MOLECULAR PHYLOGENETICS AND EVOLUTION
Vol. 10, No. 3, December, pp. 310–322, 1998
ARTICLE NO. FY980530

Tandemly Repeated Sequences in the Mitochondrial DNA Control Region and Phylogeography of the Pike-Perches *Stizostedion*

Joseph E. Faber[1] and Carol A. Stepien[2]

Department of Biology, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106-7080

Received August 22, 1997; revised March 9, 1998

DNA sequences from the mitochondrial DNA control region are used to test the phylogeographic relationships among the pike-perches, *Stizostedion* (Teleostei: Percidae) and to examine patterns of variation. Sequences reveal two types of variability: single nucleotide polymorphisms and 6 to 14 copies of 10- to 11-base-pair tandemly repeated sequences. Numbers of copies of the tandem repeats are found to evolve too rapidly to detect phylogenetic signal at any taxonomic level, even among populations. Sequence similarities of the tandem repeats among *Stizostedion* and other percids suggest concerted evolutionary processes. Predicted folding of the tandem repeats and their proximity to termination-associated sequences indicate that secondary structure mediates slipped-strand mispairing among the d-loop, heavy, and light strands. Neighbor-joining and maximum parsimony analyses of sequences indicate that the genus is divided into clades on the continents of North America and Eurasia. Calibrating genetic distances with divergence times supports the hypothesis that *Stizostedion* dispersed from Eurasia to North America across a North Pacific Beringial land bridge approximately 4 million years before present, near the beginning of the Pliocene Epoch. The North American *S. vitreum* and *S. canadense* appear separated by about 2.75 million years, and the Eurasian *S. lucioperca* and *S. volgensis* are diverged by about 1.8 million years, suggesting that speciation occurred during the late Pliocene Epoch.

INTRODUCTION

The vertebrate control region contains the only noncoding sequences of the mitochondrial (mt) genome. It is composed of a central conserved section that is flanked by highly variable left and right domains (Brown *et al.*, 1986). The central section contains a conserved sequence block-D (CSB-D; Saccone *et al.*, 1991) that is involved in heavy strand replication, including initiation by a three-stranded displacement loop (d-loop; Clayton, 1982). The left domain usually has one or more copies of the termination associated sequence (TAS; Clayton, 1984), that signals termination of the d-loop strands (Doda *et al.*, 1981). The right domain contains the origin site for heavy strand replication (Ori-H) and conserved sequence blocks (CSB 1, 2, and 3) that are believed to function in its initiation (Doda *et al.*, 1981; Walberg and Clayton, 1981; Saccone *et al.*, 1991).

In addition to these conserved sections, the control region contains intervening sequences in the left and right domains that have high levels of polymorphism, rendering them useful for addressing evolutionary relationships among populations and species (Moritz *et al.*, 1987; Palumbi, 1996). Two types of sequence variability are most commonly found in the control region of teleost fishes: nucleotide polymorphisms and variable numbers of copies of tandemly repeated sequences (Lee *et al.*, 1995; Stepien and Kocher, 1997; Cesaroni *et al.*, 1997). Nucleotide polymorphisms (point mutations) in the control region have been used to distinguish among species (e.g., Bermingham *et al.*, 1986; Stepien, 1995; Phillips and Oakley, 1997; Turner, 1997; Stepien *et al.*, 1998) and population groups of fishes (e.g., Stepien, 1995; Sturmbauer *et al.*, 1997; Nielsen *et al.*, 1997; Stepien *et al.*, 1998).

Tandemly repeated sequences are found in the control region of many vertebrates, although their phylogenetic utility is not resolved (Hoelzel *et al.*, 1994). In some studies, nucleotide substitutions (point mutations) within the tandem repeats of the control region appeared to reflect phylogeny (e.g., among species of ursid bears, Hoelzel *et al.*, 1994; acipenserid sturgeon, Brown *et al.*, 1996; and the minnow *Cyprinella spiloptera*, Broughton and Dowling, 1997). In other studies, they did not provide phylogenetic signal (e.g., among species of mammals, Hoelzel *et al.*, 1994). Variable numbers of copies of tandem repeats provided reliable phylogenies in some studies (e.g., among populations of white sturgeon *Acipenser transmontanus;* Brown *et al.*,

[1] Present Address: West Virginia University, Parkersburg, WV 26101.

[2] To whom correspondence and reprint requests should be addressed. E-mail: cas20@pop.cwru.edu.

1055-7903/98 $25.00

1992) and were uninformative in others (e.g., among populations of the Atlantic cod *Gadus morhua*, Arnason and Rand, 1992; and populations of the walleye *Stizostedion vitreum*, Stepien and Faber, 1998).

The teleost fish genus *Stizostedion* (Teleostei: Percidae) provides an opportunity to study the evolution of both types of polymorphisms in the mtDNA control region. *Stizostedion* comprises five species in the Northern hemisphere, four of which are distributed widely and are highly valued as game fishes. The natural ranges of the walleye *S. vitreum* and the sauger *S. canadense* extend throughout much of eastern and northern North America (Fig. 1). The native ranges of zander *S. lucioperca*, the Volga zander *S. volgensis*, and the rare sea pike-perch *S. marinum* are in central and western Eurasia (Fig. 1; Collette and Banarescu, 1977). Fossil evidence suggests that the Percidae originated in western Eurasia (Collette and Banarescu, 1977). *Stizostedion* is believed to have dispersed from Eurasia to North America, and its dispersal route and time of divergence are disputed (reviewed by Collette and Banarescu, 1977). The primary hypothesis (hypothesis A of Fig. 1) is that *Stizostedion* crossed a North Pacific land bridge during the Neogene Period (about 1.8 to 24 million years before present (mybp); Collette and Banarescu, 1977). Alternatively, present-day distributions in eastern North America and western Eurasia suggest that the genus may have dispersed using a North Atlantic route. This may have occurred either via rivers across a land bridge at the end of the Cretaceous or the beginning of the Eocene, 37 to 144 mybp (hypothesis B of Fig. 1; Svetovidov and Dorofeeva, 1963; Balon *et al.*, 1977) or by using brackish water at the edge of retreating glaciers during the late Pleistocene, about 15,000 ybp (hypothesis C of Fig. 1; Cihar, 1975).

Evolutionary relationships among these species have not been tested extensively using genetics. Whole mtDNA RFLPs and allozyme polymorphisms confirmed evolutionary divergences among North American walleye and sauger and the Eurasian zander and Volga zander (Billington *et al.*, 1990, 1991). RFLPs identified simple sequence polymorphisms and length heterogeneity in walleye due to insertions/deletions in the vicinity of the control region (Billington *et al.*, 1990, 1991). Phylogenies of relationships among *Stizostedion* by Billington *et al.* (1990, 1991) were based entirely on presence/absence of restriction sites and did not include sequence data.

The objectives of the present investigation are to (1) discern systematic and phylogeographic relationships among the species of *Stizostedion* and test the three hypotheses of transcontinental dispersal, (2) compare repeated and nonrepeated sequences and elucidate the structure of the mtDNA control region, and (3) evaluate the possible mechanism governing the formation and maintenance of repeats.

MATERIALS AND METHODS

Samples Collected

Collection sites for four species of *Stizostedion* are shown in Fig. 1. Samples of 199 walleye (*S. vitreum*) from the Great Lakes region and the Ohio River (which were sequenced for population genetic studies reported in Stepien, 1995; Faber and Stepien, 1997; and Stepien and Faber, 1998) and four sauger (*S. canadense*) from the Ohio River were analyzed (collections are described in Faber and Stepien, 1997; and Stepien and Faber, 1998). Zander (*S. lucioperca*, $N = 2$) from England and Volga zander (*S. volgensis*, $N = 1$) from the Danube River, Czech Republic also were sequenced. Specimens were collected by seine, electroshocking, or hook and line. Whole individuals or tissues (fin, muscle, eggs, or liver) were either immediately frozen at −80°C or preserved in 95% ethanol in the field. Frozen samples

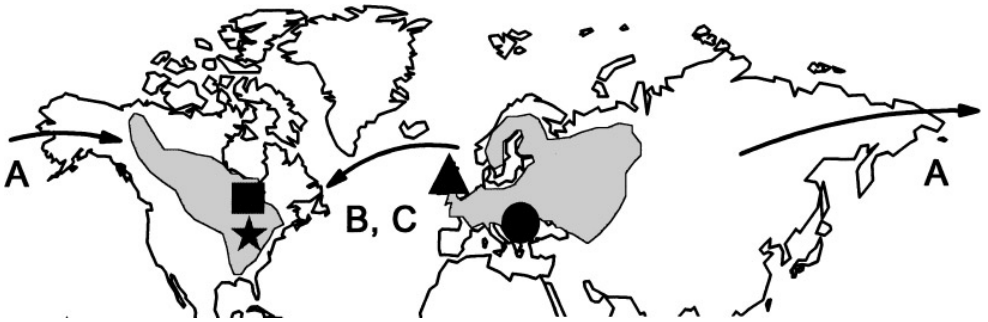


FIG. 1. Geographic ranges, sampling sites, and hypotheses for dispersal of *Stizostedion* from Eurasia to North America. Endemic geographic distributions are outlined. Samples include; the walleye *S. vitreum* ■ from the North American Great Lakes and Ohio River ($N = 199$), the sauger *S. canadense* ★ from the North American Ohio River ($N = 4$), the zander *S. lucioperca* ▲ from the Ouse River, England ($N = 2$), and the Volga zander *S. volgensis* ● from the Danube River, Czechoslovakia ($N = 1$). Dispersal hypotheses include (A) across the North Pacific Beringia land bridge during the Neogene 1.8–24 mybp (Collette and Banarescu, 1977), (B) over a land bridge in the North Atlantic at the end of the Cretaceous or beginning of the Eocene, 37–144 mybp (Svetovidov and Dorofeeva, 1963; Balon *et al.*, 1977), and (C) via brackish water margins of retreating glaciers in the North Atlantic during the late Pleistocene 15,000 ybp (Cihar, 1975).

were stored at −80°C, and ethanol-preserved materials were stored at room temperature prior to DNA extraction.

