Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 9558–9562, December 1990
Genetics

Evolution of *Drosophila* mitochondrial DNA and the history of the *melanogaster* subgroup

(paleopopulation biology/maternal inheritance/evolutionary rate/population structure)

YOKO SATTA AND NAOYUKI TAKAHATA

National Institute of Genetics, Mishima 411, Japan

Communicated by Motoo Kimura, August 24, 1990

ABSTRACT **The nucleotide sequences of a common region of 15 mitochondrial DNAs (mtDNAs) sampled from the *Drosophila melanogaster* subgroup were determined. The region is 2527 base pairs long, including most of the NADH dehydrogenase subunit 2 and cytochrome oxidase subunit 1 genes punctuated by three tRNA genes. The comparative study revealed (*i*) the extremely low saturation level of transitional differences, (*ii*) recombination or variable substitution rates even within species, (*iii*) long persistence times of distinct types of mtDNA in *Drosophila simulans* and *Drosophila mauritiana*, and (*iv*) an apparent lack of within-type variations in island species. Also found was a high correlation among the transitional rate, the saturation level, and the G+C content (or codon usage). It appears that *D. simulans* and *D. mauritiana* have maintained highly structured populations for more than 1 million years. Such structures are consistent with the origination of *Drosophila sechellia* from *D. simulans*. Yet geographic isolation is so weak as to show no evidence for further speciation. Moreover, one type of mtDNA shared by *D. simulans* and *D. mauritiana* suggests either recent divergence or ongoing introgression.**

In general, mitochondrial DNA (mtDNA) in higher animals is transmitted mainly through female gametes (1). Although there are a number of copies within a cell, the population tends to be genetically homogeneous (homoplasmic), and heteroplasmic individuals produced by paternal leakage are expected to be unusual (2). An immediate consequence is that mtDNA recombinants are rare and the genome tends to be in linkage disequilibrium (3). In *Drosophila*, however, two types of mtDNA carried by a single individual differed from each other by 19 base pairs (bp) scattered over 975 bp, so the heteroplasmy could not be accounted for by a single mutational event (4). A more likely explanation is incomplete maternal inheritance of mtDNA and, in fact, a subsequent backcross experiment (5) revealed that the average paternal contribution per fertilization is as high as 0.1%. In light of these findings, it is questionable if recombination or gene conversion can be ignored for understanding the evolution of *Drosophila* mtDNA.

It was noted that the extent of mtDNA variation in *Drosophila simulans* and *Drosophila mauritiana* is much higher than that in *Drosophila melanogaster* (4, 6). In each of the former two species, there are at least two distinct types of mtDNA, the nucleotide differences being more than 1%, but in the latter there is only one type. This characteristic may be applied also to nuclear genes. Although there are no substantial differences between *D. simulans* and *D. melanogaster* in allozyme heterozygosity or the extent of amino acid changes (7), the silent nucleotide differences at the alcohol dehydrogenase locus in *D. melanogaster* (8) tend to be smaller than those in *D. simulans* (9), and this was confirmed by a recent exhaustive sequence analysis of an 11-kilobase (kb) region encompassing the locus (10). In addition, a 40-kb region around the rosy and snake loci analyzed by restriction enzymes indicated that *D. simulans* is 6 times as variable as *D. melanogaster* (11). Since the two species are similar in their geographic and climatic distributions (12), it is puzzling how such large genetic differences arose.

Striking peculiarities of *Drosophila* mtDNA are its extremely low content of G and C residues at the third codon positions (13–15) and small sequence differences among diverged species (16). It was asked whether such small differences are an indication of a low rate of mtDNA divergence or whether they are the result of the saturation of mtDNA changes at a low ceiling (16). From comparisons of distantly related species, however, it was difficult to determine which is a more likely explanation.

