Intraspecific genetic structure of white sucker (*Catostomus commersoni*) in northeastern North America as revealed by mitochondrial DNA polymorphism

Pascale Lafontaine and Julian J. Dodson

Abstract: Restriction fragment length polymorphisms in mitochondrial DNA (mtDNA) were used to study the influence of Pleistocene glaciations on the intraspecific genetic structure and distribution of the white sucker (*Catostomus commersoni*) in northeastern North America. A total of 312 white sucker from 13 populations, including a population of dwarf ecotypes (*Catostomus commersoni utawana*), were analysed. An average of 93 fragments per individual and 40 haplotypes were generated by nine restriction endonucleases. Four discrete clades were identified but the majority of the genotypes found (70%) were not associated with any of the clades. The phylogenetic continuity and the geographic admixture of some of the few clades identified are associated with the extensive distribution of the species south of the ice sheets during Pleistocene glaciation events. Hierarchical analysis of the variability in mtDNA revealed a significant regional subdivision to the northwest and southeast of the St. Lawrence drainage system and significant structure at the population level. As the dwarf form exhibited a haplotype identical to that found among normal forms, we tentatively conclude sympatric divergence as the most likely origin of the dwarf and normal ecotypes sampled in the Adirondacks.

Résumé : Nous avons utilisé les polymorphismes de longueur de fragment de restriction sur l'ADN mitochondrial (ADNmt) afin d'étudier l'influence des glaciations du Pleistocène sur la structure génétique intraspécifique et sur la distribution du meunier noir (*Catostomus commersoni*) au nord-est de l'Amérique du Nord. Un total de 312 meuniers noirs provenant de 13 populations, dont une comprenant un écotype nain (*Catostomus commersoni utawana*), ont été analysés. Une moyenne de 93 fragments par individu et 40 haplotypes ont été générés par neuf endonucléases de restriction. Quatre petits groupes distincts ont été identifiés mais la majorité des génétypes retrouvés (70%) n'ont été associés à aucun de ces groupes. La continuité phylogénétique et le chevauchement de la distribution de quelques un de ces groupes identifiés sont associés à la large distribution de l'espèce au sud de la calotte de glace lors des événements glaciaires du Pléistocène. L'analyse hiérarchique de la variabilité de l'ADNmt révèle une subdivision régionale significative entre le nord-ouest et le sud-est du bassin de drainage du fleuve St-Laurent ainsi qu'une structure significative au niveau de la population. Comme la forme naine présentait un haplotype identique à celui retrouvé parmi la forme normale, nous suggérons une divergence sympatrique comme explication la plus probable de l'origine de l'écotype nain et normal échantilloné dans les Adirondacks.

Introduction

Historical biogeographical events have played an important role in shaping contemporary intraspecific genetic structure (Bermingham and Avise 1986). Geological and other historical events have changed the distributional areas of organisms, built up or suppressed geographic barriers, and gathered or isolated genetic populations from the rest of the species. The Pleistocene glaciations caused important changes in the distribution and genetic structure of North American fishes. Glaciation is believed to have decreased genetic diversity through population-bottleneck effects and to have confined populations of northern temperate fishes in separate glacial refugia where they evolved independently for thousands of years, resulting in the development of genetically distinct intraspecific groups (Bernatchez 1995; McAllister et al. 1986). When retreating, the ice cover left huge postglacial lake and river connections that provided formidable opportunities for the dispersal and admixture of these previously isolated groups (Bernatchez and Dodson 1991).

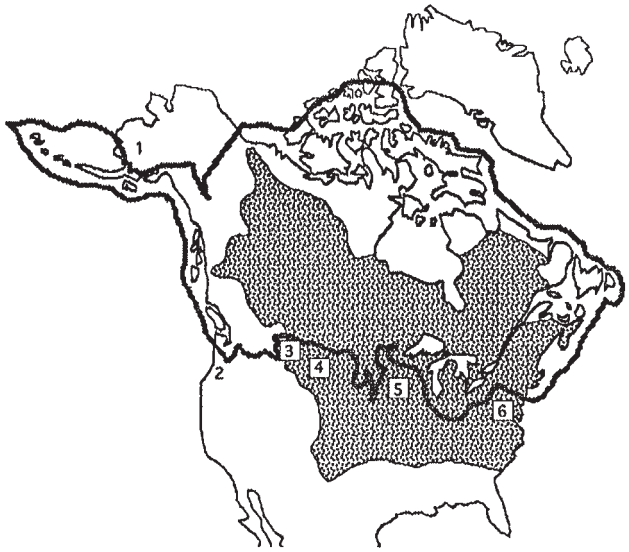
Studies of restriction fragment length polymorphisms (RFLPs) in the mitochondrial DNA (mtDNA) of northern fishes have clearly illustrated the impact of glaciation on intraspecific genetic structure. Haploid, maternally inherited, and devoid of crossing over, mtDNA is a powerful instrument to detect genetic variation within and between fish populations. The ability of mtDNA to retain a history of past isolation, in spite of contemporary admixture, is well demonstrated in a variety of species (Bermingham and Avise 1986; Avise et al. 1987; Billington and Hebert 1991; Avise 1992). Among northern fishes, mtDNA analyses of walleye (*Stizostedion vitreum*) (Billington et al. 1992), lake trout (*Salvelinus namaycush*) (Grewe and Hebert 1988), whitefish (*Coregonus clupeaformis*) (Bernatchez and Dodson 1991), cisco (*Coregonus artedii)* (Bernatchez and Dodson 1990*a*), and rainbow smelt (*Osmerus mordax*) (Bernatchez 1995) have all revealed the presence of major, intraspecific genetic discontinuities related to the

Received April 26, 1996. Accepted August 28, 1996.
J13435

P. Lafontaine and J.J. Dodson.[1] Département de biologie, Université Laval, Ste-Foy, QC G1K 7P4, Canada.

[1] Author to whom all correspondence should be addressed. e-mail: julian.dodson@bio.ulaval.ca

Fig. 1. Maximum extent of the Wisconsin glaciation (solid line), distribution of white sucker (stippled area), and location of the principal glacial refugia: 1, Beringian; 2, Columbian; 3, Banff–Jasper; 4, Missourian; 5, Mississipian; 6, Atlantic.



genetic divergence of intraspecific groups of fish occupying different glacial refugia at various times during the Pleistocene.

White sucker, *Catostomus commersoni*, is one of the most abundant and widely distributed fish species in North America (Scott and Crossman 1974). Much of the present distributional area of white sucker was covered by the ice cap during the Pleistocene, except for the southern part, which remained ice free (Fig. 1). During the last glaciation, white sucker probably survived in refugia located south of the ice sheets that may have favoured the evolution of intraspecific genetic groups in allopatry. Thus, recolonization of the newly deglaciated areas of North America may have involved several distinct intraspecific genetic groups. White sucker found in the northeastern part of North America may have originated from one or all of three glacial refugia: the Mississippi, Missouri, and Atlantic (Crossman and McAllister 1986). Hudson Bay drainages may have been recolonized from the Mississippi and (or) the Atlantic glacial refugium. White sucker from the Great Lakes may have come from the Mississippi and (or) from the Missouri glacial refugium. The St. Lawrence valley may have been invaded by fishes coming from the Mississippi and the Atlantic glacial refugia (Underhill 1986). Although only one genetic group of white sucker would be expected along the east coast of North America, Schmidt (1986) proposed that, on the basis of the wide distribution of the species, it may be possible to find more than one genetic group in the Atlantic region. All of these scenarios remain speculative. Therefore, the major objective of this study was to document the influence of glaciation on the intraspecific genetic structure of white sucker by relating intraspecific genetic divergence among contemporary populations to their distribution in northeastern North America.

We also investigated a special ecotype of the white sucker: a dwarf form (*Catostomus commersoni utawana*) lives in sympatry with the normal ecotype in the Adirondack Mountains (New York state, U.S.A.). Dwarf white sucker apparently live deep in lakes; the spawning period is a few weeks later than the spawning period of normal white sucker and the length at maturity ranges between 127 and 203 mm, compared with 310 and 460 for the normal ecotype. Dwarf white sucker grow slowly but mature 1 year before the normal ecotype, which matures at 4 years of age in this area (Dence 1948; Beamish 1970).

