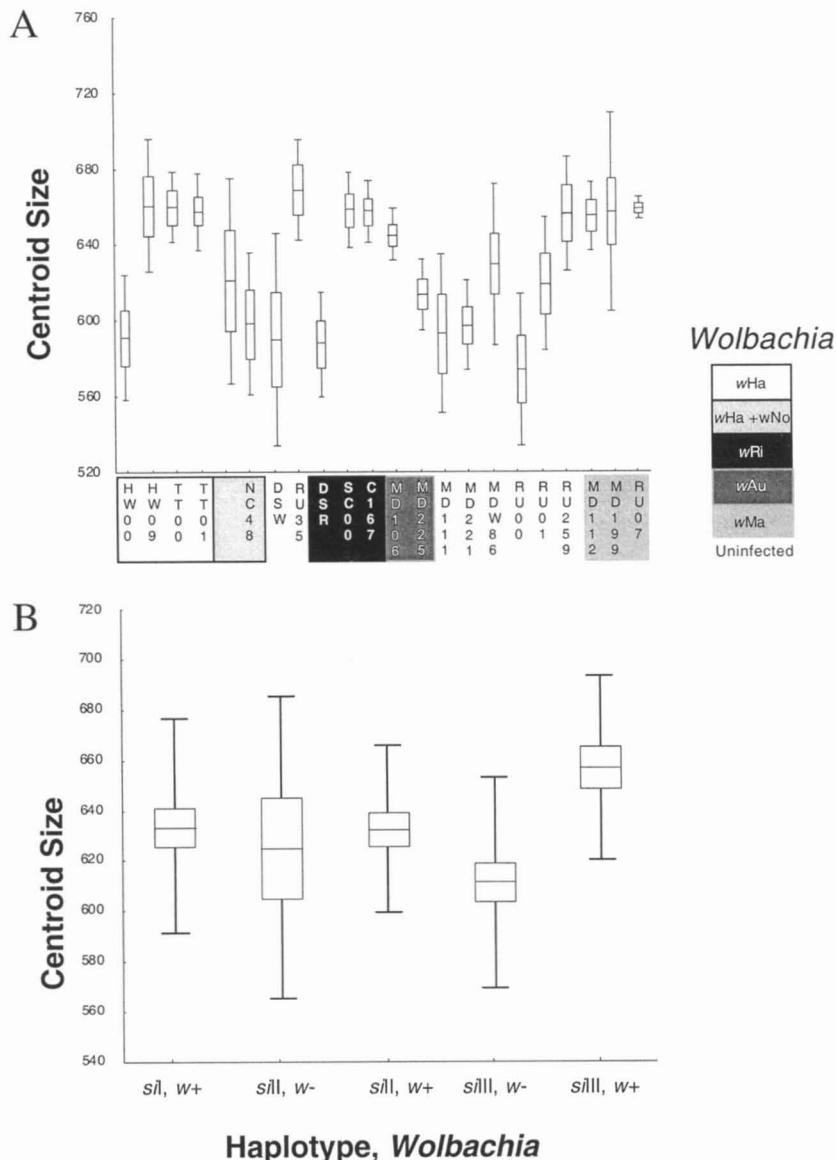


FIG. 3. (A) Box plots of centroid size by isofemale line. (B) Centroid size by mitochondrial haplotypes, with and without *Wolbachia* infection. For each group, the mean (central horizontal line) plus or minus one standard error (box) and one standard deviation (terminal horizontal lines) is shown.

Significant differences in size (Fig. 3B) were associated with haplotype and *Wolbachia* infection ( $F_{1,4} = 3.74$ ,  $P < 0.001$ ). Tukey's honest significant difference post hoc procedure shows that the ANOVA results are due to the significant difference ( $P < 0.003$ ) between *siIII* lines that are infected with *Wolbachia* (656.9 units<sup>2</sup>) or uninfected (611.2 units<sup>2</sup>).

#### Shape

Genital arch shapes fell into two groups that had little relation to isofemale line, mitochondrial haplotype, or infection by *Wolbachia*. The shape variation among scatters of PC scores revealed no obvious groupings. However, different isofemale lines do occupy different portions of the morphospace (e.g., Fig. 4).

The third PC (variance = 11.4%) captures an allometric effect (Fig. 5): Larger genital arches are more arched than smaller ones. Because we fail to reject the null hypothesis of homogeneity of within-group slopes ( $P = 0.69$ ), the significant difference among mean shapes of isofemale lines is independent of size (ANCOVA,  $F_{20,90} = 2.2$ ,  $P < 0.01$ ). The ANCOVA results also fail to reject the null hypothesis of homogeneity among the mitochondrial haplotype–infection groups ( $F_{4,103} = 2.01$ ,  $P = 0.07$ ).

Differences in shape among isofemale lines are not associated with either the mitochondrial haplotype or the presence of *Wolbachia*. A MANOVA on the remaining 77 PC scores (i.e., not including PC3) rejects the null hypothesis of homogeneity among the isofemale lines (Wilks' lambda,  $F_{51428,632} = 1.176$ ,  $P < 0.01$ ; Hotelling-Lawley trace,

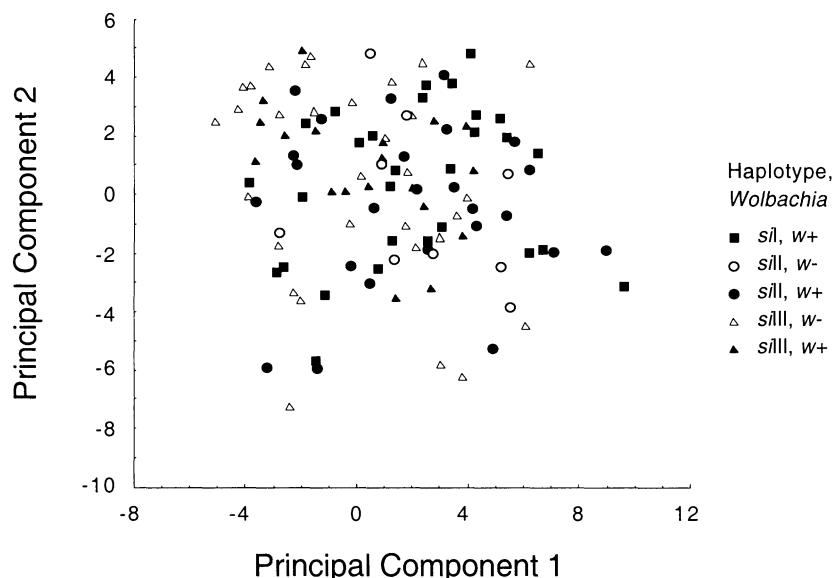


FIG. 4. Scatter of individual scores on the first two principal components of the elliptical Fourier descriptors. Symbols: *siI* (boxes), *siII*, infected by *Wolbachia* (shaded), uninfected (open).

$F_{S1428,464} = 1.232$ ,  $P < 0.005$ ). A subsequent MANOVA found no differences in mean shape among the mitochondrial-infection groups (Wilks' lambda,  $F_{S280,159} = 0.954$ ,  $P = 0.64$ ; Hotelling-Lawley trace,  $F_{S280,150} = 0.940$ ,  $P = 0.67$ ).

#### Classification

The classifications of size and shape into maximum-likelihood groups were independent of each other and of mtDNA haplotypes and isofemale lines. We compared the maximum-likelihood classifications for size and shape against the three mtDNA haplotypes. The results of  $G$ -tests (Table 5) fail to reject the null hypotheses of independence between the haplotype and either the size ( $P = 0.69$ ) or shape ( $P = 0.96$ ) classifications. Because haplotypes *siII* and *siIII* form a monophyletic group, we compared the size and shape clas-

sifications with a two-haplotype designation (*siI* and *siII + siIII*). The values of  $G$  are 1.27 and 0.005 for size and shape, respectively, and are less than the  $\chi^2$  critical value of 3.841 ( $P = 0.05$ ). We also failed to reject the null hypothesis that the size classification is independent of the shape classification ( $P = 0.36$ ).

