Molecular Ecology (2001) **10**, 987–1001

Mitochondrial DNA phylogeography of lake cisco (*Coregonus artedi*): evidence supporting extensive secondary contacts between two glacial races

J. TURGEON* and L. BERNATCHEZ
GIROQ, Département de biologie, Université Laval, Ste-Foy, Québec, Canada G1K 7P4

Abstract

The comparative molecular phylogeography of regional fish fauna has revealed the wide distribution of young clades in freshwater fishes of formerly glaciated areas as well as interspecific incongruences in their refugial origins and recolonization routes. In this study, we employed single-strand conformation polymorphism (SSCP) and sequence analyses to describe mitochondrial DNA (mtDNA) polymorphism among 27 populations of the lake cisco (*Coregonus artedi*) from its entire range of distribution in order to evaluate the hypothesis of dual glacial refuges proposed by Bernatchez & Dodson against the traditional view that this species is solely of Mississippian origin. Results indicate that this taxon is composed of two closely related groups that are widely distributed and intermixed over most of the sampled range. The estimated level of divergence (0.48%), the contrast in the geographical distribution of each group, as well as the general distribution of *C. artedi* in North America together support the hypothesis that one group dispersed from a Mississippian refuge via the proglacial lakes, while the other is of Atlantic origin and also took advantages of earlier dispersal routes towards eastern Hudson Bay drainages. However, the signal of past range fragmentation revealed by a nested clade analysis was weak, and did not allow to formally exclude the hypothesis of a single Mississippian origin for both lineages. Comparisons with the phylogeographic patterns of other Nearctic freshwater fishes suggest that the salinity tolerance and thermal sensitivity of lake cisco may have been determinant for its extensive postglacial dispersal. The presence or co-occurrence of sympatric or allopatric eco/morphotypes were not found to be necessarily associated with the presence of both haplotype groups.

Keywords: *Coregonus artedi*, colonization, dispersal, ecotype, refuge, mtDNA, phylogeography

Received 7 August 2000; revision received 9 November 2000; accepted 9 November 2000

Introduction

It is now well recognized that the Pleistocene glaciations have had profound effects on the extant distribution of Holarctic biota, and that these effects were particularly severe in North America (NA) due to the extent of the last Wisconsinan ice sheet (Pielou 1991). Effects of glaciations were felt beyond the ice margin as the general climate became drier and affected the hydrological networks in southern nonglaciated areas. These, however, were evidently both more direct and drastic in formerly glaciated areas of north temperate NA. In these regions, the restrictive dispersal requirements of freshwater fishes has resulted in distributional patterns that are intimately related to the location of glacial refuges and the variable continental hydrology characterizing the period of ice retreat at the end of the last (Wisconsinan) glaciation. The immense proglacial lakes then provided formidable dispersal avenues for those species able to withstand their changing nature and has resulted in typically larger range sizes for fishes of formerly glaciated areas (McAllister *et al*. 1986). Inferences on the refugial origin and postglacial dispersal routes of north temperate freshwater fishes of NA have been obtained by comprehensive biogeographical studies using distribution, geographical variation, fossils and parasites as sources of information (Hocutt & Wiley 1986).

Correspondence: Julie Turgeon. *Present address: Department of Biology, Dartmouth College, Hanover, NH 03755 USA. Fax: 603 6461347; E-mail: julie.turgeon@dartmouth.edu

More recently, molecular phylogeography based on the analysis of mitochondrial DNA (mtDNA) variation has supplied independent data that have shed new light on the postglacial history of several widely distributed NA freshwater fishes. The first studies used restricted geographical surveys that confirmed the presence of distinct lineages whose geographical distribution and estimated time of divergence corresponded to dispersal from different refuges and divergence during the Pleistocene (e.g. Billington & Hebert 1988; Grewe & Hebert 1988; Bernatchez & Dodson 1990a,b). The power of this approach, however, was better substantiated by extensive surveys of entire species range of distribution (Bernatchez & Dodson 1991; Billington *et al*. 1992; Wilson *et al*. 1996; Bernatchez 1997; Danzmann *et al*. 1998; Wilson & Hebert 1998; Taylor *et al*. 1999). These comprehensive studies confirmed the identity of glacial races and refugial origins inferred from traditional biogeography, but they additionally demonstrated new refugial origins and suggested unsuspected long dispersal events.

