

# Mitochondrial DNA Evolution in the Drosophila nasuta Subgroup of Species

Hwei-yu Chang,\* Degui Wang, and Francisco J. Ayala

Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92717, USA

Summary. The Drosophila nasuta group consists of about 12 closely related species distributed throughout the Indo-Pacific region. They are of great interest because of their evolutionary idiosyncrasies including little morphological differentiation, the ability to intercross in the laboratory often producing fertile offspring, and substantial chromosomal evolution. Studies of metric traits, reproductive isolation, and chromosomal and enzyme polymorphisms have failed to resolve the phylogeny of the species. We report the results of a survey of the mitochondrial DNA (mtDNA) restriction patterns of the species. The phylogeny obtained is consistent with other available information and suggests that D. albomicans may represent the ancestral lineage of the group. The amount of polymorphism in local populations ( $\pi = 1.0\%$  per site) is within the typical range observed in other animals, including Drosophila. The degree of differentiation between species is, however, low: the origin of the group is tentatively dated about 6-8 million years ago. This study confirms the usefulness of mtDNA restriction patterns for ascertaining the phylogeny of closely related species.

**Key words:** Molecular evolution – Genetic polymorphism – Phylogeny – Population genetics – Speciation

## Introduction

Restriction patterns of mitochondrial DNA (mt-DNA) have become a powerful tool in evolutionary studies of closely related organisms (e.g., Avise et al. 1979, 1984; Brown 1980; Brown et al. 1982; DeSalle et al. 1986; Hale and Beckenbach 1985; Latorre et al. 1986, 1988; Solignac et al. 1986; Cann et al. 1987). The mtDNA is maternally transmitted in animals; hence, it evolves as a clone by gradual accumulation of mutations. The reconstruction of phylogenetic (matrilineal) lineages is facilitated by the absence of recombination, which complicates the evolution of nuclear DNA. The rapid evolution of animal mtDNA (Avise and Lansman 1983; Brown 1983, 1985; but see Powell et al. 1986; Vawter and Brown 1986) makes it particularly suitable for ascertaining the phylogeny of closely related species or even populations of the same species.

The Drosophila nasuta group consists of about a dozen species widely distributed throughout the Indo-Pacific region (Wilson et al. 1969; Kitagawa et al. 1982). They are classified as the nasuta subgroup within the D. immigrans group. The nasuta species exhibit features that make them evolutionarily interesting, such as extensive polymorphism (even within a single species) in the metaphase configuration of the chromosomes. Absence of ethological isolation (at least in the laboratory) has brought into question the specific status of some taxa (Wilson et al. 1969; Kitagawa et al. 1982; Ramachandra and Ranganath 1985, 1986). Extensive morphometric, chromosomal, and reproductive isolation studies have failed to provide a coherent picture of the phylogenetic relationships among the species-a situation that handicaps the interpretation of the numerous evolutionary studies of this set of species.

We report herein a survey of the mtDNA restriction patterns of 13 taxa of the *nasuta* group, plus *D. immigrans* as an "outgroup" or reference species, seeking to ascertain the phylogeny of this group of species.

Offprint requests to: F.J. Ayala

<sup>\*</sup> Current address: Department of Entomology and Plant Pathology, National Taiwan University, Taipei, Taiwan 11529



**Fig. 1.** Geographic location of the 31 strains of 13 species and subspecies of the *Drosophila nasuta* subgroup used in the survey of mitochondrial DNA. The abbreviations used are: albo = D. albomicans; kepu = D. kepulauana; kohk = D. kohkoa; nasu = D. nasuta; nive = D. niveifrons; palli = D. pallidifrons; pula = D. pulaua; sulf-a = D. sulfurigaster albostrigata; sulf-b = D. s. bilimbata; sulf-n = D. s. sulfurigaster; F and H are the symbols for two new species not yet formally described or named. The survey also includes one strain of D. *immigrans* collected in Winters, California, not shown in the map.

### **Materials and Methods**

Drosophila Strains. The geographic origins of the 32 strains studied are listed in Table 1 (see also Fig. 1). They comprise 12 named taxa (nine species and four subspecies) plus 2 taxa identified as distinct species but not yet formally described or named (taxa Fand H). Strains 2–8, 10, 11, and 13–16 were collected by Dr. Fei-Jann Lin of the Institute of Zoology, Academia Sinica, Taipei; strain 32 was collected by Dr. Wyatt W. Anderson of the University of Georgia, Athens; all other stocks were provided by Dr. Osamu Kitagawa, Tokyo Metropolitan University. The geographic origins of the 32 strains cover the known distribution of the species, from Hawaii to eastern Africa, and from tropical Australia to southern Japan, China, and India.

Each strain is derived from a single female collected in the wild. We maintained each strain in six (6-dram) culture vials. The progenies of the six vials were intermingled each generation before subculturing them in six fresh vials, which was done every 20 days.

Extraction of mtDNA. The method was modified from Coen et al. (1982) as follows. Fifteen young flies were gently homogenized in a 1.5-ml microcentrifuge tube with 320  $\mu$ l of solution I (10 mM Tris, 60 mM NaCl, 5% sucrose wt/vol, 10 mM EDTA, pH 7.8). Four hundred microliters of solution II (300 mM Tris, 1.25% sodium dodecyl sulfate, 5% sucrose, 100 mM EDTA, 0.8% freshly mixed diethyl pyrocarbonate, pH 9) were then added. The mixture was incubated at 65°C for 30 min, after which 120 ml of 3 M sodium acetate was added and the mixture was kept on ice for 45 min. After centrifugation for 10 min in an Eppendorf centrifuge, DNA was precipitated by isopropanol and resuspended in 250  $\mu$ l of double-distilled water with 0.6  $\mu$ l diethyl pyrocarbonate. The mixture was left at room temperature for 20 min, after which 250  $\mu$ l of double-distilled water were added. The DNA was then precipitated with 70% ethanol. Residual ethanol was removed by drying the precipitate in a desiccator for 30 min, after which the DNA was dissolved in an appropriate amount of 10 mM Tris, 10 mM EDTA, pH 8. An enriched fraction of mtDNA was obtained with this extraction procedure that gave well-resolved bands after restriction enzyme digestion.