We explored the relationship of *Wolbachia* infection to the maximum-likelihood classifications for size and shape by  $G$ -tests (Table 6). Following Williams' adjustment of the  $G$ -statistic (Sokal and Rohlf 1995) and sequential Bonferroni correction (Rice 1989), the null hypothesis of independence is rejected (Table 6). This is due to the larger genital arches in *Wolbachia*-infected *siIII* lines. In contrast, we conclude that the mitochondrial-infection classification is independent of the two-group shape classification (Table 6) after sequential Bonferroni correction ( $P = 0.26$ ).

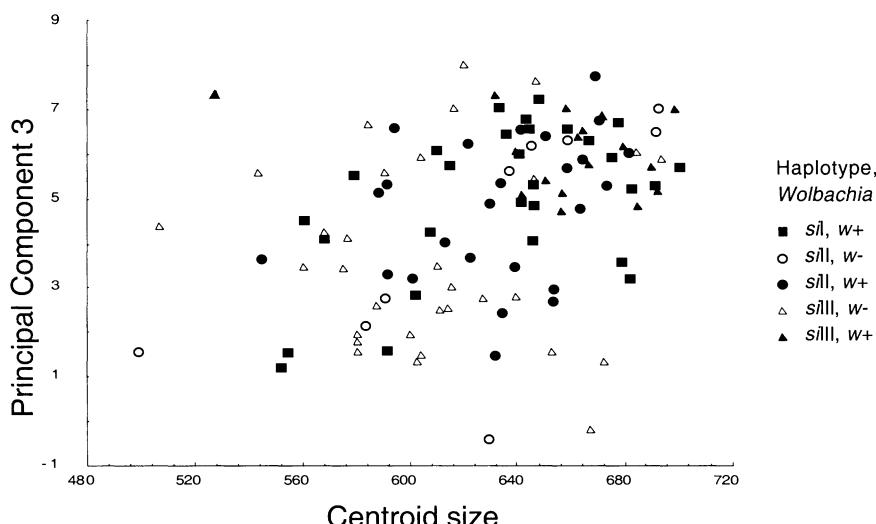


FIG. 5. Scatter of individual scores on the third principal component of elliptical Fourier descriptors versus values of centroid size. Symbols: *siI* (boxes), *siII* (circles), *siIII* (triangles), infected by *Wolbachia* (shaded), uninfected (open).

TABLE 5. *G*-tests of independence of mitochondrial haplotypes from two-group size and shape classifications.

Haplotype	Size		Shape	
	1	2	1	2
siI	10	19	13	16
siII	19	24	16	18
siIII	15	26	22	28
<i>G</i>		0.83 <sup>†</sup>		0.08 <sup>†</sup>

<sup>†</sup> The critical value of  $\chi^2$  at  $P = 0.05$  is 5.991. Because *G* is less than 5.991, the Williams' adjustment was not used because it lowers the value of *G*. Because  $P > 0.05$ , the sequential Bonferroni correction was not used because it raises the value of *P* toward 1.0.

### Trees from morphology

The lack of congruence among the gene trees and the size or shape networks suggests that there is no consistent morphological subdivision among the isofemale lines. The unrooted minimum evolution trees generated from generalized

TABLE 6. *G*-tests of independence for five haplotype  $\times$  *Wolbachia* infection groups against two-group size and shape classifications, respectively.

Haplotype, <i>Wolbachia</i> <sup>†</sup>	Size		Shape	
	1	2	1	2
siI, <i>w</i> +	10	19	13	16
siII, <i>w</i> -	4	5	5	4
siII, <i>w</i> +	11	14	11	14
siIII, <i>w</i> -	22	9	19	12
siIII, <i>w</i> +	2	17	3	16
<i>G</i> <sub>adj</sub>			19.83**	10.72

<sup>†</sup> *w*+ and *w*- indicate the presence and absence of *Wolbachia* infection, respectively.

\*\*  $P < 0.01$ ; *P*-values are corrected by the sequential Bonferroni procedure.

distances among the isofemale lines for size and shape do not convey the hierarchical structure evident in the mitochondrial haplotype tree or the group structures present in the nuclear genes (Figs. 2, 6, 7). On the size tree, there are

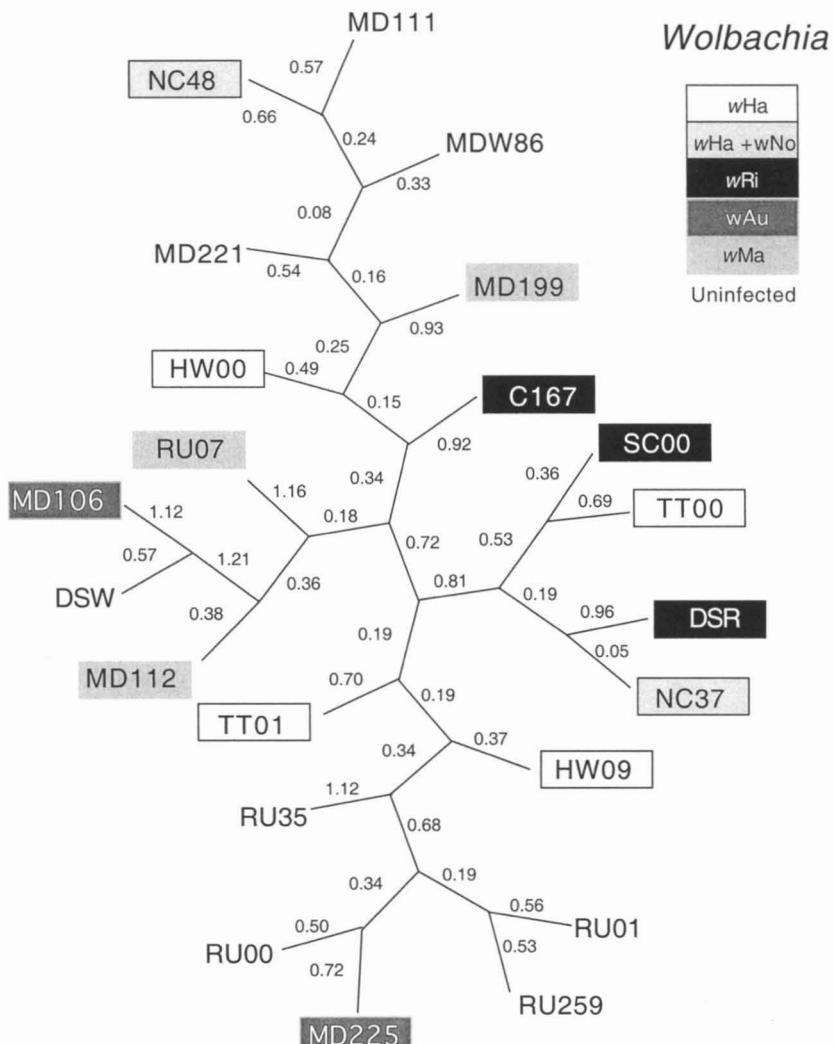


FIG. 6. Unrooted minimum evolution network of Euclidean distance matrix among isofemale lines calculated from centroid sizes. Branch lengths are shown by each branch. The isofemale lines contained within boxes (shaded or not) are those infected with *Wolbachia*; minimum evolution score = 20.60.

