Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 8813–8817, November 1986
Population Biology

Extensive variation and heteroplasmy in size of mitochondrial DNA among geographic populations of *Drosophila melanogaster*

(restriction enzymes/sequence variation/natural selection/gene flow/population structure)

LAWRENCE R. HALE AND RAMA S. SINGH

Department of Biology, McMaster University, Hamilton, ON, Canada, L8S 4K1

Communicated by Robert K. Selander, July 10, 1986

ABSTRACT **Size variation and heteroplasmy in mitochondrial DNA (mtDNA) are relatively common in natural populations of *Drosophila melanogaster*. Of 92 isofemale lines of flies obtained from various geographic regions throughout the world, 75 lines were homoplasmic and showed a total of 12 different mtDNA size classes. The remaining 17 lines were heteroplasmic, each line carrying two different mtDNAs, and, in all but one case, the mtDNAs in these heteroplasmic lines differed in size; a total of nine size classes was represented among them. In cases where one type was predominant within an individual, it was usually the smaller mtDNA. This finding parallels what was observed in homoplasmic lines, in that the smaller mtDNAs were much more common than the larger variants in most populations. The data suggest a high rate of mutational occurrence of mtDNA size variants and some natural selection against them.**

In studies of variation and evolution in mitochondrial DNA (mtDNA), two assumptions have been made about the way this molecule changes. First, it has been assumed that all the molecules inherited by an individual organism from its maternal parent are the same; i.e., that the individual is homoplasmic for mtDNA (1, 2). It is this characteristic that allows unambiguous maternal lineages to be inferred. Second, of the three types of mutational changes—base substitution, addition, and deletion—occurring in mtDNA, base substitution is assumed to occur much more frequently than addition or deletion (3–5). Recently, several cases of heteroplasmy, in which two abundant mtDNA variants occur abundantly in an individual, have been reported for both vertebrate and invertebrate taxa (6–11). In two of these cases, experiments demonstrated that the mtDNA heteroplasmy was relatively stable and was passed on from parent to offspring for a number of generations (7, 9).

Size variation in mtDNA appears to be much more common in some organisms than in others. Densmore *et al.* (6) detected 37 different size variants in 92 individuals of the parthenogenetic lizard *Cnemidophorus tesselatus*, and some individuals were found to be heteroplasmic. In contrast, the frequency of sequence divergence, as reflected by variation in restriction fragment pattern, was very low (0.06%). Extensive studies of mtDNA variation in rodents have shown appreciable sequence variation but no size polymorphism (12–14). Human mtDNA, on the other hand, has shown appreciable amounts of both sequence and size variation (3, 4).

We here report the results of an analysis of size variation in mtDNA in a world-wide collection of *Drosophila melanogaster*. We have found extensive variation in size and also have identified a number of lines that were heteroplasmic, all but one of which carried two mtDNAs differing in size. Possible models for the origin and maintenance of size polymorphism and heteroplasmy in natural populations of *D. melanogaster* are considered.

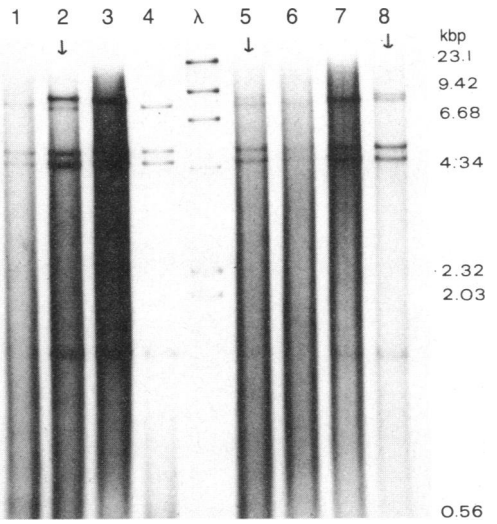

FIG. 1. *Cla* I digests of mtDNA from single isofemale lines from eight geographical populations of *D. melanogaster*, demonstrating size variation and three cases of heteroplasmy for size-variant mtDNAs. Lanes with arrows (lanes 2, 5, and 8) are instances of heteroplasmic lines, as shown by the doublet at the position of the top fragment. *Hin*dIII-cut λ DNA is the size standard. Numbers at right denote sizes of λ DNA fragments. The gel was 0.6% agarose.