Restriction Endonuclease Digestion and Electrophoresis. Six endonucleases were used that recognize 6-bp sequences: SalI, BamHI, PstI, XbaI, EcoRI, and Hind III. The endonucleases were bought from Boehringer Mannheim, New England Biolabs, and Pharmacia. The mtDNA was dissolved in 20  $\mu$ l doubledistilled water and digested according to the procedure described by Maniatis et al. (1982) and following the supplier's recommendations.

The digestion fragments were separated in 0.8% agarose by gel electrophoresis using a Tris-acetate buffer (4 mM Tris, 2 mM acetic acid, 0.2 mM EDTA, pH 8.1). After electrophoresis, the gel was stained with ethidium bromide (0.1  $\mu$ g/ml) and the DNA bands were visualized under UV light. The fragment size was estimated by graphical interpolation using  $\lambda$ DNA digested with *Hind* III as a size standard.

Table 1. Geo	graphic origin,	stock number, a	and mtDNA	morph of the	Drosophila strains
--------------	-----------------	-----------------	-----------	--------------	--------------------

Species	Strain number	Locality	Composite morph
D. albomicans	1	Wulai. Taiwan	
	2	Wulai, Taiwan	I
	3	Wulai, Taiwan	v
	4	Shanhutan, Taiwan	IV
	5	Shanhutan, Taiwan	I
	6	Funglin, Taiwan	Ι
	7	Nakhon Nayok, Thailand	I
	8	Nakhon Nayok, Thailand	II
	9	Chiangmai, Thailand	Ι
	10	Chiangmai, Thailand	III
	11	Chiangmai, Thailand	III
	12	Nago, Japan	II
	13	Nago, Japan	II
	14	Nago, Japan	II
	15	Ishigaki-jima, Japan	II
	16	Ishigaki-jima, Japan	VI
D. kepulauana	17	Palawan, Philippines	х
D. kohkoa	18	Sarawak, Malaysia	XI
	19	Sarawak, Malaysia	XII
D. nasuta	20	Kandy, Sri Lanka	II
	21	Mysore, India	II
	22	Mombasa, Kenya	IX
D. pallidifrons	23	Ponape, Caroline Island	XIII
D. pulaua	24	Sarawak, Malaysia	VIII
D. sulfurigaster albostrigata	25	Kota Kinabaru, Malaysia	I
D. sulfurigaster bilimbata	26	Honolulu, Hawaii	VII
D. sulfurigaster neonasuta	27	Mysore, India	VII
D. sulfurigaster sulfurigaster	28	Papua, New Guinea	VII
Taxon F	29	Sarawak, Malaysia	VIII
D. niveifrons	30	Papua, New Guinea	XIV
Taxon H	31	Papua, New Guinea	XV
D. immigrans	32	Winters, California	XVI

Taxons F and H have been identified as new species by Dr. Osamu Kitagawa, Tokyo Metropolitan University, but have not yet received formal scientific names

Dendrogram Clustering Methods. A parsimony network of the 16 composite mtDNA morphs was produced by minimizing the number of restriction site changes required over the whole network. In addition, we have constructed a phylogenetic tree using the computer program KITSCH from Dr. Joseph Felsenstein of the University of Washington, Seattle. The input data are the pairwise genetic distances between species (the matrix displayed in Table 5, except that the four *D. sulfurigaster* subspecies are combined). This clustering method assumes that the distance values are independent and additive, and that the rate of evolution is constant (see Felsenstein 1984 for a discussion of the assumptions). Other computer programs (e.g., FITCH, also prepared by Dr. Felsenstein, that does not assume the hypothesis of an "evolutionary clock") gave identical or fairly similar clusterings.

# Results

We estimate the length of the *D. nasuta* mtDNA to be 16.5 kb, the same with all six restriction enzymes. We have not detected any length variation within or between the species.

The six endonucleases yield a total of 33 mtDNA restriction sites in the 32 strains. SalI does not yield

any restriction sites. *Bam*HI yields 1 digestion site, that is polymorphic; *Pst*I yields 2, *Xba*I 5, *Eco*RI 10, and *Hind*III 15 (Fig. 2). Only four sites are shared by all strains (one site each for *Xba*I and *Hind*III, and two for *Eco*RI); the other 29 (88%) sites are polymorphic. The restriction fragments are numbered consecutively in Fig. 2 from largest to smallest, starting with 1, separately for each enzyme. The restriction sites and the patterns obtained are arbitrarily lettered for reference.

The digestion patterns obtained with each enzyme have been arranged in Fig. 2 so as to minimize the number of sites that need to be gained or lost in a network that connects them all. The transition from one pattern to another requires a single site change (gain or loss) in 10 (53%) cases (including the transitions from an undigested pattern to patterns A and B); five (26%) transitions require two mutational steps; and four (21%) require three steps. Two- or three-step transitions may exist because the intermediates are not represented in our samples or because they have been lost through evolutionary time. Both alternatives may pertain in our case. The

	Bamhi		<b>Pst</b> I	XbaI						EcoRI					
Morph	<u>a</u>	b	С	d	e	f	g	h	i	j	k	1	m	n	0
I	_	_		+	+	_	+	_	+	+	-	-	+	_	
II	_	-	_	+	+	~	+	_	+	+	_	_	+	-	~
III	_	_		+	+	~	+	-	+	+	-	_	+		-
IV	_	_	_	+	+	~	+	-	+	+	_	-	+	-	-
v	_	-	-	+	+	~	+	-	+	+		_	+	-	~
VI	+	_	-	+	+	~	+	_	+	+	-	-	+	-	
VII	_	+		+	+		+		+	+	_	_	+	_	
VIII	-	+	_	+	+		÷	_	+	+	_	-	+	-	_
IX	_	_	_	+	+	-	+	_	+	+	_	-	-	_	-
Х	_	_	_	+	+		+	_	+	+	_	_	+	_	
XI	_	-	-	+	+	~	+	-	+	+		-	+	_	~
XII	_	-	-	+	+	-	+	-	+	+	-	-	+	-	
XIII	-	-	_	+	+		+	-	+	+	-	_	+	-	+
XIV	_	-		+	_	+	-	÷	+	+	-	_	+	-	+
XV	+	_	-	+	+		+	_	+	+	-	+	_	-	~
XVI	-	+	+	+	-	~	-	_	+	+	+	-	_	÷	-

