Length and Restriction Site Heteroplasmy in the Mitochondrial DNA of American Shad (*Alosa sapidissima*)

Paul Bentzen, William C. Leggett and Gregory G. Brown

Department of Biology, McGill University, Montreal, Quebec H3A 1B1, Canada

Manuscript received February 3, 1987
Revised copy accepted November 12, 1987

ABSTRACT

Restriction endonuclease analysis was used to assess mitochondrial DNA (mtDNA) variation in American shad (*Alosa sapidissima*) collected from 14 rivers ranging from Florida to Quebec. Two types of heteroplasmy were observed, one involving a major length polymorphism and the other a single restriction site. Shad mtDNA occurred in two principal size classes, 18.3 and 19.8 kb. Of 244 shad examined, 30 were heteroplasmic and carried both size classes of mtDNA in varying proportions; the remainder were homoplasmic for the smaller size class of mtDNA. The large mtDNA variant occurred most frequently at the southern end of the range, and except for two individuals from Nova Scotia, was not detected among shad from rivers north of the Delaware. In contrast, ten shad heteroplasmic for a *Sal*I restriction site originated from rivers ranging from South Carolina to Nova Scotia. DNA mapping and hybridization experiments indicated that the length polymorphism is in the D-loop-containing region and consists of a tandemly repeated 1.5-kb DNA sequence occurring in two and three copies, respectively, in the two major size classes of shad mtDNA. Continuous length variation up to approximately 40 bp occurs among copies of the repeat both within and among individuals. Restriction site data support the conclusion that both forms of heteroplasmy in shad mtDNA have originated more than once.

THE mitochondrial genome of animals is compact in organization, highly variable in sequence and ranges in size from 15700 to about 23000 base pairs (bp) (BROWN 1983, KESSLER and AVISE 1985, MORITZ and BROWN 1986). Numerous studies (summarized in AVISE and LANSMAN 1983; BROWN 1983, 1985; AVISE 1986) based on nucleotide sequencing and restriction endonuclease analysis have led to two general conclusions regarding the variability of animal mitochondrial DNA (mtDNA): (1) Polymorphism is common among conspecifics, but individuals usually appear homoplasmic, that is, they exhibit only a single mtDNA genotype; and (2) At least among conspecifics and closely related species, most of this polymorphism takes the form of silent base substitutions or minor (few bp) insertions or deletions.

These generalizations are based primarily on data from mammalian mtDNA. Recent observations suggest that they may not be valid for at least some nonmammalian groups. Major length polymorphisms (several hundred bp or more), in some cases associated with heteroplasmy, have been reported for several species of Drosophila (FAURON and WOLSTENHOLME 1980b; REILLY and THOMAS 1980; SOLIGNAC, MONNEROT and MOUNOLOU 1986; HALE and SINGH 1986), Gryllus crickets (HARRISON, RAND and WHEELER 1985), Cnemidophorus lizards (DENSMORE, WRIGHT and BROWN 1985; MORITZ and BROWN 1986), Rana and Hyla frogs (MONNEROT, MOUNOLOU and SOLIGNAC 1984; BIRMINGHAM, LAMB and AVISE 1986), and *Amia calva*, a fish (BIRMINGHAM, LAMB and AVISE 1986).

In this study we report on another species, the American shad (*Alosa sapidissima*). We find that mtDNA variation in shad also differs markedly from the generalizations above. *A. sapidissima* is a widely distributed anadromous fish that spawns in rivers ranging from Florida to Quebec. The juveniles remain in their natal rivers until autumn, then undertake lengthy seasonal migrations in the sea for 4–6 yr before homing to their natal rivers to spawn (MANSUETI and KOLB 1953). We have been studying mtDNA polymorphism in *A. sapidissima* primarily to assess the extent of genetic differentiation among river populations. We address here an unusual aspect of mtDNA variation in shad, that is the occurrence of frequent and widespread heteroplasmy. We have observed two types of heteroplasmy in samples of shad mtDNA, one involving a large length polymorphism, the other a single restriction site. Below we present data on the molecular character of the length polymorphism in shad mtDNA. We also report on the frequencies of heteroplasmy among samples collected from 14 populations spanning the range of the species. Finally, we describe restriction site data that lead to the conclusion that both forms of heteroplasmy in shad mtDNA have arisen more than once.

MATERIALS AND METHODS

Female shad were collected in 1985 and 1986 during their upstream spawning migrations in 14 rivers ranging from Florida to southern Quebec (Figure 1). Ovaries were removed from the shad at the collection sites and held on ice for 1–7 days before use. Mitochondrial DNA was isolated from 5 g of ovarian tissue (consisting almost entirely of mature or nearly mature oocytes) from each fish. The method used was a modified version of that described by CHAPMAN and POWERS (1984). Samples were homogenized in 4–5 volumes of TEK buffer (50 mM Tris, 10 mM EDTA, 1.5% KCl, pH 7.5) with a motor driven glass Teflon homogenizer. The homogenate was then transferred to a 50-ml polypropylene centrifuge tube and underlayed with a 15% sucrose-TEK solution using a long stem pasteur pipette. The homogenate was then centrifuged at 1,000 × *g* for 10 min. Following centrifugation a relatively sharp boundary was usually visible at the top of the sucrose-TEK layer. The upper layer was carefully drawn off the sucrose cushion and transferred to another 50 ml centrifuge tube, then centrifuged at 12,000 × *g* for 10 min. The supernatant was then decanted and the remaining crude mitochondrial pellet was resuspended in 10 ml of TEK, transferred to a 15 ml glass centrifuge tube, and centrifuged again at 12,000 × *g* for 10 min. The supernatant was poured off and the pellet was resuspended in 1.8 ml of TEK, then 0.2 ml of 10% non-idet-TEK solution was added to the sample. The sample was shaken to distribute the non-idet, then left for 5–10 min to allow complete lysis of the mitochondria. The lysate was then centrifuged at 12,000 × *g* for 10 min and the resulting pellet discarded. The supernatant was extracted a minimum of three times with phenol/chloroform, following which mtDNA was precipitated with the addition of two volumes of ethanol.

Mitochondrial DNA from one shad was purified by CsCl gradient centrifugation (LANSMAN *et al.* 1981). *Pst*I fragments collectively encompassing the entire mitochondrial genome of this fish were cloned in pBR322 using techniques described by MANIATIS, FRITSCH and SAMBROOK (1982). The same method was also used to clone a *Pst*I fragment encompassing the position of a variant *Sal*I site in two other shad.