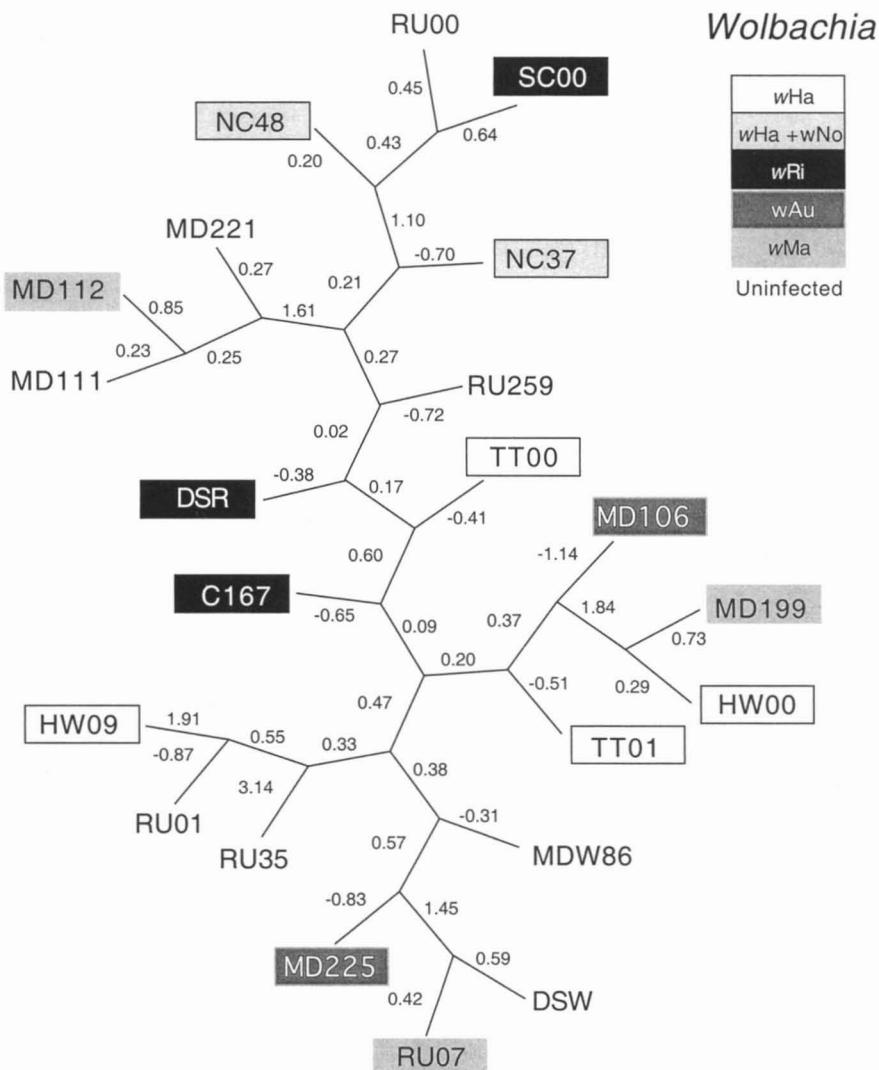


FIG. 7. Unrooted minimum evolution network of generalized distance matrix among isofemale lines calculated from principal component scores of shape data (elliptical Fourier descriptors). Branch lengths are shown by each branch. The isofemale lines contained within boxes (shaded or not) are those infected with *Wolbachia*; minimum evolution score = 21.34.

TABLE 7. Number of mating pairs by mitochondrial haplotype and *Wolbachia* infection.

Set A

Male haplotype	Female haplotype		
	HW09 siI, wHa	MD225 siII, wAu	MD199 siIII, wMa
HW09 siI, wHa	18	34	24
MD225, siII, wAu	5	12	7
MD199 siIII, wMa	9	20	23

Set B

Male haplotype	Female haplotype		
	TT01 siI, wHa	MD106 siII, wAu	RU07 siIII, wMa
TT01 siI, wHa	20	12	9
MD106 siII, wAu	19	16	8
RU07 siIII, wMa	21	9	12

two potential groupings of isofemale lines that possess haplotype *siIII*, although each group contains a single line with a different haplotype. On the shape tree a relatively small group (MD112, MD111, MD221) possesses a single haplotype (*siIII*). The short internodes and longer external branches, however, cast doubt the branching topology of the networks because they approach unresolved polychotomies.

#### Mating Behavior

There is no evidence of assortative mating among the *D. simulans* mitochondrial haplotypes. Approximately one-sixth of the flies mated during our experimental period, 152 and 126 mating pairs of 900 on each day, respectively (300 in each of three cages). The haplotype of each fly that mated was unambiguously identifiable using color marking.

The total numbers of matings, by female and male haplotype, are presented in Table 7. There was no evidence of assortative mating (set A:  $G_4^2 = 2.84$ ,  $P = 0.59$ ; set B:  $G_4^2 = 2.93$ ,  $P = 0.57$ ).

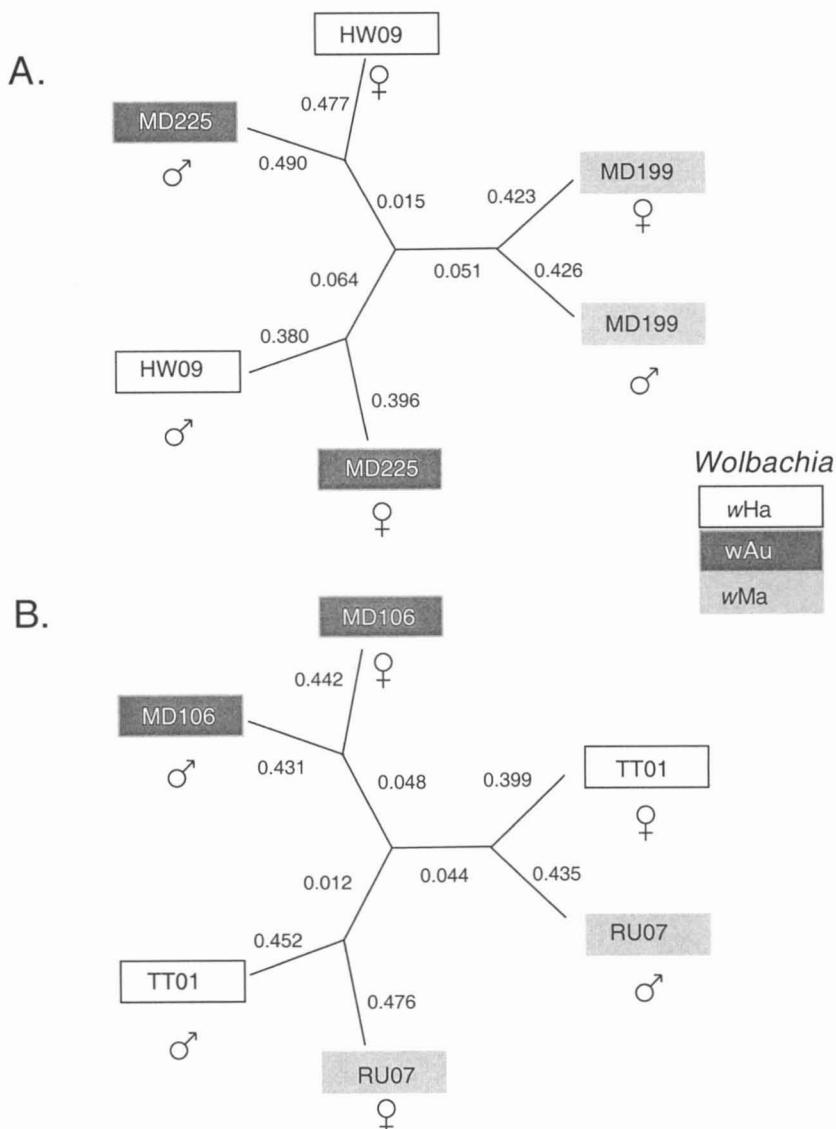


FIG. 8. Unrooted minimum evolution network of the mating data. Branch lengths are shown by each branch. **Set A:** minimum evolution score = 2.722. **Set B:** minimum evolution score = 2.738.

