Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 4198–4201, June 1990
Evolution

Mitochondrial DNA evolution in experimental populations of *Drosophila subobscura*

(neutral haplotypes/fitness of mtDNA/nuclear–cytoplasmic coadaptation/natural selection)

MARIANO FOS, M. ANGELES DOMÍNGUEZ, AMPARO LATORRE, AND ANDRÉS MOYA*

Departamento de Genética, Facultad de Biología, Universitat de Valencia, Doctor Moliner, 50, 46100 Burjasot, Valencia, Spain

Communicated by Francisco J. Ayala, March 5, 1990 (received for review December 15, 1989)

ABSTRACT **When two mitochondrial DNA (mtDNA) haplotypes of *Drosophila subobscura* compete in experimental populations with discrete generations, one or the other approaches fixation, depending on the nuclear background with which they are associated. The approach to fixation, however, is strongly dependent on the effective number of females in the population, N_f. Whether or not the ultimate fate of a given mtDNA haplotype is determined by random genetic drift depends on N_f as well as on the relative fitnesses. Our experimental results show that the mtDNA polymorphisms observed in natural populations are affected by interactions among nuclear polymorphisms, random genetic drift, and direct selection on the mtDNA haplotypes.**

Most studies concerning mitochondrial DNA focus on determining the levels of polymorphism in natural populations of an ever-increasing number of species. The study of animal mitochondrial DNA (mtDNA) plays an important role because of this molecule's usefulness as a genetic marker in population and evolutionary biology (1, 2). Implicit in much of this work is the assumption that mtDNA variation is selectively neutral, while little effort has been directed to the experimental estimation of the parameters contributing to maintaining the observed levels of polymorphism.

Tests of possible selective differences mediated by mtDNA have been performed either (*i*) by measuring conditional fitnesses of nuclear genetic variants in different cytoplasmic backgrounds (3, 4) or (*ii*) by studying the competition between mtDNA lines that differ in restriction patterns (5). We have carried out a selection experiment involving two well-characterized mtDNA haplotypes of *Drosophila subobscura* in controlled nuclear backgrounds. This species exhibits in most of its chromosomes a rich inversion polymorphism that interferes with recombination among chromosomes carrying different inversion sequences. We have taken advantage of this property of *D. subobscura* to achieve the necessary control of the nuclear genetic background for the mtDNA competition experiments.

MATERIALS AND METHODS

Experimental Lines. We have used two *D. subobscura* strains from two very different geographic origins: Helsinki, Finland (H), and Las Raíces, Canary Islands, Spain (R). By means of a tedious set of crosses (6), two lines homozygous for all chromosomes were obtained. One, named HI, carries the standard sequence for the A, J, U, E, and O chromosomes and haplotype I for mtDNA; the other one, named RVIII, is homozygous for the A_2, J_1, U_{1+2}, O_{3+4}, and $E_{1+2+9+12}$ inversions and carries mtDNA haplotype VIII (7).

Two additional lines are used in the present work, RI and HVIII, which were obtained by a series of crosses that assume a completely maternal inheritance of mtDNA. These lines have the complementary nuclear-mtDNA composition of the original two; that is, HVIII has the standard sequences for the five chromosomes but mtDNA haplotype VIII, whereas RI has the chromosomal set of inversions mentioned above combined with mtDNA haplotype I. It is worth noticing that in the process of obtaining these complementary lines little or no recombination between the chromosomes from the two geographic origins occurs, so that the nuclear backgrounds of the original strains are preserved. Two additional lines were obtained, both nuclear backgrounds hybrid A_{st}/A_2, J_{st}/J_1, U_{st}/U_{1+2}, O_{st}/O_{3+4}, and $E_{st}/E_{1+2+9+12}$, but one with mtDNA haplotype I, the other with haplotype VIII; these are named HRI and HRVIII, respectively.

Discrete Generation Experiments in Population Cages. Three population cages were started at the same time, each with 1000 individuals (500 males and 500 females), and haplotypes I and VIII at equal frequencies. The nuclear background was HR, H, and R for population cages C1, C2, and C3, respectively. Cages were started with 12 food cups and kept at 19°C. Egg laying lasted 6 days, after which the food cups with eggs and larvae were moved to a new cage. When the F_1 appeared, 12 more food cups were added and left 6 days for egg laying; this set of food cups was then moved to a new population cage. All cages were on a similar cycle of discrete generations. One hundred fertilized females were sampled each generation and placed in separate cultures that were used for extracting mtDNA and estimating haplotype frequencies.