In this paper, we address these questions based on sequence information on 15 mtDNAs, each about 2.5 kb long, sampled from the *D. melanogaster* subgroup.*

MATERIAL AND METHODS

We analyzed four closely related species of the *melanogaster* subgroup. *D. melanogaster* and *D. simulans* are cosmopolitan, while *D. mauritiana* and *Drosophila sechellia* are endemic to Mauritius and the Seychells. It is known from the restriction enzyme analysis (6) that while there exists only one distinct type of mtDNA in each of *D. melanogaster* and *D. sechellia*, there are three in *D. simulans* and two in *D. mauritiana*. We designate these seven types by L1 for *D. melanogaster*, L2 for *D. sechellia*, L3, L4†, and L5 for *D. simulans*, and L4 and L6 for *D. mauritiana*. We use similar symbols, L4† and L4, to indicate the fact that L4† occurring in La Réunion and Madagascar and L4 occurring in Mauritius are virtually identical (4).

We chose only one strain for each of types L1, L3, and L6. These strains were established from single inseminated females collected from Japan, Hawaii, and Mauritius, respectively. For L2, L4, L4†, and L5, we sampled several strains with different geographic origins to examine the within-type variation. Two strains of L2 came from the Seychells, three of L4† came from Réunion and Madagascar, three of L5 came from Réunion, Madagascar, and Kenya, and four of L4 came from Mauritius. The total number of strains sampled was thus 15. In the subsequent data analysis, we also used the known sequences of *D. melanogaster* (13, 14) and *Drosophila yakuba* (15).

We isolated mtDNA from each of the above strains and employed the standard strategies of cloning and sequencing

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: ND2, NADH dehydrogenase subunit 2; COI, cytochrome oxidase subunit 1.

*The sequences from which the results reported in this paper have been derived have been deposited in the GenBank data base (accession nos. M57907 through M57912).

(e.g., ref. 17). Three different subclones used for sequencing were derived from the 4.8-kb *Hin*dIII cloned fragment and covered position 38 (*Hin*dIII site) to 1056 (*Tha* I site), 1056 either to 1966 (*Xho* I site) or to 1780 (*Sac* I site), and 1966 (or 1780) to 2564 (*Pvu* II site). The numbering of nucleotides is that of de Bruijn (13). Thus, the region sequenced is 2527 bp long and contains most of the NADH dehydrogenase subunit 2 (ND2) (822 bp) and cytochrome oxidase subunit 1 (COI) (1494 bp) genes punctuated by a noncoding region (20 bp) and three tRNA genes (186 bp): $tRNA^{Trp}$, $tRNA^{Cys}$, and $tRNA^{Tyr}$.

RESULTS

In total, there are 207 variable sites in the common region of 2527 bp. Among these variable sites, two are located at the putative initiation codon of COI consisting of four nucleotides (17), three are in the three tRNA genes, and two are in the 20-bp noncoding region. In what follows, we shall focus on the remaining 200 variable sites located in the 822-bp ND2 and 1494-bp COI regions.

Two remarkable features were immediately clear. First, there is not even a single nucleotide change within types L2, L4, $L4^{\dagger}$, and L5. Restriction enzyme analyses for a larger sample consistently revealed the same result. The analyses also showed that L6 does not have any derivative, although L3 has minor variants (18). Thus, this monomorphism is a common feature in *D. sechellia*, *D. simulans*, and *D. mauritiana*. Second, $L4^{\dagger}$ in *D. simulans* has a sequence identical to L4 in *D. mauritiana* except at position 2153. At this third position of a tryptophan codon, $L4^{\dagger}$ carries a G residue while the remaining types, including L4, have an A residue, so that it is most likely that a synonymous transition substitution occurred in the lineage of $L4^{\dagger}$.

If we exclude $L4^{\dagger}$, there are six distinguishable types of mtDNA in our sample, which differ at 199 nucleotide sites (Fig. 1). In any pairwise comparison, the proportion of synonymous changes is about 90% and there is no correlation of this with the total number of nucleotide changes. It is therefore expected that most changes, if not all, are likely to be selectively neutral (19, 20). However, the COI gene is more prone to synonymous changes than is the ND2 gene,