#### Tree from mating behavior

Unrooted minimum evolution networks of the mating data graphically show that there is no evidence for assortative mating among *D. simulans* haplotypes (Fig. 8). In set A, males and females of MD199 (*siIII*) were more likely to mate with each other than with males and females of either HW09 (*siI*) or MD225 (*siII*; Fig. 8A). In set B, males and females of MD106 (*siII*) were more likely to mate with each other than with males and females of TT01 (*siI*) or RU07 (*siIII*; Fig. 8B). Inspection of the branch lengths in both replicates, however, shows that the terminal branches are an order of magnitude longer than the internal branches. These results suggest that both networks approach unresolved polychotomies.

#### DISCUSSION

Mitochondrial genes have been widely used in population studies both for convenience and owing to the general belief

that gene frequencies are governed primarily by migration and genetic drift and that most of the variation within a species is selectively neutral. However, factors other than genetic drift are expected to be important determinants governing the fate of mutations. Recent studies have suggested that mtDNA may recombine in some cases (Saville et al. 1998; Awadalla et al. 1999; Eyre-Walker et al. 1999; Hagelberg et al. 1999; Kajander et al. 2000), however, the lack of normal recombination in mitochondria means that each genome has a single genealogical history and all genes will share that history. Any evolutionary force acting at any one site will equally affect the history of the whole molecule. Thus, the fixation of an advantageous mutation by selection, for example, will cause the fixation of all other polymorphisms by a process known as genetic hitchhiking (Maynard Smith and Haigh 1974). Also, the removal of deleterious mutations may depress linked silent and synonymous variation via background selection (Charlesworth et al. 1993, 1995; Charlesworth

1994). Even the quickly evolving noncoding origin of replication region cannot be assumed to have neutral allele frequencies because of its linkage to the rest of the genome.

We now consider our results in the light of two central questions: (1) Is mtDNA subdivision indicative of genome wide differentiation? (2) Is *Wolbachia*-induced incompatibility a mechanism of divergence or speciation in *D. simulans*? To further study differentiation across the genome, we investigate the processes that may affect the evolution of mtDNA independently from the rest of the genome and question whether divergence of mtDNA should be used as the sole basis for recognition of species.

#### *Is Differentiation of Mitochondrial DNA Indicative of Differentiation across the Genome?*

Many authors have assumed explicitly or implicitly that differentiation within a character system is indicative of organismal differentiation or history. From the standpoint of mtDNA, this has focused attention on whether mtDNA differentiation is indicative of species trees or gene trees. The former represents cases in which differentiation in mtDNA corroborates differences in other aspects of the genome. Overall, our data lead us to caution the extrapolation of mtDNA divergence to conclusions about organismal differentiation without additional support from independent data.

Data presented here show that the distinct mitochondrial subdivision observed in the mtDNA of *D. simulans* is not well corroborated by nuclear DNA sequences, genital arch morphology, or mating behavior. The nuclear loci failed to demonstrate patterns consistent with mtDNA haplotypes or with each other. The genital arches display significant size and shape differences, but these are independent of each other as well as patterns produced by other aspects of the genome. The only significant morphological difference that was associated with mtDNA involved the differences within *siIII* for lines that did or did not carry *Wolbachia*. The mating behavior fails to show fidelity among haplotypes.

In other studies there has been little or ambiguous corroboration between mtDNA and other aspects of the genome. In studies of the rat snake, *Elaphe obsoleta*, Burbrink et al. (2000) found that both mtDNA and morphology demonstrated strong patterns of geographic differentiation in North America. However, the geographic patterns were not congruent. In an intriguing study, Lee (2000) showed that, although mating incompatibilities were found among populations of the copepod *Eurytemora affinis* with distinct 16S rRNA and cytochrome oxidase I mitochondrial haplotypes, these were not congruent with population differences based upon morphology.

However, there have been many studies in which there is corroboration across differentiated parts of the genome. DeSalle and Giddings (1986) suggested that mtDNA appears to be sensitive in establishing the sequence of evolutionary events responsible for the present geographic distribution of the drosophilids *D. differens*, *D. planitibia*, *D. silvestris*, and *D. heteroneura*. Brust (1998) found that divergence in the cytochrome oxidase II region of mtDNA corroborated results from the ITS region of the rDNA, from male genitalia, and from hybrid crossing in the mosquito *Aedes wardangensis*.

Collin (2000) found that sequence divergence of mtDNA corroborated pigmentation and life-history differences between sympatric species of slipper limpets, *Crepidula atrasolea*, *C. depressa*, and *C. plana*.

#### *What processes may affect the mitochondrial DNA evolution independently from the rest of the genome?*

Mitochondrial variation may not correlate with other aspects of the genome because of selection, introgression, poor resolution of the data, or lineage sorting (Avise 1987). In independent and complementary reviews, Nachman (1998) and Rand and Kann (1998) tested whether mtDNA is evolving in a manner consistent with a neutral equilibrium model in a broad range of taxa. Both studies investigated the published data and screened for datasets that included multiple individuals from single species and sufficient intraspecific variation so that statistical tests could be employed. The results presented in both studies suggest that slightly deleterious mutations accumulate within species but do not go to fixation among them. These findings suggest that the patterns first reported for *Drosophila* (Ballard and Kreitman 1994; Rand et al. 1994) and house mice (Nachman et al. 1994) are not unique to these taxa but extend to species that are not commensal with humans. Hey (1997) suggested that mitochondrial variation in humans has been shaped by natural selection and may not be ideal for some questions concerning the genetic structure of populations including the origin of modern humans.

The *simulans* clade of the *D. melanogaster* subgroup illustrates the ambiguity caused by the introgression of mtDNA and the retention of ancestral DNA. The mtDNA data implies that *D. mauritiana* is paraphyletic relative to *D. simulans* and *D. sechellia* and that *D. simulans* is paraphyletic to *D. sechellia* (Solignac and Monnerot 1986; Satta et al. 1987; Ballard 2000b). First, it is likely that there has been introgression of *D. simulans* *siIII* mtDNA into *D. mauritiana* (Ballard 2000c). Second, it seems likely there has been retention of ancestral mtDNA polymorphism in *D. sechellia* or there was an ancient introgression event. In contrast, the overall picture gathered from multiple nuclear genes is of two allopatric speciation events that occurred quite near one another in time. Kliman et al. (2000) investigated the origins and divergence of *D. simulans*, *D. mauritiana*, and *D. sechellia* using the patterns of DNA sequence variation at 14 different genes. *Drosophila sechellia* appears to have had a reduced effective population size for some time and is accumulating slightly deleterious mutations as a result. *Drosophila simulans* and *D. mauritiana* are both highly polymorphic and the two species share many polymorphisms, probably since the time of common ancestry.

Distinct selective forces may be acting on the mitochondrial and nuclear genomes. Ballard (2000a) applied a comparative genomics approach to investigate the nucleotide variation within and among mitochondrial and autosomal DNA of *D. simulans*. Overall, at least three lines of evidence suggested that significantly different forces are influencing the evolution of *D. simulans* mtDNA and nuclear autosomal DNA. First, the mtDNA haplotypes are geographically subdivided, whereas the genealogy of *Adhr* clearly indicates ex-

tensive gene flow between the geographic regions. Second, there is a significant deficiency of mitochondrial variation in each *D. simulans* haplotype relative to the autosomal variation. Third, the ratio of synonymous to nonsynonymous substitutions is not equal in all branches of the well-resolved phylogeny. The obvious question that must be addressed, then, is what evolutionary mechanism can influence the mtDNA so dramatically but yet have no detectable effect on mating behavior, morphology, or nuclear genealogies. Adaptation of the mitochondrial genome and of specific nucleomitochondrial gene complexes to the local environment may influence the evolution of mtDNA. One extrinsic force that has been shown to influence mtDNA evolution in natural populations is *Wolbachia* (Turelli and Hoffmann 1991, 1995).