Restriction endonuclease digestions were carried out under the conditions specified by the vendor (Bethesda Research Laboratories). Some digestions were treated with 1–2 μg of RNase A to facilitate the visualization of small (<500 bp) restriction fragments that would otherwise by obscured by RNA present in samples prepared by the CHAPMAN and POWERS (1984) protocol.

Restriction fragments were separated in 0.6 and 1% horizontal agarose gels, and visualized by ethidium bromide staining. A 1-kb DNA ladder purchased from BRL served as a weight marker. The stoichiometric relationships of some bands were evaluated by scanning 10.2 × 12.7 cm (Kodak Tri-x Pan) photographs of gels with a (Bio-Rad model 1650) densitometer and integrating absorbtion peaks. The intensity of each band was assumed to be directly proportional to the amount of DNA present. Corrections were made for the size of the DNA fragments in each band when evaluating the molar relationships of different bands.

The DNA fragments in some gels were transferred to (Bio-Rad Zeta-Probe) nylon membranes using the REED and MANN (1985) modification of the SOUTHERN (1975) technique. These were hybridized to DNA probes labeled with [^{32}P]dCTP by nick translation (RIGBY *et al.* 1977). The hybridization probes comprised cloned *Pst*I fragments of shad mtDNA, a 1.5-kb *Eco*RI fragment isolated from cloned shad mtDNA by polyacrylamide electrophoresis (MANIATIS, FRITSCH and SAMBROOK 1982), and two cloned fragments of *Rattus norvegicus* mtDNA corresponding to known portions of the rat mitochondrial genome (BROWN *et al.* 1986; G. G. BROWN, unpublished data). Hybridizations were carried out under the conditions specified by Bio-Rad, except that for those involving the rat probes, the temperature of the final wash was dropped from 55° to 25°. The results were visualized by autoradiography.

FIGURE 1.—Origin of shad samples. Names in parentheses, and the positions of letters on the map denote approximate locations of collection sites. A, St. John R. (Welaka and Mayport, Florida); B, Altamaha R. (Darien, Georgia); C, Cooper R. (Bonneau Dam, South Carolina); D, Waccamaw R. (Georgetown, South Carolina); E, Pamlico R. (Washington, North Carolina); F, York R., Virginia); G, Delaware R. (Lambertville, New Jersey); H, Hudson R. (Claverack, New York); I, Connecticut R. (Lyme, Connecticut); J, Annapolis R. (Annapolis Royal, Nova Scotia); K, St. John R. (Cambridge Narrows, New Brunswick); L, Miramichi R. (Newcastle, New Brunswick); M, St. Lawrence R. (Trois Rivieres, Quebec); N, Richelieu R. (Chambly, Quebec).

Both whole mtDNA and cloned *Pst*I fragments were used to map the relative positions of cleavage sites recognized by the endonucleases *Kpn*I, *Sca*I, *Hpa*I, *Eco*RV, *Pvu*II, *Sst*I, *Sst*II, *Pst*I, *Sal*I and *Eco*RI (Figure 2). The positions of the sites were determined by the double digest method (MANIATIS, FRITSCH and SAMBROOK 1982).

RESULTS

Length polymorphism and heteroplasmy: *Sal*I restriction analysis of mtDNA samples from 244 *A.*

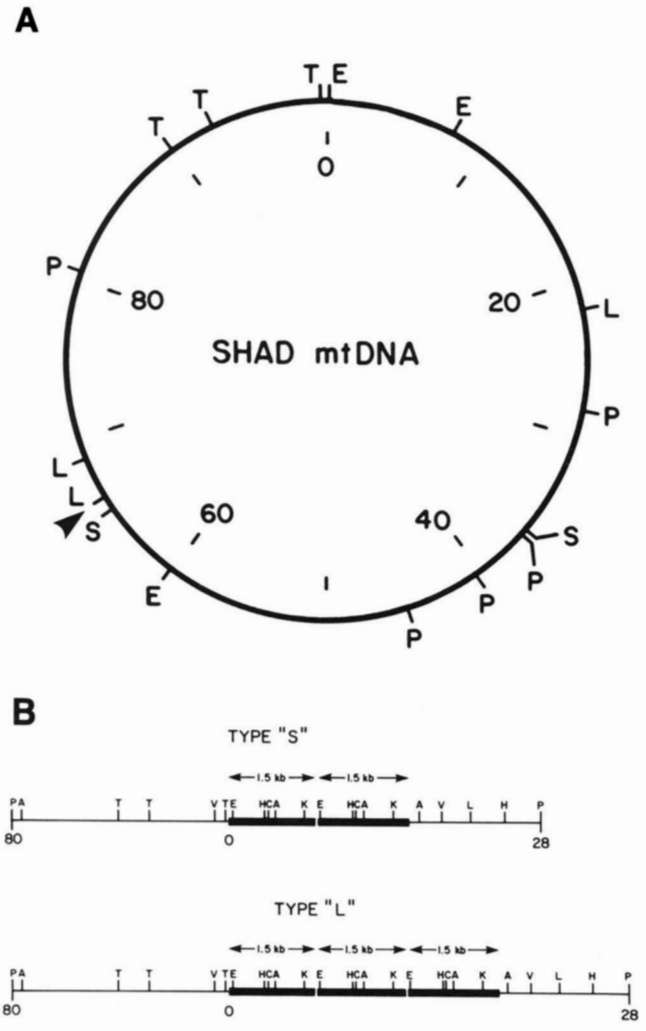

FIGURE 2.—Restriction site map of shad mtDNA. A, *Sca*I; C, *Eco*RV; E, *Eco*RI; H, *Hpa*I; K, *Kpn*I; L, *Sal*I; P, *Pst*I; S, *Sst*I; T, *Sst*II; V, *Pvu*II. A, Partial map of entire molecule of type S form. The arrow indicates the *Sal*I site that is absent in Sal-B mtDNA. B, Detailed map of region (map units 80–28) containing tandem repeats (denoted by thickened segments).