Discrete Generation Experiments in Culture Bottles. Eight different experimental lines, named L1–L8, were started in half-pint culture bottles, each with 125 ml of food. The nuclear background was always H, with mtDNA haplotype I or VIII. Lines L1–L4 were started each generation with 20 couples. The initial frequency of haplotype I was 0.7, 0.7, 0.3, and 0.3, for L1, L2, L3, and L4, respectively. Lines L5–L8, on the other hand, were started with 160 couples in each of the first six generations and thereafter with 80 couples. The initial frequency of haplotype I was 0.7, 0.7, 0.3, and 0.3 for lines L5, L6, L7, and L8, respectively.

To avoid intense competition, each line was made up of three culture bottles in the case of lines started with 20 couples, but three sets of two bottles each for the lines started each generation with 160 couples (only three bottles were used after generation seven, when only 80 couples were set up each generation), so that 80 couples were placed in one bottle and the other 80 in the other. The flies used to set up each line each generation were transferred at 5-day intervals to new bottles for egg laying, and they were discarded at the end of the third 5-day period.

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.

*To whom reprint requests should be addressed.

Part of the F_1 progenies emerging in each bottle were used to initiate the next generation and part were used to estimate mtDNA haplotype frequencies. Seven, 6, and 7 (20 in total) couples were randomly chosen from the three bottles to start the next generation in lines L1–L4. In the case of lines L5–L8, 25, 25, 30, 30, 25, and 25 (160 in total) couples were randomly selected from the six bottles for the first six generations, and 25, 30, and 25 (80 in total), from the three bottles in the remaining generations. The sample size used for estimating haplotype frequency was always 24, 8 per bottle in the three-bottle system, and 4 per bottle in the six-bottle system.

mtDNA Analysis. mtDNA was extracted according to ref. 7 and digested with restriction enzymes that yield different digestion pattern for haplotypes I and VIII. The haplotypes I and VIII of *D. subobscura* are the most divergent ones among those known (8). They show different digestion patterns with restriction endonucleases *Hin*dIII, *Hae* III, and *Hpa* II, so that three restriction site polymorphisms are obtained, one per enzyme, that are localized in the genes of the NADH complex of the physical map of *Drosophila yakuba* (9), which has been completely sequenced. The estimated divergence between the two haplotypes (according to ref. 10, p. 104) is 0.012. Assuming a size of 16,500 base pairs (bp), the haplotypes differ by approximately 203 bp. Only one restriction enzyme (*Hae* III) was used to distinguish between the two haplotypes.

Statistical Procedures. The intensity of selection acting on the mtDNA was evaluated by two methods. First we use the method of Fisher (11, 12) for testing whether the changes in haplotype frequencies can be explained as the result of genetic drift alone. Fisher and Ford have shown that the effective size of a population in each generation, and the sample taken from it, can be used to generate an expected matrix of covariances between the gene frequencies observed in the various generations. This covariance matrix can be used to test, by means of a χ^2, the hypothesis that random drift alone accounts for the observed changes in gene frequency. This method is sensitive only to fairly large differences in selective values—i.e., of the magnitude $s = 0.05$ (12).

The second method is based on a model of linear frequency change due to selection, which may be used to detect selection coefficients as small as 0.01 (12). The magnitude of the selection differential per generation is measured by a linear parameter, Y. The statistical significance for such a linear trend as well as nonlinear deviations is evaluated by χ^2 tests with 1 and $g - 2$ degrees of freedom, respectively, where g is the number of generations elapsed between the samples.

Both tests use gene frequencies arcsin transformed. We have also transformed the haplotype frequencies, but with an arcsin transformation that also corrects for small sample size (13). If x_i and n_i are the frequency and sample size of haplotype I at generation i, then

$$a_i = \arcsin[x_i/(n_i + 1)]^{1/2} + \arcsin[(x_i + 1)/(n_i + 1)]^{1/2}.$$

This statistic is asymptotically normal with a mean of 2arcsin $p^{1/2}$, and variance $1/n_i$, where p (estimated by x_i/n_i) is the true frequency of haplotype I in the population. Note that the sample variance of the estimator a_i is determined entirely by the sample size.

We have considered an effective population size, N_e, equal to the number of females, N_f, used to start each generation, which for organelle genes seems quite reasonable (14).