The information recently provided by fish molecular phylogeography, however, extends well beyond confirming and/or identifying refuges and dispersal routes, and has been greatly enriched by a comparative approach at various taxonomic and geographical scales (Avise 1994; Hewitt 1996; Turner *et al*. 1996; Zink 1996; Bernatchez & Wilson 1998; Hewitt 1999). For example, evolutionary insight was acquired by the combined results of several studies indicating that increased levels of intraspecific polymorphisms defining sympatric ecomorphotypes is often, but not necesarily, associated with zones of secondary contacts between distinct refugial races (e.g. Bernatchez & Dodson 1991; Bernatchez 1997; Taylor & McPhail 1999). In turn, a global comparison of the phylogeographic structures of regional fish fauna from southern (nonglaciated) and northern (formerly glaciated) areas revealed striking differences (Bernatchez & Wilson 1998). On the one hand, southern species were generally composed of highly divergent clades showing highly concordant but limited distributional ranges and contact zones. In contrast, northern species appeared to be composed of relatively young clades (averaging 0.5–2% sequence divergence over the entire mtDNA genome) dominated by fewer but more widely dispersed haplotypes, and they globally exhibited reduced interspecific concordance in terms of clade distribution, contact zone size and location. This absence of concordance in patterns is likely due to differences in ecological factors related to each species tolerance to the changing physical environments (e.g. current, salinity, turbidity, and temperature) and competitive ability during the dispersal and colonization processes, as well as to demographic factors impinging on the survival chances of each lineage. Therefore, information on the phylogenetic structure of several wide ranging fish species from the north temperate regions are required to identify which factors, ecological and/or demographic, are truly predictive of phylogeographic structure of young but distinct evolutionary lineages.

Lake cisco (*Coregonus artedi*) is an important planktivorous forage fish of North American cold freshwater ecosystems and is among the most widely distributed freshwater fishes in NA (Scott & Crossman 1973; Lee *et al*. 1980). Lake cisco is characterized by extensive phenotypic polymorphisms that have resulted in an early taxonomic frenzy, with some 39 species being described early in the last century (reviewed in Clarke 1973; Scott & Crossman 1973). The origin of these polymorphisms is still unknown, and the global phylogenetic and taxonomic relationships of *C. artedi* with other endemic ciscoes are unresolved (Bodaly *et al*. 1991; Bernatchez & Dodson 1994; Reed *et al*. 1998; Turgeon *et al*. 1999). *C. artedi* is the most widely distributed cisco taxon in NA, and it is generally recognized that it represents one of the ancestral lineage(s) of ciscoes in North American interior lakes (Smith & Todd 1984). The extant distribution range of *C. artedi* is almost entirely comprised within the limits of maximal glacial coverage during the last glaciation (Fig. 1). It ranges from the upper Mackenzie River system (Great Bear Lake) through most of central Canada and into the Laurentian Great Lakes watershed. *C. artedi* is also present on both sides of Hudson Bay, around Ungava Bay, in the western part of the Quebec Peninsula, and extends in river systems of the lower St. Lawrence River (Ottawa River to Saguenay River). It reaches its southern limit of distribution in the upper Mississippi watershed of north central United States. It is absent from Beringia and eastern Canada, but occurs in the upper parts of the Susquehanna and Hudson rivers watersheds in New England (Hocutt *et al*. 1986; Schmidt 1986).

On the basis of these distributional data, most authors recognized the Mississippian refuge as the unique origin of postglacial dispersal for lake cisco (McPhail & Lindsey 1970; Crossman & McAllister 1986; Underhill 1986; but see Mandrak & Crossman 1992). However, a study on mtDNA restriction fragment length polymorphisms (RFLP) of *C. artedi* identified two divergent groups of haplotypes that exhibited contrasting geographical distributions in the James/Hudson Bay area (Bernatchez & Dodson 1990a). The level of divergence between these clades (mean sequence divergence of 0.52%) is comparable to what has been documented among refugial races in other north temperate freshwater fishes (reviewed in Bernatchez & Wilson 1998). Moreover, one group was more abundant in James Bay while the other dominated in Hudson Bay, suggesting that they may represent distinct refugial races having colonized the region by alternative southern and northern routes, respectively. However, this survey was geographically restricted and therefore limited the extent of inferences on the postglacial colonization history of that species.