Surprisingly widespread, dwarfism is found among many fish species, especially salmonids. Studies have shown that two mtDNA scenarios are associated with this phenomenon. In the case of whitefish, Bernatchez and Dodson (1990*b*) demonstrated that the sympatric occurrence of the dwarf and normal ecotypes in the Allagash basin, Maine, U.S.A, resulted from a secondary contact between two phylogenetic groups of allopatric origin. The same scenario was presented to explain the origin of different ecotypes of whitefish in the Squanga Creek drainage, Yukon (Bernatchez et al. 1996).

Dwarfism is also present among limnetic smelt, *Osmerus mordax*, but Taylor and Bentzen (1993) presented evidence that sympatric divergence was responsible for the presence of the two ecotypes. In fact, dwarf and normal smelt belong to the same mtDNA genetic clade and must have diverged from a common ancestor. Their morphological differences are characters important in foraging: dwarf smelt have more gill rakers and larger eyes, but shorter upper jaws and shallower heads than normal and anadromous smelt. According to these morphological traits, dwarf smelt are planktivorous whereas normal and anadromous smelt are macrophagous and commonly piscivorous (Taylor and Bentzen 1993). Therefore, the second objective of this study was to determine the phylogenetic relationship of dwarf and normal white sucker. The occurrence of phylogenetically distinct mtDNA haplotypes discriminating dwarf white suckers from normal ecotypes would be consistent with a scenario of allopatric divergence.

Materials and methods

Sampling

Between 2 and 68 fish were sampled from each of 13 sites situated in the northeastern part of the white sucker's distribution (Table 1 and Fig. 2). Dwarf white sucker were sampled in Wolf Lake, New York, where normal sucker are usually found in sympatry. Biologists from the Adirondack Ecological Center also sampled normal white sucker in Catlin Lake, adjacent to Wolf Lake, where dwarf sucker are also present. Other sites were sampled either by the authors or by collaborators. Fish were immediately dissected to remove liver and eggs, which were subsequently frozen or kept on wet ice prior to mtDNA extraction.

Mitochondrial DNA extraction and analysis

Mitochondrial DNA was successfully extracted and analysed for 312 fish. Tissues were extracted using the method of Bernatchez et al. (1988). Aliquots of mtDNA were digested separately with two hexameric (*Dra*I, *Xba*I), four multihexameric (*Acc*I, *Ava*I, *Ban*II, *Hin*cII), and three multipentameric (*Ava*II, *Ban*I, *Nci*I) restriction endonucleases. Mitochondrial DNA fragments were electrophoretically separated on 0.8, 1.0, or 1.2% agarose gels for 4 h at 75 V.

Ethidium bromide staining revealed restriction fragments when mtDNA was obtained in sufficient quality and quantity. In cases of low or poor-quality yield, mtDNA fragments were transferred to nylon membranes under alkaline conditions and hybridized with a pure radiolabelled *Catostomus commersoni* mtDNA pool probe prepared

Table 1. Sampling locations, sample sizes, and number of mtDNA haplotypes for populations of *Catostomus commersoni*.

	Sampling site	Sample size	Number of haplotypes
a	Lake Manitoba (Manitoba)	26	8
b	Jack Fish Bay (Lake Superior, Ont.)	24	8
c	Gatineau River (Quebec)	73	12
d	St-Maurice River (Quebec)	68	12
e	Sans nom River (Mastigouche, Que.)	17	1
f	Great Whale River (North Quebec)	2	2
g	St. Lawrence River (Quebec City)	18	8
h	Yoyo Lake (Forestville, Que.)	22	1
i	Amédée River (Baie-Comeau, Que.)	3	3
j	Cliff Lake (Maine)	13	2
k	Catlin Lake (New York state)	15	4
l	Wolf Lake (New York state)	12	3
m	St. Andrews (New Brunswick)	19	1

Note: Letters (a–m) correspond to sampling sites illustrated in Fig. 2.

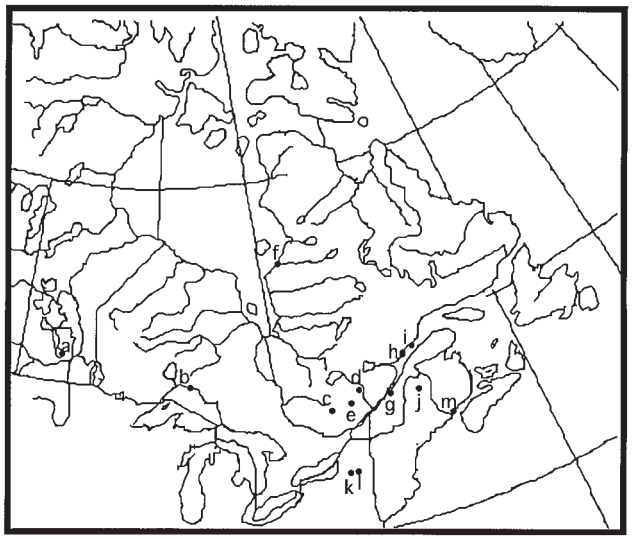
with the multiprimed (Amersham) DNA labelling reaction. Membranes were submitted to stringent washings at 68°C and were then autoradiographed using an intensifying screen for 5–36 h at –70°C.

Sizes of fragments were estimated by running simultaneously into the agarose gel digests of phage lambda DNA with *Hin*dIII and *Eco*RI–*Hin*dIII double digest. Some fragments less than 300 base pairs (bp) in length could not be scored. Distinct single endonuclease patterns were identified by a specific letter. Each fish was assigned a multiletter code that described its composite mtDNA genotype.

Restriction-site differences between genotypes were inferred from changes in fragment patterns because such changes could be readily explained by specific site gains or losses. Estimates of nucleotide sequence divergence (p) among mtDNA haplotypes calculated by the site approach of Nei and Li (1979) were used to estimate genetic distance among haplotypes. The resulting distance matrix served to construct a phenogram using the unconstrained branch-length clustering method of Fitch and Margoliash (1967) (program FITCH in the PHYLIP 3.4 computer package, Felsenstein 1992). In addition, a site presence–absence matrix was used to generate phylogenetic trees according to Wagner parsimony criteria using the MIX program from PHYLIP. Symplesiomorphic and autapomorphic characters were omitted from the site matrix. Confidence in parsimony trees was assessed by bootstrap resampling of the restriction site presence–absence matrix (N = 10 000) using SEQBOOT of PHYLIP. A majority-rule consensus tree was built using CONSENSE of PHYLIP. Phylogenetic trees were rooted using a haplotype of longnose sucker (*Catostomus catostomus*), sampled in the St. Lawrence River (Quebec City). As a measure of inter- and intra-population genetic diversity, we computed the maximum likelihood estimation of the average number of nucleotide substitutions per site within populations (nucleotide diversity) and between populations (nucleotide divergence) (Nei and Tajima 1981, 1983; Nei 1987). The neighbour-joining program of PHYLIP was used to cluster populations according to nucleotide divergence.

We estimated the likelihood of detecting mitochondrial DNA diversity at the population level using the combinatorial approach of Hebert et al. (1988) as applied by Bernatchez et al. (1989) to mtDNA data. The relationship between the number of fish sampled and the number of mtDNA genotypes detected was estimated for each population by an incremental random choice of individuals (increments of 4 fish to a maximum of 312 or 204 fish). The procedure was repeated 10 times for each sampling intensity to generate an estimate of random sampling variance. This procedure was done on all of the 312 fish of the sampled area and also on the 204 fish from the St. Lawrence drainage system. The relationship between the number of genotypes

Fig. 2. Location map of *Catostomus commersoni* sampling sites named in Table 1.



detected as a function of genome sampling was estimated for each population by incremental random choice of restriction enzymes. Results were obtained by an incremental choice of one enzyme repeated 10 times for each incremental step. As different restriction enzymes produced different numbers of fragments, we attributed the fragment numbers to their respective enzymes and summed the fragments for any enzyme combination.

Geographical heterogeneity in the frequency of mtDNA RFLP haplotypes among sampling sites in northeastern north America was analyzed using analysis of molecular variance (AMOVA) (Excoffier et al. 1992). We set interhaplotypic distances as equal for the analysis of molecular variance. Two geographical groups were formed with the sampled populations: northwest of the St. Lawrence River (Lake Manitoba, Jack Fish Bay, Gatineau River, St. Maurice River, Sans nom River, Great Whale River, Yoyo Lake, and Amedée River) and southeast of the St. Lawrence River (Cliff Lake, Catlin Lake, Wolf Lake, and St. Andrews). The St. Lawrence population was eliminated because of its central position in this distribution. F_{st} and associated p values ($p < 0.05$) were calculated within and between the two geographical groups.