MATERIALS AND METHODS

***Drosophila* Stocks.** We studied 92 isofemale lines of *D. melanogaster* that were collected in various geographic locations around the world (Table 1). These lines had been maintained in the laboratory from 4 to 5 years at 18°C on cornmeal medium (10 g of agar, 45 g of sucrose, 30 g of malt powder, 17 g of brewer's yeast, 100 ml of corn syrup, 71 g of cornmeal, 10 g of soy meal, and 5 ml of propionic acid per liter of water).

DNA Isolation. mtDNA obtained from either second-instar larvae or adult flies was purified by differential centrifugation according to the protocols of Lansman *et al.* (15), with modifications described by Hale and Beckenbach (16).

Restriction Analysis. Nine restriction enzymes were used on each mtDNA isolate. Restriction fragments were end-labeled with Klenow DNA polymerase I (Boehringer Mannheim), as described by Maniatis *et al.* (17), except that the reaction was allowed to proceed for only 20 min. Labeled mtDNA was electrophoresed in 0.5% and 2.0% agarose gels. The gels were vacuum-dried onto Whatman 3MM paper, and the DNA fragments were visualized by autoradiography.

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: mtDNA, mitochondrial DNA; kbp, kilobase pair(s).

Table 1. Isofemale lines showing mtDNA size variants that are homoplasmic in geographic populations of *D. melanogaster*

Population	No. of lines with mtDNA size class (kbp)*												
	18.1	18.3	18.6	18.7	18.8	18.9	19.0	19.1	19.2	19.3	19.4	19.6	19.9
Ottawa			5	1	1	1	1			1			
Hamilton			2										
Vancouver			1†										
California			3						1				
Florida				2									
Texas			4		1	1		1				1	1
Argentina			3										
England			2	1			1						
France			3	2		1		2					
West Africa	1		3					2			1		
Central Africa			2										
Korea			1		1			1					
Japan			1	3									
Taiwan			3										
Vietnam			3		1				1				
Australia			3	1		1							
India			2			2							
Total	1	0	41	10	4	6	2	6	2	1	1	1	1

*Size of mtDNA (in kilobase pairs, kbp) is indicated above each column.
†Heteroplasmic for mtDNAs differing by the presence or absence of a *Hae* III site.

RESULTS

Identification of Heteroplasmic Lines. Potential cases of heteroplasmy were identified when a restriction digest produced a double band where only one was expected. The reproducibility of the patterns under conditions of excess enzyme and long incubation periods (12–16 hr) ruled out the possibility of partial restriction digests. Fig. 1 shows *Cla* I digests of eight isofemale lines, three of which were heteroplasmic for mtDNAs differing in size. Lines were designated as heteroplasmic only if the size variants appeared in digests of all enzymes used in this study.

Because some of the lines studied had been maintained in the laboratory from 4 to 5 years, we made two tests to determine whether the apparent high frequency of heteroplasmy was caused by contamination of lines by a few flies carrying a different-sized mtDNA. First, we scrutinized restriction digests from each putative heteroplasmic line for evidence of heterogeneity in restriction pattern. If the heteroplasmy in size was caused by contamination, the lines in question should show heterogeneity not only for size variants but also for any site variants for which the contaminating flies may have differed. Long autoradiographic exposures of the filters increased the probability of detecting low-frequency site-variant fragments. None of the lines that were putatively heteroplasmic for size variants was found to be heterogenous for site variants. A second test involved restarting each of four putative heteroplasmic lines from a single female and examining the mtDNA in three subsequent generations. The heteroplasmy was manifested in these test lines, which means that the original lines are heteroplasmic. We did not note any changes in the relative proportions of the two mtDNA types over three generations.