Table 2. Restriction pattern of the 16 mtDNA morphs

Each restriction site is identified by a letter; + and - refer to the presence or absence of the restriction site



Fig. 2. Diagram of the restriction mtDNA patterns obtained with the five endonucleases manifesting variation among the strains. The arrows indicate the direction in which restriction sites are gained or lost as indicated above each arrow. For each endonuclease, fragments are numbered consecutively starting with the largest one. The lengths in kilobases for each fragment (numbered in parentheses) are as follows. *Bam*HI: 16.5 (1). *Pst*I: 16.5 (1), 8.9 (2), 7.6 (3). *Xba*I: 16.5 (1), 8.3 (2), 6.7 (3), 6.3 (4), 3.5 (5), 1.5 (6). *Eco*RI: 7.2 (1), 5.5 (2), 4.3 (3), 4.1 (4), 4.0 (5), 3.9 (6), 3.9 (7), 3.8 (8), 3.0 (9), 3.0 (10), 2.1 (11), 1.8 (12), 1.7 (13), 1.1 (14). *Hind* III: 16.5 (1), 13.8 (2), 12.0 (3), 9.4 (4), 8.5 (5), 8.0 (6), 7.0 (7), 6.5 (8), 5.6 (9), 5.2 (10), 5.1 (11), 5.0 (12), 4.9 (13), 4.7 (14), 4.5 (15), 3.3 (16), 2.7 (17), 2.6 (18), 2.6 (19), 2.5 (20), 2.4 (21), 2.0 (22), 1.9 (23), 1.8 (24), 1.4 (25), 0.7 (26).

number of strains sampled is small, so that some existing patterns may not have been included in the samples. But one species, *D. immigrans*, is not a close relative of the others and, hence, intermediates connecting it to the other species may have become extinct. Indeed, the transitions from *D. immigrans* to the *nasuta* species require more than one mutational step for all three enzymes in which multi-step

	EcoRI Hind III																
p	q	r	s	t	u	v	w	x	У	z	а	b	с	d	е	f	g
+	-	+	+	_	_		_		+	-	_	_		+	_	_	+
+	_	+	+		-		_	-	+	_	_		_	+		+	+
+	—	+	+	-	-	-	_		+	-	_	-	_	+	-	+	+
+	_	+	+	-	-	-	-		+	-	+	-	-	+	-	-	+
-	+	_	+	_	-	-	—	-	+	-	-	-	-	+	_	+	+
+		+	+	_	_	-	-	-	+	-	-	-	-	+	_	-	+
+	-	+	+	_	_	_	_	+	_	_	-	—	-	—	-		+
+		+	+	-	-			~	_	+	-	_	-	-	-	+	+
+	-	+	+	-	-	-	-		+	_	-	_	-	+	-	+	+
+	_	+	+	-	-	-	-		+	-		-	+	-		-	+
+	_	+	+	-	-	+	_	~	-	-	-	-	-	-	-	-	+
+	—	+	+	-	-	-	_	~	-	-	-	-	-	-	-	-	+
+	_	+	+	-	-	_	_		+	-	+	_	-	+	-	-	+
+	-	+	+	+	-	-	+	-	-	-	-	-	-	-	+	-	+
+	-	+	+	_	-		_	-	-	-	-	_	_	+	-	-	+
-	-	+	+	-	+	-	_	+	_	-	-	+	-	-	-	-	+

transitions are required (*XbaI*, *Eco*RI, and *Hind* III) and account for four of the nine transitions requiring more than one mutational step.

Table 2 lists the 16 composite mtDNA patterns ("morphs" or haplotypes) obtained by combining the patterns separately obtained with each endonuclease. The table shows the restriction sites present in each morph. Table 1 (last column) shows the mtDNA morph found in each of the 32 strains surveyed.

The 16 mtDNA morphs can be interconnected in the "most parsimonious" network displayed in Fig. 3. The minimum number of mutational steps required for the transition from one to another morph is represented in the figure by the number of short dashes crossing the connecting lines. Thus, only one mutational step is required for the transition from morph I to any one of morphs II, IV, and VI, or vice versa; two are required for the transition from morph I to X or vice versa; four from I to either VII or XV; and so on.

As shown in Fig. 3 and Table 1, three morphs are each shared by two species: morph I is shared by *D. albomicans* and *D. sulfurigaster albostrigata*, morph II is shared by *D. albomicans* and *D. nasuta*, and morph VIII by *D. pulaua* and taxon *F*. All other morphs are species-specific. On the whole, the mtDNAs of the *nasuta* species are fairly similar and separated by few mutational steps. The most distinctive in Fig. 3 is morph XVI, present in *D. immigrans*, that does not belong to the *nasuta* subgroup although it is a close relative; morph XVI differs by a minimum of 10 mutations from morph VII, the most similar to it. Among the *nasuta* species, the most different one is morph XIV (*D. niveifrons*), separated by six mutations from the nearest one, morph XII (*D. kohkoa*). The network shown in Fig. 3 represents a possible evolutionary pathway in the evolution of the mtDNA of the *nasuta* species, but the network does not have direction; it does not by itself tell us which one is the most ancestral morph.

The genetic similarity between two mtDNA morphs may be quantified by the proportion of the restriction sites, S, or of the fragments, F, they share in common (Avise et al. 1979; Lansman et al. 1981). Table 3 gives, above the diagonal, the S values for all pairwise comparisons between the 16 morphs. The fraction of restriction sites shared by any two morphs ranges from 0.960 (between morphs II and III) to 0.333 (between morphs V and XVI). Morph XVI, specific of *D. immigrans*, is of course the most different, sharing on the average only 0.436 of the sites with any other morph.

If it is assumed that changes in restriction sites are due to nucleotide substitutions rather than to insertions or deletions, it becomes possible to estimate the frequency, p, of nucleotide differences per nucleotide site between any two morphs (Upholt 1977; Nei and Li 1979; Nei and Tajima 1981; Nei 1987). Table 3 gives, below the diagonal, the p value for all pairwise comparisons between the mtDNA morphs. The number of nucleotide substitutions per site ranges from 0.7% (between morphs I and II, and four others) to 18.3% (between morphs V and XVI). On the average, the incidence of nucleotide substitutions between D. immigrans (morph XVI) and any other species is 14.0%.