Dissimilar evolutionary forces may be acting on the mitochondrial genome and on morphology. Schneider et al. (1999) compared mitochondrial and morphological divergence in eight populations of the widespread leaf-litter skink, *Carlia rubrigularis*, to investigate the relative importance of natural selection and geographic isolation in generating phenotypic diversity in this species. There is deep mtDNA divergence (~12%) on either side of the Black Mountain Corridor, but there are many shared haplotypes among samples from paired rainforest and open-forest sites (Joseph et al. 1995). In contrast, morphology is remarkably consistent among populations on either side of the range but differs among rainforest and open forest. The authors hypothesize that the morphological differences may be driven by a strong selection of lizard-eating avian predators in the open forest.

#### *Should divergence of mitochondrial DNA be used as the sole basis for recognition of species?*

Some recent studies regard mtDNA divergence as a sufficient basis for differentiation of species (e.g., the case with the Indonesian coelacanth; Holder et al. 1999). Our results shed important light on this topic. A complete discussion of species concepts is beyond the scope of this paper, but many researchers recognize this area to be contentious (Paterson 1985; Coyne and Orr 1998; Filchak et al. 2000; Goldstein and DeSalle 2000).

From the standpoint of biological concepts (Dobzhansky 1935; Mayr 1942) or specific mate recognition (Paterson 1985), the data suggest that *D. simulans* is a single species. The differentiation within the mtDNA is not indicative of reproductive incompatibility and would, therefore, not contribute to species recognition. The behavioral data fail to show mating preferences among the tested lines. The large variation in the size and shape of the genital arch, apparently, is insufficient to prevent random mating. There also is no evidence to suggest a reduction in nuclear gene flow among the distinct mtDNA haplotypes.

From the viewpoint of a phylogenetic species concept, the most strongly patterned and hierarchical signals are displayed within the mtDNA (Fig. 2). Indeed, a parsimony analysis of all the DNA data strongly supports the monophyly of the three mtDNA haplotypes (three equally most parsimonious trees of 1509 steps, CI = 0.81). Using tree-based, unique character diagnosable (Goldstein and DeSalle 2000) or minimum evolutionary unit (Moritz 1994) concepts, researchers

would recognize three species or evolutionary units, each with a separate mitochondrial haplotype. From this perspective, researchers would argue that the lack of reproductive isolation is retention of primitive (ancestral) features and therefore is uninformative (Rosen 1979, pp. 275–278). Nonetheless, the example of *D. simulans* also shows that the terminal taxa based on the three haplotypes and their relationships to each other are not based on the phylogenetic history within *D. simulans* sensu lato. The hierarchical structure and history is correlated with *Wolbachia* infection, not phylogeny (James and Ballard 2000). Thus, the apparent evidence for organismal subdivision of the three *D. simulans* haplotypes inferred from mtDNA sequence data is not so improbable given alternate explanations that do not assume species status (Faith and Trueman 2001).

The different species concepts have very different implications for the interpretation of the data generated for this study of *D. simulans*. Nonetheless, there is a commonality present in the concepts that species are independent lineages; whether for biological species vis-à-vis reproduction or for phylogenetic species vis-à-vis characters. Whatever concept, it is inherent that when used as terminal taxa in a phylogenetic or phylogeographic analysis, the “species” must reflect organismal history. Thus, use of only mtDNA as a basis for species definition, or any single character system for that matter, would fail in its mission to the extent that organismal history is not embodied in that character set.

#### *Is Wolbachia-Induced Incompatibility a Mechanism of Divergence or Speciation in Invertebrates?*

In a recent paper, Bordenstein et al. (2001) suggested that *Wolbachia*-induced cytoplasmic incompatibility between the parasitic wasps *Nasonia giraulti* and *N. longicornis* occurred before the evolution of other postmating isolating mechanisms (including hybrid inviability and sterility). This led Zimmer (2001) to speculate that *Wolbachia*-induced incompatibility may be an important mechanism of speciation in invertebrates. *Wolbachia*-induced bidirectional incompatibility occurs between the *D. simulans* mitochondrial types and between *N. giraulti* and *N. longicornis*. In the following section, we compare the similarities and differences in the two systems. Overall, these data do not support the hypothesis that *Wolbachia* is a mechanism of speciation or of pangenomic divergence in *D. simulans*.

The observed population subdivision in *D. simulans* and *Nasonia* make it difficult to determine whether these taxa are species in the biological sense. *Drosophila simulans* *sI* and *sII* are allopatric. The *sII* haplotype is sympatric with *sIII* in Madagascar, Reunion, and East Africa. *Nasonia giraulti* and *N. longicornis* are allopatric in North America. One sensitive method to assess biological species is assortative mating studies testing for premating isolation. In neither case is there any evidence of premating isolation. We show that *D. simulans* females infected with different strains of *Wolbachia* mate randomly (Table 7). *Nasonia giraulti* females show no mate discrimination, but, somewhat surprisingly, *N. longicornis* females preferentially mate with *N. giraulti* males relative to homospecific males (Bordenstein et al. 2001).

In this study, the mating experiments do not support the

hypotheses that differences in genital arch size or shape are associated with premating incompatibilities in *D. simulans*. We do note that *Wolbachia* infection is associated with a significant difference in genital-arch morphology: *siIII* lines that are infected with *Wolbachia* have significantly larger and more highly arched genitalia than *siIII* uninfected lines. The comparison of the morphologies of *siIII* *Wolbachia*-infected to other haplotypes, whether infected or not, is not significantly different. This result is very interesting because unpublished results (ACJ, JWOB) indicate that *wMa* infection causes an increase in fecundity. Hariri et al. (1998) found a line of *Sphyracephala beccarii* (a stalk-eyed fly) in which males infected with *Wolbachia* had higher fertility. Returning to the comparison with *Nasonia*, we note that the morphological differences among *N. longicornis* and *N. giraulti* were calculated from minimal sample sizes: 10 males and females from across their entire geographic ranges (Darling and Werren 1990). The fact that all measures overlap among species eliminates the morphology as being unique or diagnostic of particular species.

In both *D. simulans* and *Nasonia*, there is postmating incompatibility, but it is not clear that it significantly inhibits gene flow in either case. If all else is equal, the taxa with the highest incompatibility are expected to show the lowest gene flow—in this case *D. simulans* *wHa* and *wRi* individuals. In the mitochondrial genome of *D. simulans*, the geographic subdivision (Ballard 2000a) is correlated with *Wolbachia* infection status (Fig. 2). In contrast to this subdivision, *ND75* (Fig. 2) and intron 1 of the *Adhr* locus (Ballard 2000a) are assorting randomly with regard to cytotype, and *Wolbachia* infection has not produced genetically distinct lineages. In *Nasonia*, there are less data currently available. In subfamily II, the *N. longicornis* transposable elements do not clearly resolve into a separate lineage from the *N. giraulti* elements, although the two *N. giraulti* elements form a distinct clade (McAllister and Werren 1997).

### Conclusions

Mitochondrial DNA is useful for studying a variety of evolutionary questions. However, we suggest that investigators test whether data gathered from mtDNA can be corroborated prior to making any assumptions about organismal processes and histories. Also, we make an appeal (*sensu* Ballard and Kreitman 1995) to include phylogenetic and statistical tests of neutrality as a supplement to phylogenetic and population genetic analyses when using mtDNA as an evolutionary marker. Failing to reject the null hypothesis of neutrality would support the generality of evolutionary hypotheses generated from mtDNA.