Distribution of Heteroplasmic and Size-Variant mtDNAs. In a sample of 92 isofemale lines of *D. melanogaster* from geographically diverse populations, 75 were homoplasmic and 17 were heteroplasmic. Of the 17 heteroplasmic lines, 16 carried two size-variant mtDNAs and one carried two site-variant mtDNAs. Tables 1 and 2 show the distribution of homoplasmic and heteroplasmic lines, respectively, in samples from various geographic populations. There were 13 size

Table 2. mtDNA size variants in heteroplasmic isofemale lines from geographic populations of *D. melanogaster**

Population	No. of lines	mtDNA size classes (kbp) detected												
		18.1	18.3	18.6	18.7	18.8	18.9	19.0	19.1	19.2	19.3	19.4	19.6	19.9
Vancouver	2			+ᵖ			+							
Massachusetts	2			+							+			
Texas	1			+					+					
Argentina	1			+					+ᵖ					
England	1				+ᵖ						+			
West Africa	1			+			+							
	1			+ᵖ			+							
Korea	1			+ᵖ		+								
	1			+						+ᵖ				
Taiwan	1				+			+						
Vietnam	1		+	+ᵖ										
Australia	1				+		+							
India	1			+					+					
	1			+ᵖ					+					

Superscript p indicates predominant type within line.
*In addition, one line from Vancouver that did not show intra-line size variation was heteroplasmic for mtDNAs differing by the presence or absence of a *Hae* III site (see Table 1).

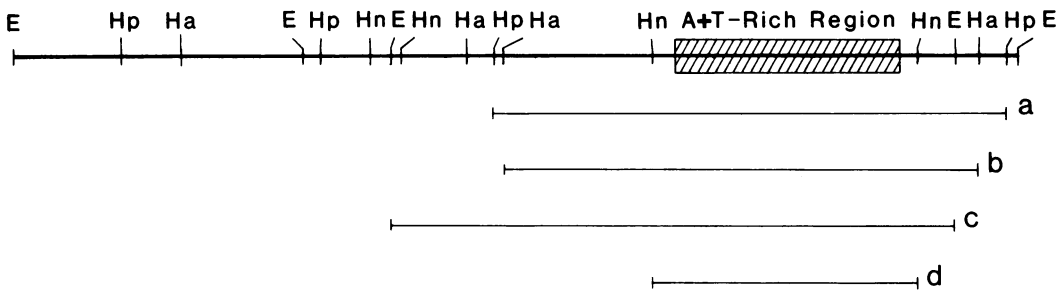

FIG. 2. Restriction map of *D. melanogaster* mtDNA, as described by previous studies (18–20). The map is for four restriction enzymes: *Hae* III (Ha), *Hin*dIII (Hn), *Hpa* II (Hp), and *Eco*RI (E). Fragments involved in size variation in homoplasmic and heteroplasmic lines are shown: a (*Hpa* II), b (*Hae* III), c (*Eco*RI), and d (*Hin*dIII).

variants within the resolution power of our gels (0.1 kbp). In most populations and in the species as a whole, the 18.6-kbp variant was the predominant type (Table 1).

Table 2 shows the composition of heteroplasmic lines. Where one size class was visibly more abundant (greater quantity of DNA) than the other in a line, the predominant one is denoted by a superscript (p). Most heteroplasmic lines carried the 18.6-kbp variant, which was the most common type in the homoplasmic lines. One line from Vancouver (shown in Table 1) was homoplasmic for size but heteroplasmic for two variants differing by the presence or absence of a *Hae* III recognition site.

Localization of Size Variants. All size variants map to the A+T-rich region of the mtDNA molecule, as determined by the comparison of our results with previously published restriction maps of *D. melanogaster* mtDNA (18–20). Fig. 2 shows a representation of size variants based on four enzymes. The size variants occur exclusively within the *Hin*dIII fragment defined by the two *Hin*dIII sites that closely flank the boundaries of the A+T-rich region. All fragments that showed size variation overlap in their map position, but, unfortunately, because of the unavailability of maps for other enzymes, we were not able to localize more precisely the sites of the size alterations. The difference between the smallest and largest molecule was about 1.1 kbp, resulting in a range for the mtDNA molecule of 18.1–19.9 kbp, which is consistent with other reports of size variation in this species (20, 21).

DISCUSSION

The amount of size variation in *D. melanogaster* populations is surprisingly large, especially when compared with restriction site variation. In several of the populations studied, more size variants than restriction site variants were detected (Table 3), and virtually all composite restriction (clonal) types varied in size (Table 4). In particular, two common clonal types that differed by only two restriction sites each showed seven size variants (Table 4). The observation that a given size variant occurred in a number of clonal types that were not necessarily closely related suggests that size variants have had multiple origins and postdate the divergence of many clonal types.