Results

The nine endonucleases generated a total of 191 restriction fragments with a mean of 93 fragments per individual. Seven out of nine enzymes were polymorphic and discriminated a total of 40 mtDNA haplotypes among the 312 fish analysed from 13 sampling sites (Tables 1 and 2; Fig. 2). Fragment patterns generated by each enzyme and restriction-site differences between genotypes inferred from changes in fragment patterns are presented in Figs. 3 and 4, respectively. The 150 restriction sites so identified included 59 synapomorphic sites. The mean size of the mtDNA genome, estimated by averaging the sum of all digestion patterns, was 17 769 bp (SD = 1295), similar to the length estimates obtained for a variety of other fish species (Billington and Hebert 1991). The mean pairwise sequence divergence was 0.70 ± 0.34% and ranged from 0.1 to 2.4% among all white sucker lineages.

An analysis of sequence divergence between haplotypes of white sucker using the site method and clustered with Fitch revealed no major divisions in the tree (Fig. 5). Small clusters

Table 2. Definitions of composite haplotypes, absolute frequency (*n*), and distribution of the mtDNA haplotypes resolved among *Catostomus commersoni*.

Composite haplotypes										*n*	Locality
1	A	A	A	A	A	A	A	A	A	3	c
2	A	A	A	A	A	A	A	A	D	1	i
3	A	A	A	A	A	A	A	B	D	1	g
4	A	A	A	A	A	A	B	A	E	2	d
5	A	A	A	A	A	A	B	A	H	1	k
6	A	A	A	A	A	A	B	B	E	34	g, j, m
7	A	A	A	A	A	A	C	B	E	1	j
8	A	A	A	A	A	B	A	A	A	73	c, d, e
9	A	B	A	A	A	B	A	A	A	1	c
10	A	A	A	A	A	B	A	A	D	1	d
11	A	A	A	A	A	B	B	A	A	1	c
12	A	A	A	A	B	A	A	A	A	1	c
13	A	A	A	A	B	A	A	A	B	13	b, c, d
14	A	A	A	A	B	A	A	A	C	5	c, d, g, i
15	A	A	A	A	B	A	A	A	E	2	a, b
16	A	B	A	A	B	A	A	A	E	9	a
17	A	A	A	A	C	A	A	A	B	2	d, f
18	A	A	A	A	C	A	A	A	C	45	b, c, d, f, g
19	A	A	A	A	C	A	A	A	D	2	d, g
20	A	A	A	A	C	A	A	A	E	2	a
21	A	A	A	A	C	A	A	A	F	2	b, d
22	A	A	A	A	C	A	A	A	G	4	d
23	B	B	A	A	C	A	A	A	E	1	c
24	A	C	A	A	C	A	A	A	E	1	g
25	A	A	A	A	D	A	A	A	A	7	c
26	A	A	A	A	D	A	A	A	B	2	d
27	A	A	A	A	D	A	A	A	D	1	k
28	A	A	A	A	E	A	A	A	E	22	a, b
29	A	A	A	A	E	A	A	A	H	1	a
30	A	A	A	A	E	A	A	A	I	2	b
31	A	A	A	A	F	A	A	A	D	56	c, d, g, h, i, k, l
32	A	A	A	A	F	A	D	A	D	2	l
33	A	A	A	A	G	B	A	A	A	1	c
34	A	A	A	A	H	A	A	A	A	1	g
35	A	B	A	A	H	A	A	A	E	1	a
36	A	B	A	A	I	A	A	A	J	1	a
37	A	B	A	A	J	A	A	A	E	3	a, b
38	A	A	A	A	K	A	A	A	D	1	k
39	A	A	A	A	L	A	A	A	D	1	l
40	A	A	A	A	M	A	A	A	E	2	b

Note: Restriction enzymes are in order *Xba*I, *Dra*I, *Acc*I, *Ava*I, *Ava*II, *Ban*I, *Ban*II, *Hin*cII, *Nci*I. Capital letters identify fragment patterns presented in Fig. 3. Small letters refer to sample locations illustrated in Fig. 2.

of haplotypes were observed throughout the tree, but no discontinuities divided the tree into major subgroups.

The majority-rule consensus tree also failed to reveal significant division of the species into major clades. Rather, four small clades, also visible in the Fitch phenogram, were resolved (Fig. 6). These clades constituted only 30% of the haplotypes. Seventy percent of the genotypes were poorly resolved and could not be associated with a distinct clade. Haplotype 36, from Lake Manitoba, was a highly divergent haplotype distinct from all others.

Parsimony analysis of the *Ava*II and *Nci*I restriction fragment patterns revealed sites that were hypervariable. Such homoplasy could mask some important groupings among genotypes by reducing the resolution of the phylogenetic trees. However, removing these sites from the input matrix and producing a new majority-rule consensus tree did not alter tree topology. Thus, we concluded that the absence of major discontinuities in the tree was not due to homoplasy.

Group A

This clade comprised two haplotypes: No. 13 found in Jack Fish Bay, the Gatineau River, and the St. Maurice River and No. 17 from the St. Maurice River and Great Whale River (Tables 2 and 3). The clade was supported at the 58% level in the majority-rule consensus tree. It was restricted to the west of the St. Lawrence – Great Lakes axis (Fig. 7). The pairwise sequence divergence within this group was 0.2%. Restriction patterns B and C of *Ava*II and B of *Nci*I were diagnostic for this group (Table 2).

Group B

This clade comprised two haplotypes: No. 31 from the Gatineau River, St. Maurice River, St. Lawrence River, Yoyo Lake, Amédée River, Catlin Lake, and Wolf Lake and No. 32 from Wolf Lake (Tables 2 and 3). It was supported at the 82% level in the majority-rule consensus tree. This clade was principally located to the east of the St. Lawrence River (Fig. 7). *Ava*II restriction pattern F was diagnostic for the group (Table 2). The pairwise sequence divergence within this group was 0.2%. The clade comprised the majority (11 out of 12 fish) of the dwarf white suckers (*Catostomus commersoni utawana*) sampled in Wolf Lake (Table 1).

Group C

This clade comprised five haplotypes: No. 4 from the St. Maurice River, No. 5 from Catlin Lake, No. 6 from the St. Lawrence River, Cliff Lake, and St. Andrews, No. 7 from Cliff Lake, and No. 11 from the Gatineau River (Tables 2 and 3). This was supported at the 53% level in the majority-rule consensus tree. Clade C showed the most easterly distribution of the clades resolved in this study (Fig. 7). The mean pairwise sequence divergence within this group was 0.38 ± 0.25%. *Ban*II restriction fragments B and C were diagnostic for the group (Table 2).

Group D

This group comprised three haplotypes: Nos. 28 and 29 from Lake Manitoba and Jack Fish Bay and No. 30 from Jack Fish Bay (Tables 2 and 3). This was supported at the 67% level in the majority-rule consensus tree. This clade presented the most westerly distribution among sampled haplotypes (Fig. 7). The mean pairwise sequence divergence within this group was 0.13 ± 0.06%. *Ava*II fragment pattern E was diagnostic for this group (Table 2).

The maximum likelihood estimation of the average number of nucleotide substitutions per site within populations (nucleotide diversity) and between populations (nucleotide divergence) did not clearly reveal population groupings. The population sampled near Quebec City was separated from all of the others by 0.16% sequence divergence (Fig. 8). This population also presented the highest nucleotide diversity (0.57 ± 0.06%).

Fig. 3. Fragment size estimates (in base pairs) of all fragment patterns observed among white sucker and longnose sucker (Z). Asterisks identify fragments that were not observed but were assumed under the criterion of the minimum mutational steps involved in fragment changes. The presentation of two 1825-bp fragments for *Ban*II haplotype D indicates that two superimposed 1825-bp fragments were observed. The presentation of two or more fragments of the same size for different haplotypes of the same restriction enzyme indicates different origins of similar-sized fragments.