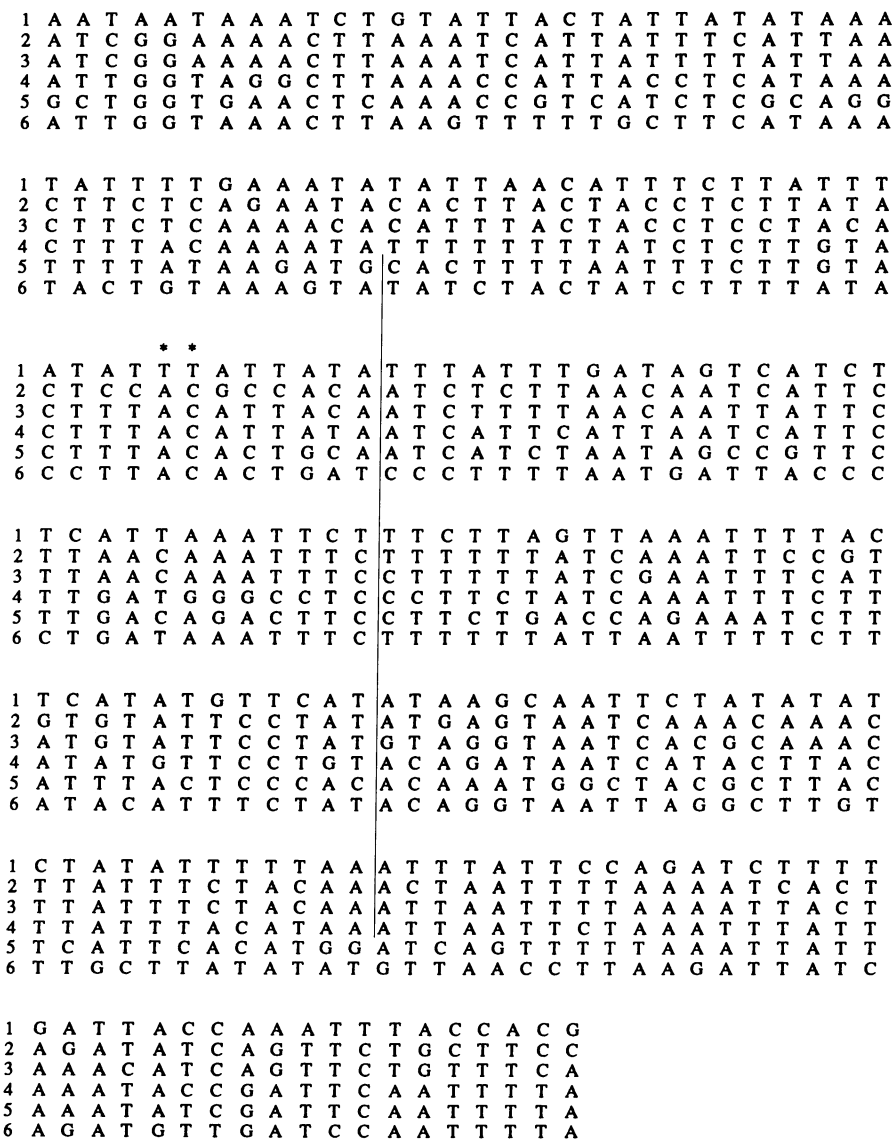

FIG. 1. The 199 variable sites among six homologous segments of 2316 bp (most of the ND2 and COI genes) that were taken from distinct types of mtDNA in *D. melanogaster* (L1), *D. sechellia* (L2), *D. simulans* (L3 and L5), and *D. mauritiana* (L4 and L6). Types $L4^{\dagger}$ and L4 correspond to siII and maI, L5 to siI, and L6 to maII in ref. 6. The asterisks mark the boundary between the ND2 and COI genes. The nucleotide positions of these 199 variable sites are available from the authors upon request and the entire nucleotide sequences are available from GenBank.