RESULTS

The changes in frequency of haplotype I are shown in Fig. 1 for population cages C1–C3 and in Fig. 2 for bottle experiments L1–L8. In population cage 1, when the two haplotypes were monitored on a mixed nuclear background, the frequencies changed dramatically and reached fixation of haplotype I by generation 5. The same tendency was observed in cage 2, where haplotype I with its own nuclear background was nearly fixed (96%) by generation 5. Cage 3 behaved in the opposite way; by generation 5 the frequency of haplotype I was only 9%. Taking into account the large effective population size (about 2000 breeding females in each generation) and how rapid and directional the haplotype frequency changes were, natural selection must be invoked as the causative process for the observed changes. The χ^2 values for testing (*i*) random genetic drift acting alone and (*ii*) linear directional change are shown in Table 1. There is no evidence

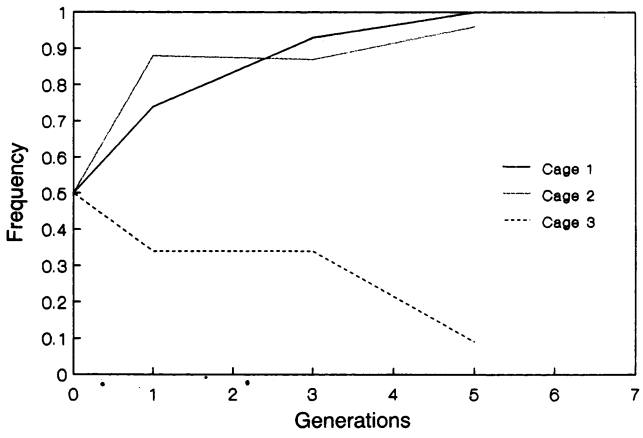

FIG. 1. Frequency changes of mtDNA haplotype I of *D. subobscura* in population cages.

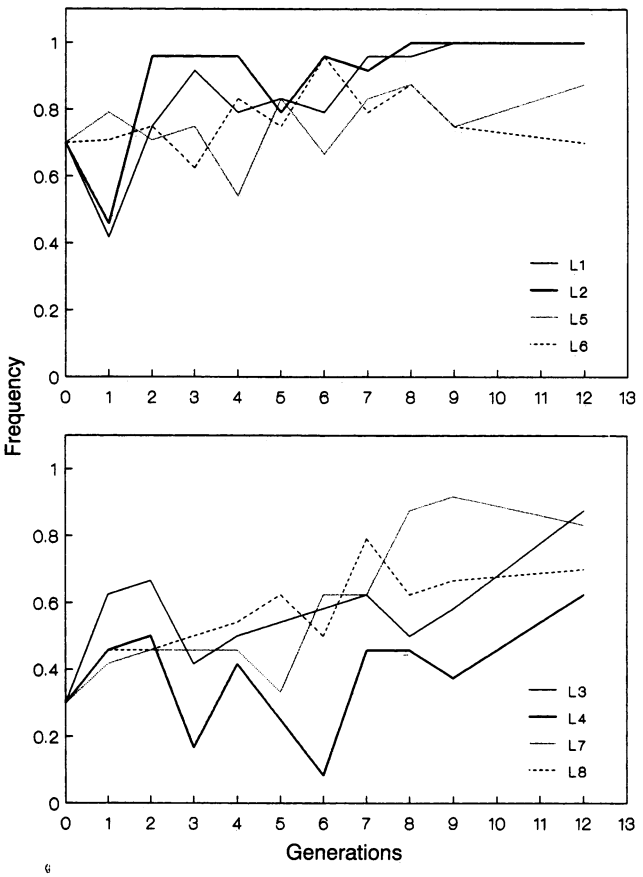

FIG. 2. Frequency changes of mtDNA haplotype I of *D. subobscura* in experimental lines in bottles.

Table 1. χ^2 values under either selective neutrality (random drift) or linear directional change for the three population cages (C1–C3) and eight culture bottle lines (L1–L8) with *D. subobscura*

Population	Random drift χ^2	df	Linear model χ^2	df	χ^2 deviation	df	Y	s
C1	45.3**	2	45.0**	1	0.3	1	0.487	0.0581
C2	5.6	2	3.5	1	2.1	1	0.319	0.0048
C3	25.3**	2	18.8**	1	6.5*	1	−0.314	0.0244
L1	11.3	9	4.3	1	7.0	8	0.148	0.0055
L2	14.0	9	2.7	1	11.3	8	0.118	0.0035
L3	9.6	9	0.8	1	8.8	8	0.064	0.0010
L4	16.2	9	0.3	1	16.0*	8	0.038	0.0004
L5	12.9	9	0.1	1	12.8	8	0.010	0.0000
L6	11.4	9	0.7	1	10.7	8	0.030	0.0002
L7	19.7*	9	9.7**	1	9.5	8	0.116	0.0034
L8	14.6	9	3.5	1	11.1	8	0.070	0.0012