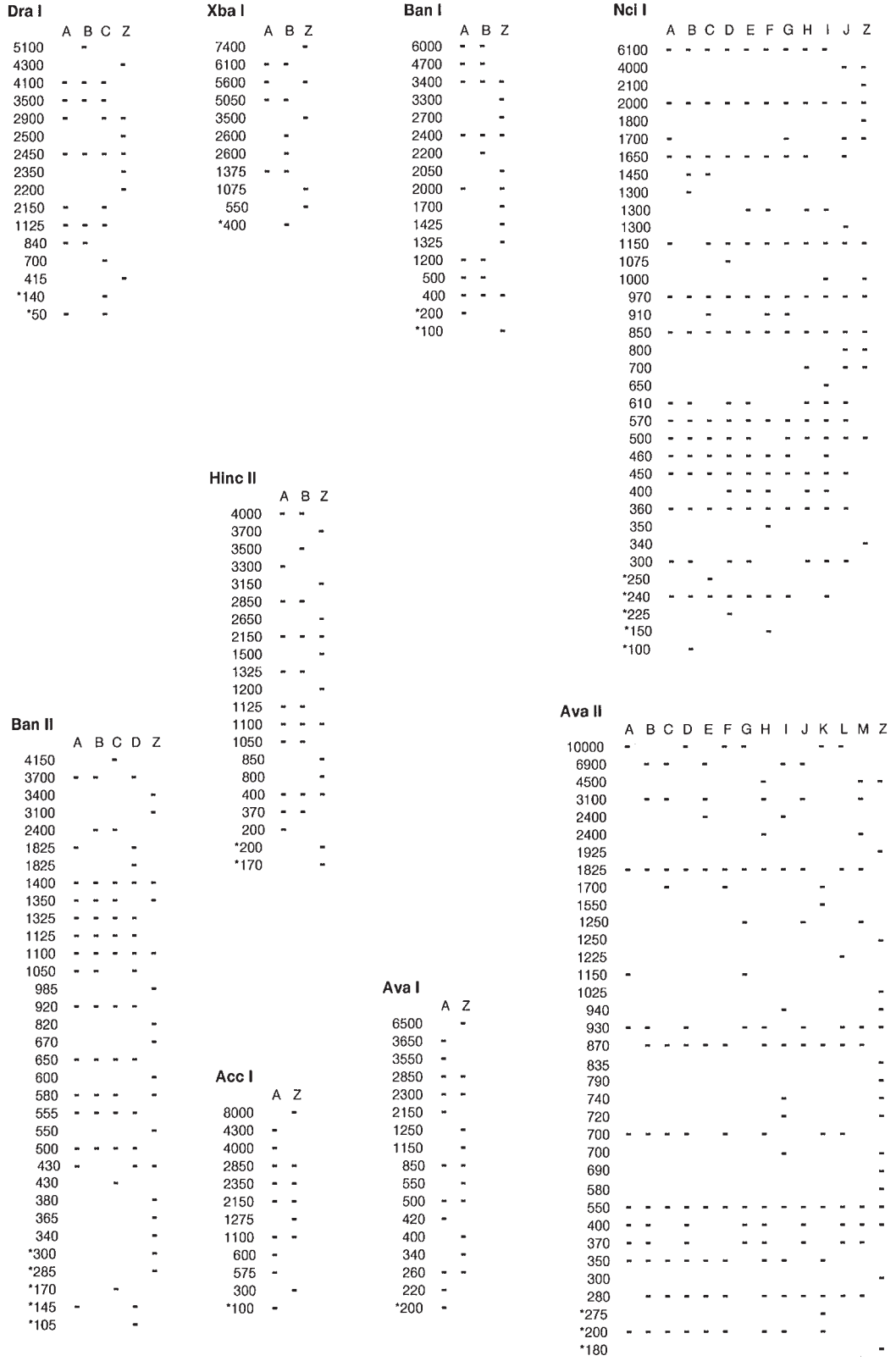
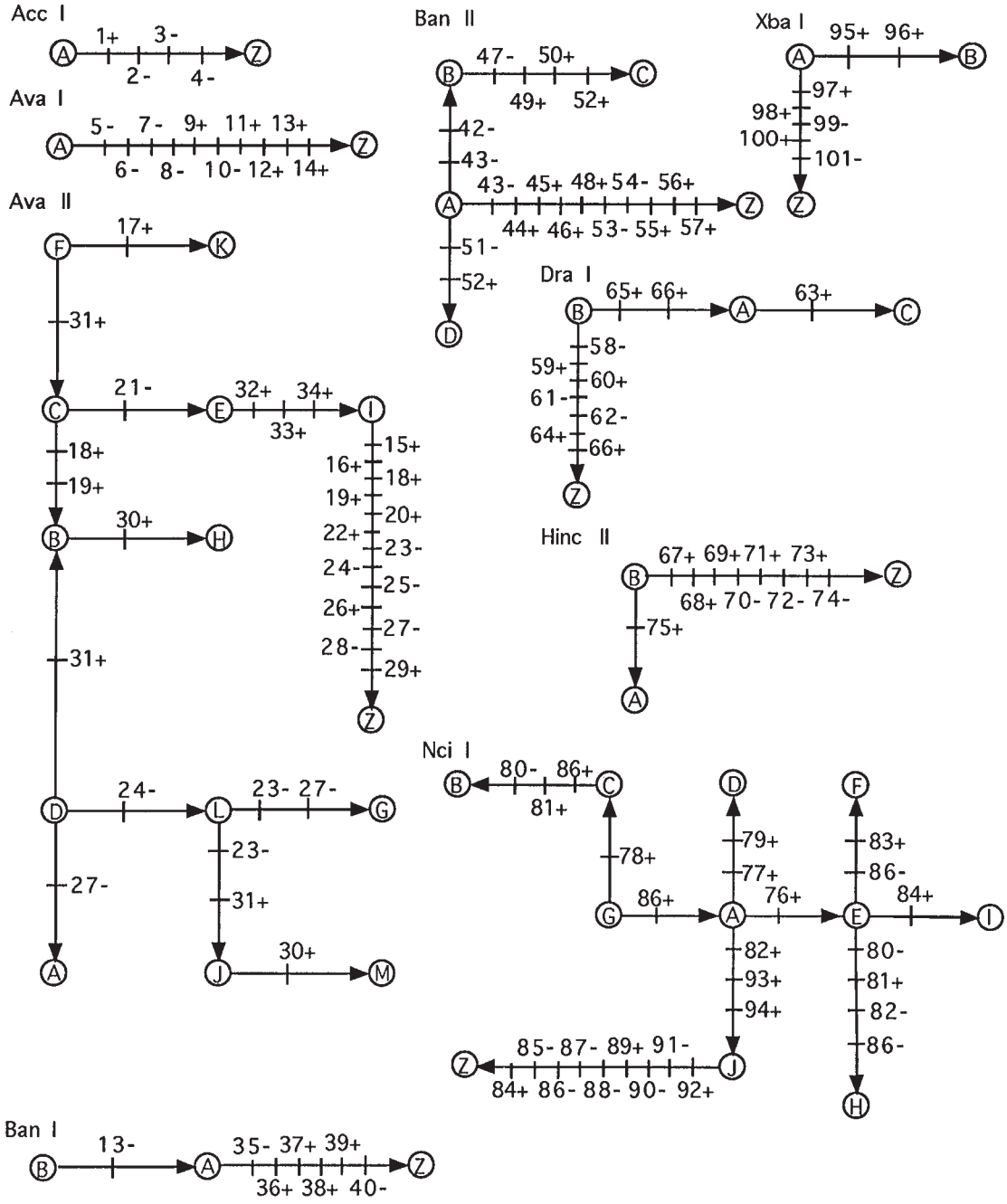


Fig. 4. Parsimonious, unrooted networks illustrating relationships among fragment patterns observed for all restriction enzymes described in Fig. 3. Site changes (losses (–) or additions (+)) involved in moving between fragment patterns (letters) are numbered along branches (arrows indicate direction of changes) and were parsimoniously deduced from observed fragment patterns.