Table 1. Numbers of nucleotide differences in 822-bp ND2 (above the diagonal) and 1494-bp COI (below the diagonal) genes

	L1	L2	L3	L4	L5	L6	L7
L1		21 (15)	22 (13)	23 (13)	21 (20)	18 (13)	43 (17)
L2	30 (38)		2 (6)	14 (6)	17 (13)	12 (12)	43 (22)
L3	27 (39)	3 (18)		14 (6)	17 (15)	12 (12)	43 (18)
L4	30 (35)	14 (35)	11 (29)		9 (13)	10 (11)	43 (19)
L5	32 (51)	16 (44)	13 (39)	8 (35)		13 (19)	46 (24)
L6	35 (43)	19 (39)	18 (33)	15 (34)	17 (42)		35 (24)
L7	67 (41)	62 (46)	61 (45)	62 (45)	60 (60)	63 (52)	

Numbers not in parentheses exclude the transitions at the third codon positions. Transitional changes at the third codon positions are in parentheses.

and Satta and Chigusa (21) discussed the relationships between substitution rates and biases in codon usage. As usual, most synonymous changes are caused by transitions at the third codon positions, but the extent of this bias is extreme in *Drosophila* and mammalian mtDNA. To examine the saturation level of transitional differences, we used the *D. yakuba* sequence (15), finding that the level may be as low as 5–7% in the case of ND2 and 7–9% in the case of COI (Table 1). This was confirmed by the COI sequence of honeybee (22), which clearly exhibits that the transitional differences between the *melanogaster* subgroup and *D. yakuba* reached a plateau.

The distance matrix based on all 199 variable sites indicated that L4 is almost equidistant from L5 and L6. This peculiarity suggested that the evolutionary process was influenced by recombination or gene conversion, or that different mitochondrial types even within a species evolved at different rates. Making various window sizes and sliding intervals, we classified the ND2 and COI region into two subregions in which topologically the same tree was obtained: In one subregion, consisting of two separate blocks (position 316–1516 and 2003–2564), L4 and L5 are more closely related, while in the other subregion, consisting of the other two blocks (38–315 and 1517–2002), L4 and L6 are. However, since the topological relationships between L4 and L1, L2, or L3 remain unchanged throughout the whole region, the peculiarity appeared to be caused by L5 and L6.

Through the present analysis and previous studies (4, 6) it became clear that there are remarkable differentiations among the different types of mtDNA in *D. simulans* and *D. mauritiana*. For instance, among 2316 bp, there are 84 nucleotide differences between L3 and L5 in *D. simulans*, the proportion being 3.6%. Similarly, there are 70 nucleotide differences between L4 and L6 in *D. mauritiana*, the proportion being 3.0%. It is therefore interesting to see how much variation exists within a given type or within a species

Table 2. Percent nucleotide differences (d) within each of seven types of mtDNA estimated from restriction enzyme and sequence analyses

	Restriction analysis			Sequence analysis		
mtDNA	n	m	d	n	m	d
L2	16	140	0	2	2527	0
L3	40[a]	200	0.03	1	2527	—
L5	20	72–96	0	3	2527	0
L4†	15	70	0	3	2527	0
L4	18	178	0	4	2527	0
L6	6	156	0	1	2527	—
L1	81	168	0.2	2[b]	2527	0.6
				7[c]	976	0.2

n = sample size, m = the number of nucleotide sites examined.
[a]After ref. 18.
[b]One in the present study and the other after ref. 13.
[c]After ref. 24.

Table 3. Weighted averages of percent nucleotide differences within and between species

	D. simulans	*D. sechellia*	*D. mauritiana*	*D. melanogaster*
D. simulans	1.6	3.1	2.8	5.0
D. sechellia		0	3.1	4.5
D. mauritiana			1.1	4.5
D. melanogaster				0.2

The frequencies of L3, L4†, and L5 in the natural population of *D. simulans* are 0.278, 0.083, and 0.639, respectively (18), and those of L4 and L6 in the natural population of *D. mauritiana* are 0.75 and 0.25.

as a whole. Table 2 summarizes the results: whereas in *D. melanogaster* only one type exists and the within-type variation is 0.2%, there is no variation within each type in *D. sechellia*, *D. simulans*, and *D. mauritiana*, as mentioned earlier. Taking into account the relative frequencies of types in natural populations, we computed the weighted average of percent nucleotide differences within a species (Table 3). *D. sechellia* is the least variable species (0%), *D. melanogaster* (0.2%) the second, *D. mauritiana* (1.1%) the third, and *D. simulans* (1.6%) the most.