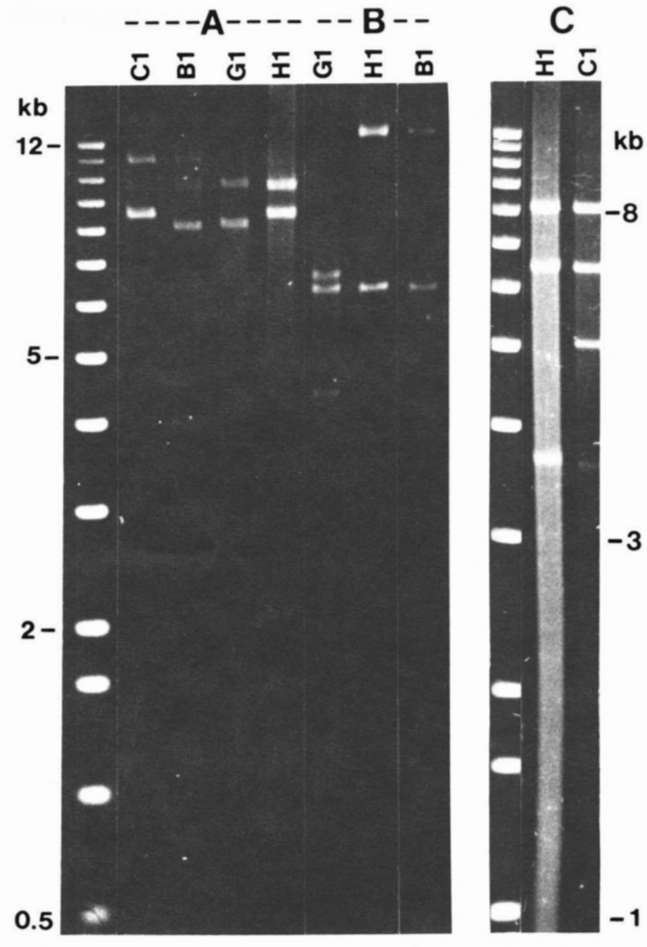

FIGURE 3.—Length heteroplasmy and restriction site polymorphisms in shad mtDNA. Samples of shad mtDNA used in digestions are named according to river of origin. A, *Sal*I digests on 0.6% agarose gel stained with EtBr. Individuals G1 and H1 are homoplasmic for type S mtDNA; C1 and B1 are heteroplasmic and contain both type S and type L mtDNA (visible in top band of digest). G1 and H1 are Sal-A and Sal-B genotypes, respectively. A 0.47-kb band present in the G1 digest is not visible on this gel. B, *Kpn*I digests on same gel as A. G1 and H1 are Kpn-A and Kpn-B genotypes, respectively. B1 is Kpn-B genotype; type L mtDNA in this sample is again visible in top band on the gel. C: *Pvu*II digests on 1% gel. The type L mtDNA in C1 is visible in the 5-kb band.

sapidissima revealed two size classes of mtDNA that differed by approximately 1.5 kb. *Sal*I cleaved the mtDNA of most shad at three locations, yielding fragments of approximately 0.47, 8.1 and 9.7 kb (Figures 2 and 3). In a minority of shad, however, the "normal" 9.7-kb band was replaced by two clearly substoichiometric bands, one 9.7 kb as before, and another of approximately 11.2 kb (Figure 3). The relative intensities of the two substoichiometric bands varied widely among individuals.

This pattern of length variation and heteroplasmy was verified with other restriction endonucleases. Concordant patterns of variation were observed when representative samples were digested with a variety of enzymes including *Pvu*II and *Kpn*I (Figure 3) as well as *Sst*I, *Sst*II and *Xba*I (not shown).

Only one size class of mtDNA (approximately 18.3 kb) was detected in 214 out of 244 samples of shad mtDNA digested with *Sal*I and visualized by EtBr staining. The remaining 30 shad were heteroplasmic: these fish carried mtDNA of both 18.3 and 19.8 kb (henceforth referred to as type S and type L mtDNA, respectively). Most of the shad carrying type L mtDNA originated from rivers in the southern part of their range (Table 1). Indeed, fish bearing type L mtDNA occurred at higher frequency (55%) among the St. John River (Florida) samples from the extreme southern end of the range of shad, than anywhere else. Only two shad (from the Annapolis River, Nova Scotia) bearing type L mtDNA originated from north of the Delaware River.

Restriction site mapping and DNA hybridization experiments revealed that types S and L mtDNA

TABLE 1

Geographic distributions of shad mitochondrial genotypes

River	Genotype			*N*
	Size	SAL	KPN	
(A) St. John (Florida)	S	A	A	13
	S,L	A	A	15
	S,L	A	B	1
				29
(B) Altamaha	S	A	A	8
	S	A	B	5
	S	B	B	2
	S,L	A	B	2
	S,L	B	B	1
				18
(C) Cooper	S	A	A	8
	S	A	B	2
	S	A,B	B	2
	S,L	B	B	1
				13
(D) Waccamaw	S	A	A	11
	S	A	B	1
	S	B	B	1
	S,L	A	A	4
				17
(E) Pamlico	S	A	A	9
	S	A,B	A	1
	S	A	B	1
	S	B	B	7
	S	B	A	1
	S,L	A	A	1
				20
(F) York	S	A	A	12
	S	A	B	3
	S	A,B	B	2
	S	B	B	2
	S	A	C	1
	S,L	A	A	2
				22
(G) Delaware	S	A	A	11
	S	A,B	A	1
	S	A,B	B	1
	S	B	B	5
	S,L	A	A	1
				19
(H) Hudson	S	A	A	17
	S	B	B	3
	S	A,B	A	1
				21
(I) Connecticut	S	A	A	10
	S	B	B	1
				11
(J) Annapolis	S	A	A	12
	S	A	B	2
	S	B	A	1
	S	A,B	A	1
	S	A,B	B	1
	S,L	A	A	1
	S,L	A	B	1
				19
(K) St. John (N.B.)	S	A	A	6
	S	A	B	3
	S	B	B	1
				10
(L) Miramichi	S	A	A	12
	S	A	B	2
	S	B	A	3
				17
(M) St. Lawrence	S	A	A	16
	S	A	B	4
	S	B	A	1
				21
(N) Richelieu	S	A	A	7
				7