In this study, we show that *Wolbachia*-induced bidirectional cytoplasmic incompatibility occurs within a single species of *Drosophila* but does not appear to act as an important isolating mechanism. There is no evidence of a reduction in nuclear gene flow between the *D. simulans* mitochondrial haplotypes (*siI*, *siII*, and *siIII*) that are infected with distinct strains of *Wolbachia* (*wAu*, *wHa*, *wMa*, *wNo*, and *wRi*). These data do not support the work of Zimmer (2001), who suggested that *Wolbachia*-induced cytoplasmic incompatibility might be a general mechanism of speciation in invertebrates.

### ACKNOWLEDGMENTS

We thank K. Ballard, M. Dean, D. Faith, A. Hoffmann, M. McMahon, M. Noor, and two anonymous reviewers for comments; K. Ballard and A. Berry for help collecting flies; ICIPE for assistance in obtaining collection permits; M. Ashburner, A. Hoffmann, K. Kaneshiro, R. Russell, M. Solignac, and C.-I. Wu for providing isofemale lines; J. Kirchman for help with sequencing; and R. Bradley and R. Baker for an unpublished manuscript. All sequencing was done in the Pritzker Laboratory of Molecular Systematics and Evolution, Field Museum. Funds were provided by National Science Foundation grant no. DEB-9702824, Comer Science and Education Foundation, J. Fahn, and the Field Museum Marshall Field Fund.

### LITERATURE CITED

- Atlan, A., H. Merçot, C. Landre, and C. Montchamp-Moreau. 1997. The sex-ratio trait in *Drosophila simulans*: geographical distribution of distortion and resistance. *Evolution* 51:1886–1895.
- Avise, J. C. 1987. Identification and interpretation of mitochondrial DNA stocks in marine species. Pp. 105–136 in H. Kumph and E. L. Nakamura, eds. *Proceedings of the stock identification workshop*. National Oceanographic and Atmospheric Administration, Panama City, FL.
- Avise, J. C., R. M. Arnold, R. M. Ball, E. Birmingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Sanders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* 18: 489–522.
- Awadalla, P., A. Eyre-Walker, and J. Maynard Smith. 1999. Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science* 286:2524–2525.
- Baba-Aissa, F., M. Solignac, N. Dennebouy, and J. R. David. 1988. Mitochondrial DNA variability in *Drosophila simulans*: quasi absence of polymorphism within each of the three cytoplasmic races. *Heredity* 61:419–426.
- Ballard, J. W. O. 2000a. Comparative genomics of mitochondrial DNA in *Drosophila simulans*. *J. Mol. Evol.* 51:64–75.
- . 2000b. Comparative genomics of mitochondrial DNA in members of the *Drosophila melanogaster* subgroup. *J. Mol. Evol.* 51:48–63.
- . 2000c. When one is not enough: introgression of mitochondrial DNA in *Drosophila*. *Mol. Biol. Evol.* 17:1126–1130.
- Ballard, J. W. O., and M. D. Dean. 2001. The mitochondrial genome: mutation, selection and recombination. *Curr. Opin. Genet. Dev.* 11:667–672.
- Ballard, J. W. O., and M. Kreitman. 1994. Unraveling selection in the mitochondrial genome of *Drosophila*. *Genetics* 138:757–772.
- . 1995. Is mitochondrial DNA a strictly neutral marker? *Trends Ecol. Evol.* 10:485–488.
- Ballard, J. W. O., M. P. Zalucki, and K. E. Basford. 1991. Colonization of perspex strips by larvae of *Austrosimulium bancrofti* (Taylor) (Diptera: Simuliidae), near Ipswich, Queensland. *Hydrobiologia* 218:255–263.
- Ballard, J. W. O., J. Hatzidakis, T. L. Karr, and M. Kreitman. 1996. Reduced variation in *Drosophila simulans* mitochondrial DNA. *Genetics* 144:1519–1528.
- Begun, D. J., and C. F. Aquadro. 1993. African and North American populations of *Drosophila melanogaster* are very different at the DNA level. *Nature* 365:548–550.
- Bookstein, F. L. 1991. *Morphometric tools for landmark data: geometry and biology*. Cambridge Univ. Press, New York.
- Bordenstein, S. R., F. P. O'Hara, and J. H. Werren. 2001. *Wolbachia*-induced incompatibility precedes other hybrid incompatibilities in *Nasonia*. *Nature* 409:707–710.
- Bourtzis, K., and S. L. O'Neill. 1998. *Wolbachia* infections and arthropod reproduction. *BioScience* 48:287–293.
- Bradley, R. D., and R. J. Baker. 2001. A test of the genetic species