The amount of mtDNA polymorphism within a population or species may be estimated by  $\pi$ , a measure of nucleotide diversity that takes into account the frequency of the various mtDNA morphs in the sample (Nei and Tajima 1981). We have a sample



Fig. 3. Network of the 16 composite mtDNA morphs of the *D. nasuta* group. The morphs are connected in a way that minimizes the total number of restriction site changes required. The lines connecting the morphs are crossed by short lines, each representing a restriction site change required to go from one morph to the next. Morphs belonging to the same species have been enclosed in dash-line rectangles. Three morphs (I, II, and VIII) are each found in two species.

of 16 strains from seven localities of one species in our study, D. albomicans. Table 4 gives along the diagonal the value of  $\pi$  for each of the six localities from which two or more strains were sampled. Because only two or three strains are available from each locality, the  $\pi$  values are subject to large possible errors. The average of the six values is more reliable:  $\pi = 1.0\%$ . This value suggests a fair amount of polymorphism within local populations of D. al*bomicans*. This tentative conclusion is supported by the related observation that two different morphs are present in each of five localities (out of a total of six) in which either two or three strains were sampled. The only other species for which two strains were available from the same locality is D. kohkoa (Table 1); the two strains carry different mtDNA morphs and the value of  $\pi$  is 0.9%. This suggests that mtDNA polymorphism may also be extensive in this species.

The nucleotide diversity, or polymorphism, for D. albomicans as a whole may be obtained by combining all 16 strains and treating them as a single sample. The value of  $\pi$  for this pooled sample is 1.0%, coincidentally the same as the average value of  $\pi$  for the local samples. This suggests that a large fraction of the total mtDNA polymorphism of the species is present in each population. This may be quantified using the method of Takahata and Palumbi (1985) for calculating *I* and *J*, the intrademe and interdeme identity probabilities (per restriction site), respectively. The mean values and standard errors are  $\overline{I} = 0.887 \pm 0.35$  and  $\overline{J} = 0.841 \pm 0.013$ . The fraction of the mtDNA variation attributable to interdeme differences (Takahata and Palumbi 1985) is  $G_{ST} = 0.257$ ; that is, 74.3% of the mtDNA polymorphism of *D. albomicans* is found within local populations, whereas only about 25% of the total is due to interdeme heterogeneity.

Table 4 also gives, above the diagonal, the value of  $d_{xy}$ , or nucleotide differentiation, between any two populations of D. albomicans (Nei 1987; see Nei and Li 1979; Nei and Tajima 1981). The first three localities are in Taiwan, the next two in Thailand, and the last two in southern Japan. The  $d_{xy}$  values do not manifest any clear differentiation between the three regions. Wulai is the one locality that appears most different from all others, which reflects the presence in this locality of morph V that differs from all other D. albomicans morphs by 3-5 mutations (see Fig. 3). Some suggestion of regional differentiation may be seen in Table 1, where it appears that morph I is the most common in Taiwan but is absent from Japan, whereas morph II is the most common in Japan but is absent from Taiwan; both morphs I and II exist in Thailand.

The average nucleotide divergence between any

Table 3. Genetic differentiation between the 16 mtDNA morphs

Com- posite morph	I	II	III	IV	v	VI	VII	VIII	IX	x	хі	XII	XIIi	XIV	xv	XVI
 I	_	0.957	0.917	0.957	0.818	0.957	0.818	0.783	0.909	0.909	0.857	0.900	0.917	0.667	0.818	0.417
п	0.007	_	0.960	0.917	0.870	0.917	0.783	0.833	0.957	0.870	0.818	0.857	0.880	0.640	0.783	0.400
III	0.015	0.007	_	0.880	0.833	0.880	0.750	0.800	0.917	0.833	0.783	0.818	0.923	0.692	0.750	0.462
IV	0.007	0.015	0.021	_	0.783	0.917	0.783	0.750	0.870	0.870	0.818	0.857	0.960	0.640	0.783	0.400
v	0.033	0.023	0.030	0.041	-	0.783	0.636	0.696	0.818	0.727	0.667	0.700	0.750	0.500	0.636	0.333
VI	0.007	0.015	0.021	0.015	0.041	_	0.783	0.750	0.870	0.870	0.818	0.857	0.880	0.640	0.870	0.400
VII	0.033	0.041	0.048	0.041	0.075	0.041	_	0.870	0.727	0.818	0.857	0.900	0.750	0.667	0.727	0.583
VIII	0.041	0.030	0.037	0.048	0.060	0.048	0.023	-	0.783	0.783	0.818	0.857	0.720	0.640	0.696	0.480
IX	0.016	0.007	0.015	0.023	0.033	0.023	0.053	0.041		0.818	0.857	0.800	0.833	0.583	0.909	0.417
x	0.016	0.023	0.030	0.023	0.053	0.023	0.033	0.041	0.033	-	0.857	0.900	0.833	0.667	0.727	0.417
XI	0.026	0.033	0.041	0.033	0.068	0.033	0.026	0.033	0.026	0.026	_	0.947	0,783	0.696	0.762	0.435
XII	0.018	0.026	0.033	0.026	0.059	0.026	0.018	0.026	0.037	0.018	0.009	_	0.818	0.727	0.800	0.455
XIII	0.015	0.021	0.013	0.007	0.048	0.021	0.048	0.055	0.030	0.030	0.041	0.033	-	0.692	0.750	0.462
XIV	0.068	0.074	0.061	0.074	0.116	0.074	0.068	0.074	0.090	0.068	0.060	0.053	0.061	-	0.583	0.462
xv	0.033	0.041	0.048	0.041	0.075	0.023	0.053	0.060	0.016	0.053	0.045	0.037	0.048	0.090	-	0.417
XVI	0.146	0.153	0.129	0.153	0.183	0.153	0.090	0.122	0.146	0.146	0.139	0.131	0.129	0.129	0.146	

Above the diagonal: S, proportion of restriction sites shared. Below the diagonal: p, estimated proportion of nucleotide pair differences between the morphs