The average sequence divergence of mtDNA in *D. melanogaster*, based on restrictions polymorphisms in our data, was 0.008–0.014 (22, 23). This value is low compared to divergence values for vertebrate mtDNA (0.017–0.018; refs. 3 and 12) but is substantially larger than that reported for *Cnemidophorus* (0.0006; ref. 6). The above observations strongly suggest that in populations of *D. melanogaster*, as in those of *Cnemidophorus*, size-altering mutations in mtDNA are frequent.

There is a greater chance of underestimating the occurrence of size variants than of site variants. In this study, fragment sizes of 4.0 kbp and above were measured to the nearest 0.1 kbp. This is a conservative resolution limit because our gels did not allow a greater degree of accuracy. Because of limited resolution in DNA fragment separation, two different size-variant mtDNAs might not be distinguished if they were closely similar in size. This sort of underestimation is not so likely with site variants, which are visibly more discrete (24). Thus it is possible that the size classes presented in Fig. 3 are heterogenous and contain many more smaller mtDNA size variants.

A peculiar feature of mtDNA size variation in *D. melanogaster* is that the distribution of size classes is strongly skewed. Fig. 3 shows the frequencies of sizes in all *D. melanogaster* lines studied. Note that the 18.6-kbp size is the most common type and that the frequencies of other variants drop off sharply as one moves away from the predominant size in either direction in the distribution. In light of the fact that there were several size variants larger than the predominant 18.6-kbp class, it is somewhat surprising that very few variants were smaller than 18.6 kbp. One line from West Africa was homoplasmic for a mtDNA variant of 18.1 kbp, and one line from Vietnam was heteroplasmic, containing 18.3- and 18.6-kbp variants. In 14 of the 18 populations sampled, the 18.6-kbp size variant was the predominant type.

The observation that most populations had the same size variant as the predominant type is in complete contrast to that for the site variants. Most of the populations studied were found to be highly differentiated for restriction site variants (25). For example, no clonal types were shared between England and France, West Africa and France, or Japan and Taiwan. Therefore, the similarity of size-variant distribution among populations is not likely to have been caused by mitochondrial gene flow. On the other hand, it also seems likely that if all these size variants were selectively neutral, then random fixation within each population would result in some populations having a predominant size variant other than the 18.6-kbp type, which is not the case.

Table 3. Size and restriction site variation in selected populations of *D. melanogaster*

Population	No. of lines studied	No. of size variants*	No. of site variants	
			Mean	Range
Australia	6	4	1.75	1–2
Ottawa	10	6	1.50	1–2
Texas	10	6	2.00	1–3

The number of size variants is condensed from Tables 1 and 2. The number of site variants is the average number of restriction enzyme cleavage patterns observed for the four polymorphic enzymes (*Ava* II, *Hae* III, *Taq* I, and *Mbo* I).

*The same number of size variants was shown by each polymorphic enzyme.

Table 4. Number of isofemale lines (homoplasmic/heteroplasmic) showing size variation in different mitochondrial clonal types in *D. melanogaster*

	Lines with mtDNA size class (kbp), no. homoplasmic/no. heteroplasmic												
Clonal type	18.1	18.3	18.6	18.7	18.8	18.9	19.0	19.1	19.2	19.3	19.4	19.6	19.9
ABAA			2/1					1/0	0/1				
ABBA			12/3	6/0	1/0	2/0	1/0	0/1		0/2			
AAAA			3/3	0/1	1/1	0/2	1/1	2/0	1/0				
AABA			5/0	1/1		0/1			1/0				1/0
AAEA			3/1		1/0	2/0		1/1		1/0			
AAHA		0/1	0/1										
AAAG			1/0			1/0							
AAAB				0/1						0/1			
AACA				1/0									
AACI			0/2			0/2							
AAAD			0/1					1/0					
ABAD			2/1	1/0									
CAAB				1/0		1/0		1/0					
AAGA			1/0										
ADAA	1/0												
AAAC			1/0								1/0		
DCAC			1/0										
ABCA			2/0										
AAIH			1/0										
AADA			0/1					0/1					
AAEF			1/0										
ABBE			2/0		1/0							1/0	
BBFJ			0/1					0/1					
BBBJ			3/0										

Clonal types are composites of restriction site variants for four polymorphic enzymes (*Ava* II, *Hae* III, *Taq* I, and *Mbo* I).