contain two and three copies, respectively, of a tandemly repeated 1.5-kb DNA sequence corresponding almost exactly to an *Eco*RI fragment of the same size. They also revealed that each of the tandemly repeated sequences varies by as much as 40 bp among individual shad as well as among copies of the repeat within individuals. These assertions are based on the following lines of evidence: (1) Five restriction enzymes generate a single 1.5-kb fragment in type S mtDNA. The restriction sites delimiting the 1.5-kb fragments occur in a repetitive pattern in a single 3-kb stretch of DNA (Figure 2). (2) When samples containing type L mtDNA were digested with any of the five enzymes, the resulting restriction fragment patterns were similar or identical to those obtained with type S mtDNA. Either a single 1.5-kb band or else two closely spaced bands near 1.5 kb in size were observed (*e.g.*, fish A4 and A3, respectively, in Figure 4). When only a single 1.5-kb band was seen, densitometry scans revealed an excess of DNA in the band over that expected from a single fragment. (3) The 1.5-kb *Eco*RI fragment corresponding to the proposed repeat was isolated from cloned (type S) mtDNA and used to probe a Southern blot bearing *Eco*RI digests. The probe hybridized to two fragments in a sample homoplasmic for type S mtDNA, and three in a sample containing type L mtDNA (fish A2 and A3, respectively, in Figure 4). (4) Similar patterns of fragment size variation were observed with several enzyme combinations, including *Pvu*II/*Kpn*I and *Pvu*II/*Eco*RV (Figure 5) and *Sca*I (not shown) that divide the region containing the tandem repeats in types S and L mtDNA into two or three fragments, respectively (Figure 2).

The fact that the region encompassing the major length polymorphism maps adjacent to an *Sst*II site

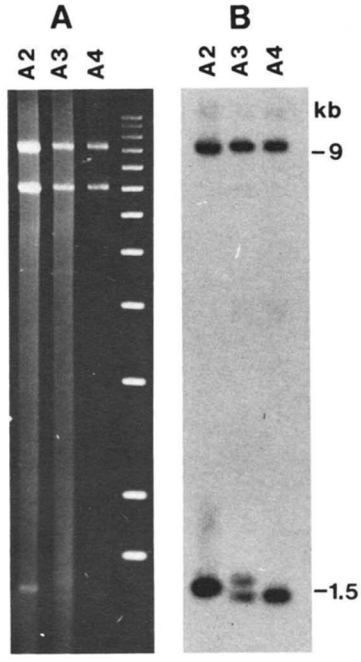

FIGURE 4.—Evidence of 1.5-kb tandem repeats in shad mtDNA. A, *Eco*RI digests on 0.6% gel stained with EtBr. B, Autoradiograph of Southern blot of same gel probed with radioactively labeled 1.5-kb *Eco*RI fragment derived from cloned type S mtDNA. Individual A2 is homoplasmic for type S mtDNA; A3 and A4 are heteroplasmic and contain type L mtDNA. Due to length variation that is prevalent among copies of the repeat, the two 1.5-kb fragments formed by the digest of A3 appear as discrete bands, whereas they appear as a doublet in the digest of A4. The 1.5-kb probe has also hybridized to another copy of the repeat present in the largest fragment in each digest (see Figure 2).

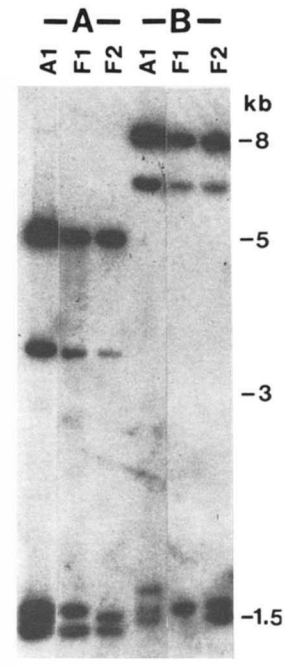

FIGURE 5.—Length variation among copies of the tandem repeat. Autoradiograph of Southern blot of 1% gel probed with radioactively labelled cloned *Pst*I fragment (map units 80–28). A, *Pvu*II/*Kpn*I; B, *Pvu*II/*Eco*RV. Individual A1 is heteroplasmic and contains type L mtDNA; F1 and F2 are homoplasmic for type S mtDNA. Each enzyme combination cleaves the region containing the tandem repeats in types S and L mtDNA into two and three segments, respectively (Figure 2). Variation in the size of homologous fragments is evident in both enzyme combinations.

suggests that it might be located in the D-loop containing region, since *Sst*II cleaves the mtDNA of widely diverse vertebrates at two locations in the rRNA genes, close to the D-loop (GLAUS *et al.* 1980; BROWN and SIMPSON 1981; MORITZ and BROWN 1986). *Sst*II cleaves shad mtDNA at three locations (Figure 2); two of these presumably represent the conserved rRNA sites. Two hybridization experiments were performed to test this conclusion and orient the shad mtDNA restriction map. Digests of shad mtDNA were probed with clones of rat mtDNA specific to two adjacent regions of the molecule: (1) the 16S rRNA gene and (2) the 12S rRNA gene, the D-loop containing region and part of the cytochrome *b* gene (Brown *et al.* 1986; G. G. BROWN, unpublished data). The first probe hybridized strongly to a shad *Pst*I/*Sst*II fragment spanning map units 80–90 in Figure 2. The second probe hybridized to an *Sst*II fragment spanning map units 93–10, as well as to an adjacent *Sst*II/*Pst*I fragment spanning map units 0–28. These observations indicate that the organization of shad mtDNA in the region extending from the rRNA genes into the cytochrome *b* gene is similar to that of mammalian mtDNA, and confirms that the major length variation occurs in the D-loop-containing region.

Site heteroplasmy: Although the mtDNA of most shad contained three *Sal*I restriction sites (henceforth referred to as Sal-A mtDNA), a second mitochondrial genotype (Sal-B) was also observed. Mapping experiments indicated that Sal-B mtDNA was characterized by the loss of one *Sal*I restriction site (Figure 2). *Sal*I digests of Sal-B mtDNA yielded two fragments of approximately 8600 and 9700 bp (*e.g.*, fish L2 in Figure 6). Samples that contained Sal-B mtDNA were geographically widespread in origin. They were observed in 12 of the 14 shad populations sampled, and amounted to 40 (16%) of the 244 shad assayed with *Sal*I (Table 1).