**Table 4.** Genetic variation within local populations ( $\pi$ , along the diagonal) and genetic differentiation between populations ( $d_{xy}$ ) of *D. albomicans* 

Population	Wulai	Shanhutan	Funglin	Nakhon	Chiangmai	Nago	Ishigaki	$\bar{d}_{xy}$
Wulai	0.022	0.014	0.011	0.011	0.018	0.013	0.015	0.014
Shanhutan		0.007	0.004	0.007	0.013	0.012	0.012	0.010
Funglin			_	0.004	0.010	0.007	0.007	0.007
Nakhon				0.007	0.008	0.004	0.008	0.007
Chiangmai					0.010	0.007	0.011	0.011
Nago						0.000	0.008	0.009
Ishigaki							0.015	0.010

two local populations of *D. albomicans* is  $\bar{d}_{xy} = 1.0\%$ , the same value obtained for the average diversity within a local population ( $\pi = 1.0\%$ ), consistent with the observation made earlier that most of the mtDNA polymorphism of the species is within local populations, with limited geographic or "race" differentiation.

Table 5 gives the per nucleotide difference,  $d_{xy}$ , for all pairwise comparisons between the 14 taxa surveyed. If *D. niveifrons* and *D. immigrans*, as well as taxon *H* are excluded, the matrix is fairly homogeneous, in that the average  $d_{xy}$  value between one species and all others is about the same for all species (about 3%). *Drosophila niveifrons* and taxon *H* are more differentiated from the other taxa in the *nasuta* subgroup (and from each other). *Drosophila immigrans*, which is in a different subgroup, is the most differentiated; on the average, the per site nucleotide differentiation between *D. immigrans* and any *nasuta* species is  $d_{xy} = 13.0\%$ .

Figure 4 gives the phylogeny of the species obtained from the data displayed in Table 5 (but combining the information for the four subspecies of D. sulfurigaster) following the procedure described in the Materials and Methods. Drosophila nasuta and D. albomicans appear as very closely related to each other, and both closely related to D. pallidifrons. These three are in turn closely related to the triad made up by D. kepulauana, D. kohkoa, and D. sulfurigaster. Drosophila pulaua (and taxon F) are also closely related to the other named species, whereas taxon H is more different, and D. niveifrons still more. As expected, the outgroup reference species, D. immigrans, is the most differentiated. These relationships agree, on the whole, with those shown in Fig. 3, which is based on a more robust approach that does not depend on questionable statistical assumptions (such as the independence and additivity of the distance value) or on the constancy of evolutionary rates.



**Fig. 4.** Phylogeny of the *Drosophila nasuta* group based on the matrix of genetic distances between the species. The clustering method used (KITSCH computer program) assumes a constant rate of evolution and that genetic distance values are independent and additive.

### Discussion

The species of the Drosophila nasuta subgroup exhibit distinctive characteristics that make them unusually interesting for evolutionary studies. These characteristics include: little morphological differentiation between species in spite of their distribution over an enormous territory; pairs of species with none or little sexual isolation and able to produce fertile hybrids; and substantial evolution in the configuration of the metaphase chromosomes, with polymorphism within single species. Extensive morphological, cytological, reproductive, and allozymic studies have provided little discriminant information about the evolutionary relationships among the species. The present investigation of the mtDNA restriction patterns is a first step toward resolving the phylogeny of this group of species.

The D. nasuta subgroup consists of about a dozen known species, classified within the D. immigrans group. The first described species of the subgroup, D. nasuta, was first collected in the Seychelles Islands, off the eastern coast of equatorial Africa (Lamb 1914), but its known distribution extends from Kenya and Madagascar to Sri Lanka and India. [Drosophila nasuta has recently been found on the west coast of Africa, in Dahomey, Cameroun, and Congo, but it seems likely that its presence there is due to recent introductions associated with human activity (see David and Tsacas 1980).]

The geographic distribution of the *nasuta* group extends throughout the Indo-Pacific region, from Hawaii through Southeast Asia and India to eastern Africa, and from tropical Australia and New Guinea to the Ryukyu islands of Japan and to southern China (Kitagawa et al. 1982 and references therein). The center of the distribution is in Southeast Asia,



Fig. 5. MtDNA digests obtained with the endonuclease *Hind* III. Starting from the left, lane 1 is pattern M (see Fig. 2); lanes 2–6 are, respectively, patterns W, P, U, S, and O; lane 7 is  $\lambda$ DNA that provides a size standard; lane 8 is pattern S.

Malaysia, Borneo, and the Philippines, where several species are sympatric and very abundant.

The *nasuta* species are morphologically quite similar, so that individuals of different species have sometimes been classified under a single name. (The opposite has also occurred; members of the same species from geographically distant regions have received different species names.) Kitagawa et al. (1982) have conducted morphometric studies of 15 traits. These, however, have yielded little useful information for species discrimination purposes, even after sophisticated statistical analysis: the most definitive result is that two species, *D. nasuta* and *D.* 

**Table 5.** Genetic differentiation  $(d_{xy})$  between taxa of the *D. nasuta* group

Taxon	2	3	4	5	6	7	8	9	10	11	12	13	14	Aver- age
<ol> <li>D. albomicans</li> <li>D. kepulauana</li> <li>D. kohkoa</li> <li>D. nasuta</li> <li>D. pallidifrons</li> <li>D. pulaua</li> <li>D. s. albostri- gata</li> <li>D. s. bilimbata</li> <li>D. s. sulfuri- gaster</li> <li>Taxon F</li> <li>D. niveifrons</li> </ol>	0.023	0.030 0.022	0.010 0.026 0.030	0.019 0.030 0.037 0.023	0.039 0.041 0.030 0.034 0.055	0.007 0.016 0.022 0.010 0.015 0.041	0.041 0.033 0.022 0.045 0.048 0.023 0.033	9 0.041 0.033 0.022 0.045 0.048 0.023 0.033 0.000	0.041 0.033 0.022 0.045 0.048 0.023 0.033 0.000 0.000	0.039 0.041 0.030 0.034 0.055 0.000 0.041 0.023 0.023	0.073 0.068 0.057 0.079 0.061 0.074 0.068 0.068 0.068 0.068	0.040 0.053 0.041 0.033 0.048 0.060 0.033 0.053 0.053 0.053 0.060 0.090	0.150 0.146 0.135 0.151 0.129 0.122 0.146 0.090 0.090 0.090 0.122 0.129	age 0.043 0.043 0.038 0.043 0.047 0.043 0.037 0.037 0.037 0.037 0.043 0.075
<ul><li>13. Taxon H</li><li>14. D. immigrans</li></ul>								_					0.153	0.059 0.127

*albomicans* have, on the average, larger bodies than the others.