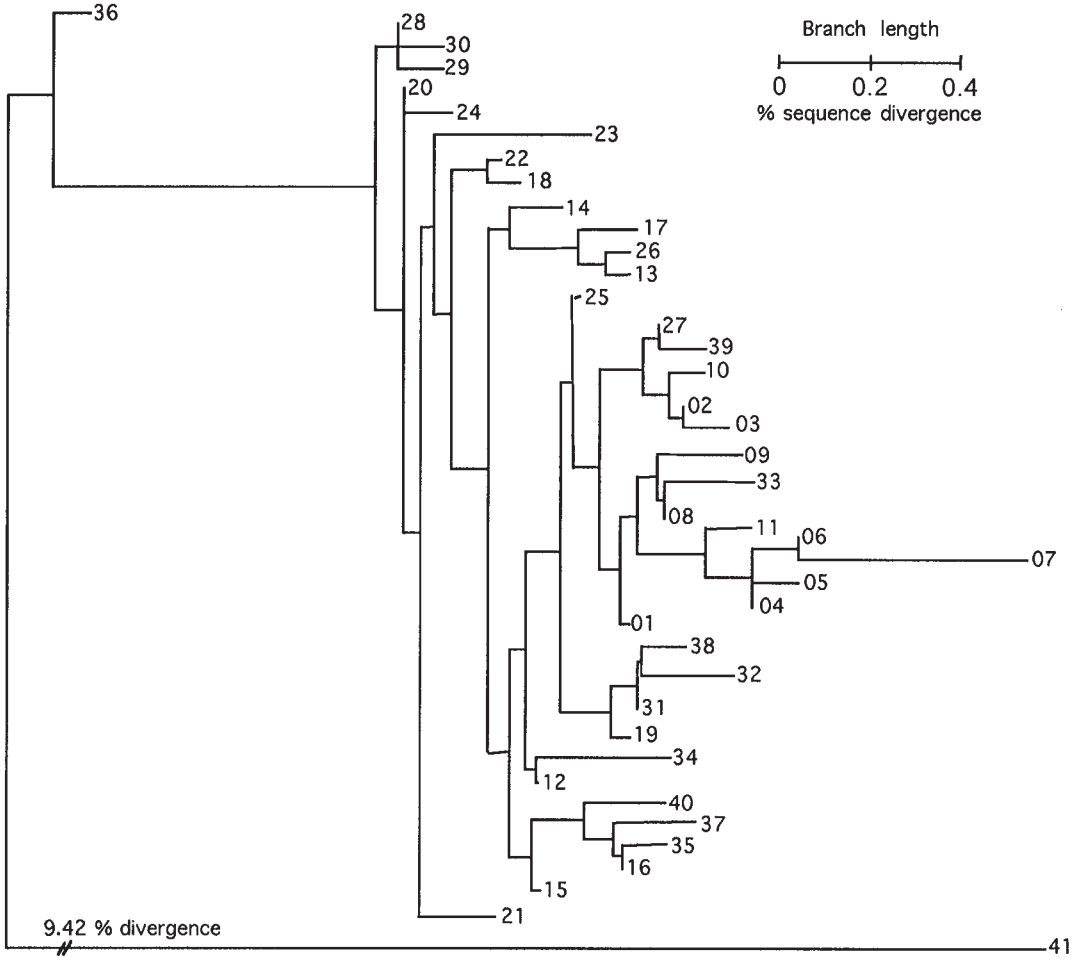
Hierarchical analysis of the total variation in mtDNA of white sucker revealed a significant regional subdivision in eastern North America and significant structure at the population level. Divergence between groups sampled to the northwest and southeast of the St. Lawrence River was significant and accounted for 14.30% of the variation ($p < 0.031$). Populations within regions were significantly differentiated and accounted for 35.63% of the variation ($p < 0.001$). The largest fraction of total variation was found within populations (50.07%, $p < 0.001$).

The number of mtDNA genotypes detected as a function of the sample size of specimens and restriction fragments increased rapidly for the whole area studied (Fig. 9). A sample size of more than 312 specimens and 96 restriction fragments was inadequate to detect all mtDNA variants within the sampled area (Figs. 9, 10). In the case of the St. Lawrence drainage, sample sizes of up to 204 specimens were not sufficient to detect all mtDNA variants (Fig. 11).

Discussion

The mean pairwise intraspecific sequence divergence of

Fig. 5. Fitch phenogram clustering the distance matrix, on the basis of restriction sites, of percent sequence divergence among 40 mtDNA haplotypes described for *Catostomus commersoni*. Numbers refer to haplotypes defined in Table 2. The tree was rooted with haplotype 41 represented by *Catostomus catostomus*.



0.70 ± 0.34%, ranging from 0.1 to 2.4%, is the greatest observed to date among northern fishes: *Coregonus clupeaformis* (0.36 ± 0.16%; range 0.07–0.86%), *Stizostedion vitreum* (0.47 ± 0.20%; range 0.15–0.60%), *Coregonus artedii* (0.52 ± 0.22%; range 0.08–1.03%), and *Salvelinus fontinalis* (0.62 ± 0.34%; range 0.11–1.32%). These four species and *Salvelinus namaycush* found in northeastern North America are composed of two or three genetic clades separated by many mutational steps and display heterogeneous geographic distributions (Grewe and Hebert 1988; Billington and Hebert 1988; Bernatchez and Dodson 1990*a*, 1990*b*; Bernatchez and Danzmann 1993). The most commonly accepted explanation for major genetic discontinuities with heterogeneous geographic distributions involves long-term barriers to gene flow, such that conspecific populations occupy easily recognizable branches on an intraspecific phylogenetic tree (Avise et al. 1987). In northern North America, such long-term barriers to gene flow are associated with the refugia created during Pleistocene glaciation events. Genetic groups with more easterly distributions are generally attributed to any number of Atlantic refugia (eastern coastal plains and exundated banks), whereas those with a more westerly distribution are attributed to the Mississippi or Missouri refugium. These genetic groups are deduced from phenograms and cladograms exhibiting genetic breaks or discontinuities that form clusters of haplotypes.

In the present study, a major genetic discontinuity was observed between haplotype 36 and all other haplotypes in Fitch and the majority-rule consensus tree. However, within the majority cluster, no major discontinuities were observed. Rather, four small clades separated by few mutational steps from other haplotypes were defined. Haplotypes of clade D are clearly associated with a westerly distribution, being found only in Manitoba and Lake Superior. Clade C is mostly located in the Atlantic region, but some haplotypes (Nos. 4, 5, 11) are found in more central regions. Clades A and B exhibit a west–east distribution relative to the St. Lawrence system, with some intermingling to the west of the St. Lawrence (Table 3). Most haplotypes (70%) were poorly resolved by the majority-rule consensus tree. The majority of fish characterized by these haplotypes were found west of the St. Lawrence River. Although a neighbour-joining phenogram showed no significant population groupings, an AMOVA based uniquely on the frequency distributions of haplotypes revealed that dividing the distribution of white suckers to the east and west of the St. Lawrence River accounted for a significant portion of the total

Fig. 6. Majority-rule consensus tree clustering 40 mtDNA haplotypes of *Catostomus commersoni* and their geographical distribution. Numbers at forks indicate the percentage of time the group consisting of the haplotypes located to the right of the fork occurred among the trees, out of 10 000 trees. The tree was rooted with *Catostomus catostomus*, haplotype 41. Numbers at the end of the tree forks indicate haplotypes defined in Table 2. Letters (a–m) refer to sampling sites identified in Table 1 and numbers below sampling sites are absolute abundances. The principal mtDNA clades are identified with vertical lines.