Ten of the samples that carried Sal-B mtDNA also appeared to contain some Sal-A DNA. *Sal*I digests of these samples yielded fragments of both 8600 and 8100 bp in sub-stoichiometric quantities (*e.g.*, fish F3, H2 and J1 in Figure 6). The possibility that this apparent heteroplasmy was caused by a 500-bp length polymorphism was eliminated by the observation that no sub-stoichiometric bands were seen when the same samples were digested with *Pst*I, *Sst*I and *Sst*II. In addition, a Southern blot of *Sal*I digests of these

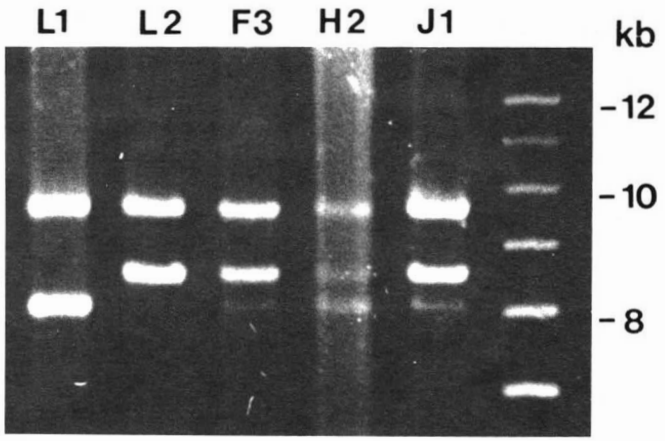

FIGURE 6.—Restriction site heteroplasmy in shad mtDNA. Restriction fragment patterns obtained on 0.6% gel after exhaustive digestion with *Sal*I (details in text). Individuals L1 and L2 are homoplasmic for Sal-A and Sal-B mtDNA, respectively; F3, H2 and J1 are heteroplasmic for the two genotypes. A 0.47-kb band generated by the Sal-A genotype is not visible on this gel.

samples revealed the presence of a 470-bp fragment in substoichiometric quantities. Such electrophoretic banding patterns could have resulted from incomplete digestion of Sal-A mtDNA, but two lines of evidence countered this possibility: (1) Five of the apparently heteroplasmic samples were digested several times with increasing amounts of enzyme and increasing incubation times. Even when these samples were exposed to 50 units of *Sal*I and incubated for sixteen hours, the resulting electrophoretic banding patterns remained unchanged from those initially observed (Figure 6). These digestion conditions represented a many-fold increase in the amount of enzyme and incubation time over that usually required to obtain complete digestion with *Sal*I. (2) If the apparently heteroplasmic mtDNA samples were really incompletely digested Sal-A mtDNA, the 8100-bp band would be expected to be more intense than the 8600-bp band, which would then represent the residual presence of an intermediate digestion product. In fact, in all but one of the samples in question (see sample H2 in Figure 6), the 8600 bp band was more intense than the 8100-bp band. Moreover, none of the *Sal*I digests of these samples exhibited any other evidence of incomplete digestion, such as additional restriction fragments corresponding to the sum of other fragments.

Finally, the possibility that the *Sal*I site polymorphism and associated heteroplasmy might have been the product of an artifact such as postmortem methylation of the DNA was ruled out by a cloning experiment. The variant *Sal*I site in question was observed in DNA cloned from an individual with Sal-A mtDNA, but not in corresponding DNA fragments cloned from two individuals with Sal-B mtDNA, thus indicating that the site polymorphism is the result of an alteration in the primary sequence of the DNA.

***Kpn*I Data:** Whether heteroplasmic shad are the products of persistent heteroplasmy stemming from unique mutations, or alternatively, result from recurrent mutations, can be addressed only by restriction site data that delineate separate lines of descent. For the samples heteroplasmic for type L mtDNA, this issue is partially resolved by the *Sal*I site data, since although most samples bearing type L mtDNA were Sal-A, two were Sal-B (Table 1). Thus, the L form appears to have arisen at least twice, once in a Sal-A lineage and once in a Sal-B lineage. For the shad heteroplasmic for the *Sal*I restriction site, however, data from at least one other informative enzyme are needed.

A preliminary survey of several shad from each population with 16 hexanucleotide restriction enzymes revealed very few polymorphisms (P. BENTZEN, unpublished data). Only one enzyme, *Kpn*I, revealed a polymorphism that was both widespread and common; hence it was used to further analyze the 244 shad mtDNA samples previously surveyed with *Sal*I.

Three *Kpn*I genotypes were observed. Kpn-A was characterized by four restriction sites generating fragments of approximately 1.5, 4.3, 6.1 and 6.4 kb (Figures 2 and 3). In Kpn-B both restriction sites within the tandem repeats were absent, resulting in fragments of 6.1 and 12.3 kb (Figure 3). In Kpn-C only one restriction site within the tandem repeats was missing; the remaining sites generated fragments of 7.9, 6.1 and 4.3 kb. Among the samples surveyed with *Kpn*I, 187 were Kpn-A and 56 were Kpn-B. Only one sample from the York River, Virginia, was Kpn-C. The two common genotypes were also geographically widespread: Kpn-A was observed in all, and Kpn-B in all but one, of the 14 populations (Table 1). All four possible combinations of the *Sal*I and *Kpn*I genotypes were observed. Of the samples that were heteroplasmic Sal-A,B, four were Kpn-A and six were Kpn-B. Samples heteroplasmic for type L mtDNA were divided among three composite genotypes. The majority (24) were Sal-A/Kpn-A, but four were Sal-A/Kpn-B and two were Sal-B/Kpn-B (Figure 3, Table 1). Samples containing type L mtDNA that were Kpn-B lacked the variant *Kpn*I sites in all three copies of the tandem repeat. These observations support the view that both forms of heteroplasmy have arisen more than once in shad mtDNA (see DISCUSSION).

DISCUSSION

Length polymorphism: The mapping and hybridization data indicate that the 1.5-kb length polymorphism in shad mtDNA is generated by a 1.5-kb

tandemly repeated sequence that occurs in two copies in type S and three copies in type L mtDNA. The occurrence of a sequence triplication of this scale is a novel feature of the shad mtDNA length polymorphism. The only other major length polymorphism that has been characterized in vertebrate mtDNA involves a single tandem repeat (MORITZ and BROWN 1986).

Each of the copies of the 1.5-kb repeat is subject to minor (<40 bp) length variation both within and among individuals. The tandem repeats are in the D-loop containing region and map adjacent to an *Sst*II site located within the 12S rRNA gene. This suggests that the polymorphism may also encompass the gene for phenylalanyl tRNA, and possibly part of the 12S rRNA gene, assuming that shad mtDNA shares the gene arrangement common to other vertebrates (BROWN 1983, 1985).