Genetic Analysis

Approximately 1–2 g tissue from each sample was frozen and ground in liquid nitrogen using a cylindrical stainless steel mortar and pestle (Stepien, 1995). DNA was extracted in guanidine thiocyanate buffer and purified using proteinase K, RNase, phenol, and chloroform, following standard protocols (Sambrook *et al.*, 1989; Stepien *et al.*, 1993; Stepien, 1995). The entire mtDNA control region was amplified in three sections using conserved primers (Kocher *et al.*, 1989; Meyer *et al.*, 1990) and the polymerase chain reaction (PCR; Mullis and Faloona, 1987). The light chain 5′ end or 'left' domain of the control region (from the proline tRNA gene to the central conserved section) was amplified using the oligonucleotide primers L15926, 5′-TCA AAG CTT ACA CCA GTC TTG TAA ACC-3′ (Kocher *et al.*, 1989) and H16498, 5′-CCT GAA GTA GGA ACC AGA TG-3′ (Meyer *et al.*, 1990). The light chain 3′ end or 'right' domain of the control region (from the central conserved section to the phenylalanine tRNA gene) was amplified with the light strand complement of H16498, L16498 5′-CAT CTG GTT CCT ACT TCA GG-3′ and H503 5′-GCA CGA GAT TTA CCA ACC C-3′ (Titus and Larson, 1995). The central conserved section (167 base pairs (bp) in length) was amplified with custom primers designed from sequences conserved among species of the family Percidae; L16378, 5′-AAT GTA GTA AGA GCC TA-3′ and H16578, 5′-GGG TAA CGA GGA GTA TG-3′ (Faber and Stepien, 1997). Heavy chain primers were end-labeled with biotin (Hultman *et al.*, 1989) and the DNA strands were separated using Dynabeads M-180 streptavidin (Dynal Corp., Oslo, Norway) for single-strand sequencing (Hultman *et al.*, 1989; Uhlen, 1989). Both strands were sequenced separately using diluted PCR primers and ^{35}S-labeled dATP with Sanger di-deoxy sequencing (Sanger *et al.*, 1977) and Sequenase Version 2.0 Sequencing kits (Amersham/U.S. Biochemical Corp., Cleveland, OH, product No. 70770). Sequencing reactions were run on 6% polyacrylamide gels for two, five, and eight h in order to resolve approximately 600 bp from the primer, and bands were visualized by autoradiography.

Data Analysis

DNA sequences were read into a Macintosh computer using an IBI/Kodak digitizer, aligned using MacVector-AssemblyLIGN software (IBI, 1992), and manually checked. Structural features in the control region were identified by comparison with sequences in other teleost fishes (Lee *et al.*, 1995; Stepien, 1995; Stepien *et al.*, 1998). Locations of tandemly repeated sequences were compared with those of other vertebrates following Fumagalli *et al.* (1996). Relative proportion of polymorphic nucleotides (p_n) was calculated following Nei (1987). Secondary folding structures and their free energies were calculated for nucleotide sequences of repeats using the MFOLD program, vers. 3.2 (Zuker, 1989; SantaLucia *et al.*, 1996; Zuker, 1996). Calculations were conducted at a temperature of 10°C, near the physiological optimum for percid fishes (Craig, 1987). The most energetically stable (lowest free energy) structure is reported for each species.

Phylogenies were analyzed using two methods: genetic distance neighbor joining (NJ) clustering and maximum parsimony (MP) analysis of character states. Kimura two-parameter distances (Kimura, 1980) were used to correct for unequal number of transitional versus transversional substitutions often observed in mtDNA sequence data (Brown *et al.*, 1982; Kocher and Carlton, 1997). Pairwise (p) genetic distances (Nei, 1987) were calculated in addition to Kimura two-parameter distances, since the former are more commonly used to estimate divergence times (Hillis *et al.*, 1996). Standard errors of all distances were calculated following Kumar *et al.* (1993). Because there were no appreciable differences between the two distance measures (they varied by 0.001 to 0.005, and averaged 0.003), evolutionary times were estimated from the Kimura (1980) two-parameter distances. Genetic distances and possible divergence times were compared and calibrated with those estimated for walleye from RFLPs of the entire mtDNA molecule by Ward *et al.* (1989) and Billington *et al.* (1990, 1991) and with those for the percid genus *Gymnocephalus* from mtDNA control region sequences (Stepien *et al.*, 1998). In the absence of a fossil record, calibration of divergence times with other taxa and other data sets is often used (Avise, 1994). In the case of walleye and other percids, in areas of the mtDNA control region other than the repeats, the calibration of 2% sequence divergence per million years appears to be accurate (see Stepien and Kocher, 1997; Stepien *et al.*, 1998), corresponding to an average rate for the mtDNA molecule (Avise, 1994; Palumbi, 1996). This rate is slower than that of the mtDNA control region in mammals (reviewed in Avise, 1994), but average for fishes, whose mtDNA evolves more slowly, apparently related to poikilothermy (Martin and Palumbi, 1993; Avise, 1994).

A distance neighbor-joining tree (NJ; Saitou and Nei, 1987) was constructed using MEGA (Molecular Evolutionary Genetics Analysis, vers 1.01; Kumar *et al.*, 1993), and support for the individual nodes was tested with 1000 bootstrap replicates (Rzhetsky and Nei, 1992). Most parsimonious relationships among species and genera were evaluated using the branch and bound algorithm (Hendy and Penny, 1982) in PAUP 3.1.1 (Phylogenetic Analysis Using Parsimony vers. 3.1.1; Swofford, 1993), and support for nodes of the cladograms was estimated from 1000 bootstrap replications (Felsenstein, 1985). DNA sequences from other percids, including the ruffe *Gymnocephalus cernuus* (GenBank Accession No.

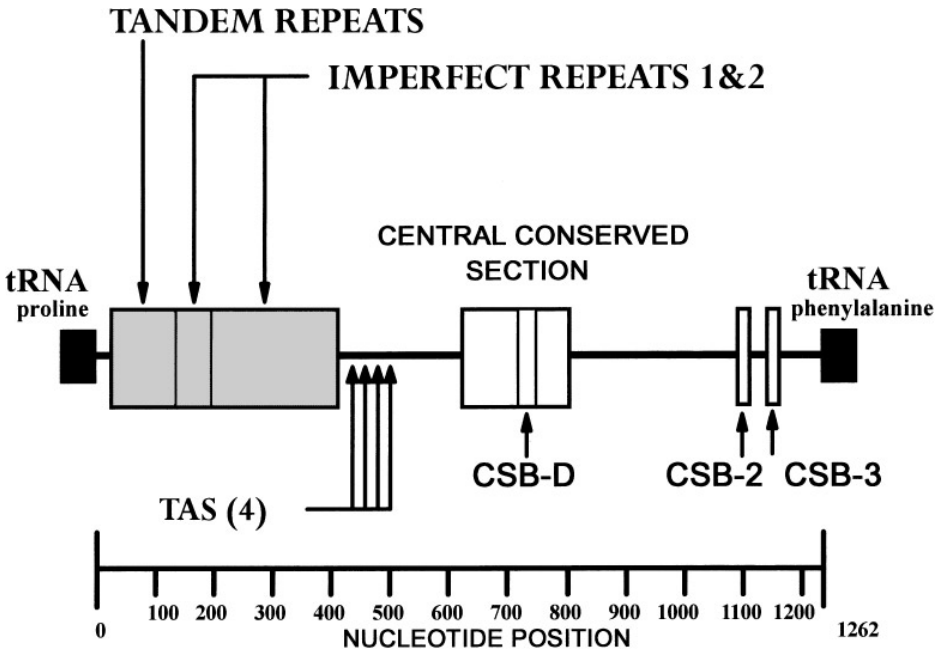


FIG. 2. Structure and sites of variability in the mitochondrial DNA control region of *Stizostedion*. The control region is flanked by sequences that code for transfer RNAs. The control region consists of conserved sections, including the termination associated sequence (TAS; Doda *et al.*, 1981; Southern *et al.*, 1988), the central conserved section (Lee *et al.*, 1995), and conserved sequence boxes D, 2, and 3 (CSB-D, CSB-2, and CSB-3; Southern *et al.*, 1988; Saccone *et al.*, 1991), intervened by variable sections that include tandem repeats and base substitutions. The length of the control region varies from 918 bp in the zander *S. lucioperca* to 1262 bp in the walleye *S. vitreum*. Nonrepeated sequences range from 853 bp in zander to 857 bp in the sauger *S. canadense*. Length heterogeneity is due primarily to variable numbers of tandem repeats, ranging from 65 bp in the zander and Volga zander *S. volgensis* to 414 bp in *S. vitreum*. The control region of walleye varies from 1181 to 1262 bp due to variable numbers of repeats (326 to 414 bp in length).