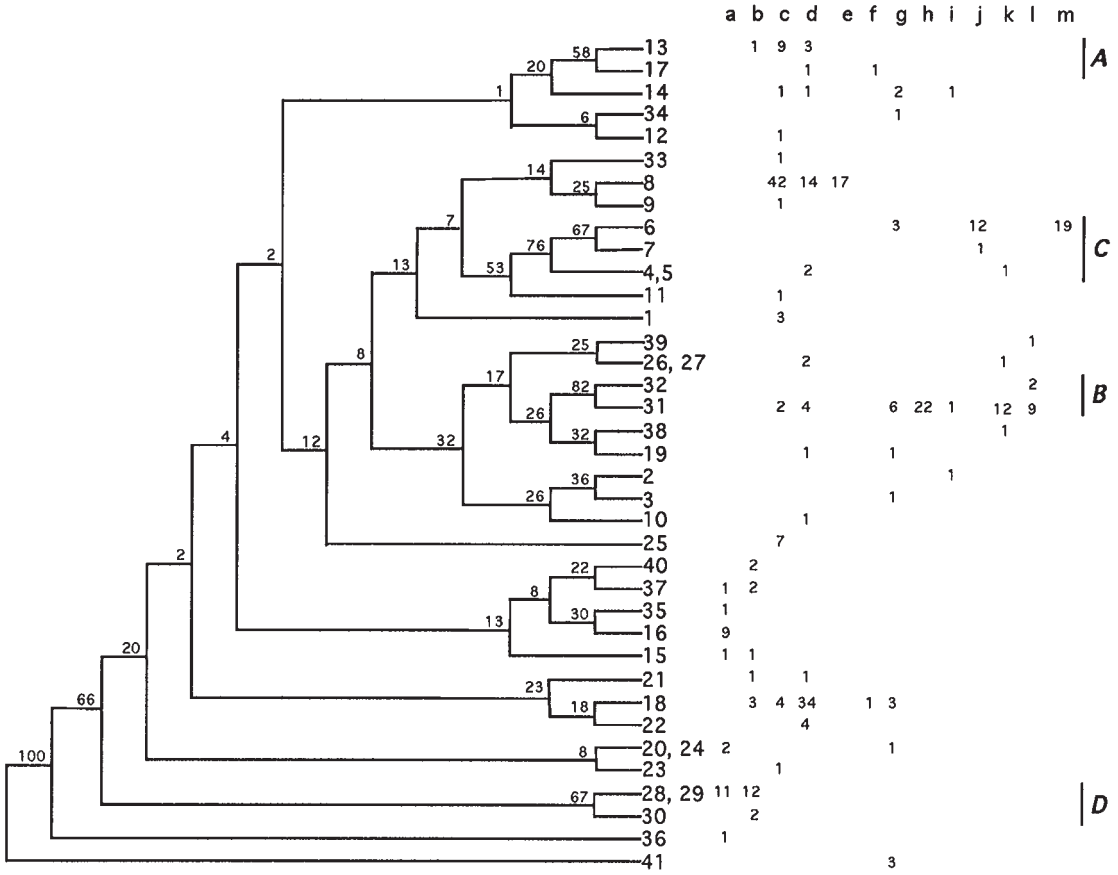


Table 3. Raw data for the geographic distribution of fish classified according to the four mtDNA clades and the poorly resolved haplotypes (PRH).

Clade/site	a	b	c	d	e	f	g	h	i	j	k	l	m
A	0	1	9	4	0	1	0	0	0	0	0	0	0
B	0	0	2	4	0	0	6	22	1	0	12	11	0
C	0	0	1	2	0	0	3	0	0	13	1	0	19
D	11	14	0	0	0	0	0	0	0	0	0	0	0
PRH	15	9	61	58	17	1	9	0	2	0	2	1	0

Note: Capital letters refer to genetic groups described in Fig. 7 and lowercase letters refer to sampling sites described in Fig. 2.

genetic variation. The St. Lawrence sample stands apart, largely owing to high genetic diversity (Figs. 8, 11).

The existence of many closely related haplotypes that are only partially geographically localized has been associated with species or subsets of species with historically intermediate levels of gene flow between geographic populations (phylogeographic category V, Avise et al. 1987). In this scenario, ancestral haplotypes may be dispersed over a wide area whereas more recent mutations are confined to specific areas (e.g., *Amia calva*, Bermingham and Avise 1986). This explanation does not apply to the present situation. There is no doubt

Fig. 7. The geographic distribution of the four discrete mtDNA clades. Letters refer to genetic groups described in Fig. 6. Poorly resolved haplotypes (PRH) are pooled in the same category. See raw data in Table 3.

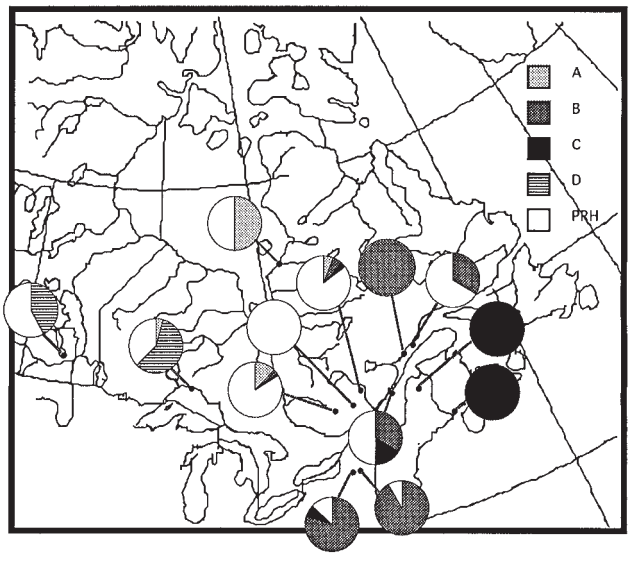


Fig. 8. Neighbour-joining phenogram clustering populations according to the distance matrix resulting from maximum likelihood estimation of the net average number of nucleotide substitutions per site between populations (nucleotide divergence). Intrapopulation nucleotide diversity is presented to the left of each population. The tree was rooted with the population of the St. Lawrence River (Quebec). Letters are populations identified in Table 1.

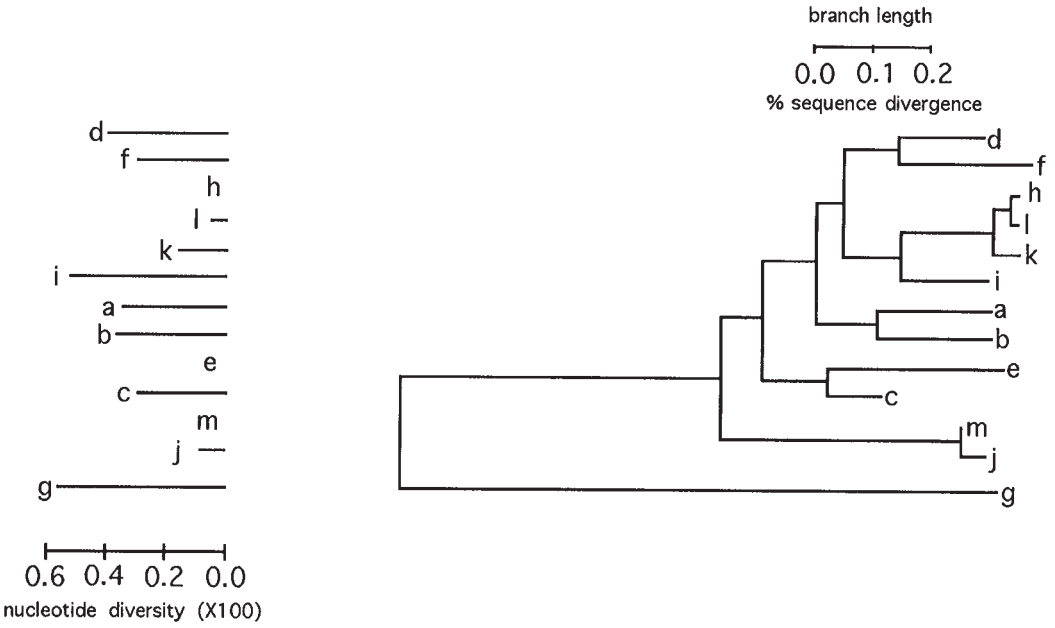


Fig. 9. Effect of fish sample size on the number of mtDNA clonal lines detected within all 13 populations of *Catostomus commersoni*. Incremental steps were stopped at 312 fish. Vertical bars are standard deviations based on 10 repetitions at each incremental step.