An explanation for the contrasting distribution patterns of size vs. site variants of mtDNA is suggested by the data in Table 2. All heteroplasmic lines contained either the 18.6- or the 18.7-kbp size variants. In cases where one type was prevalent over the other, most flies had the smaller size variant as the predominant type. (In two cases, one line from Korea and one line from Argentina, the reverse is true.) This suggests that size variants frequently arise by mutation and are subject to elimination by purifying natural selection. What effect the larger variants have on organismal fitness is not apparent, especially since the A+T-rich region is noncoding. However, the A+T-rich region contains the origin of replication (26–28), alterations to which may confer deleterious effects. Replication of DNA is a complex process requiring interaction with a number of enzymes and cofactors (29). mtDNA has a covalently closed circular supercoiled structure; size alterations may affect its tertiary conformation and, consequently, its interaction with the DNA polymerase, resulting in differential efficiency of replication.

The possible role of selection on size variation depends in part on the mechanism of generation of these variants. Despite the fact that size variants have been reported in mtDNA for several taxonomic groups, we do not know the mechanism by which they arise. As suggested by Densmore *et al.* (6), the mechanism could be of two types. Recombinational events resulting in unequal sequence exchange

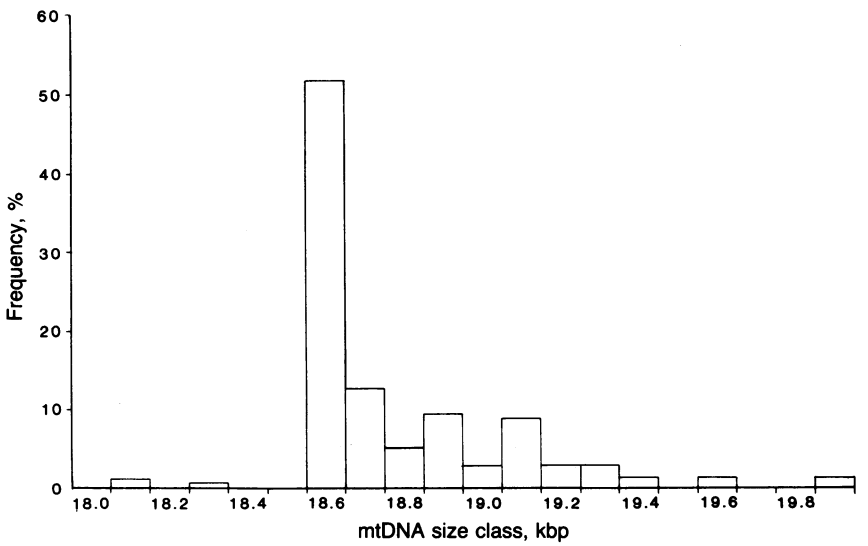

FIG. 3. Frequency distribution of mtDNA size variants in *D. melanogaster*. In calculating frequencies, the variants in heteroplasmic lines were given half the weight of those in homoplasmic lines.

between A+T-rich regions would produce mtDNAs of both smaller and larger sizes. However, size variants smaller than the prevalent 18.6-kbp type are extremely rare in *D. melanogaster*. Hence, we would need to postulate either strong purifying selection against the smaller-sized variants or a mechanism that generates only larger-sized variants. Available data on mtDNAs tend to indicate that recombination does not occur (30). We favor the hypothesis that size variation is generated by sequence duplications. We cannot judge from our results whether a single repetitive sequence has been tandemly repeated in *D. melanogaster*. In the genus *Drosophila*, only members of the melanogaster species group have A+T-rich regions larger than 1.1 kbp (31). The smallest region observed for *D. melanogaster* in this study was about 4.7 kbp. A rather large repeated sequence (0.5 kbp) has been proposed to have generated the larger A+T-rich regions of the melanogaster species (12, 32). We believe it more likely that size variants are generated by DNA duplications of a variable length during replication (6). In this case, size variants smaller than 18.6 kbp would be rare or absent simply because they were not generated.