U90620; Faber and Stepien, 1997; Stepien *et al.*, 1998) and the banded darter *Etheostoma zonale* (GenBank Accession No. U90621; Faber and Stepien, 1997), were used as outgroups in the NJ and MP analyses.

RESULTS

Sequence Polymorphisms

DNA sequences are reported in GenBank with the following Accession Nos.: walleye *S. vitreum* U90617, sauger *S. canadense* U90618, zander *S. lucioperca* U90624, and Volga zander *S. volgensis* AF007824. Alignment of mtDNA control region sequences reveals conserved sections with intervening variable sections (Fig. 2). The left domain contains four putative termination associated sequences (TAS; Table 1). A central conserved section is found in the downstream 3′ direction that includes the conserved sequence block-D (CSB-D). Conserved sequence blocks-2 and 3 (CSB-2, 3) are located in the right domain near the phenylalanine tRNA gene (Fig. 2). A single PCR band and one DNA sequence was found in every individual sampled.

Total length of the control region is similar among species of *Stizostedion:* 918 bp in the zander, 919 bp in the Volga zander, 1141 bp in the sauger, and ranging from 1184 to 1262 bp in the walleye (Fig. 2). Length heterogeneity in walleye is due to variable numbers of repeated sequences at the repeat site-1 (analogous to the R1 of other vertebrates; Fumagalli *et al.*, 1996), beginning 21 bp downstream of the proline tRNA gene (Fig. 2). Repeated sequences account for 65 bp of the control region in zander and the Volga zander, 284 bp in the sauger, and range from 326 to 414 bp in the walleye (Table 2).

Excluding the tandem repeats, the remaining lengths of the control region are similar among the species:

TABLE 1

DNA Sequences of Four Putative Termination Associated Sequences (TAS) in the Mitochondrial DNA Control Region of *Stizostedion*

Number	Nucleotide position	Sequence
TAS 1	37–50	ACATCTATATTAAC *T*
TAS 2	71–83	ACATATATGTTTT
TAS 3	105–117	ACATTCATATATC *C C*
TAS 4	136–146	ACAT--AAAGCAT

Note. A dash indicates an alignment gap and substitutions among taxa are italicized below the sequence for walleye.

TABLE 2

Sequences of Repeats from the Repeat Section-1 of the Mitochondrial DNA Control Region (R1) for Four Species of *Stizostedion* and the Outgroups *Gymnocephalus cernuus* and *Etheostoma zonale*

Taxon	Total N nucleotides	Perfect repeat		Imperfect repeat −1	
		Sequence	Number of repeats	Sequence	Number of repeats
banded darter *Etheostoma zonale* (N = 1)	107	GCAAGAGTTT	9	GCAAACGTTTACACACG	1
ruffe *Gymnocephalus cernuus* (N = 10)	182	GCAAGTATTT	17	GCAAATACAATT	1
Zander *Stizostedion lucioperca* (N = 2)	65	GCAAGTATTT	5	GCAAGCAAATACATA	1
Volga zander *S. volgensis* (N = 1)	65	GCAAGCATTT	5	GCAAGCAAATACATA	1
Sauger *S. canadense* (N = 4)	284	GCCCAAACAT	10	(a) GCACAAACAT	9
				(b) GCAAGTATTTA	4
				(c) GCAAGTATTA	1
				(d) GCAAGTATTT	3
Walleye *S. vitreum* (N = 199)	326–414	GCAAATATTTT	6–14	(e) GCAAATTTT	1
		C		(f) CGCCAATTTT	0–4
				A A	
				G	
				(g) GCCAAGCGATT	0–1

	Imperfect Repeat-2	
	Sequence	Number of repeats
Walleye *S. vitreum* (N = 199)	AGTACTCACTTAATAGTCATTTAACAAGCATTTAACA *G* GCATTTAACAACGTTTAGATGTCATTTATAGTCATTTA *A A G G* *C* CAACATTTAACAACATTTAGCAAGCAATTAGTAGTCA *G G A* CTTAATAGTCATTTAACAAACATTTAACAAACATTTA *G G* ACAAGCATTTAGTAGGCGTTTAGCAGGCATTTAACAA *A A* GCAATTACAAACATTTGATAGCACAAAATACATA *T T CC G*	1

Note. Repeats are assigned to three different groups depending on the nucleotide sequence; perfect repeat, imperfect repeat-1, and imperfect repeat-2. The three types of repeats occur in order from the 5′ to 3′ direction, starting at control region nucleotide 21. The sauger (*S. canadense*) has four imperfect repeats-1 that are lettered in the order of their occurrence (a through d) and the walleye (*S. vitreum*) has three imperfect repeats-1 that are lettered in the order of their relative occurrence (e–g). Intraspecific polymorphisms among haplotypes are italicized below the consensus sequences, in order of decreasing relative frequency.

numbering 853 bp in the zander, 854 bp in the Volga zander, 855 bp in the walleye, and 857 bp in the sauger. Lengths of the control region also are similar in the outgroups: 853 bp in the banded darter and 854 bp in the ruffe. Aligning the nonrepeated sequences among species of *Stizostedion* reveals a total of 95 polymorphic nucleotides (p_n = 0.111), with 39 sites (p_n = 0.147) in the left domain, 50 sites (p_n = 0.116) in the right domain, and 6 in the central conserved section (p_n = 0.036). The ruffe and the banded darter outgroups differ from the four species of *Stizostedion* at 59 nucleotide sites. The walleye varies intraspecifically by 22 nucleotide polymorphisms in the nonrepeated sequences (N = 199; p_n = 0.026), with 14 (p_n = 0.053) in the left domain, 8 (p_n = 0.019) in the right domain, and 0 in the central conserved section. Including all repeated and nonrepeated sequences, walleye from the geographically isolated Great Lakes and Ohio River drainage systems differ by 27 fixed nucleotides. Single-nucleotide polymorphisms identified in the nonrepeated sequences of the control region of walleye show geographic patterning, and their evolutionary significance is treated elsewhere (Faber and Stepien, 1997; Stepien and Faber, 1998). No intraspecific polymorphisms are evident in the zander (N = 2) or the sauger (N = 4) sequenced.

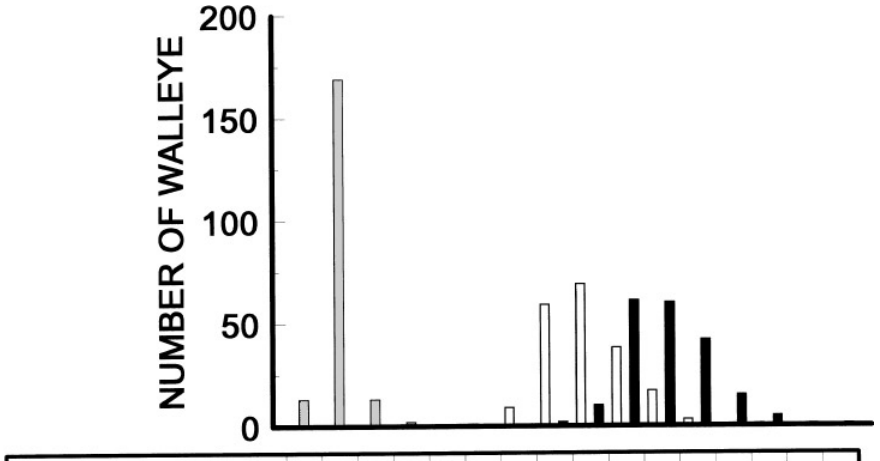
Tandem Repeats

Three types of repeats are found in the 5′ to 3′ direction at the repeat section-1 (R1; Fumagalli *et al.*, 1996): perfect repeats (PR), imperfect repeats-1 (IR-1), and imperfect repeats-2 (IR-2) (Fig. 2). PR are arrays of 10 or 11 bp that are identical in sequence and have 5 to 17 exact copies. Sequences of PR follow the basic motif GCAA(pur)(A/T)(pur)TTT(pyr/−) in *Stizostedion*, except for sauger which has the sequence GCCCAAACAT (Table 2). Only the zander and the ruffe have identical sequences of the PR among the percids sequenced here (Table 2). The sequences of PR in the zander and the Volga zander differ by a single fixed transitional substitution (Table 2). The Eurasian and North American taxa differ by one to two fixed nucleotides and a single insertion/deletion in the walleye and by six fixed nucleotides in the sauger (Table 2). The sequences of PR in the sauger and walleye differ by five fixed nucleotides and one insertion/deletion (Table 2).