The males of all the *nasuta* species but one are distinguishable from other species groups, because they have white to silvery markings on the frons; the exception is *D. pallidifrons*. These markings are the most informative trait also for discriminating among the species of the group. *Drosophila sulfurigaster* has pronounced whitish bands along the frontal orbits; *D. pulaua* has similar bands but much fainter; *D. nasuta*, *D. albomicans*, *D. kohkoa*, *D. kepulauana*, and *D. niveifrons* have a white to silvery sheen over the entire frons that is less conspicuous in *D. kohkoa* than in the rest (Wilson et al. 1969; Okada and Carson 1982).

The metaphase configuration of the *nasuta* chromosomes has, like other species of the *immigrans* group, a very long rod-shaped chromosome 3 that is composed of two of the five long rods characteristic of *Drosophila* (and which in species such as *D. melanogaster* are associated in a V-shaped chromosome, having the centromere in the middle, rather than near one end as in the *nasuta* species). In addition, the haploid chromosomal complement has a rod X (or Y), a V-shaped chromosome 2, and a dotlike chromosome 4.

The configuration of the Y chromosome is highly polymorphic among and within species. In *D. sulfurigaster*, the Y chromosome is sometimes a rod, nearly as long as the X, and with a small constriction near the centromere. However, in the subspecies *D. s. albostrigata* there are two additional types of Y chromosome, J-shaped and V-shaped, with all three types coexisting in some localities; whereas in the subspecies *D. s. sulfurigaster* the Y chromosome is V-shaped in Australia and J-shaped in New Guinea and New Ireland. A J-shaped Y chromosome is also found in some strains of *D. kepulauana*, although most have the standard rod-shaped Y. *Drosophila kohkoa* strains often have this standard rod-shaped Y but some strains have small V- or J-shaped Y chromosomes.

The fourth chromosome has some heterochromatin added in *D. kepulauana* and still more in some strains of *D. kohkoa* in which this "dot" chromosome looks like a "comma" or small J. In *D. albomicans*, the fourth chromosome looks like a rod, somewhat smaller than the other chromosomes, owing to the incorporation of substantial amounts of heterochromatin.

The most bizarre configuration of the metaphase chromosomes is found in *D. albomicans*, in which the double-length chromosome 3 is attached to the X or the Y, so that three of the rodlike elements of the *Drosophila* genome are fused into one chromosome (and with the consequence that the third chromosome genes are sex-linked). Hence, *D. albomicans* has only two autosomes: a "dot" that is largely heterochromatic and a V-shaped chromosome. The majority of the genome (about 60%) is inherited in a sex-linked fashion and is contained within one chromosome.

The reproductive relationships between species are extremely complex (Wilson et al. 1969; Nirmala and Krishnamurthy 1973, 1974; Kitagawa et al. 1982; Chang 1984). Drosophila nausta, D. albomicans, and D. kepulauana produce fertile hybrid progeny when intercrossed. Crosses between the subspecies of D. sulfurigaster produce fertile hybrids in most cases, although the hybrid males, or both males and females, are sterile in some combinations. The least reproductively isolated species is D. albomicans, whose females produce progeny with the males of all other species except D. s. neonasuta and taxon F (*D. niveifrons* and taxon H have not been tested), and whose males produce progeny with females of all other species except the two mentioned and *D. pallidifrons*. The most isolated species are taxon F and *D. pallidifrons*, closely followed by *D. nasuta* and *D. s. neonasuta*.

None of the studies just reviewed provides a clear picture of the evolutionary history of the *D. nasuta* subgroup. Nor are any significant clues forthcoming from polytene chromosome studies (Wilson et al. 1969), courtship behavior (Spieth 1969), or the limited published data on enzyme polymorphisms (Kanapi and Wheeler 1970; Kitagawa et al. 1982). The restriction patterns of the mtDNA herein reported provide, however, significant information about the phylogeny of the *nasuta* species.

Figures 3 and 4 suggest that most species (i.e., excepting *D. niveifrons* and taxon *H*) are very closely related. Indeed, some morphs (I and II) are each shared by two species, and some interspecific differences are smaller than some intraspecific ones. In particular, morph V present in *D. albomicans* from Wulai differs by four mutational steps from morph I, found in the same locality; and it differs by 3–5 mutations from any morph found in the species. The *D. s. albostrigata* mtDNA differs by four mutational steps from the morph found in any of the other three *D. sulfurigaster* subspecies.

The small differentiation between the mtDNAs of the *nasuta* species is consistent with their morphological similarity and restricted reproductive isolation. It is consistent as well with the hypothesis that speciation in the *nasuta* group may have been largley impelled by chromosomal reorganization an atypical situation in the genus *Drosophila*.

The distinct geographic characteristics of the region where the *nasuta* species are endemic are worth noticing. There are numerous small islands as well as a number of large islands (Borneo, Java, Sumatra, Philippines, etc.) and peninsulas, some amounting to subcontinents (India, Indochina). It is tempting to speculate that this sort of environmental fractionation may have facilitated the speciation by chromosomal reorganization that is distinctive of the *nasuta* group.

We surmise as a working hypothesis that *D. al*bomicans may represent the ancestral lineage of the nasuta group. We base this suggestion on the joint consideration of two facts. One is that *D. albomicans* is the least reproductively isolated species; i.e., it produces hybrids (and often fertile ones) when crossed with almost any other species. The second part is the central position of morph I in the mtDNA network (Fig. 3). The network lacks directionality by itself, so that the most ancestral morph could be any one, whether centrally or marginally located. But it is apparent that no matter which morph is arbitrarily chosen as the ancestral one, the derivation of most other morphs requires morph I as the intermediate; in fact, morph I has to be ancestral to all others, except four at the most. If we accept the hypothesis of *D. albomicans* as the ancestral lineage of the group, it would seem likely that it became separated from other *immigrans*-like species as a consequence of the fusion of the long-rod chromosome 3 with the sex chromosome. The evolution of the other species would have been accompanied then by the break of this unusual association between the sex and the third chromosomes.