Y is the coefficient estimated for the linear change per generation, and s is an estimate of the selection coefficient back-transformed from Y by $s = \sin^2(Y/2)$. *, $P < 0.05$; **, $P < 0.01$.

of a random drift effect in population cages 1 and 3. The abrupt 38% change of population cage 2 during the first generation is not detected with the test used. However, the presence of selection is evident in this population cage, which showed a dramatic change of haplotype I from frequency 0.5 to 0.96 in only five generations.

The major difference between the population cage and bottle experiments is the effective number of individuals that start each generation: population cages have an effective population size 10–100 times larger. Table 1 shows the χ^2 values for lines L1–L8 when random genetic drift and selection are tested. The bottle lines separately do not show a clear linear trend. However, the χ^2 values for lines L1, L2, L5, and L6 combined and for L3, L4, L7, and L8 combined show definite linear patterns. The rationale for this grouping has been theoretically demonstrated by Wilson (15). Each experimental line in the two sets was started at a haplotype I frequency of either 0.7 or 0.3, so that the populations in each set may be considered as replicates of the same base population. Table 2 shows the total χ^2 values. The lines started with high frequency of haplotype I are clearly more affected by random drift than the lines started with low frequency, where a linear selection trend is apparent. It is remarkable that deviations from linearity are not significant, which indicates that the tendency of haplotype I is to displace haplotype VIII until fixation is achieved. Depending, however, on the variance of the haplotype frequencies, the tendency to reach fixation appears to be higher, on the average, for populations with higher population sizes; i.e., populations with higher effective numbers will fix haplotype I in fewer generations.

DISCUSSION

The three population cages have experienced frequency changes of haplotype I (cages 1 and 2) or haplotype VIII (cage 3) on the order of 46% in just five generations. As has been previously shown by MacRae and Anderson (5) in *D. pseudoobscura*, such dramatic changes are not expected if the mtDNA haplotypes are selectively neutral. In our case, assuming that the effective population size for the organelle genes is the same as the effective number of females (14), it can be concluded that the frequency changes in the population cages are almost exclusively controlled by selection. According to Hartl and Clark (ref. 16, p. 351), alleles (haplotypes) with selection coefficients that are high in relation to

Table 2. Total χ^2 values when random drift and linear model are retested

Initial frequency	Random drift χ^2	df	Linear model χ^2	df	χ^2 deviation	df
0.7	49.6	36	7.8	4	41.8	32
0.3	60.1**	36	14.7**	4	45.4	32

Models were retested considering that L1, L2, L5, and L6 and L3, L4, L7, and L8 discrete generation experiments with bottles were started from the same base population, with frequencies of haplotype I of 0.7 and 0.3, respectively. **, $P < 0.01$.

the effective population number (i.e., $s < 1/2N_e$) are determined largely by selection. For mtDNA the effective population size for organelle genes is taken as the effective number of females, just replacing $2N_e$ by N_f (14). Estimates of selection coefficients in population cages under a linear selection model (see Table 1) range between 0.0048 and 0.0581. These selection coefficients are always larger than $1/N_f$ when N_f is approximately 2000. Two additional points deserve notice. First, haplotypes placed with their own nuclear background tend to be positively selected. Second, when the original nuclear genomes of haplotypes I and VIII are mixed, selection favors haplotype I. That is, some nuclear–cytoplasmic coadaptation must be invoked to explain the changes observed in population cages 2 and 3; when such coadaptation is broken (cage 1) and mtDNA evolves more independently, haplotype I shows higher fitness than haplotype VIII.

The results of the experimental lines with bottles indicate that either selection favoring haplotype I or random genetic drift plays significant roles. According to the inequality mentioned above, if the selection coefficient is less than 0.05 and 0.0087, respectively, for populations with N_f equal to 20 and 114 (which are the harmonic means for lines started at 80 and 160 couples), then random drift should be a major factor determining the ultimate fate of the mtDNA haplotypes. What is apparent, however, in experimental lines L3, L4, L7, and L8 (i.e., those started at 0.3 of haplotype I) is a linear trend towards higher frequency. The other lines fluctuate randomly, probably due to the fact that the equilibrium frequency (or fixation; see, for instance lines L1 and L2) is near.