The walleye has 6 to 14 PR copies, and copy number appears fixed in the other species (Table 2). The frequency of PR copies of walleye is normally distributed, and 9 copies occur most frequently (in 35%; Fig. 3). The sequence of the PR is GCAAATATTT**T** in walleye from the Great Lakes and GCAAATATTT**C** in 3 of 11 individuals from the geographically isolated Ohio River (Stepien and Faber, 1998).

The IR-1 are similar in sequence and length to the PR, but are fewer in copy number (ranging from 1 to 9; Table 2). The zander and the Volga zander share the synapomorphy of a single copy of IR-1, that is 15 bp in length and identical in sequence. Sequences and length of the IR-1 are similar between the walleye and sauger, ranging from 9 to 11 bp in length (Table 2). Four types of IR-1 (IR-1a–d; Table 2) occur in the sauger and number 17 copies, totaling 184 bp. The walleye has three types of IR-1 (IR-1e–g; Table 2) that vary in copy number among haplotypes and total 20 to 60 bp (Table 2). The number of IR-1 copies is normally distributed in the walleye (Fig. 3). The primary sequence of the IR-1f (Table 2) also varies among individual walleye. The IR-2 is based loosely on the motif AACAACATTT and occurs only in walleye, totaling 220 bp (Table 2). The sequence of IR-2 varies among individual haplotypes, and the total length of this section (220 bp) is conserved. Point substitutions are found at 17 nucleotide positions, and 203 sites are invariable.

Sequences of the R1 appear to form stable secondary folding structures (Fig. 4). The free energies of the structures are −23 kcal M^{-1} in the Volga zander, −26.8 kcal M^{-1} in the zander, −72.2 kcal M^{-1} in the sauger, and up to −126.4 kcal M^{-1} for walleye. Repeated sections of the zander and Volga zander form relatively simple structures, with PR and one IR-1 folding into a single stem. Structures in the sauger and walleye are larger and complex, with a greater number of stems and loops. The PR of sauger fold into five stems and loops, and the IR-1 folds to form a single stem and loop. The PR of walleye form a single stem and loop, and the IR-1 and IR-2 sections produce a more complex struc-



NUMBER OF REPEATS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
PERFECT □						1	9	59	69	38	17	3	0	1		
IMPERFECT-1 ▨	13	169	13	2												
PERFECT AND IMPERFECT-1 ■								2	10	61	60	42	15	5	1	1

FIG. 3. Frequency histograms for the number of copies of the perfect repeat (PR) and imperfect repeat-1 (IR-1) among 196 walleye (*Stizostedion vitreum*) sampled from the North American Great Lakes region. White bars indicate the frequency distribution of perfect repeats, hatched bars indicate the frequency distribution of imperfect repeats-1, and filled bars show the combined frequencies.

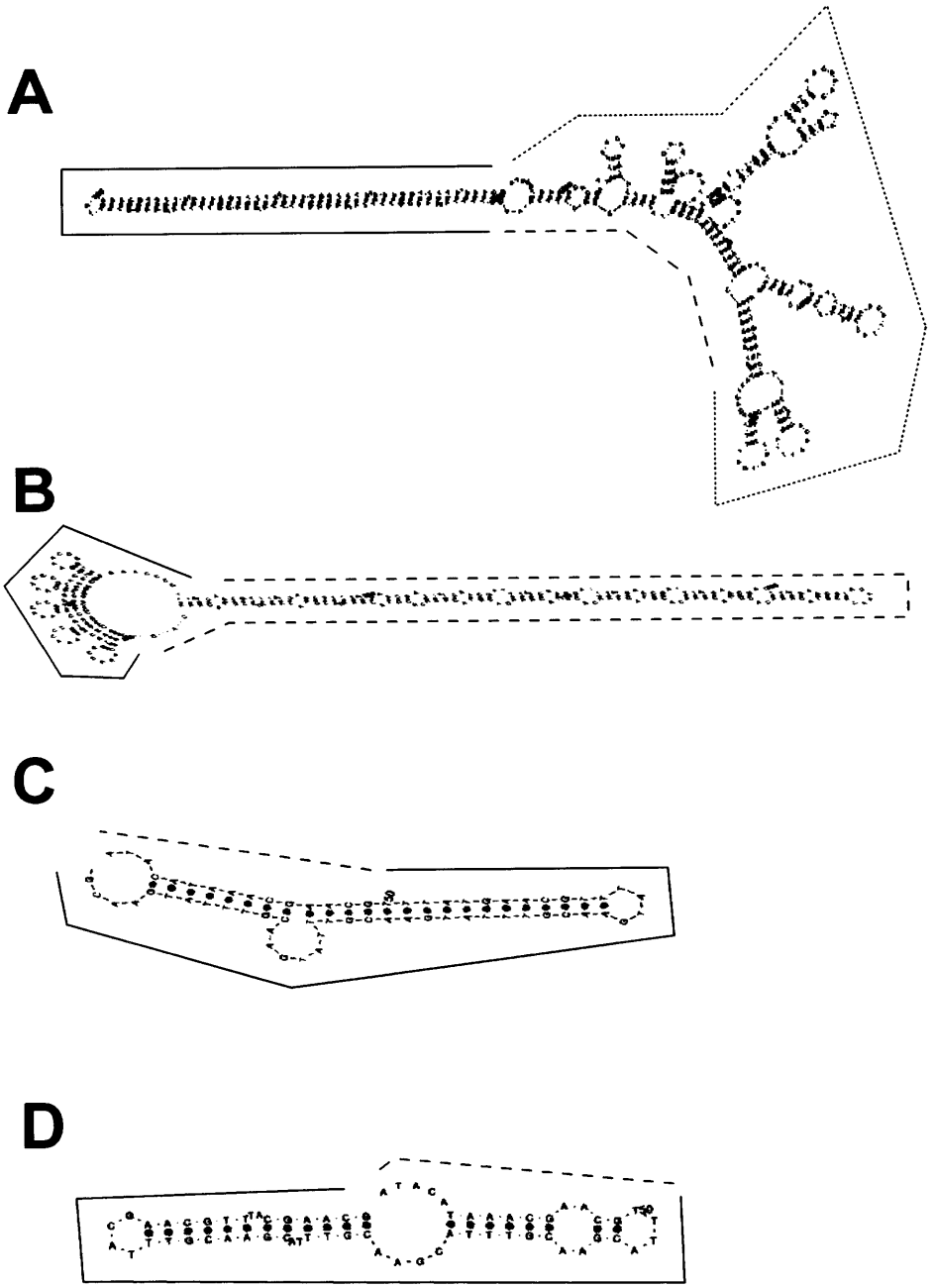


FIG. 4. Secondary folding structures of repeat sequences in the mtDNA control region of species of *Stizostedion*, determined using MFOLD vers. 2.3 (Zuker, 1996). Folding structures of tandem repeats are shown for (A) walleye, *S. vitreum*, 358 bp including nine perfect repeats, with an estimated free energy of -94.8 kcal M^{-1}; (B) sauger, *S. canadense*, 274 bases with estimated free energy of -72.2 kcal M^{-1}; (C) zander, *S. lucioperca*, 65 bp with an estimated free energy of -26.8 kcal M^{-1}; and (D) Volga zander, *S. volgensis*, 65 bp with free energy of -23 kcal M^{-1}. ——, perfect repeats; – – –, imperfect repeat 1; ·····, imperfect repeat-2.

ture (Fig. 4). Free energies associated with different numbers of PR in the walleye are shown in Table 3. Secondary structures become increasingly energetically stable as the number of perfect repeat copies increases.

Phylogenetic Relationships

Genetic distance analysis of the nonrepeated control region sequences produces the NJ tree shown in Fig. 5. Two primary bifurcations are supported, with North American walleye and sauger clustering together and Eurasian zander and Volga zander in the other clade. The North American and Eurasian taxa are separated by an average genetic distance of 0.081 ± 0.010, suggesting an evolutionary divergence of 4.05 ± 0.50 mybp. The North American sauger and walleye are separated by an average genetic distance of 0.055 ±

TABLE 3

Free Energies of Folding Structures for Variable Numbers of Perfect Repeats in the mtDNA Control Region Repeat Section of the Walleye, *S. vitreum*

Number of perfect repeats	Free energy ($-$kcal M^{-1})
6	76.4
7	82.3
8	88.9
9	94.8
10	101.4
11	107.3
12	113.9
13	119.8
14	126.4

Note. Free energies were calculated using the MFOLD program, vers 3.2 (Zuker, 1996).

0.008, supporting a divergence of 2.75 ± 0.40 mybp. Walleye endemic to the Ohio River and those in the Great Lakes are diverged by an average distance of 0.029 ± 0.005 and 1.45 ± 0.25 mybp. In Eurasia, the zander and Volga zander diverge by a genetic distance of 0.037 ± 0.006, equivalent to 1.85 ± 0.30 mybp. Combining the nonrepeated sequences and a single copy of the PR for each species increases the genetic distances among taxa, but not by a significant degree. For example, the average genetic distance between the North American and Eurasian taxa increases by 0.004 ± 0.001 (200,000 ± 50,000 years), supporting a divergence of 4.25 ± 0.5 mybp. The North American walleye and sauger are then diverged by 3.05 ± 0.4 mybp (an increase of 300,000 years), the Ohio River and Great Lakes walleye are separated by 1.65 ± 0.20 mybp (an increase of 200,000 years), and the Eurasian zander and Volga zander are separated by 1.90 ± 0.30 mybp (an increase of 100,000 years).