Apart from a single report of a major deletion involving the coding region of mouse mtDNA (BOURSOT, YONEKAWA and BONHOMME 1987) other examples of large-scale length variation in animal mtDNA have also implicated the D-loop containing region or its homologue in insects, the adenine + thymine (A + T) rich region. Large-scale length variation in the A + T rich region has been observed both among and within several species of Drosophila (FAURON and WOLSTENHOLME 1976, 1980a,b; Wolstenholme *et al.* 1979, REILLY and THOMAS 1980; HALE and SINGH 1986). Much of this length variation is generated by 470-bp tandem repeats that vary in copy number among and within species (SOLIGNAC, MONNEROT and MOUNOLOU 1986). The D-loop-containing region is the site of length variation among species of artiodactyls and primates (UPHOLT and DAWID 1977; FERRIS, WILSON and BROWN 1981), as well as minor (<15 bp) intraspecific variation in mammals (CANN and WILSON 1983; HAUSWIRTH *et al.* 1984; BROWN *et al.* 1986), and much larger variation (400–700 bp) in the frog *Rana esculenta* (MONNEROT, MOUNOLOU and SOLIGNAC 1984). Major intraspecific length variation in mtDNA has been characterized in several species of Cnemidophorus lizards. In *C. tesselatus* and *C. tigris marmoratus* mtDNA, tandem repeats of a 64-bp stretch in the D-loop containing region are primarily responsible for variation of up to 370 bp among individuals (DENSMORE, WRIGHT and BROWN 1985). In *C. exsanguis* mtDNA, a length polymorphism of at least 4.8 kb is caused by a single tandem repeat that incorporates minor internal length variation, and spans the D-loop-containing region as well as some adjacent structural genes (MORITZ and BROWN 1986).

Heteroplasmy: The survey of mitochondrial genome size in shad produced evidence of restriction site heteroplasmy in ten individuals. This is noteworthy, since very little evidence of restriction site heteroplasmy has previously been reported. This may be less a consequence of the absolute rarity of the phenomenon than of two biases that affect restriction endonuclease analysis of mtDNA (AVISE and LANSMAN 1983). First, examples of heteroplasmy involving only a single restriction site are much less likely to be detected than those associated with major length polymorphisms, since the former manifest themselves with only one enzyme whereas the latter can be detected with many enzymes. Second, cases of restriction site heteroplasmy are liable to be misinterpreted as instances of incomplete digestion, since the patterns of substoichiometric bands that result from both phenomena may be similar. Of course, this ambiguity also means that apparent cases of restriction site heteroplasmy must be viewed with caution. In this study, we addressed the possibility of incomplete digestion by considerably increasing both enzyme quantity and incubation time in repeat digests of suspect samples. Since the substoichiometric bands evident in Figure 6 showed no change in intensity from previous digests carried out under less intense conditions, the conclusion that they represent restriction site heteroplasmy is warranted.

Prior to this study, evidence of restriction site heteroplasmy has been limited to a single maternal lineage each of cows (HAUSWIRTH and LAIPIS 1985) and *Drosophila melanogaster* (HALE and SINGH 1986). The occurrence in widely scattered locales of shad heteroplasmic for a *Sal*I restriction site (Table 1) provides empirical evidence that site heteroplasmy may be a widespread phenomenon in at least some species.

All of the mtDNA examined in this study was isolated from oocytes. Since high rates of heteroplasmy have also been detected in frog oocytes (MONNEROT, MOUNOLOU and SOLIGNAC 1984), it is possible that heteroplasmy may be more prevalent in germline cells than in somatic tissues.

The heteroplasmy in shad mtDNA could have resulted from biparental inheritance, or from incomplete segregation of the variant forms of mtDNA involved. We favor the latter possibility, since all available evidence indicates that at least within the limits of detection, metazoan mtDNA is inherited solely through maternal lineages (AVISE and LANSMAN 1983; AVISE *et al.* 1984).

The various composite *Sal*I/*Kpn*I genotypes and the possible mutational pathways that link them are depicted in Figure 7. The pathways indicate that if the inheritance of mtDNA in shad is strictly maternal then at least the transition between alternate *Sal*I genotypes has occurred more than once. For example, for all instances of *Sal*I heteroplasmy to have resulted from a single mutational event would require that in some lines of descent the mtDNA population remained heteroplasmic for the *Sal*I polymorphism while undergoing a complete transition between *Kpn*I

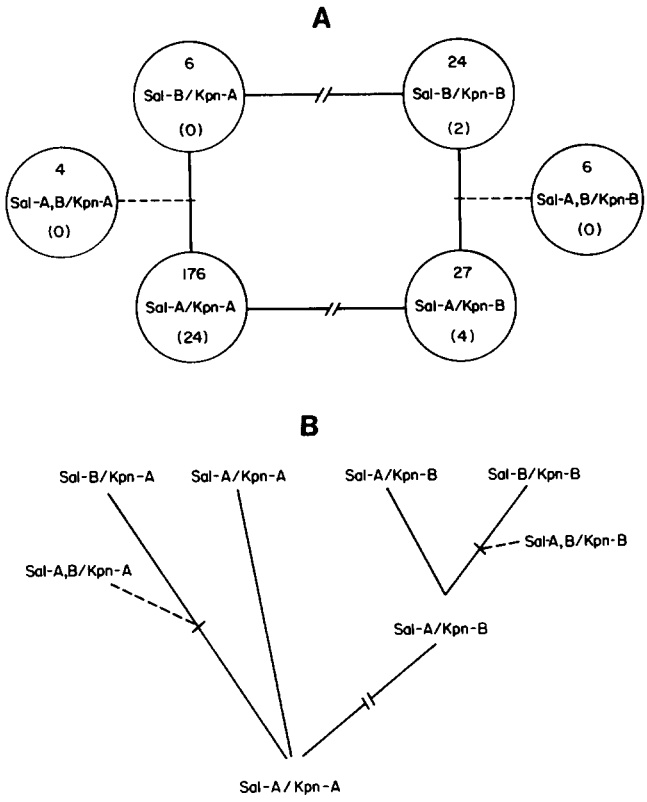

FIGURE 7.—Possible mutational pathways involved in the generation of the various composite *SalI*/*Kpn*I genotypes. The solid lines crossing the branches indicate the number of restriction site changes involved in the transition from one genotype to another. A, Parsimony network linking genotypes. The total number of individuals, and the number ( ) bearing type L mtDNA are indicated for each genotype. The transition between the two *Kpn*I genotypes need only have occurred once, but could have occurred in individuals that were either Sal-A or Sal-B. B, Dendrogram depicting one of the four possible mutational pathways indicated in A. The construction emphasizes the necessity of independent origins of Sal-A,B heteroplasmy in two different lines of descent.

genotypes involving two restriction site changes. Almost certainly then, the transition between the alternate *Sal*I genotypes occurred at least once in each of two lineages, one Kpn-A and the other Kpn-B. Similarly, since samples containing type L mtDNA were divided among three composite genotypes, shad bearing this form of heteroplasmy stemmed from at least three independent mutational events.