Clark (17) has developed a deterministic model of selection with nuclear and cytoplasmic transmission, and he has shown that, under additive, multiplicative, or symmetric viability matrices, both nuclear and cytoplasmic polymorphisms are not protected. This has been observed in the cage experiments with respect to the possibility of a stable polymorphism for the mtDNA haplotypes. Such a tendency to fixation, however, is highly mitigated in small populations.

Except for population cage 1, our experimental populations are formally similar to population cages 11 and 12 of MacRae and Anderson (5): two different haplotypes compete in the context of a common nuclear background. MacRae and Anderson have found a non-statistically-significant increase of the Bogotá (BOG) mtDNA haplotype of *D. pseudoobscura*. Taking into account the large effective number of females in their case, the net frequency increases observed (0.109 and 0.074 for cages 11 and 12, respectively) suggest that the BOG haplotype is positively selected. The nuclear background effect should, however, be evident in the rest of their population cages, particularly cage 1. Assuming that there is no linkage between nuclear and mtDNA genes (see refs. 17 and 18 for the theory of cytonuclear disequilibrium), it is possible, as demonstrated analytically by Robertson (19) and by simulation studies by Birky and Walsh (20), that selection in the nucleus interferes to some extent with selection in the mtDNA. The effect of nuclear background on

mtDNA could explain the rapid change of mtDNA frequencies during the first 12 generations of ref. 5. By generation 13, when they sampled cage 1 to initiate the other cages, the nuclear background probably was quite homogeneous, with a high proportion of the genes coming from BOG. Given this nuclear composition, an increase in the frequency of the BOG haplotype should be expected, which is observed but without statistical significance. The rapid changes observed in our population cage 1 could be also interpreted as due to the additional selective effect of nuclear background on mtDNA.

Our results are of interest for understanding the variability of mtDNA in natural populations. mtDNA haplotypes and fitness differences in the interactions with the nuclear genome lead to natural selection. In turn, genetic drift alters how haplotypes respond to selection.

This work has been supported by Grant PB86-0517 from the Dirección General Investigación Cientifica y Técnica (Spain) to A.M. and by fellowships from the Ministerio de Educación y Ciencia (Spain) to A.L. and A.M. to visit the Department of Ecology and Evolutionary Biology, University of California, Irvine.

1. Avise, J., Arnold, J., Ball, R. M., Bermingham, E., Lamp, T., Neigel, J. E., Reeb, C. & Saunders, N. C. (1987) *Annu. Rev. Ecol. Syst.* **18,** 489–522.
2. Harrison, R. G. (1989) *Trends Ecol. Evol. Biol.* **4,** 6–11.
3. Clark, A. G. (1985) *Genetics* **111,** 97–112.
4. Clark, A. G. & Lyckegaard, E. (1988) *Genetics* **118,** 471–481.
5. MacRae, A. & Anderson, W. W. (1988) *Genetics* **120,** 485–494.
6. Latorre, A., Moya, A. & De Frutos, R. (1988) *Evolution* **42,** 1298–1308.
7. Latorre, A., Moya, A. & Ayala, F. J. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 8649–8653.
8. Latorre, A., Barrio, E., Moya, A. & Ayala, F. J. (1988) *Mol. Biol. Evol.* **5,** 717–728.
9. Clary, D. O. & Wohlstenholme, D. R. (1985) *J. Mol. Evol.* **22,** 252–271.
10. Nei, M. (1987) *Molecular Evolutionary Genetics* (Columbia Univ. Press, New York).
11. Fisher, R. A. & Ford, E. B. (1947) *Heredity* **1,** 143–174.
12. Schaffer, H. E., Yardley, D. & Anderson, W. W. (1977) *Genetics* **87,** 371–379.
13. DeSalle, R. A., Templeton, A., Mori, I., Pletscher, S. & Johnston, J. S. (1987) *Genetics* **116,** 215–223.
14. Birky, C. W., Maruyama, T. & Fuerst, P. (1983) *Genetics* **103,** 513–527.
15. Wilson, S. R. (1980) *Genetics* **95,** 489–502.
16. Hartl, D. & Clark, A. (1989) *Principles of Population Genetics* (Sinauer, Sunderland, MA).
17. Clark, A. G. (1984) *Genetics* **107,** 679–701.
18. Asmussen, M., Arnold, J. & Avise, J. (1987) *Genetics* **115,** 755–768.
19. Robertson, A. (1961) *Genet. Res.* **2,** 189–194.
20. Birky, C. W. & Walsh, J. B. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 6414–6418.