DISCUSSION AND CONCLUSION

The equidistant relationship of L4 to L5 and L6 may be caused by recombination or gene conversion between L5 and L6, or by variable substitution rates among lineages. To examine these possibilities, we first assumed the molecular clock hypothesis or rough constancy of substitution rates, and therefore constructed a UPGMA tree (23) for each of the subregions. Trees A and B in Fig. 2 are statistically different from each other in the sense in ref. 9. Clearly, either recombination or gene conversion can result in mosaic structures of L5 and L6 with respect to their ancestries, but recombination is a more parsimonious explanation. It should be noted, however, that this conclusion depends on the assumption of the molecular clock hypothesis. Hence, for each subregion, we examined NJ trees (24), which take account of variable rates. Although the topology depends on whether or not the *D. yakuba* sequence is included, the substitution rate leading to L5 is always estimated about twice as fast as compared with others. This stems from an excess of transitional changes at the third codon positions of L5.

Thus, to examine the extent and rate constancy of transitional differences within the *melanogaster* subgroup, we again constructed an NJ tree after excluding the transitions at the third codon positions. Effectively, the tree is based on the transversions and on the frequent C ↔ T transitions at the first codon positions. The topology of this tree is the same as that in Fig. 2*B* and the branch lengths show rough rate constancy for those changes. Given this topology, we inferred transitional changes in each branch by a least-square method (Fig. 3). It is noteworthy that more transitional changes tend to be assigned to recent branches. This is consistent with the expected rapid rate and low saturation level of transitions but is inconsistent with the result of DNA·DNA hybridization (25). It is also clear that the transitional rate differs from lineage to lineage and between the ND2 and COI genes. These differences may be related to the G+C content at the third codon positions and/or codon choice patterns (17, 21). The G+C content also varies from lineage to lineage and, in fact, there is a striking correlation between the G+C content and the estimated transitional rate (Fig. 4). Type L1, with the lowest G+C content (3.3% for ND2 and 4.8% for COI), accumulates transitional changes most slowly, while L5, with the highest G+C content (7.6%

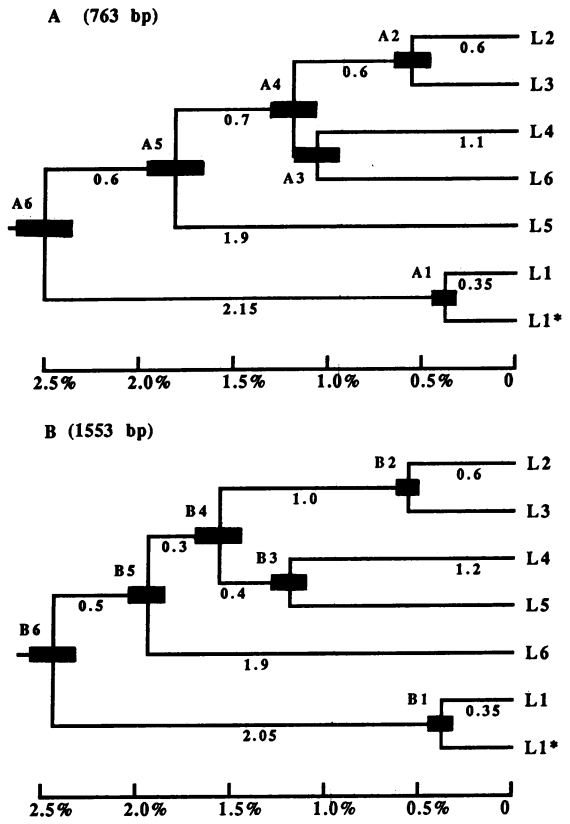

FIG. 2. UPGMA trees (23) for subregions A (position 38–315 and 1517–2002) and B (position 316–1516 and 2003–2564). A bar at each branching point shows one standard error (see ref. 9). *, based on the sequence in ref. 13.