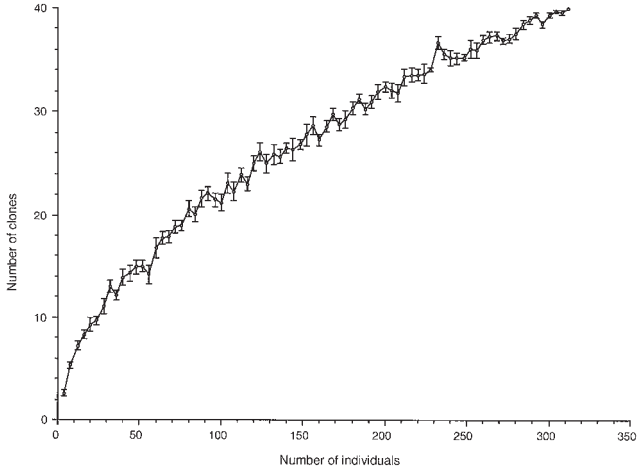
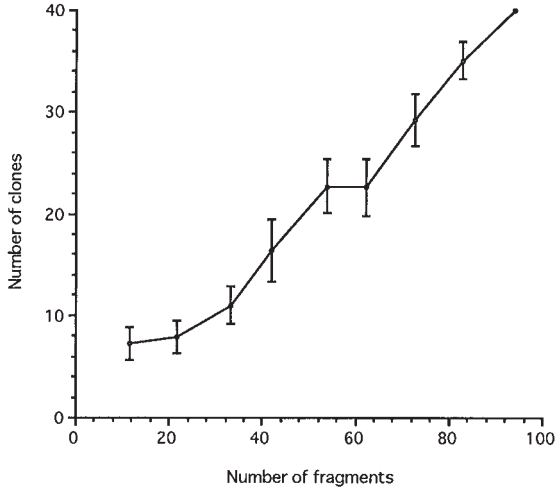


Fig. 10. Effect of number of restriction fragments assayed on the number of mtDNA clonal lines detected within all populations of white sucker studied. Vertical bars are standard deviations based on 10 repetitions at each incremental step.



that white sucker, over a large part of their present distribution, were confined to glacial refugia during Pleistocene glacial events that limited gene flow among populations. However, it is probable that white sucker were also extensively distributed to the south of the ice cap. The present distribution of the species includes areas that were well south of the maximum extent of the Laurentide ice sheet (Fig. 1). Thus, the northern part of the range may have been recolonized from multiple origins, contributing to the phylogenetic continuity of the species observed today. In addition, high genetic diversity may have been retained throughout the glacial distribution of the species that minimized the development of genetic discontinuities by stochastic lineage extinction. Avise et al. (1984) demonstrated that contemporary intraspecific mtDNA variability is influenced by stochastic lineage extinction, which in turn is a function of long-term effective population size. Today, the species is widely distributed and so abundant that it may represent more than half of the total fish biomass in some lakes (Chen and Harvey 1994). Thus, large effective population sizes in the past may have acted to buffer stochastic lineage extinction among geographically separated populations.

The four clades resolved in the majority-rule consensus tree involve haplotypes that diverged long before the last glaciation.

Fig. 11. Effect of fish sample size on the number of mtDNA clonal lines of white sucker detected within populations of St. Lawrence drainage: Gatineau River, St. Maurice River, Sans nom River, St. Lawrence River (Quebec City), Yoyo Lake, and Amédée River. Incremental steps were stopped at 204 fish. Vertical bars are standard deviations based on 10 repetitions at each incremental step.

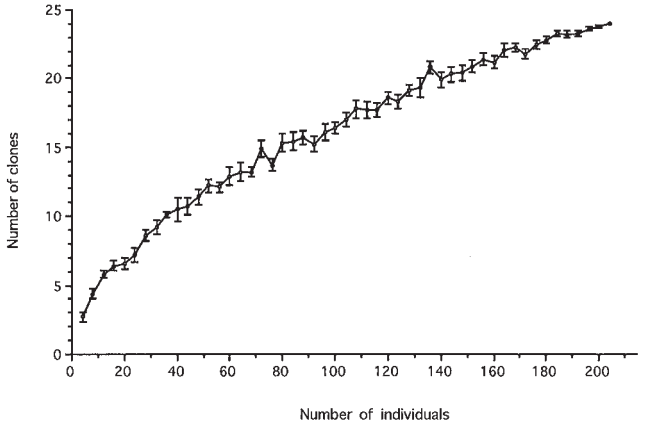
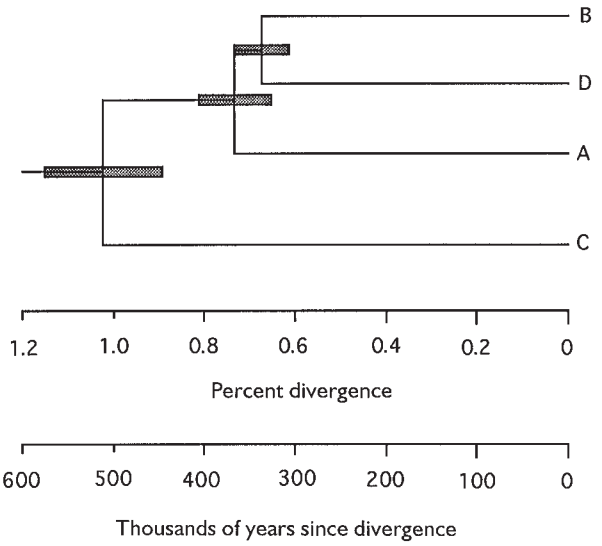


Fig. 12. Phenogram showing estimated percent sequence divergence among mtDNA of the four clades studied. Estimation of time of divergence is based on Brown's calibration of mtDNA evolution determined for mammalian species (Brown 1983). Horizontal bars represent standard deviations.



Assuming that the rate of divergence of 2% per million years calculated for mammals (Brown 1983) is a maximum rate for fish, clade C diverged from the others at least 500 000 years ago. The other clades diverged approximately 350 000 years ago (Fig. 12). However, the mitochondrial molecular clock was estimated at only 1% divergence per million years for three species of Pacific salmon (*Oncorhynchus* species) when molecular divergence estimates were calibrated with fossil evidence (Smith 1992). If this rate was applied to the present case, estimated divergence times would be doubled. In either case, the phylogenetic structure described here involves haplotype groupings that evolved during previous Pleistocene glaciations.

The dwarf white sucker population of Wolf Lake belongs to Clade B along with most other normal-sized white sucker of nearby Catlin Lake. One haplotype (No. 31) was the most common genotype characterizing both dwarf and normal ecotypes from the Adirondack sampling sites. It thus appears likely that both ecotypes were derived recently from a common ancestor. This scenario is comparable to the case of *Osmerus mordax* (Baby et al.1991; Taylor and Bentzen 1993) rather then the case of *Coregonus clupeaformis* (Bernatchez and Dodson 1991; Bernatchez et al. 1996). Unlike this latter study, however, the absence of other sampled lakes with dwarf sucker makes it impossible to assess the validity of a polyphyletic origin of dwarfism in this species. More intensive sampling in this region should be pursued to achieve this goal. It would also be important to find other regions where ecotypes are living in sympatry to determine if the divergence is truly sympatric or allopatric.

The attempt to relate haplotype groupings to recent glacial refugia is obviously hampered by the weak phylogeographic structure of the species. Clade D is most clearly related to a western distribution and may indicate recolonization from the Mississippi or Missouri refuge, as originally proposed by Crossman and McAllister (1986). The highly divergent haplotype No. 36 from Lake Manitoba suggests the existence of other divergent clades in the western part of their distribution, but this requires further sampling to confirm. Clade C appears to be Atlantic in origin, but it is present in the central part of the study region. The actual extent of this group will require a greater sampling effort. In fact, our sampling effort was not sufficient to document all of the genetic variability of the species in the study area (Fig. 9). Even with two thirds of all fish sampled coming from the St. Lawrence River drainage, sample size was not great enough to exhaust haplotype diversity (Fig. 11). Furthermore, addition of new restriction enzymes will surely reveal more polymorphic restriction patterns (Fig. 10). Nevertheless, available results illustrate that the St. Lawrence River delineates a transition zone in the recolonization of northeastern North America by white sucker, with fish of clades B and C largely restricted to areas located to the southeast of the St. Lawrence River. This suggests that these two clades were probably associated with Atlantic refugia located somewhere to the east of the Appalachians along the coastal plain of eastern North America.

In conclusion, the phylogeographic structure of the white sucker appears to be unique among fishes studied to date in northeastern North America. The lack of phylogenetic discontinuities and the intermingling of some of the few significant clades that do exist is in sharp contrast to other species sampled in this area. As this may be associated with an extensive distribution south of the ice sheet, a similar phylogeographic structure may be expected among other species with similar distributions and abundances.