- concept: cytochrome *b* sequences and mammals. *J. Mammal.* 82: 960–973.
- Brust, R. A., J. W. O. Ballard, F. Driver, D. M. Hartley, N. J. Galway, and J. Curran. 1998. Molecular systematics, morphological analysis, and hybrid crossing identify a third taxon, *Aedes* (*Halaedes*) *wardangensis* sp. nov., of the *Aedes* (*Halaedes*) *australis* species group (Diptera: Culicidae). *Can. J. Zool.* 76: 1236–1246.
- Bull, J. J., J. P. Huelsenbeck, C. W. Cunningham, D. L. Swofford, and P. J. Waddell. 1993. Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* 42:384–397.
- Burbrink, F. T., R. Lawson, and J. B. Slowinski. 2000. Mitochondrial DNA phylogeography of the polytypic North American rat snake (*Elaphe obsoleta*): a critique of the subspecies concept. *Evolution* 2000:2107–2118.
- Caggesse, C., G. Ragone, B. Perrini, R. Moschetti, V. De Pinto, R. Caizzi, and P. Barsanti. 1999. Identification of nuclear genes encoding mitochondrial proteins: isolation of a collection of *D. melanogaster* cDNAs homologous to sequences in the human gene index database. *Mol. Gen. Genet.* 261:64–70.
- Caspari, E., and G. S. Watson. 1959. On the evolutionary importance of cytoplasmic sterility in mosquitoes. *Evolution* 13: 568–570.
- Cazemajor, M., C. Landre, and C. Montchamp-Moreau. 1997. The sex-ratio trait in *Drosophila simulans*: genetic analysis of distortion and suppression. *Genetics* 147:635–642.
- Charlesworth, B. 1994. The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genet. Res.* 63:213–227.
- Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* 134:1289–1303.
- Charlesworth, D., B. Charlesworth, and M. T. Morgan. 1995. The pattern of neutral molecular variation under the background selection model. *Genetics* 141:1619–1632.
- Collin, R. 2000. Phylogeny of the *Crepidula plana* (Gastropoda: Calyptraeidae) cryptic species complex in North America. *Can. J. Zool.* 78:1500–1514.
- Coyne, J. A. 1983. Genetic basis of differences in genital morphology among the three sibling species of *Drosophila*. *Evolution* 37:1101–1117.
- . 1985. Genetic studies of three sibling species of *Drosophila* with relationship to theories of speciation. *Genet. Res.* 46: 169–192.
- . 1993. The genetics of an isolating mechanism between two sibling species of *Drosophila*. *Evolution* 47:778–788.
- Coyne, J. A., and B. Charlesworth. 1989. Genetic analysis of X-linked sterility in hybrids between three sibling species of *Drosophila*. *Heredity* 62:97–106.
- Coyne, J. A., and M. Kreitman. 1986. Evolutionary genetics of two sibling species of *Drosophila*, *D. simulans* and *D. mauritiana*. *Evolution* 40:673–691.
- Coyne, J. A., and H. A. Orr. 1998. The evolutionary genetics of speciation. *Philos. Trans. R. Soc. Lond. B* 353:287–305.
- Coyne, J. A., S. Simeonidis, and P. Rooney. 1998. Relative paucity of genes causing inviability in hybrids between *Drosophila melanogaster* and *D. simulans*. *Genetics* 150:1091–1103.
- Cunningham, C. W. 1997. Can three incongruence tests predict when data should be combined? *Mol. Biol. Evol.* 14:733–740.
- Darling, D. C., and J. H. Werren. 1990. Biosystematics of *Nasonia* (Hymenoptera: Pteromalidae): two new species reared from bird's nests in North America. *Ann. Entomol. Soc. Am.* 83: 352–370.
- DeSalle, R., and L. V. Giddings. 1986. Discordance of nuclear and mitochondrial DNA phylogenies in Hawaiian *Drosophila*. *Proc. Natl. Acad. Sci. USA* 83:6902–6906.
- Dobzhansky, T. 1935. A critique of the species concept in biology. *Philos. Sci.* 2:344–345.
- Efron, B. 1982. The jackknife, the bootstrap, and other resampling plans. Pp. 1–92 in CBMS-NSF regional conference series in applied mathematics. Vol. 38. CBMS-NSF, Philadelphia, PA.
- Eyre-Walker, A., N. H. Smith, and J. Maynard Smith. 1999. How clonal are human mitochondria? *Proc. R. Soc. Lond. B* 266: 477–483.
- Faith, D. P., and J. W. H. Trueman. 2001. Towards an inclusive philosophy for phylogenetic inference. *Syst. Biol.* 50:1–20.
- Farris, J. S., M. Källersjö, A. Kluge, and C. Bult. 1995. Testing significance of incongruence. *Cladistics* 10:315–319.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using bootstrap. *Evolution* 39:783–791.
- Ferson, S., F. J. Rohlf, and R. K. Koehn. 1985. Measuring shape variation of two-dimensional outlines. *Syst. Zool.* 34:59–68.
- Filchak, K. E., J. B. Roethle, and J. L. Feder. 2000. Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature* 407:739–742.
- Fu, Y. X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925.
- Fu, Y. X., and W. H. Li. 1993. Statistical tests of neutrality of mutations. *Genetics* 133:693–709.
- Goldstein, P. Z., and R. DeSalle. 2000. Phylogenetic species, nested hierarchies, and character fixation. *Cladistics* 16:364–384.
- Hagelberg, E., N. Goldman, P. Lio, S. Whelan, W. Schiefenhovel, J. Clegg, B., and D. K. Bowden. 1999. Evidence for mitochondrial DNA recombination in a human population of island Melanesia. *Proc. R. Soc. Lond. B* 266:485–492.
- Hale, L. R., and A. A. Hoffmann. 1990. Mitochondrial DNA polymorphism and cytoplasmic incompatibility in natural populations of *Drosophila simulans*. *Evolution* 44:1383–1386.
- Hariri, A. R., J. H. Werren, and G. S. Wilkinson. 1998. Distribution and reproductive effects of *Wolbachia* in stalk-eyed flies (Diptera: Diopsidae). *Heredity* 81:254–260.
- Hey, J. 1997. Mitochondrial and nuclear genes present conflicting portraits of human origins. *Mol. Biol. Evol.* 14:166–172.
- Hillis, D. M., J. Huelsenbeck, and C. W. Cunningham. 1994. Application and accuracy of molecular phylogenies. *Science* 264: 671–676.
- Hoelzer, G. A. 1997. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* 51:622–626.
- Hoffmann, A. A., M. Turelli, and G. M. Simmons. 1986. Unidirectional incompatibility between populations of *Drosophila simulans*. *Evolution* 40:692–701.
- Hoffmann, A. A., D. Clancy, and J. Duncan. 1996. Naturally occurring *Wolbachia* infection in *Drosophila simulans* that does not cause cytoplasmic incompatibility. *Heredity* 76:1–8.
- Holder, M. T., M. V. Erdmann, T. P. Wilcox, R. L. Caldwell, and D. M. Hillis. 1999. Two living species of coelacanths? *Proc. Natl. Acad. Sci. USA* 96:12616–12620.
- Hudson, R. R., M. Kreitman, and M. Aguade. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics* 116: 153–159.
- Hunt, G., and R. E. Chapman. 2001. Evaluating hypotheses of instar-grouping in arthropods: a maximum likelihood approach. *Paleobiology* 27:466–484.
- James, A. C., and J. W. O. Ballard. 2000. Expression of cytoplasmic incompatibility in *Drosophila simulans* and its impact on infection frequencies and distribution of *Wolbachia pipiensis*. *Evolution* 54:1661–1672.
- James, A. C., M. D. Dean, M. E. McMahon, and J. W. O. Ballard. 2002. Dynamics of double and single *Wolbachia* infections in *Drosophila simulans* from New Caledonia. *Heredity* 88: 182–189.
- Johns, G. C., and J. C. Avise. 1998. A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. *Mol. Biol. Evol.* 15:1481–1490.
- Joseph, L., C. Moritz, and A. Hugall. 1995. Molecular support for vicariance as a source of diversity in rainforest. *Proc. R. Soc. Lond. B* 260:177–182.
- Kajander, O. A., A. T. Rovio, K. Majamaa, J. Poult, J. N. Spelbrink, I. J. Holt, P. J. Karhunen, and H. T. Jacobs. 2000. Human mtDNA sublimons resemble rearranged mitochondrial genomes found in pathological states. *Hum. Mol. Genet.* 9:2821–2835.
- Kliman, R. M., P. Andolfatto, J. A. Coyne, F. Depaulis, M. Kreitman, A. J. Berry, J. McCarter, J. Wakeley, and J. Hey. 2000.