The two common *Kpn*I genotypes were Kpn-A and Kpn-B. The transition between these genotypes entails the parallel loss (or gain) of two restriction sites, one in each copy of the tandem repeat present in type S mtDNA. A third genotype (Kpn-C), in which only one of the restriction sites within the repeat is absent, was only observed in one individual. These features suggest that the 1.5-kb repeat sequences may evolve in concert, as has been suggested for the 470-bp repeats in Drosophila mtDNA (SOLIGNAC, MONNEROT and MOUNOLOU 1986).

The occurrence of continuous length variation within the 1.5-kb repeat suggests that the gain or loss of the variant *Kpn*I sites might result from sequence insertions or deletions. The fact that the length variation in the repeat sequence is widespread among individuals raises the additional possibility that the transition between the two common *Kpn*I genotypes has also occurred more than once.

Since most animals appear homoplasmic, it has generally been assumed that mitochondrial variants sort out rapidly (AVISE and LANSMAN 1983), a view also supported by theoretical arguments (UPHOLT and DAWID 1977; CHAPMAN *et al.* 1982; BIRKY *et al.* 1983). Recent breeding experiments with insects, however, indicate that heteroplasmy may persist in some species for hundreds of generations (SOLIGNAC *et al.* 1984; RAND and HARRISON 1986).

The high incidence of heteroplasmy in shad suggests that the segregation of mitochondrial variants may occur relatively slowly in this species. It is surprising, however, that we failed to encounter any shad that appeared homoplasmic for type L mtDNA. This result contrasts with other reports of heteroplasmy in which the variant forms of mtDNA in question have also been observed in the homoplasmic state, as well as with our own observations of restriction site heteroplasmy in shad. Only a minority (25%) of the shad with Sal-B mtDNA appeared heteroplasmic. The absence of shad homoplasmic for the type L variant suggests the possibility that this form of mtDNA may be under negative selection. Nevertheless, the type L variant was clearly the predominant form of mtDNA present in some individuals (*e.g.* fish C1 in Figure 3).

The geographic distribution of the type L variant also points to the possibility of selective influences on the occurrence of this genotype. Unlike the Sal-B variant, which was distributed across the whole range of the species, the type L variant was largely restricted to the southern part of the range, and was particularly prevalent in the St. John River population in Florida at the extreme southern end of the range (Table 1). It is conceivable that water temperatures or some other variable correlated with latitude might favor the occurrence of type L mtDNA, although we are unable to propose any mechanisms for such a selective effect.

The data presented here support the suggestion made by BIRMINGHAM, LAMB and AVISE (1986) that length polymorphisms and heteroplasmy may be more prevalent in the mtDNA of at least some lower vertebrate groups than in that of mammals or birds. They also provide the first evidence of geographically widespread forms of heteroplasmy that can be attributed to mutations that occurred independently in different mitochondrial lineages.

We are grateful for the help of M. J. DADSWELL, P. DUMONT,

G. JUDY, J. G. LOESCH, A. J. LUPINE, Y. MAILHOT, J. W. MCCORD, G. D. MELVIN, J. PROVOST, and the staff of the South Carolina Wildlife and Marine Resources Department in obtaining shad. This research was supported by grant A6513 from the Natural Sciences and Engineering Research Council of Canada to W.C.L. and grant EQ2177 from the Fonds pour la Formation de Chercheurs et l'Aide a la Recherche (FCAR) of the Province of Quebec to G.G.B. P.B. received fellowship support from the J. W. McConnell Foundation and the Fonds FCAR of Quebec.

LITERATURE CITED

AVISE, J. C., 1986 Mitochondrial DNA and the evolutionary genetics of higher animals. Phil. Trans. R. Soc. Lond. B **312:** 325–342.

AVISE, J. C., E. BIRMINGHAM, L. G. KESSLER and N. C. SAUNDERS, 1984 Characterization of mitochondrial DNA variability in a hybrid swarm between subspecies of bluegill sunfish (*Lepomis macrochirus*). Evolution **38:** 931–941.

AVISE, J. C., and R. A. LANSMAN, 1983 Polymorphism of mitochondrial DNA in populations of higher animals. pp. 147–164. In: *Evolution of Genes and Proteins*, Edited by M. NEI and R. K. KOEHN. Sinauer Associates, Sunderland, Mass.

BIRMINGHAM, E., T. LAMB and J. C. AVISE, 1986 Size polymorphism and heteroplasmy in the mitochondrial DNA of lower vertebrates. J. Hered. **77:** 249–252.

BIRKY, C. W., JR., T. MARUYAMA and P. FUERST, 1983 An approach to population theory for genes in mitochondria and chloroplasts, and some results. Genetics **103:** 513–527.

BOURSOT, P., H. YONEKAWA and F. BONHOMME, 1987 Heteroplasmy in mice with deletion of a large coding region of mitochondrial DNA. Mol. Biol. Evol. **4:** 46–55.

BROWN, G. G., and M. V. SIMPSON, 1981 Intra- and interspecific variation of the mitochondrial genome in *Rattus norvegicus* and *Rattus rattus:* restriction enzyme analysis of variant mitochondrial DNA molecules and their evolutionary relationships. Genetics **97:** 125–143.

BROWN, G. G., G. GADELETA, G. PEPE, C. SACCONE and E. SBISA, 1986 Structural conservation and variation in the D-loop containing region of vertebrate mitochondrial DNA. J. Mol. Biol. **193:** 503–511.

BROWN, W. M., 1983 Evolution of animal mitochondrial DNA. pp. 62–88. In: *Evolution of Genes and Proteins*, Edited by M. NEI and R. K. KOEHN. Sinauer Associates, Sunderland, Mass.