Figures 3 and 4 show that, as expected, D. immigrans is much more different from the nasuta group species than these are from each other. Figure 3 also suggests that the phylogenetic connection between D. immigrans and the nasuta species occurred through D. sulfurigaster. This suggestion must, however, be received with caution. Drosophila immigrans may also be connected to the nasuta species through D. niveifrons, by requiring two more mutations than through D. sulfurigaster. The restriction patterns used for reconstructing the network become more imprecise as the mtDNAs become more different, so that some mutations may go undetected. In any case, the connection between immigrans and the *nasuta* group through *D. sulfurigaster* is consistent with the hypothesis advanced earlier that D. albomicans and morph I are ancestral to most of the others. One question that remains completely unresolved is whether the *D. immigrans* lineage is ancestral to the *nasuta* lineages or rather derives from one of them.

The levels of mtDNA sequence polymorphism within local populations of D. albomicans (and D. *kohkoa*) is  $\pi = 1.0\%$ , comparable to the amount of polymorphism found in the Old World populations of D. subobscura ( $\pi = 1.1\%$ ; Latorre et al. 1986). This amount of polymorphism falls within the typical range found in animals (see Avise and Lansman 1983; Brown 1985; Boursot and Bonhomme 1986). Humans are toward the lower end of that range:  $\pi$ = 0.23-0.47% (Brown 1980, 1985; Cann et al. 1987; but see Aquadro and Greenberg 1983). The average differentiation between local populations of D. al*bomicans* ( $d_{xy} = 1.0\%$ ) is about three times greater than the differentiation between major human racial groups ( $d_{xy} = 0.27-0.45\%$ ; Cann et al. 1987), but within the typical range for most animals (Latorre et al. 1986; Saunders et al. 1986). Examples of animals with none or little geographic differentiation are the American eel Anguilla rostrata (Avise et al. 1986) and D. simulans (Baba-Aissa and Solignac 1984).

Whether there is a molecular evolutionary clock

for mtDNA is an unsettled issue. There is now convincing evidence that the rates of evolution of mitochondrial and nuclear DNA are fairly similar in some groups of organisms, such as Drosophila and sea urchins, but very different in others, such as birds and primates (Powell et al. 1986; Vawter and Brown 1986). The conclusion follows that the rate of evolution of either the mitochondrial or the nuclear DNA (or both) is not constant among groups of organisms. Estimates of the rate of mtDNA evolution are heterogeneous even within a group of organisms. For example, the estimated rate of nucleotide substitution in mtDNA ranges from 0.25 to 2.5% per site per million years for mammals (Brown et al. 1982; Nei 1982, 1985; Hasegawa et al. 1985), and from 0.24 to 6.2% for Drosophila (Solignac et al. 1986).

The heterogeneities just mentioned cast doubt on any attempt to infer the time of divergence between species on the basis of mtDNA differences. It is often assumed nevertheless that a rate of 1% nucleotide substitution per site per million years may be roughly correct. Taking this rate as a working hypothesis and for discussion purposes, we conclude that D. nasuta and D. albomicans diverged about 500,000 years ago (see Table 5; notice that with the assumed rate, the number of million years since the divergence of two species is one-half the value of  $d_{xy}$  in percent). Similarly, D. niveifrons diverged from the other species of the nasuta group about 3-5 million years (Myr) ago; whereas most lineages of the subgroup would have diverged from each other between 1 and 2 Myr ago. The divergence between the nasuta species and D. immigrans occurred about 6-8 Myr ago.

The mtDNA nucleotide diversity  $(d_{xy})$  between sibling species of the D. melanogaster group ranges from 4.9 to 8.8% (Solignac et al. 1986) that corresponds to divergence times between 2.5 and 4.4 million years. Interspecific DNA hybridization between the mtDNAs of D. melanogaster and its sibling D. yakuba yields a divergence time between 3 and 4 million years (Powell et al. 1986). Most of the D. nasuta species are morphologically more readily distinguishable from each other than the sibling species of the D. melanogaster group; yet, their mtDNAs are less differentiated on the average. If we assume that the rates of mtDNA evolution are the same in both groups, we must conclude that morphological evolution has been slower in the D. melanogaster group. Alternatively, we may infer that morphological evolution has proceeded no faster in the D. nasuta than in the D. melanogaster group, but mtDNA evolution has been much faster among the melanogaster species. Be that as it may, comparison of the two groups leads to the conclusion

that morphological and molecular evolution do not always proceed at proportional rates—a discrepancy that has been noted previously for other sets of data (King and Wilson 1975; Cherry et al. 1978).

Acknowledgments. We are greatly indebted to Dr. Fei-Jann Lin for most of the *D. albomicans* strains; to Dr. Osamu Kitagawa for three strains of *D. albomicans* and all other strains of the various *nasuta* species, as well as for comments on this manuscript; to Dr. Wyatt W. Anderson for the *D. immigrans* strain; and to Dr. Masatoshi Nei for statistical advice. The research was supported by US Department of Energy Contract AS 03-76 EV 70200.