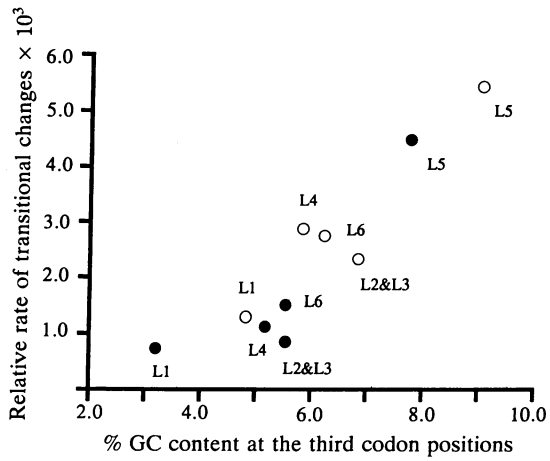

FIG. 4. Relationship between the transitional rate at the third codon positions and the G+C content for five lineages of mtDNA. The relative rate is defined by the ratio of transitional changes per third position to the total nucleotide differences except the transitions, which are estimated in Fig. 3. ○, COI; ●, ND2. To avoid any correlation produced by shared internodal branches, the ratio is computed from a tip to the nearest branching point. However, since the numbers of transversions for the branches leading to L2 and L3 are too small, they are clustered. The ratio (denoted by L2&L3) is computed for the branch that separates cluster L2 and L3 from L4 and L5.

for ND2 and 9.0% for COI), does most rapidly. Such a correlation can be expected (26) since, for a given G+C frequency (w), the transitional rate and the saturation level are in proportion to $w(1 - w)$. Under compositional biases, even a small fluctuation of G+C content can change the transitional rate significantly while still keeping the saturation level very low. Obviously, if the transitional rate varies not only among different types of mtDNA within species (L3, L4†, and L5 or L4 and L6) but also between subregions within a type, the phylogenetic peculiarity among L4, L5, and L6 disappears.

We think the above interpretation likely, and if this is the case, we have had a good explanation for why molecular clocks sometimes run erratically and for why the use of transitional or synonymous changes in molecular phylogeny should be limited. Nonetheless, we could not totally dismiss the hypothesis of recombination, which seems more parsimonious. Further sequencing will provide more information on this matter.

Figs. 2 and 3 are also informative about the history of the *melanogaster* subgroup. To estimate the divergence times of six types of mtDNA, we used the transversional differences at the third codon positions for which a molecular clock is most evident. Relative to the distance from *D. yakuba*, branching point A6 (or B6) is 34%, A3, A4, and A5 are in the range of 12–18%, and A2 is about 3%. If *D. yakuba* and *D. melanogaster* diverged 13–17 million years ago (see ref. 17 and references therein), A6 amounts to 4.4–5.8 million years ago. Despite some uncertainty about this calibration, it is remarkable that *D. simulans* has been carrying the three distinct ancestral lineages L3, L4†, and L5 over 1 million years and that *D. mauritiana* has been carrying the two L4 and L6. In contrast to humans, the generation time of *Drosophila* is very short, and 1 million years may amount to 10 million generations. The persistence times of these lineages are very long, comparable (in terms of generation numbers) with those observed at the major histocompatibility complex loci in mice and rats (27) or in humans and African apes (28, 29). If the long-term effective size is of the order of 10^6 in *Drosophila* and thus that for mtDNA is about one quarter as large (4), the divergence time of these distinct types is too long for neutral lineages to persist in a randomly mating population (30). Furthermore, it is surprising that there are no descendant lineages differentiated from such types, particularly in light of high transitional rates at the third codon positions, which may be estimated as $3–4 \times 10^{-8}$ per site per year from the comparison of L2 and L3.

Although some forms of natural selection may be compatible with the data, it is important to recall that most substi-