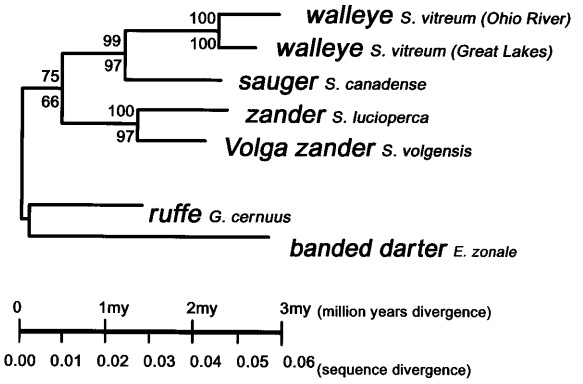


FIG. 5. The neighbor joining tree (MEGA; Kumar *et al.*, 1993) from mtDNA control region nonrepeated sequences for four species of *Stizostedion*, using two-parameter pairwise genetic distances (Kimura, 1980). The tree is rooted to the percid ruffe *Gymnocephalus cernuus* and the banded darter *Etheostoma zonale*. A single most parsimonious tree (121 steps) supporting identical relationships among *Stizostedion* was found using the branch and bound algorithm in PAUP (Vers 3.1.1; Swofford, 1993). Values for 1000 bootstrap replications from the neighbor joining analysis (Rzhetsky and Nei, 1992) are listed above the nodes, and those for parsimony (Felsenstein, 1985) are below.

Cladistic analysis yields a single most parsimonious tree that is congruent with the NJ tree. The most parsimonious tree is 121 steps long with a consistency index excluding uninformative characters of 0.75. There is a single next-most parsimonious tree of 123 steps that rearranges the banded darter as the sister group to the clade containing the walleye and the sauger. All relationships in the NJ and MP trees are supported by significant bootstrap values (Fig. 5).

DISCUSSION

Structure of the Control Region

Identification of a central conserved section, variable left and right domains, putative TAS sequences (Doda *et al.*, 1981), and CSB-D, 2, and 3 (Southern *et al.*, 1988; Saccone *et al.*, 1991) indicate that the sequences required for heavy strand replication in vertebrate mtDNA also are conserved in *Stizostedion*. Their similarity in the Atlantic cod *Gadus morhua* (Arnason and Rand, 1992), the fish family Percidae (Faber and Stepien, 1997; Stepien *et al.*, 1998; Nesbo *et al.*, 1998) the salmonids (Shedlock *et al.*, 1992), and several other taxa of bony fishes (Lee *et al.*, 1995), suggests that the mode of DNA replication is conserved. The putative TAS sequences of *Stizostedion* are somewhat variable (Table 1) and are not found in the tandem repeats (Table 2), unlike those of some other groups (e.g., acipenserid sturgeon, Buroker *et al.*, 1990; Brown *et al.*, 1996; and pleuronectiform flatfishes, Lee *et al.*, 1995). A sequence for Ori-H was not found in *Stizostedion* and does not occur in most other teleosts that have been studied (Lee *et al.*, 1995). This suggests differences in replication of teleost mtDNA from other vertebrates (Clayton, 1982). Simple point mutations (polymorphic nucleotides) in the control region of *Stizostedion* are interspersed among arrays of the repeats and the CSBs. Most of the variability is located in the left and right domains, similar to other fishes (e.g., Lee *et al.*, 1995; Stepien, 1995; Stepien *et al.*, 1998) and vertebrates (Brown, 1986; reviewed by Moritz *et al.*, 1987).

Phylogeography of Stizostedion

Genetic distance and cladistic analyses based on simple point mutations in nonrepeated sequences support a primary evolutionary divergence between the North American and Eurasian species of *Stizostedion* (Fig. 5). Time estimates suggest 4.05 ± 0.50 mybp of divergence between the North American and Eurasian taxa, supporting hypothesis A (Fig. 1) that *Stizostedion* dispersed from Eurasia to North America via Beringia during the early Pliocene Epoch. These results refute the hypotheses of dispersal across the North Atlantic

near the end of the Cretaceous or the beginning of the Eocene (hypothesis B of Fig. 1; Svetovidov and Dorofeeva, 1963; Balon *et al.*, 1977) and during the late Pleistocene (hypothesis C of Fig. 1; Cihar, 1975).

Patterns and estimates of evolutionary divergence between the North American and Eurasian taxa are in general agreement with those of Billington *et al.* (1990, 1991), who used allozyme polymorphisms and whole mtDNA RFLPs to hypothesize that intercontinental dispersal of *Stizostedion* occurred via Beringia. Our genetic distances are smaller than those of Billington *et al.* (1991) at the higher taxonomic levels, suggesting a calibration discrepancy among the different molecular data sets. Divergence differences among the data sets may reflect a large number of third codon position changes in protein coding genes and differences in stem and loop regions in rDNA encoding regions in the Billington *et al.* (1990, 1991) whole mtDNA RFLP analyses (Palumbi, 1996; Stepien and Kocher, 1997). Alternatively, it may be that the control region evolves more slowly in *Stizostedion* than does the mtDNA molecule as a whole.

Taxa from North America and Eurasia were estimated by Billington *et al.* (1990, 1991) to have diverged during the middle to late Miocene Epoch, about 8 to 10.5 mybp from allozyme polymorphisms and 7.86 ± 1.15 mybp from mtDNA RFLPs. Our results and calibration suggest a more recent dispersal during the Pliocene Epoch. All of these divergence estimates are plausible, as the Beringial Isthmus was present during low sea levels in the middle to late Miocene and again in the early Pliocene (Stanley, 1986). Obtaining data on the remaining European *Stizostedion* species (*marinum*) is important for resolving this issue, as this taxon is considered to be morphologically intermediate between the European and North American taxa (Svetovidov and Dorofeeva, 1963; Billington *et al.*, 1991).

In addition to the Pliocene and Miocene Epochs, the fossil record indicates that several opportunities existed for fishes to disperse between the Eurasian and North American continents. For example, the Cyprinidae (minnows), the Cottidiae (sculpins), and the Salmonidae (salmon) may have dispersed by either an Atlantic or Pacific route during the Oligocene Epoch, 25 to 38 mybp (Cavender, 1986). Genetic distances from DNA sequences suggested that the whitefish genus *Coregonus* dispersed via Beringia during the Pleistocene (200,000 years to 1.2 mybp; Bernatchez, 1995).

Our divergence estimate of 2.75 ± 0.40 mybp between the North American walleye and sauger is similar to that calculated from mtDNA RFLPs and allozymes (3.0 ± 1.3 mybp; Billington *et al.*, 1990). Billington *et al.* (1991) estimated divergence between the European zander and Volga zander of 2 to 6 mybp, overlapping the upper range of our estimate of 1.80 ± 0.30 mybp. Further genetic studies of variation across the ranges of *Stizostedion* species in Eurasia and North America are needed to elucidate their dispersal routes and patterns of geographic divergence.

Phylogenetic Utility of Repeated Sequences

Repeated sequences in the mtDNA control region have been reported for several groups of unrelated fishes. Repeated sequences are found at two sites in the vertebrate control region: the R1 (near the stop site for the d-loop, adjacent to the proline tRNA gene) and the R2 (near the start site for the d-loop, near OriH) (Fumagalli *et al.*, 1996). Presence of tandem repeats in the R1 of the sauger was corroborated by Turner (1997) who examined only a portion of the left domain and whose sequences appear incorrect for the repeats. PR that are 10- or 11-bp-long and similar in length and sequence to those in *Stizostedion* are found at the R1 site in other percids, including *Crystallaria*, *Etheostoma*, *Gymnocephalus*, *Perca*, and *Percina* (Faber and Stepien, 1997). The PR in percids are shorter (by 30 to 62 bp) than those found in other fishes. Longer repeats in the R1 site were found in the Atlantic cod *Gadus morhua* (four copies of a 40-bp repeat, Arnason and Rand, 1992) and in species of sturgeon *Acipenser* (one to six copies of a 78- to 82-bp repeat; Brown *et al.*, 1992, 1996; Miracle and Campton, 1995).

Repeats in the mtDNA control region of fishes have typically been identified in conjunction with heteroplasmy—having two or more mtDNA molecules of different sizes within a cell. Heteroplasmy of repeats in the control region is known in the Atlantic cod *Gadus morhua* (Arnason and Rand, 1992), acipenserid sturgeon (Brown *et al.*, 1992, 1996), American shad *Alosa sapidissima* (Bentzen *et al.*, 1988), the European sea bass *Dicentrarchus labrus* (Cesaroni *et al.*, 1997), and the minnow *Cyprinella spiloptera* (Broughton and Dowling, 1997). Our finding of only one PCR product and a single DNA sequence per individual suggests that heteroplasmy may not occur in species of *Stizostedion*. Repeat sequence heteroplasmy was found by cloning the European yellow perch *Perca fluviatilis* (Nesbo *et al.*, 1998). Heteroplasmy of repeats larger in size (greater than 40 bp) than those in our study has been detected using our methods of agarose electrophoresis and ethidium bromide staining of PCR products (Fumagalli *et al.*, 1996). It is possible that heteroplasmy of the small 10- to 11-bp repeat units in *Stizostedion* may not have been visualized due to small molecular weight differences among PCR products. This hypothesis may be tested by amplifying the repeat section, followed by acrylamide electrophoresis and Southern blotting (Southern, 1975) with repeat-specific probes (following Arnason and Rand, 1992). Heteroplasmy has no known effect on the DNA sequences of the repeats and has little bearing on the phylogenies estimated from repeated sequences in this study.