We are grateful to Drs. W. Brown, D. Rand, R. K. Selander, and D. Wolstenholme for their comments on an earlier version of this manuscript and to R. Rhem for his assistance with the figures. This research was supported by an operating grant to R.S.S. (A0235) from the Natural Sciences and Engineering Research Council of Canada.

1. Francisco, J., Brown, G. G. & Simpson, M. V. (1979) *Plasmid* **2**, 426–436.
2. Lansman, R. A., Avise, J. C. & Huettel, M. D. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1969–1971.
3. Aquadro, C. F. & Greenberg, B. D. (1983) *Genetics* **103**, 287–312.
4. Cann, R. L. & Wilson, A. C. (1983) *Genetics* **104**, 699–711.
5. Brown, W. M., George, M., Jr., & Wilson, A. C. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1967–1971.
6. Densmore, L. D., Wright, J. W. & Brown, W. M. (1985) *Genetics* **110**, 689–707.
7. Solignac, M., Monnerot, M. & Mounolou, J.-C. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6942–6946.
8. Solignac, M., Genermont, J., Monnerot & Mounolou, J.-C. (1984) *Mol. Gen. Genet.* **197**, 183–188.
9. Harrison, R. G., Rand, D. M. & Wheeler, W. C. (1985) *Science* **228**, 1446–1448.
10. Monnerot, M., Mounolou, J.-C. & Solignac, M. (1984) *Biol. Cell* **52**, 213–218.
11. Bermingham, E., Lamb, T. & Avise, J. C. (1985) *Genetics* **110**, s77 (abstr.).
12. Avise, J. C., Giblin-Davidson, C., Laerm, J., Patton, J. C. & Lansman, R. A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 6694–6698.
13. Lansman, R. A., Avise, J. C., Aquadro, C. F., Shapira, J. F. & Daniel, S. W. (1983) *Evolution* **37**, 1–16.
14. Ferris, S. D., Sage, R. D., Prager, E. M., Ritter, U. & Wilson, A. C. (1983) *Genetics* **105**, 681–721.
15. Lansman, R. A., Shade, R. O., Shapira, J. F. & Avise, J. C. (1981) *J. Mol. Evol.* **17**, 214–226.
16. Hale, L. R. & Beckenbach, A. T. (1985) *Can. J. Genet. Cytol.* **27**, 357–364.
17. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
18. Shah, D. M. & Langley, C. H. (1979) *Nature (London)* **281**, 696–699.
19. Fauron, C. M.-R. & Wolstenholme, D. R. (1980) *Nucleic Acids Res.* **8**, 2439–2452.
20. Reilly, J. G. & Thomas, C. A., Jr. (1980) *Plasmid* **3**, 109–115.
21. Fauron, C. M.-R. & Wolstenholme, D. R. (1980) *Nucleic Acids Res.* **8**, 5391–5410.
22. Kaplan, N. & Risko, K. (1981) *J. Mol. Evol.* **17**, 156–162.
23. Nei, M. & Li, W.-H. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5269–5273.
24. Templeton, A. R. (1983) *Evolution* **37**, 221–244.
25. Hale, L. R. & Singh, R. S. (1985) *Genetics* **110**, S42 (abstr.).
26. de Bruijn, M. H. L. (1983) *Nature (London)* **304**, 234–241.
27. Goddard, J. M. & Wolstenholme, D. R. (1980) *Nucleic Acids Res.* **8**, 741–757.
28. Goddard, J. M. & Wolstenholme, D. R. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3886–3890.
29. Clayton, D. A. (1982) *Cell* **28**, 693–705.
30. Hayashi, J.-I., Tagashira, Y. & Yoshida, M. C. (1985) *Exp. Cell Res.* **160**, 387–395.
31. Wolstenholme, D. R., Goddard, J. M. & Fauron, C. M.-R. (1979) in *Proceedings of the Eighth Annual ICN–UCLA Symposium on Molecular and Cellular Biology: Extrachromosomal DNA*, eds. Cummings, D., David, I. B., Borst, P., Weissman, S. & Fox, F. (Academic, New York), pp. 409–425.
32. Fauron, C. M.-R. & Wolstenholme, D. R. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3623–3627.