## References

- Aquadro CF, Greenberg BD (1983) Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. Genetics 103:287–312
- Avise JC, Lansman RA (1983) Polymorphism of mitochondrial DNA in populations of higher animals. In: Nei M, Koehn RK (eds) Evolution of genes and proteins. Sinauer, Sunderland MA, pp 147–164
- Avise JC, Lansman RA, Shade RO (1979) The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. I. Population structure and evolution in the genus *Peromyscus*. Genetics 92:279-295
- Avise JC, Bermingham B, Kessler LG, Saunders NC (1984) Characterization of mitochondrial DNA variability in a hybrid swarm between subspecies of bluegill sunfish (Lepomis macrochirus). Evolution 38:931-941
- Avise JC, Helfman GS, Saunders NC, Stanton Hales L (1986) Mitochondrial DNA differentiation in North Atlantic eels: population genetic consequences of an unusual life history pattern. Proc Natl Acad Sci USA 83:4350–4354
- Baba-Aissa, Solignac M (1984) La plupart des populations de Drosophila simulans ont probablement pour ancetre un femelle unique dans un passe recent. CR Seances Acad Sci 299: 289-292
- Boursot P, Bonhomme F (1986) Genetique et evolution du genome mitochondrial des metazoaires. Genet Sel Evol 18: 73-98
- Brown WM (1980) Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. Proc Natl Acad Sci USA 77:3605–3609
- Brown WM (1983) Evolution of animal mitochondrial DNA. In: Nei M, Koehn RK (eds) Evolution of genes and proteins. Sinauer, Sunderland MA, pp 62–88
- Brown WM (1985) The mitochondrial genome of animals. In: MacIntyre RJ (ed) Molecular evolutionary genetics. Plenum, New York, pp 95-130
- Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial DNA sequences of primates: tempo and mode of evolution. J Mol Evol 18:225-239
- Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. Nature 325:31-36
- Chang H (1984) Study of hybridization and phylogenetic relationship between a pair of sibling species, *Drosophila nasuta* Lamb and *D. albomicans* Duda. Dissertation, University of California, Davis CA
- Cherry LM, Case SM, Wilson AC (1978) Frog perspective on the morphological difference between humans and chimpanzees. Science 200:209-211
- Coen ES, Thoday JM, Dover G (1982) Rate of turnover of

structural variants in the rDNA gene family of Drosophila melanogaster. Nature 195:564-568

- David JR, Tsacas L (1980) Cosmopolitan, subcosmopolitan and wide-spread species: different strategies within the drosophilid family (Diptera). CR Seances Soc Biogeogr 57:11-26
- DeSalle R, Val Giddings L, Kaneshiro KY (1986) Mitochondrial DNA variability in natural populations of Hawaiian Drosophila. II. Genetic and phylogenetic relationships of natural populations of D. silvestris and D. heteroneura. Heredity 56: 87–96.
- Felsenstein J (1984) Distance methods for inferring phylogenies: a justification. Evolution 38:16-24
- Hale LR, Beckenbach AT (1985) Mitochondrial DNA variation in *Drosophila pseudoobscura* and related species in Pacific Northwest populations. Can J Genet Cytol 27:357-364
- Hasegawa M, Kishino H, Yano T (1985) Dating of the humanape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174
- Kanapi CG, Wheeler MR (1970) Comparative isozyme patterns in three species of the *Drosophila nasuta* complex. Texas Rep Biol Med 28:261–278
- King MC, Wilson AC (1975) Evolution at two levels: molecular similarities and biological differences between humans and chimpanzees. Science 188:107-116
- Kitagawa O, Wakahama K, Fuyama Y, Shimada Y, Takanashi E, Hatsumi M, Uwabo M, Mita Y (1982) Genetic studies of the *Drosophila nasuta* subgroup, with notes on distribution and morphology. Jpn J Genet 57:113-141
- Lamb CG (1914) Diptera: Heteroneuridae, Ortalidae, Trypetidae, Sepsidae, Micropezidae, Drosophilidae, Geomyzidae, Milichiidae of the Seychelles. Trans Linn Soc London 16: 307–372
- Lansman RA, Shade RO, Shapira JF, Avise JC (1981) The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. J Mol Evol 17: 214-226
- Latorre A, Moya A, Ayala FJ (1986) Evolution of mitochondrial DNA in *Drosophila subobscura*. Proc Natl Acad Sci USA 83:8649–8653
- Latorre A, Barrio E, Moya A, Ayala FJ (1988) Mitochondrial DNA evolution in the *Drosophila obscura* group. Mol Biol Evol 50:717-728
- Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor NY
- Nei M (1982) Evolution of human races at the gene level. In: Bonne-Tamir B, Cohen P, Goodman M (eds) Part A: the unfolding genome. Alan R Liss, New York, pp 167–181
- Nei M (1985) Human evolution at the molecular level. In: Ohta T, Aoki K (eds) Population genetics and molecular evolution. Japan Scientific Societies Press, Tokyo, pp 41–64

- Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York
- Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269-5273
- Nei M, Tajima F (1981) DNA polymorphism detectable by restriction endonucleases. Genetics 97:145-163
- Nirmala SS, Krishnamurthy NB (1973) Drosophila neonasuta, a new species of Drosophila from Mysore (Diptera: Drosophilidae). Orient Insects 7:267-270
- Nirmala SS, Krishnamurthy NB (1974) Endophenotypic variability in natural populations of *Drosophila nasuta*. Egypt J Genet Cytol 3:211-228
- Okada T, Carson H (1982) Drosophilidae associated with flowers in Papua, New Guinea. III. Zingiberales. Kontyu, Tokyo 50:396-410
- Powell JR, Caccone A, Amato GD, Yoon C (1986) Rates of nucleotide substitution in *Drosophila* mitochondrial DNA and nuclear DNA are similar. Proc Natl Acad Sci USA 83:9090– 9093
- Ramachandra NB, Ranganath HA (1985) Supernumerary chromosomes in Drosophila nasuta albomicana. Experientia 41: 680–681
- Ramachandra NB, Ranganath HA (1986) The chromosomes of two Drosophila races: D. nasuta nasuta and D. nasuta albomicana. Chromosome 93:243-248
- Saunders NC, Kessler LG, Avise JC (1986) Genetic variation and geographic differentiation in mitochondrial DNA of the horseshoe crab, *Limulus polyphemus*. Genetics 112:613-627
- Solignac M, Monnerot M, Mounolou JC (1986) Mitochondrial DNA evolution in the *melanogaster* species subgroup of *Dro-sophila*. J Mol Evol 23:31–40
- Spieth HT (1969) Courtship and mating behavior of the Drosophila nasuta subgroup of species. Univ Texas Publ 6918: 255-270
- Takahata N, Palumbi SR (1985) Extranuclear differentiation and gene flow in the finite island model. Genetics 109:441– 457
- Upholt WB (1977) Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. Nucleic Acids Res 4:1257-1265
- Vawter L, Brown WM (1986) Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. Science 234:194–196
- Wilson FD, Wheeler MR, Harget M, Kambysellis M (1969) Cytogenetic relations in the Drosophila nasuta subgroup of the immigrans group of species. Univ Texas Publ 6918:207– 253

Received August 23, 1987/Revised August 22, 1988