DNA sequences of the PR have limited applicability for resolving phylogenies among percids. Divergences

from the repeated sequences in *Stizostedion* approximate those calculated from the nonrepeated sequences. The repeated sequences of sauger and walleye are very different, indicating that the repeated and nonrepeated sequences do not evolve in tandem. Sequences of the PR are identical in the ruffe and the zander (this study) and in 12 other percids reported by Turner (1997), including *Crystallaria asprella*, five species of *Percina*, and six species of *Etheostoma*. Faber and Stepien (1997) tested for phylogenetic signal in the tandem repeats of walleye, sauger, and zander, concluding that the numbers of copies of repeats are uninformative.

Polymorphisms in the PR sequences are useful for resolving relationships at the intraspecific level. A single transitional polymorphism separates walleye from the Great Lakes and the Ohio River (Table 2), and Stepien *et al.* (1998) found geographic patterning in a polymorphism of the PR in Eurasian ruffe. Sequences of the PR have discerned patterns of divergence across large geographic areas in the minnow (*Cyprinella spiloptera;* Broughton and Dowling, 1997), shrews (*Sorex;* Stewart and Baker, 1994), and crickets (Rand, 1992). Recurrent substitutions appear to result in homoplasy and/or convergence among repeated sequences, obscuring relationships among all but the most recently derived taxa. This type of homoplasy may occur more often in organisms with short repeats, such as percids, in which case mutations at a few nucleotide positions may obscure the phylogenetic signal. Sequencing more individuals of each species of *Stizostedion* may reveal additional useful intraspecific characters in the PR.

The sequences of IR also have phylogenetic utility (Stewart and Baker, 1994). Identical IR-1 in the Eurasian species of *Stizostedion* and corresponding similarity between the North American species supports phylogenetic division between the continents. Occurrence of an IR-2 in walleye appears unique among vertebrate taxa. The IR-2 differs in sequence from the PR and IR-1 of walleye and sauger, suggesting that this polymorphism evolved after their evolutionary split. High levels of intraspecific polymorphisms are found in the IR-2 of walleye and appear to be useful population markers (Stepien and Faber, 1998). The American plaice (*Hippoglossoides platessoides*) possesses a somewhat similar IR-2 structure at the R2 near the CSB-2 (Lee *et al.*, 1995). Unlike walleye, the plaice has no other repeated sequences in the control region and its IR-2 sequence is not polymorphic.

Frequencies of genotypes of varying lengths did not reveal patterns of population divergences in the cod *Gadus morhua* (Arnason and Rand, 1992) or walleye *Stizostedion vitreum* (Stepien and Faber, 1998). In contrast, length frequencies suggested broad scale patterns of biogeographic divergence among populations of the sea bass *Dicentrarchus labrax* (Cesaroni *et al.*, 1997). A finite number of alleles (number of repeats) and a high mutation rate (addition or loss of repeats) may act to homogenize length frequencies among populations of some species (Arnason and Rand, 1992). In other words, rapid evolutionary change and turnover of the number of repeats may obscure divergence among populations. The numbers of repeats (PR and IR-1) are normally distributed in the walleye, Atlantic cod (Arnason and Rand, 1992), some species of sturgeon (Brown *et al.*, 1996), and European sea bass (Cesaroni *et al.*, 1997), suggesting that a common process may stabilize their frequencies. Stabilizing selection and/or random deletions and additions of repeats may occur, and the contributions of each remain to be evaluated (Arnason and Rand, 1992).

Formation of Repeats

Point mutations are shared among all copies of PR in species of *Stizostedion*, suggesting a process of concerted evolution. Several mechanisms may mediate concerted evolution, including recombination and transposition (Hasson *et al.*, 1984; Rand and Harrison, 1989), unequal crossing over or gene conversion (Slightom *et al.*, 1980; Hoelzel *et al.*, 1993), and strand slippage (Levinson and Gutman, 1987; Buroker *et al.*, 1990; Wilkinson and Chapman, 1991). Because vertebrate mtDNA does not appear to recombine (Hayashi *et al.*, 1985; Moritz *et al.*, 1987; Birky, 1991), strand slippage is the more probable mechanism. Slipped-strand mispairing (SSM) involves the denaturation of a section of DNA duplex, followed by mispairing of strands and their subsequent repair or replication (Fresco and Alberts, 1960). Depending on how the mispaired structure (i.e., the tandem repeat) is resolved—either by inserting sequence to complement the mispairing or by deletion to eliminate the mispairing—SSM can result in gains or losses of repeats (Fresco and Alberts, 1960). This process appears to be a common cause of repeated sequences and concerted evolution in both the nuclear and mitochondrial genomes (Levinson and Gutman, 1987; Broughton and Dowling, 1997).

Other strand slippage models involve the SSM mechanism and are specific for the formation of repeats in the mtDNA control region. Repeats at either the R1 or the R2 positions near the beginning and end of the d-loop strand occur in diverse vertebrate taxa, suggesting that competition among the d-loop, heavy, and light chain strands for sequence pairing may cause misalignment and replication slippage (Fumagalli *et al.*, 1996). The illegitimate elongation (IE) model of Buroker *et al.* (1990) suggests that mispairing and slippage near the 5′ end of the d-loop promotes the formation of repeats at the R1 site. If repeats contain a TAS, then d-loop strands of different lengths may extend beyond the R1 section, allowing competitive displacement between the d-loop strands and the heavy strand for pairing with the light strand. Displacement may facilitate secondary folding of the repeated sequences, promoting mispairing and losses/gains of these arrays in either the heavy or light

strands. R1 sequences of sturgeon (Brown *et al.*, 1996) meet the predictions of the Buroker *et al.* (1990) model, including the concerted evolution of repeats (repeat sequences are conserved within individuals and divergent among species), a minimum of three repeats (required for secondary structures), and variations of sequences of the repeated copies at the 5′ and 3′ ends. A variation of this model proposed by Wilkinson and Chapman (1991) for R1 repeats in the evening bat also invokes SSM mediated by secondary folding of repeats, except that deletions or additions occur only on the heavy strand near the 5′ end of R1 sequences. Unidirectional replication of the heavy strand DNA produces perfect repeats at the 5′ end of the R1 and allows sequence divergence in the 3′ repeats. This pattern of concerted evolution also is observed in shrews (Fumagalli *et al.*, 1996).

Repeats in *Stizostedion* are most consistent with the predictions of the Wilkinson and Chapman (1991) model. Repeats appear to have evolved in a concerted manner, with identical repeated sequence motifs occurring within species and divergence among species. IR are found only at the 3′ ends of R1 sections adjacent to the PR, suggesting that replication and deletion of repeats, and their homogenization, occurs only at the 5′ end of the R1. Energetically stable folding structures in the repeats of *Stizostedion* may cause misalignment of the d-loop and the heavy strands with the light strand during replication, thereby promoting concerted evolution of the repeats. Finally, normality of the distribution of number of repeated copies suggests that they are randomly added and deleted, as predicted by the SSM.

Some features of the repeated sequences in *Stizostedion* are inconsistent with these models. The models of Buroker *et al.* (1990) and Wilkinson and Chapman (1991) require that repeated sequences contain TAS motifs, thereby terminating d-loop strands in the R1 section. The four TAS arrays in *Stizostedion* and other percids (Faber and Stepien, 1997) are located 17 to 114 bp downstream from the R1 site, suggesting that d-loop strands may not terminate in the R1 section. The relationship between the location of the TAS and the location of d-loop termination is unknown for most nonmammalian vertebrates, and terminations of d-loop strands occur 24 to 63 nucleotides from the TAS in mice and 51 to 53 nucleotides from the TAS in humans (Doda *et al.*, 1981). D-loop strands of *Stizostedion* and other percids may terminate in the R1 section, satisfying the models of Buroker *et al.* (1990) and Wilkinson and Chapman (1991). The repeats of *Stizostedion* are unusual in their apparent lack of TAS sequences, since R1 repeats identified to date in most other groups possess TAS arrays, except for the cod *Gadus morhua* (Arnason and Rand, 1992). The occurrence of the IR-2 in walleye is another unique feature that is unknown from the R1 section of other vertebrates. The IR-2 may be a section that is no longer homogenized by the process of concerted evolution. This hypothesis would explain its apparent sequence similarity to the repeats, the divergences among its repeated motifs, and the high level of polymorphism. The PR and IR-1 are more palindromic than the IR-2 and form more energetically stable secondary structures (Fig. 4), which may promote illegitimate elongation and concerted evolution.