Acknowledgements

We thank the many people who assisted in obtaining samples for this study: Mr. Fernand Saurette, Canadian Department of Fisheries and Oceans (DFO), Winnipeg, Manitoba; Dr. Kelly Munkittrick, DFO, Great Lakes Laboratory for Fisheries and Aquatic Sciences, Burlington, Ontario; Dr. Pierre Magnan, Université du Québec à Trois Rivières, Trois-Rivières, Québec; Dr. Louis Bernatchez, Université Laval, Ste-Foy, Québec; Mr. Eric Desbiens, Association de la chasse et de la pêche de

Forestville, Québec; Mr. Michel Breault, ministère de l'Environnement et de la Faune, Baie Comeau, Québec; Mr. Richard Sage, Jr., Adirondack Ecological Center, Newcomb, New York; and Mr. Fred Purton, Huntsman Marine Sciences Center, St. Andrews, New Brunswick. We thank Drs. L. Bernatchez, R. Danzman, and D. Pallotta and an anonymous referee for providing critical reviews of the paper and Mrs. Françoise Colombani and Mr. Dominique Gagné for assistance in the field and laboratory. This study was funded by grants from the Natural Sciences and Engineering Research Council of Canada to J.J.D. (individual research grant) and to J.J.D., Peter Hodson, and Catherine Couillard (strategic grant).

References

Avise, J.C. 1992. Molecular population structure and biogeographic history of a regional fauna: a case history with lessons for conservation biology. Oikos, **63**: 62–76.

Avise, J.C., Neigel, J.E., and Arnold, J. 1984. Demographic influences on mitochondrial DNA survivorship in animal populations. J. Mol. Evol. **20**: 99–105.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., and Saunders, N.C. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. **18**: 489–522.

Baby, M.-C., Bernatchez, L., and Dodson, J.J. 1991. Genetic structure and relationships among anadromous and landlocked populations of rainbow smelt, *Osmerus mordax* Mitchill, as revealed by mtDNA restriction analysis. J. Fish Biol. **39**(Suppl. A): 61–68.

Beamish, R.J. 1970. Factors affecting the age and size of the white sucker *Catostomus commersoni* at maturity. Ph.D. thesis, University of Toronto, Toronto, Ont.

Bermingham, E., and Avise, J.C. 1986. Molecular zoogeography of freshwater fishes in the southeastern United States. Genetics, **113**: 939–965.

Bernatchez, L. 1995. A role for molecular systematics in defining evolutionary significant units in fishes. Am. Fish. Soc. Symp. **17**: 114–132.

Bernatchez, L., and Danzmann, R.G. 1993. Congruence in control-region sequence and restriction-site variation in mitochondrial DNA of brook charr (*Salvelinus fontinalis* Mitchell). Mol. Biol. Evol. **10**: 1002–1014.

Bernatchez, L., and Dodson, J.J. 1990*a*. Mitochondrial DNA variation among anadromous populations of cisco (*Coregonus atedii*) as revealed by restriction analysis. Can. J. Fish. Aquat. Sci. **47**: 533–543.

Bernatchez, L., and Dodson, J.J. 1990*b*. Allopatric origin of sympatric populations of lake whitefish (*Coregonus clupeaformis*) as revealed by mitochondrial-DNA restriction analysis. Evolution, **44**: 1263–1271.

Bernatchez, L., and Dodson, J.J. 1991. Phylogeographic structure in mitochondrial DNA of the lake whitefish (*Coregonus clupeaformis*) and its relation to Pleistocene glaciations. Evolution, **45**: 1016–1035.

Bernatchez, L., Savard, L., Dodson, J.J., and Pallotta, D. 1988. Mitochondrial DNA sequence heterogeneity among James–Hudson Bay anadromous coregonines. Finn. Fish. Res. **9**: 17–26.

Bernatchez, L., Dodson, J.J., and Boivin, S. 1989. Population bottlenecks: influence on mitochondrial DNA diversity and its effect in coregonine stock discrimination. J. Fish Biol. **35**(Suppl. A): 233–244.

Bernatchez, L., Vuorinen, J.A., Bodaly, R.A., and Dodson, J.J. 1996. Genetic evidence for reproductive isolation and multiple origins of sympatric trophic ecotypes of whitefish (*Coregonus*). Evolution, **50**: 624–635.

Billington, N., and Hebert, P.D.N. 1988. Mitochondrial DNA variation in Great Lakes walleye (*Stizostedion vitreum*) populations. Can. J. Fish. Aquat. Sci. **47**: 1093–1102.

Billington, N., and Hebert, P.D.N. 1991. Mitochondrial DNA diversity in fishes and its implications for introductions. Can. J. Fish. Aquat. Sci. **48**(Suppl. 1): 80–94.

Billington, N., Barrette, R., and Hebert, P.D.N. 1992. Management implications of mitochondrial DNA variation in walleye stocks. N. Am. J. Fish. Manage. **12**: 276–284.

Brown, W. 1983. Evolution of animal mitichondrial DNA. *In* Evolution of genes and proteins. *Edited by* M. Nei and R.K. Koehn. Sinauer Associates, Sunderland, Mass. pp. 62–88.

Chen, Y., and Harvey, H.H. 1994. Maturation of white sucker, *Catostomus commersoni*, populations in Ontario. Can. J. Fish. Aquat. Sci. **51**: 2066–2076.

Crossman, E.J., and McAllister, D.E. 1986. Zoogeography of freshwater fishes of the Hudson Bay drainage, Ungava Bay and the Arctic Archipelago. *In* The zoogeography of North American freshwater fishes. *Edited by* C.H. Hocutt and E.O. Wiley. J. Wiley and Sons, New York. pp. 53–104.

Dence, W.A. 1948. Life history, ecology and habits of the dwarf sucker, *Catastomus commersonni utawana* Mather, at the Huntington Wildlife Station. Roosevelt Wildlife Bull. **8**: 81–150.

Excoffier, L., Smouse, P.E., and Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics, **131**: 479–491.

Felsenstein, J. 1992 PHYLIP (phylogeny inference package). Version 3.4. Department of Genetics, SK-50, University of Washington, Seattle, Wash.

Fitch, W.M., and Margoliash, E. 1967. Construction of phylogenetic trees. Science (Washington, D.C.), **155**: 279–289.

Grewe, P.M., and Hebert, P.D.N. 1988. Mitochondrial DNA diversity among broodstocks of the lake trout, *Salvelinus namaycush*. Can. J. Fish. Aquat. Sci. **45**: 2114–2122.

Hebert, P.D.N., Ward, R.D., and Weider, L.J. 1988. Clonal-diversity patterns and breeding system variation in *Daphnia pulex*, an asexual–sexual complex. Evolution, **42**: 147–159.

McAllister, D.E., Platania, S.P., Schueler, F.W., Baldwin, M.E., and Lee, D.S. 1986. Ichtyofaunal patterns on a geographic grid. *In* The zoogeography of North American freshwater fishes. *Edited by* C.H. Hocutt and E.O. Wiley. J. Wiley and Sons, New York. pp. 18–50.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nei, M., and Li, W.H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. U.S.A. **76**: 5269–5273.

Nei, M., and Tajima, F. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics, **967**: 145–163.

Nei, M., and Tajima, F. 1983. Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. Genetics, **105**: 207–217.

Schmidt, R.E. 1986. Zoogeography of the northern Appalachians. *In* The zoogeography of North American freshwater fishes. *Edited by* C.H. Hocutt and E.O. Wiley. J. Wiley and Sons, New York. pp. 137–159.

Scott, W.B., and Crossman, E.J. 1974. Poissons d'eau douce du Canada. Bull. Fish. Res. Board Can. No. 184F.

Smith, G.R. 1992. Introgression in fishes: significance for paleontology, cladistics and evolutionary rates. Syst. Biol. **41**: 41–57.

Taylor, E.B., and Bentzen, P. 1993. Evidence for multiple origins and sympatric divergence of trophic ecotypes of smelt (*Osmerus*) in northeastern North America. Evolution, **47**: 813–832.

Underhill, J.C. 1986. The fish fauna of the Laurentian Great Lakes, the St. Lawrence Lowlands, Newfoundland and Labrador. *In* The zoogeography of North American freshwater fishes. *Edited by* C.H. Hocutt and E.O. Wiley. J. Wiley and Sons, New York. pp. 105–136.