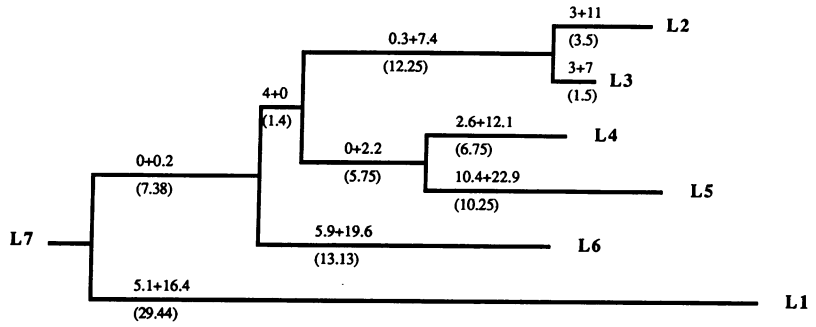

FIG. 3. NJ tree (24) based on the nucleotide differences at the first and second codon positions and transversional differences at the third positions (branch lengths are given in parentheses). *D. yakuba*, designated by L7, is used as an outgroup species. The number of transitional differences at the third positions for each branch was estimated by a least-square method similar to that in ref. 20 (p. 77). Two numbers along each branch stand for ND2 and COI, respectively.

tutions are synonymous and that mtDNA is essentially haploid. Of course, the latter automatically excludes the possibility of overdominant selection, as it does for *Escherichia coli* (31). Thus, the data lead us to conclude that they should be explained by the neutral hypothesis (19, 20). The only possibility under this hypothesis is to assume that these island species have not mated at random and have had highly structured populations for a considerable length of time (32). Each subpopulation must have been small enough to fix one type and the extent of gene flow must have been limited. The fact that types L3, L4†, and L4 are found only in islands in the West Indian or Pacific Ocean may reflect such a mode of isolation retained until quite recently. These characteristics are in sharp contrast to those in *D. melanogaster*, a more widespread and older species than *D. simulans* and *D. mauritiana*. The present-day similarity between *D. melanogaster* and *D. simulans* in their geographic and climatic distribution (12) is possibly recent, so that the genomes could not inscribe such short histories. It is possible that the global effective population size (N_e) relevant to neutral mutations was larger in *D. simulans* or *D. mauritiana* than in *D. melanogaster* because of the difference in population structure: with limited gene flow, the value of N_e for the entire population can be much larger than the actual number of individuals (33), resulting in higher levels of nuclear and mtDNA variation (7, 11).

The above model of population structure agrees with the speculation that *D. sechellia* was derived from *D. simulans*, probably less than 0.6 million years ago, and that *D. mauritiana* was as well, probably less than 1 million years ago. However, L4† in *D. simulans* and L4 in *D. mauritiana* differ only at one site in 2527 bp. The probability, P, that two DNA segments of r length are different by at most one site through the divergence time t with the per-site rate β is $(1 + 2\beta rt)\, e^{-2\beta rt}$. For $r = 2527$, $\beta = 10^{-8}$ (substitution rate averaged over all codon positions), and $t = 10^6$, the P value is negligibly small, implying that the divergence time assumed is improbable. It must therefore be concluded that t is at most 10^5 years. In the absence of introgression between *D. simulans* and *D. mauritiana*, this estimate of t sets an upper bound of the divergence time of the two species, suggesting their recent speciation (4). On the other hand, if the speciation indeed occurred around 1 million years ago, we must conclude that this period was not sufficient for these island species to develop a complete reproductive barrier, despite the long-lasting highly structured population. In this case, the shared type L4 ≈ L4† of mtDNA must be due to ongoing introgression.

As epitomized in ref. 34, *the history of the earth is recorded in the layers of its crust; the history of all organisms is inscribed in the chromosomes*. The introduction of DNA technology into the study of population biology has allowed scientists, among other things, to decipher the history of organisms. For it to be more effective, what is indispensable, we believe, is a proper sampling design and consideration from the viewpoint of population genetics.

We thank Drs. M. Kimura and S. I. Chigusa for their encouragement and help. We are grateful to Drs. J. F. Crow, R. H. Crozier, and A. C. Wilson for their useful comments on an early version of this paper. This work is supported in part by grants from the Japanese Ministry of Education, Culture and Science.