BROWN, W. M., 1985 The mitochondrial genome of animals. pp. 95–130. In: *Molecular Evolutionary Genetics*, Edited by R. J. MACINTYRE. Plenum Press, New York.

CANN, R. L., and A. C. WILSON, 1983 Length mutations in human mitochondrial DNA. Genetics **104:** 699–711.

CHAPMAN, R. W., and D. A. POWERS, 1984 A method for the rapid isolation of mitochondrial DNA from fishes. Tech. Rep. No. UM-SG-TS-84-05. Maryland Sea Grant Program, College Park, Md.

CHAPMAN, R. W., J. C. STEPHENS, R. A. LANSMAN and J. C. AVISE, 1982 Models of mitochondrial DNA transmission genetics and evolution in higher eucaryotes. Genet. Res. **40:** 41–57.

DENSMORE, L. D., J. W. WRIGHT and W. M. BROWN, 1985 Length variation and heteroplasmy are frequent in mitochondrial DNA from parthenogenetic and bisexual lizards (genus *Cnemidophorus*). Genetics **110:** 689–707.

FAURON, C. M.-R., and D. R. WOLSTENHOLME, 1976 Structural heterogeneity of mitochondrial DNA molecules within the genus *Drosophila*. Proc. Natl. Acad. Sci. USA **73:** 3623–3627.

FAURON, C. M.-R., and D. R. WOLSTENHOLME, 1980a Extensive diversity among *Drosophila* species with respect to nucleotide sequences within the adenine + thymine-rich region of mitochondrial DNA molecules. Nucleic Acids Res. **8:** 5391–5410.

FAURON, C. M.-R., and D. R. WOLSTENHOLME, 1980b Intraspecific diversity of nucleotide sequences within the adenine + thymine-rich region of mitochondrial DNA molecules of *Drosophila mauritiana, Drosophila melanogaster, Drosophila simulans*. Nucleic Acids Res. **8:** 5391–5410.

FERRIS, S. D., A. C. WILSON and W. M. BROWN, 1981 Evolutionary tree for apes and humans based on cleavage maps of mitochondrial DNA. Proc. Natl. Acad. Sci. USA **78:** 2431–2436.

GLAUS, K. R., H. P. ZASSENHAUS, N. S. FECHEIMER and P. S. PERLMAN, 1980 Avian mtDNA: structure, organization and evolution. pp. 131–135. In: *The Organization and Expression of the Mitochondrial Genome*, Edited by A. M. KROON and C. SACCONE. Elsevier/North Holland Biomedical Press, Amsterdam.

HALE, L. R., and R. S. SINGH, 1986 Extensive variation and heteroplasmy in size of mitochondrial DNA among geographic populations of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **83:** 8813–8817.

HARRISON, R. G., D. M. RAND and W. C. WHEELER, 1985 Mitochondrial DNA size variation within individual crickets. Science **228:** 1446–1448.

HAUSWIRTH, W. W., and P. J. LAIPIS, 1985 Transmission genetics of mammalian mitochondria: a molecular model and experimental evidence. pp. 49–59. In: *Achievements and Perspectives of Mitochondrial Research Volume II: Biogenesis*, Edited by E. QUAGLIARELLO, E. C. SLATER, F. PALMIERI, C. SACCONE and A. M. KROON. Elsevier Science Publishers, Amsterdam.

HAUSWIRTH, W. W., M.-J. VAN DE WALLE, P. J. LAIPIS and P. D. OLIVO, 1984 Heterogenous mitochondrial DNA D-loop sequences in bovine tissue. Cell **37:** 1001–1007.

KESSLER, L. G., and J. C. AVISE, 1985 A comparative description of mitochondrial DNA differentiation in selected avian and other vertebrate genera. Mol. Biol. Evol. **2:** 109–125.

LANSMAN, R. A., R. O. SHADE, J. F. SHAPIRA and J. C. AVISE, 1981 The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. III. Techniques and potential applications. J. Mol. Evol. **17:** 214–226.

MANIATIS, T., E. F. FRITSCH and J. SAMBROOK, 1982 *Molecular Cloning*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

MANSUETI, R., and H. KOLB, 1953 A historical review of the shad fisheries of North America. Md. Board Natur. Resources, Dep. Res. Educ., Publ. 97, Chesapeake Biol. Lab., Solomons, Md.

MONNEROT, M., J.-C. MOUNOLOU and M. SOLIGNAC, 1984 Intra-individual length heterogeneity of *Rana esculenta* mitochondrial DNA. Biol. Cell. **52:** 213–218.

MORITZ, C., and W. M. BROWN, 1986 Tandem duplication of D-loop and ribosomal RNA sequences in lizard mitochondrial DNA. Science **23:** 1425–1427.

RAND, D. M., and R. G. HARRISON, 1986 Mitochondrial DNA transmission genetics in crickets. Genetics **114:** 955–970.

REED, K. C., and D. A. MANN, 1985 Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res. **13:** 7207–7221.

REILLY, J. G., and C. A. THOMAS, JR., 1980 Length polymorphisms, restriction site variation, and maternal inheritance of mitochondrial DNA of *Drosophila melanogaster*. Plasmid **3:** 109–115.

RIGBY, P. W. J., M. DIECKMANN, C. RHODES and P. BERG, 1977 Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase. I. J. Mol. Biol. **113:** 237–251.

SOLIGNAC, M., M. MONNEROT and J.-C. MOUNOLOU, 1986 Concerted evolution of sequence repeats in *Drosophila* mitochondrial DNA. J. Mol. Evol. **25:** 53–60.

SOLIGNAC, M., J. GENERMONT, M. MONNEROT and J.-C. MOUNOLOU,

1984 Genetics of mitochondria in Drosophila: mtDNA inheritance in heteroplasmic strains of *D. mauritiana*. Mol. Gen. Genet. **197:** 183–188.

SOUTHERN, E. M., 1975 Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:** 503–517.

UPHOLT, W. B., and I. B. DAWID, 1977 Mapping of mitochondrial DNA of individual sheep and goats: rapid evolution in the D loop region. Cell **11:** 571–583.

WOLSTENHOLME, D. R., J. M. GODDARD and C. M.-R. FAURON, 1979 Structure and replication of mitochondrial DNA from the genus Drosophila. In: *Extrachromosomal DNA*, Edited by D. J. CUMMINGS, P. BOERST, I. B. DAWID, S. M. WEISSMAN and C. F. FOX. Academic Press, New York.

Communicating editor: J. AVISE