CONCLUSIONS

Patterns of variation of mtDNA control region sequences support the hypothesis of colonization of North America by *Stizostedion* over the Beringial Isthmus during the early Pliocene Epoch. The North American and Eurasian species are genetically divergent, and nonrepeated sequences have utility for discerning intra- and interspecific phylogenies. Repeated sequences in the R1 of *Stizostedion* apparently evolve via the process of illegitimate elongation, and are probably mediated by the mechanism of slipped-strand mispairing. Rapid evolutionary change in the nucleotide sequences of tandem repeats obscure relationships among distantly related genera and are most useful for identifying biogeographically isolated population groups. Frequencies of copies of tandem repeats evolve rapidly, obscuring evolutionary relationships at even the population level.

ACKNOWLEDGMENTS

This research was supported by grants to C. Stepien from the N.O.A.A. Sea Grant Program in Ohio (Project R/LR-3-PD, Grant NA90AA-D-SG496; 1994–5) and the Lake Erie Protection Fund (LEPF-07-94; 1995–7). Sequencing of the ruffe *Gymnocephalus cernuus* was supported by a grant to C. Stepien from the N.O.A.A. Sea Grant Program for Nonindigenous Species (Project R/NIS-1, Grant NA46RG0482; 1995–8). J. Faber was supported by a graduate research assistantship from the Department of Biology, Case Western Reserve University. We thank the following individuals for providing samples: N. Billington (University of Southern Illinois), R. Mercker and R. Woodruff (Bowling Green State University), M. White (Ohio University), T. Bader, C. Baker, G. Isbell, C. Knight, R. Knight, K. Paxton, and M. Turner (Ohio Division of Wildlife), D. Einhouse (New York Department of Environmental Conservation), R. Haas and M. Thomas (Michigan Division of Wildlife), S. Nepszy and L. Halyk (Ontario Ministry of Natural Resources), D. Busch (U.S. Fish and Wildlife Service), M. Burnham-Curtis and J. Selgeby (Biological Resources Division of the U.S. Geological Survey), J. Gunderson and D. Jensen (Minnesota Sea Grant), and the Dunnville Fish and Hunt Club (Ontario).

REFERENCES

Arnason, E., and Rand, D. M. (1992). Heteroplasmy of short tandem repeats in mitochondrial DNA of Atlantic cod, *Gadus morhua*. *Genetics* **132:** 211–220.

Avise, J. C. (1994). "Molecular Markers, Natural History and Evolution," Chapman & Hall, New York.

Balon, E. K., Momot, W. T., and Regier, H. A. (1977). Reproductive guilds of percids: Results of the paleogeographical history of and ecological successions. *J. Fish. Res. Board Can.* **34:** 1910–1921.

Bentzen, P., Leggett, W. C., and Brown, G. G. (1988). Length and restriction site heteroplasmy in the mitochondrial DNA of American shad (*Alosa sapidissima*). *Genetics* **118:** 509–518.

Bermingham, E., Lamb, T., and Avise, J. C. (1986). Size polymorphism and heteroplasmy in the mitochondrial DNA of lower vertebrates. *J. Hered.* **77:** 249–252.

Bernatchez, L. (1995). A role for molecular systematics in defining evolutionarily significant units in fishes. *In* "Evolution and the Aquatic Ecosystem: Defining Unique Units in Population Conservation" (J. L. Nielsen, Ed.), pp. 114–132, Am. Fish. Soc. Symp. 17. American Fisheries Society, Bethesda, MD.

Billington, N., Hebert, P. D. N., and Ward, R. D. (1990). Allozyme and mitochondrial DNA variation among three species of *Stizostedion* (Percidae): Phylogenetic and zoogeographical implications. *Can. J. Fish. Aquat. Sci.* **47:** 1093–1102.

Billington, N., Danzmann, R. G., Hebert, P. D. N., and Ward, R. D. (1991). Phylogenetic relationships among four members of *Stizostedion* (Percidae) determined by mitochondrial DNA and allozyme analyses. *J. Fish Biol.* **39(suppl. A):** 251–258.

Birky, C. W. (1991). Evolution and population genetics of organelle genes: Mechanisms and models. *In* "Evolution at the Molecular Level" (R. K. Selander, A. G. Clark, and T. S. Whittam, Eds.), pp. 112–134, Sinauer, Sunderland, MA.

Broughton, R. E., and Dowling, T. E. (1997). Evolutionary dynamics of tandem repeats in the mitochondrial DNA control region of the minnow *Cyprinella spiloptera*. *Mol. Biol. Evol.* **14:** 1187–1196.

Brown, G. G., Gadaleta, G., Pepe, G., Saccone, C., and Sbisa, E. (1986). Structural conservation and variation in the d-loop-containing region of vertebrate mitochondrial DNA. *J. Mol. Biol.* **192:** 503–511.

Brown, J. R., Beckenbach, A. T., and Smith, M. J. (1992). Mitochondrial DNA length variation and heteroplasmy in populations of white sturgeon (*Acipenser transmontanus*). *Genetics* **132:** 221–228.

Brown, J. R., Beckenbach, K., Beckenbach, A. T., and Smith, M. J. (1996). Length variation, heteroplasmy and sequence divergence in the mitochondrial DNA of four species of sturgeon (*Acipenser*). *Genetics* **142:** 525–535.

Brown, W. M., Prager, E. M., Wang, A., and Wilson, A. C. (1982). Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *J. Mol. Evol.* **18:** 225–239.

Buroker, N. E., Brown, J. R., Gilbert, T. A., O'Hara, P. J., Beckenbach, A. T., Thomas, W. K., and Smith, M. J. (1990). Length heteroplasmy of sturgeon mitochondrial DNA: An illegitimate elongation model. *Genetics* **124:** 157–163.

Cavender, T. M. (1986). Review of the fossil history of North American freshwater fishes. *In* "The Zoogeography of North American Freshwater Fishes" (C. H. Hocutt and E. O. Wiley, Eds.), pp. 699–724, Wiley, New York.

Cesaroni, D., Venazetti, F., Allegrucci, G., and Sbordoni, V. (1997). Mitochondrial DNA length variation and heteroplasmy in natural populations of the European sea bass, *Dicentrarchus labrax*. *Mol. Biol. Evol.* **14:** 560–568.

Cihar, J. (1975). Geographic and ecological variability of perch (*Perca fluviatilis* (Linnaeus)) and history of its distribution from Eurasia to North America. *Acta Musei Nat. Prague* **31B:** 57–89.

Clayton, D. A. (1982). Replication of animal mitochondrial DNA. *Cell* **28:** 693–705.

Clayton, D. A. (1984). Transcription of the mammalian mitochondrial genome. *Annu. Rev. Biochem.* **53:** 573–594.

Collette, B. B., and Banarescu, P. (1977). Systematics and zoogeography of the fishes of the family Percidae. *J. Fish. Res. Board Can.* **34:** 1450–1463.

Craig, J. F. (1987). "The Biology of Perch and Related Fish," Timber Press, Portland.

Doda, J. N., Wright, C. T., and Clayton, D. A. (1981). Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. *Proc. Natl. Acad. Sci. USA* **78:** 6116–6120.

Faber, J. E., and Stepien, C. A. (1997). The utility of mitochondrial DNA control region sequences for analyzing phylogenetic relationships among populations, species, and genera of the Percidae. *In* "Molecular Systematics of Fishes" (T. D. Kocher and C. A. Stepien, Eds.), pp. 125–139, Academic Press, San Diego.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39:** 783–791.

Fresco, J. R., and Alberts, B. M. (1960). The accommodation of noncomplementary bases in helical polyribonucleotides and deoxyribonucleic acids. *Proc. Natl. Acad. Sci. USA* **46:** 311–321.

Fumagalli, L., Taberlet, P., Favre, L., and Hausser, J. (1996). Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews. *Mol. Biol. Evol.* **13:** 31–46.

Hasson, J. F., Mougneau, E., Cuzin, F., and Yaniv, M. (1984). Simian virus 40 illegitimate recombination occurs near short direct repeats. *J. Mol. Biol.* **177:** 53–68.

Hayashi, J. I., Tagashira, I., and Yoshida, M. C. (1985). Absence of extensive recombination between inter- and intra-species mitochondrial DNA in mammalian cells. *Exp. Cell. Res.* **160:** 387–395.

Hendy, M. D., and Penny, D. (1982). Branch and bound algorithms to determine minimal evolutionary trees. *Math. Biosci.* **59:** 277–290.

Hillis, D. M., Mable, B. K., and Moritz, C. (1996). Applications of molecular systematics. *In* "Molecular Systematics" (D. M. Hillis, C. Moritz, and B. K. Mable, Eds.), pp. 515–544, Sinauer, Sunderland, MA.

Hoelzel, A. R., Hancock, J. M., and Dover, G. A. (1993). Generation of VNTRs and heteroplasmy by sequence turnover in the mitochondrial control region of two elephant seal species. *J. Mol. Evol.* **37:** 190–197.

Hoelzel, A. R., Lopez, J. V., Dover, G. A., and O'Brien, S. J. (1994). Rapid evolution of a heteroplasmic repetitive sequence in the mitochondrial DNA control region of carnivores. *J. Mol. Evol.* **39:** 191–199.