- The population genetics of the origin and divergence of the *Drosophila simulans* complex species. *Genetics* 156:1913–1931.
- Lassy, C. W., and T. L. Karr. 1996. Cytological analysis of fertilization and early embryonic development in incompatible crosses of *Drosophila simulans*. *Mech. Dev.* 57:47–58.
- Latter, B. D. H., and J. C. Mulley. 1995. Genetic adaptation to captivity and inbreeding depression in small laboratory populations of *Drosophila melanogaster*. *Genetics* 139:255–266.
- Lee, C. E. 2000. Global phylogeography of a cryptic copepod species complex and reproductive isolation between genetically proximate “populations.” *Evolution* 54:2014–2027.
- Liu, J., J. M. Mercer, L. F. Stam, G. C. Gibson, Z. B. Zeng, and C. C. Laurie. 1996. Genetic analysis of a morphological shape difference in the male genitalia of *Drosophila simulans* and *D. mauritiana*. *Genetics* 142:1129–1145.
- Maynard Smith, J., and J. Haigh. 1974. The hitchhiking effect of a favorable gene. *Genet. Res.* 23:23–25.
- Mayr, E. 1942. *Systematics and the origin of species*. Columbia Univ. Press, New York.
- McAllister, B. F., and J. H. Werren. 1997. Phylogenetic analysis of a retrotransposon with implications for strong evolutionary constraints on reverse transcriptase. *Mol. Biol. Evol.* 14:69–80.
- McDonald, J. H., and M. Kreitman. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* 351:652–654.
- McLachlan, G. J., and K. E. Basford. 1988. *Mixture models: inference and applications to clustering*. Marcel Dekker, New York.
- McLachlan, G. J., and P. N. Jones. 1988. Fitting mixture models to grouped and truncated data via the EM algorithm. *Biometrics* 44:571–578.
- McLachlan, G. J., and T. Krishnan. 1997. *The EM algorithm and extensions*. Wiley and Sons, New York.
- McLachlan, G. J., and D. Peel. 1998. MIXFIT: an algorithm for the automatic fitting and testing of normal mixture models. Pp. 553–557 in *Proceedings of the 14th international conference on pattern recognition*. IEEE Computer Society, Los Alamitos, CA.
- Merçot, H., and D. Poinsot. 1998. *Wolbachia* transmission in a naturally bi-infected *Drosophila simulans* strain from New Caledonia. *Entomol. Exp. Appl.* 86:97–103.
- Mickevich, M. F., and J. S. Farris. 1981. The implications of congruence in *Menidia*. *Syst. Biol.* 30:351–370.
- Montchamp-Moreau, C., V. Ginhoux, and A. Atlan. 2001. The Y chromosomes of *Drosophila simulans* are highly polymorphic for their ability to suppress sex-ratio drive. *Evolution* 55:728–737.
- Moore, W. S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial gene trees versus nuclear-gene trees. *Evolution* 49:718–726.
- . 1997. Mitochondrial-gene trees versus nuclear-gene trees, a reply to Hoelzer. *Evolution* 51:627–629.
- Moritz, C. 1994. Defining ‘evolutionary significant units’ for conservation. *Trends Ecol. Evol.* 9:373–375.
- Nachman, M. W. 1998. Deleterious mutations in animal mitochondrial DNA. *Genetica* 103:61–69.
- Nachman, M. W., S. N. Boyer, and C. F. Aquadro. 1994. Nonneutral evolution at the mitochondrial NADH dehydrogenase subunit 3 gene in mice. *Proc. Natl. Acad. Sci. USA* 91:6364–6368.
- Nigro, L., and T. Prout. 1990. Is there selection on RFLP differences in mitochondrial DNA? *Genetics* 125:551–555.
- O’Neill, S. L., and T. L. Karr. 1990. Bidirectional incompatibility between conspecific populations of *Drosophila simulans*. *Nature* 348:178–180.
- O’Neill, S. L., R. Giordano, A. M. Colbert, T. L. Karr, and H. M. Robertson. 1992. 16S rRNA phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic incompatibility in insects. *Proc. Natl. Acad. Sci. USA* 89:2699–2702.
- Partridge, L., A. A. Hoffmann, and J. S. Jones. 1987. Male size and mating success in *Drosophila melanogaster* and *D. pseudoobscura* under field conditions. *Anim. Behav.* 35:468–476.
- Paterson, H. E. H. 1985. The recognition concept of species. Pp. 21–29 in E. Vrba, ed. *Species and speciation*. Transvaal Museum, Pretoria.
- Perez, D. E., C. I. Wu, N. A. Johnson, and M. L. Wu. 1993. Genetics of reproductive isolation in the *Drosophila simulans* clade: DNA marker-assisted mapping and characterization of a hybrid-male sterility gene, *Odysseus* (*Ods*). *Genetics* 134:261–275.
- Pouyaud, L., S. Wirjoatmodjo, I. Rachmatika, A. Tjakrawidjaja, R. Hadiaty, and W. Hadie. 1999. A new species of coelacanth: genetic and morphologic proof. *C. R. Acad. Sci. Ser. III Sci. Vie* 322:261–267.
- Rand, D. M., and L. M. Kann. 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. *Mol. Biol. Evol.* 13:735–748.
- . 1998. Mutation and selection at silent and replacement sites in the evolution of animal mitochondrial DNA. *Genetica* 103:393–407.
- Rand, D. M., M. Dorfsman, and L. M. Kann. 1994. Neutral and non-neutral evolution of *Drosophila* mitochondrial DNA. *Genetics* 138:741–756.
- Rice, W. R. 1989. Analyzing tables of statistical tests. *Evolution* 43:223–225.
- Rohlf, F. J. 1998. TPSDIG. State University of New York at Stony Brook, Stony Brook, NY.
- . 2000. NTSYSPC. Applied Biostatistics, New York.
- Rohlf, F. J., and J. W. Archie. 1984. A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Syst. Biol.* 33:302–317.
- Rosen, D. E. 1979. Fishes from the uplands and intermontane basins of Guatemala: revisionary studies and comparative geography. *Bull. Am. Mus. Nat. Hist.* 162:267–376.
- Rozas, J., and R. Rozas. 1997. DnaSP version 3.0: a novel software package for extensive molecular population genetics analysis. *Comput. Appl. Biosci.* 13:307–311.
- Satta, Y., H. Ishiwa, and S. I. Chigusa. 1987. Analysis of nucleotide substitutions of mitochondrial DNAs in *Drosophila melanogaster* and its sibling species. *Mol. Biol. Evol.* 4:638–650.
- Saville, B. J., Y. Kohli, and J. F. Anderson. 1998. MtDNA recombination in a natural population. *Proc. Natl. Acad. Sci. USA* 95:1331–1335.
- Schneider, C. J., T. B. Smith, B. Larison, and C. Moritz. 1999. A test of alternative models of diversification in tropical rainforests: ecological gradients vs. rainforest refugia. *Proc. Natl. Acad. Sci. USA* 96:13869–13873.
- Sokal, R. R., and F. J. Rohlf. 1995. *Biometry*. 3rd ed. W. H. Freeman, New York.
- Solignac, M., and M. Monnerot. 1986. Race formation and introgression within *Drosophila simulans*, *D. mauritiana*, and *D. sechellia* inferred from mitochondrial DNA analysis. *Evolution* 40:531–539.
- Solignac, M., M. Monnerot, and J. C. Mounolou. 1986. Mitochondrial DNA evolution in the melanogaster species subgroup of *Drosophila*. *J. Mol. Evol.* 23:31–40.
- Sumner, C. J. 1991. Nucleotide polymorphism in the *alcohol dehydrogenase duplicate* locus of *Drosophila simulans*: implications for the neutral theory. *Ecology and evolution*. Ph.D. diss., Princeton University, Princeton, NY.
- Swofford, D. L. 1998. PAUP\*: phylogenetic analysis using parsimony (\*and other methods). Sinauer Associates, Sunderland, MA.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- Ting, C. T., S. C. Tsaur, and C.-I. Wu. 2000. The phylogeny of closely related species as revealed by the genealogy of a speciation gene, *Odysseus*. *Proc. Natl. Acad. Sci. USA* 97:5313–5316.
- True, J. R., J. Liu, L. F. Stam, Z.-B. Zeng, and C. C. Laurie. 1997. Quantitative genetic analysis of divergence in male secondary sexual traits between *Drosophila simulans* and *Drosophila mauritiana*. *Evolution* 51:816–832.
- Turelli, M., and A. A. Hoffmann. 1991. Rapid spread of an inherited incompatibility factor in California *Drosophila*. *Nature* 353:440–442.
- . 1995. Cytoplasmic incompatibility in *Drosophila simulans*: dynamics and parameter estimates from natural populations. *Genetics* 140:1319–1338.
- Uenoyama, T., and Y. Inoue. 1995. Genetic studies on premating

- isolation in *Drosophila simulans*. I. A *D. simulans* line highly crossable with *D. melanogaster*. *Jpn. J. Genet.* 70:365–371.
- Werren, J. H. 1997. Biology of *Wolbachia*. *Annu. Rev. Entomol.* 42:587–609.
- Wu, C.-I., H. Hollocher, D. J. Begun, C. F. Aquadro, Y. Xu, and M. L. Wu. 1995. Sexual isolation in *Drosophila melanogaster*: a possible case of incipient speciation. *Proc. Natl. Acad. Sci. USA* 92:2519–2523.
- Zhou, W., F. Rousset, and S. L. O'Neill. 1998. Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. *Proc. R. Soc. Lond. B* 265:509–515.
- Zimmer, C. 2001. *Wolbachia*: a tale of sex and survival. *Science* 292:1093–1095.

Corresponding Editor: M. Noor