1. Gillham, N. W. (1978) *Organelle Heredity* (Raven, New York).
2. Lansman, R. A., Avise, J. C. & Huettl, M. D. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1969–1971.
3. Takahata, N. (1985) in *Population Genetics and Molecular Evolution*, eds. Ohta, T. & Aoki, K. (Japan Sci. Soc. Press, Tokyo/Springer, Berlin), pp. 195–212.
4. Satta, Y., Toyohara, N., Ohtaka, C., Tatsuno, Y., Watanabe, T., Matsuura, E. T., Chigusa, S. I. & Takahata, N. (1988) *Genet. Res. Camb.* **52,** 1–6.
5. Kondo, R., Satta, Y., Matsuura, E. T., Ishiwa, H., Takahata, N. & Chigusa, S. I. (1990) *Genetics*, in press.
6. Solignac, M., Monnerot, M. & Mounolou, J.-C. (1986) *J. Mol. Evol.* **23,** 31–40.
7. Choudhary, M. & Singh, R. S. (1987) *Genetics* **117,** 697–710.
8. Kreitman, M. (1983) *Nature (London)* **304,** 412–417.
9. Stephens, J. C. & Nei, M. (1985) *J. Mol. Evol.* **22,** 289–300.
10. Mukai, T. (1991) in *New Aspects of The Genetics of Molecular Evolution*, eds. Kimura, M. & Takahata, N. (Japan Sci. Soc. Press, Tokyo/Springer, Berlin), in press.
11. Aquadro, C. F., Lado, K. M. & Noon, W. A. (1988) *Genetics* **119,** 875–888.
12. Lachaise, D., Cariou, M.-L., David, J. R., Lemeunier, F., Tsacas, L. & Ashuburner, M. (1988) *Evol. Biol.* **22,** 159–225.
13. de Bruijn, M. H. L. (1983) *Nature (London)* **304,** 234–241.
14. Garesse, R. (1988) *Genetics* **118,** 649–663.
15. Clary, D. O. & Wolstenholme, D. R. (1985) *J. Mol. Evol.* **22,** 252–271.
16. DeSalle, R., Freedman, T., Prager, E. M. & Wilson, A. C. (1987) *J. Mol. Evol.* **26,** 157–164.
17. Satta, Y., Ishiwa, H. & Chigusa, S. I. (1987) *Mol. Biol. Evol.* **4,** 638–650.
18. Baba-Aïssa, F., Solignac, M., Dennebouy, N. & David, J. R. (1988) *Heredity* **61,** 419–426.
19. Kimura, M. (1968) *Nature (London)* **217,** 624–626.
20. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, London).
21. Satta, Y. & Chigusa, S. I. (1991) in *New Aspects of The Genetics of Molecular Evolution*, eds. Kimura, M. & Takahata, N. (Japan Sci. Soc. Press, Tokyo/Springer, Berlin), in press.
22. Crozier, R. H., Crozier, Y. C. & Mackinlay, A. G. (1989) *Mol. Biol. Evol.* **6,** 399–411.
23. Sokal, R. R. & Sneath, P. H. A. (1963) *Principles of Numerical Taxonomy* (Freeman, New York).
24. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425.
25. Sharp, P. M. & Li, W.-H. (1989) *J. Mol. Evol.* **28,** 398–402.
26. Takahata, N. & Kimura, M. (1981) *Genetics* **98,** 641–657.
27. Figueroa, F., Günther, E. & Klein, J. (1988) *Nature (London)* **335,** 265–267.
28. Lawlor, D. A., Ward, F. E., Ennis, P. D., Jackson, A. P. & Parham, P. (1988) *Nature (London)* **335,** 267–271.
29. Mayer, W. E., Jonker, M., Klein, D., Ivanyi, P., van Seventer, G. & Klein, J. (1988) *EMBO J.* **7,** 2765–2774.
30. Takahata, N. (1990) in *Population Biology of Genes and Molecules*, eds. Takahata, N. & Crow, J. F. (Baifukan, Tokyo/Macmillan, Basingstoke), pp. 267–286.
31. Selander, R. K. & Levin, B. R. (1980) *Science* **210,** 545–547.
32. Takahata, N. (1988) *Genet. Res. Camb.* **52,** 213–222.
33. Crow, J. F. & Kimura, M. (1970) *An Introduction to Population Genetics Theory* (Harper & Row, New York).
34. Kihara, H. (1947) *Ancestors of Common Wheat* (Sōgensha, Tokyo).