Hultman, T., Stahl, S., Hornes, E., and Uhlen, M. (1989). Direct solid phase sequencing of genomic and plasmid DNA using magnetic beads as solid support. *Nucleic Acids Res.* **17:** 4937–4946.

IBI (International Biotechnologies, Inc.). (1992). Assembly LIGN Sequence Assembly Software, Kodak, Inc., Rochester, NY.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16:** 111–120.

Kocher, T. D., and Carleton, K. L. (1997). Base substitution in fish mitochondrial DNA: Patterns and rates. *In* "Molecular Systematics of Fishes" (T. D. Kocher and C. A. Stepien, Eds.), pp. 13–24, Academic Press, San Diego.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Paabo, S., Villablanca, F. X., and Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* **86:** 6196–6200.

Kumar, S., Tamura, K., and Nei, M. (1993). "MEGA, Molecular Evolutionary Genetics Analysis" Vers. 1.01. Institute of Molecular Evolutionary Genetics, The Pennsylvania State University, University Park.

Lee, W. J., Conroy, J., Howell, W. H., and Kocher, T. D. (1995). Structure and evolution of teleost mitochondrial control regions. *J. Mol. Evol.* **41:** 54–66.

Levinson, G., and Gutman, G. A. (1987). Slipped-strand mispairing: A major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4:** 203–221.

Martin, A. P., and Palumbi, S. R. (1993). Body size, metabolic rate, generation time and the molecular clock. *Proc. Natl. Acad. Sci. USA* **90:** 4087–4091.

Meyer, A., Kocher, T. D., Basasibwaki, P., and Wilson, A. C. (1990). Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature* **347:** 550–553.

Miracle, A. L., and Campton, D. E. (1995). Tandem repeat sequence variation and length heteroplasmy in the mitochondrial DNA d-loop of the threatened Gulf of Mexico sturgeon, *Acipenser oxyrhynchus desotoi. J. Hered.* **86:** 22–27.

Moritz, C., Dowling, T. E., and Brown, W. M. (1987). Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. *Annu. Rev. Ecol. Syst.* **18:** 269–292.

Mullis, K. B., and Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaction. *Methods Enzymol.* **155:** 335–350.

Nei, M. (1987). "Molecular Evolutionary Genetics," Columbia Univ. Press, New York.

Nesbo, C. L., Arab, M. O., and Jakobsen, K. S. (1998). Heteroplasmy, length and sequence variation in the mtDNA control regions of three percid fish species (*Perca fluviatilis, Acerina cernua, Stizostedion lucioperca*). *Genetics* **148:** 1907–1919.

Nielsen, J. L., Fountain, M. C., and Wright, J. M. (1997). Biogeographic analysis of Pacific trout (*Oncorhynchus mykiss*) in California and Mexico based on mitochondrial DNA and nuclear microsatellites. *In* "Molecular Systematics of Fishes" (T. D. Kocher and C. A. Stepien, Eds.), pp. 53–74, Academic Press, San Diego.

Palumbi, S. R. (1996). Nucleic acids II: The polymerase chain reaction. *In* "Molecular Systematics" (D. M. Hillis, C. Moritz, and B. K. Mable, Eds.), 2nd ed., pp. 205–221, Sinauer, Sunderland, MA.

Phillips, R. B., and Oakley, T. H. (1997). Phylogenetic relationships among the Salmoninae based on nuclear and mitochondrial DNA sequences. *In* "Molecular Systematics of Fishes" (T. D. Kocher and C. A. Stepien, Eds.), pp. 145–162, Academic Press, San Diego.

Rand, D. M. (1992). RIPping and RAPping at Berkeley. *Genetics* **132:** 1223–1224.

Rand, D. M., and Harrison, R. G. (1989). Molecular population genetics of mtDNA size variation in crickets. *Genetics* **121:** 551–569.

Rzhetsky, A., and Nei, M. (1992). A simple method for estimating and testing minimum-evolution trees. *Mol. Biol. Evol.* **9:** 945–967.

Saccone, C., Pesole, G., and Sbisa, E. (1991). The main regulatory region of mammalian mitochondrial DNA: Structure–function model and evolutionary pattern. *J. Mol. Evol.* **33:** 83–91.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4:** 406–425.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74:** 5463–5467.

SantaLucia, J., Jr., Allawi, H. T., and Seneviratne, P. A. (1996). Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry* **35:** 3555–3562.

Shedlock, A. M., Parker, J. D., Crispin, D. A., Pietsch, T. W., and Burmer, G. C. (1992). Evolution of the salmonid mitochondrial control region. *Mol. Phylogenet. Evol.* **1:** 179–192.

Slightom, J. L., Blechl, A. E., and Smithies, O. (1980). Human fetal G and A-globin genes: Complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. *Cell* **21:** 627–638.

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

Southern, S. O., Southern, P. J., and Dizon, A. E. (1988). Molecular and phylogenetic studies with a cloned dolphin mitochondrial genome. *J. Mol. Evol.* **28:** 32–42.

Stanley, S. M. (1986). "Earth and Life Through Time," Freeman, New York.

Stepien, C. A. (1995). Population genetic divergence and geographic patterns from DNA sequences: Examples from marine and freshwater fishes. *In* "Evolution and the Aquatic Ecosystem: Defining Unique Units in Population Conservation" (J. L. Nielsen, Ed.), pp. 263–287, Am. Fish. Soc. Symp. 17, American Fisheries Society, Bethesda, MD.

Stepien, C. A., Dillon, A. K., and Chandler, M. D. (1998). Evolutionary relationships, phylogeography, and genetic identity of the ruffe *Gymnocephalus* in the North American Great Lakes and Eurasia from mtDNA control region sequences. *J. Great Lakes Res.* **24:** 361–378.

Stepien, C. A., Dixon, M. T., and Hillis, D. M. (1993). Evolutionary relationships of the fish families Clinidae, Labrisomidae, and Chaenopsidae: Congruence between DNA sequence and allozyme data. *Bull. Mar. Sci.* **52:** 873–896.

Stepien, C. A., and Faber, J. E. (1998). Population genetic structure, phylogeography, and spawning philopatry in walleye (*Stizostedion vitreum*) from mtDNA control region sequences. *Mol. Ecol.*, in press.

Stepien, C. A., and Kocher, T. D. (1997). Molecules and morphology in studies of fish evolution. *In* "Molecular Systematics of Fishes" (T. D. Kocher and C. A. Stepien, Eds.), pp. 1–12, Academic Press, San Diego.

Stewart, D. T., and Baker, A. J. (1994). Patterns of sequence variation in the mitochondrial d-loop region of shrews. *Mol. Biol. Evol.* **11:** 9–21.

Sturmbauer, C., Verheyen, E., Ruber, L., and Meyer, A. (1997). Phylogeographic patterns in populations of cichlid fishes from rocky habitats in Lake Tanganyika. *In* "Molecular Systematics of Fishes" (T. D. Kocher and C. A. Stepien, Eds.), pp. 97–112, Academic Press, San Diego.

Svetovidov, A. N., and Dorofeeva, E. A. (1963). Systematics, origin, and history of the distribution of the Eurasian and North American perches and pikeperches (genera *Perca, Lucioperca,* and *Stizostedion*). *Voprosy Icktiologii* **3:** 625–651. (Translated from Russian by NMFS Systematics Lab., No. 28).

Swofford, D. L. (1993). "PAUP (Phylogenetic Analysis Using Parsimony) vers. 3.1.1 for MacIntosh Computers," Ill. Nat. Hist. Surv., Champaign, IL.

Titus, T. A., and Larson, A. (1995). A molecular phylogenetic perspective on the evolutionary radiation of the salamander family Salamandridae. *Syst. Biol.* **44:** 125–151.

Turner, T. F. (1997). Mitochondrial DNA control region sequences and phylogenetic systematics of darters (Teleostei: Percidae). *Copeia* **1997:** 319–338.

Uhlen, M. (1989). Magnetic separation of DNA. *Nature* **340:** 733–734.

Walberg, M. W., and Clayton, D. A. (1981). Sequence and properties of the human KB cell and mouse L cell d-loop regions of mitochondrial DNA. *Nucleic Acids Res.* **9:** 5411–5420.

Ward, R. D., Billington, N., and Hebert, P. D. N. (1989). Comparison of allozyme and mitochondrial variation in groups of walleye, *Stizostedion vitreum*. *Can. J. Fish. Aquat. Sci.* **46:** 2074–2084.

Wiley, E. O. (1992). Phylogenetic relationships of the Percidae (Teleostei: Perciformes): A preliminary hypothesis. *In* "Systematics, Historical Ecology, and North American Freshwater Fishes" (R. L. Mayden, Ed.), pp. 247–267, Stanford Univ. Press, Palo Alto, CA.

Wilkinson, G. S., and Chapman, A. M. (1991). Length and sequence variation in evening bat d-loop mtDNA. *Genetics* **128:** 607–617.

Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. *Science* **244:** 48–52.

Zuker, M. (1996). MFOLD vers. 3.2 online. Http://www.ibc.wustl.edu/~zuker/dna.