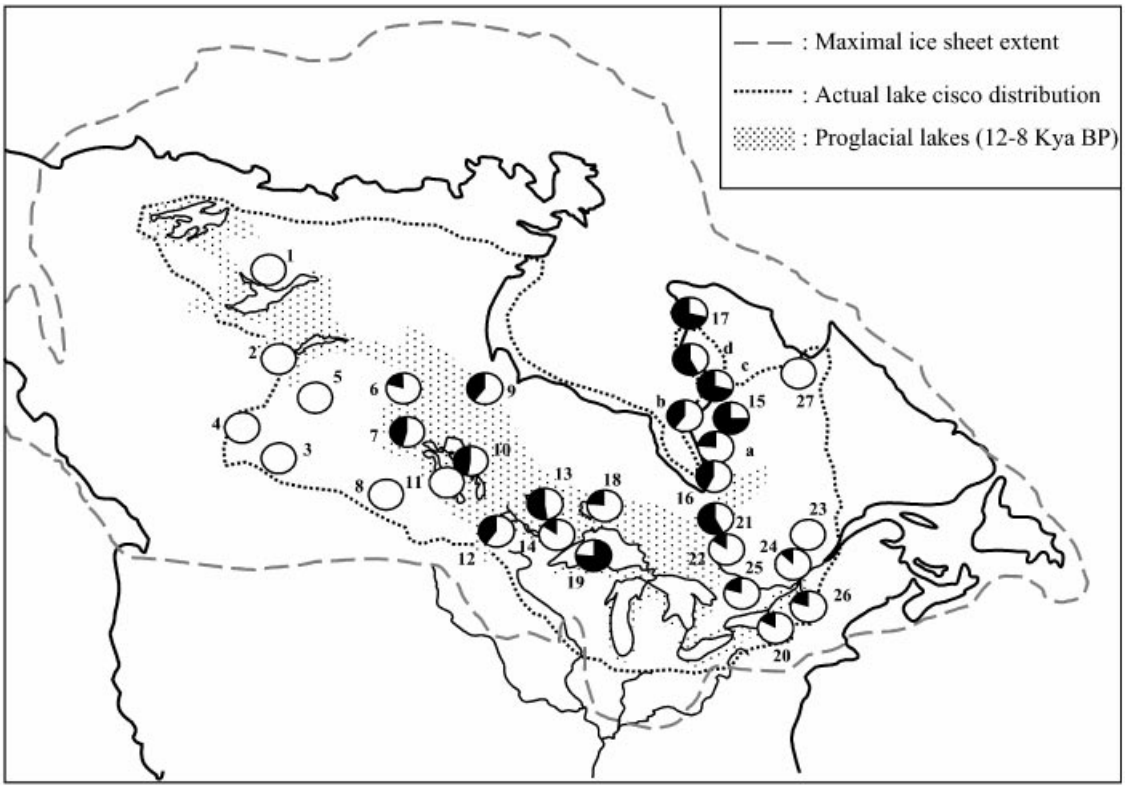

Fig. 1 Sampling sites for *Coregonus artedi* and geographical distribution of the two major mtDNA clades defined by SSCP (A: open, B: filled). Sites are numbered as in Table 1, morphotypes are pooled for sites 2, 15, and 17 (see text). Data from Bernatchez & Dodson (1990a) are as follows: a, Eastmain River; b, La Grande River; c, Great Whale River; and d, Richmond Gulf. The maximum extent of the Wisconsinan ice sheet (dashed line) and proglacial lakes (dotted area), and the extant distribution of *C. artedi* (dotted line) are also outlined.

In this paper, we report the geographical distribution of mitochondrial polymorphisms in 27 populations of *C. artedi* distributed over most of the species range in order to test the hypothesis of Bernatchez & Dodson (1990a) that there are two glacial races of lake cisco in NA. Single-strand conformation polymorphisms (SSCP) corresponding to point mutations are detected in a segment of the mitochondrial control region that has proven suitable to detect glacial races (Angers & Bernatchez 1998). Conformant sequences are extended to additional mitochondrial segments (ATPase and ND1) to establish the phylogenetic relationships and divergence levels among D-loop variants, and their geographical distribution is examined in relation to potential refugial origin and dispersal routes. The relationship between molecular and phenotypic variation is also examined in cases of allopatric and sympatric eco/morphotypes in order to gain insight on the origin of phenotypic diversity reported in *C. artedi*. The global phylogeographic structure of *C. artedi* is also compared to that of other postglacial freshwater fishes in order to assess the combined evolutionary influences of historical and ecological factors.

Materials and methods

Fish heads or tissue samples were obtained from 27 locations throughout the distribution range of *Coregonus artedi* (Table 1). Sympatric morphotypes were obtained from the eastern, central and western portions of the species range. These were large and normal adult individuals from La Grande Reservoir, Quebec (total fish length of 399 mm ± 42 and 140 mm ± 15, respectively; $n = 71$, t-test: $P < 0.001$), normal and dwarf individuals from Cormorant Lake, Manitoba (mean total number of gill rakers of 52.8 ± 2.05 and 48.3 ± 3.1, respectively; $n = 50$, t-test: $P < 0.0001$), and normal (high gill raker count) and large (low gill raker count) individuals from Barrow Lake, Alberta. The unique spring-spawning population of Lac des Écorces, QC (Pariseau *et al*. 1983; Hénault & Fortin 1989), was also analysed and compared to a downstream fall-spawning population of Poisson Blanc Reservoir. Populations of the James Bay/Hudson area are anadromous (Morin & Dodson 1986), and individuals from Rupert and Povungnituk rivers are those used for the mtDNA RFLP analysis of Bernatchez & Dodson (1990a). Other allopatric size and gill raker count polymorphisms

Table 1 Collection sites for *Coregonus artedi*, showing drainage, geographical locations, and reported phenotypes

Site	Waterbody	Drainage	Sub-drainage	Latitude (°N. min)	Longitude (°W. min)	Phenotype
1	Great Slave Lake, NWT	Mackenzie	Mackenzie	61.23	115.38	large
2a	Barrow Lake, AB		Athabasca	59.15	111.13	normal
2b						large-LGR*
3	Peerless Lake, AB		Peace River	56.40	114.30	dwarf
4	Utikuma Lake, AB		Peace River	55.50	115.30	large
5	Lake La Biche, AB		Athabasca	54.45	112.05	normal
6	Lake La Ronge, SK	Hudson (West)	Churchill	55.14	104.57	normal
7a	Lake Cormorant, MB		Churchill	54.30	100.35	normal
7b						dwarf
8	Pasqua Lake, SK		Red	51.47	103.58	large
9	Nelson River, MB		Nelson	56.50	91.00	anadromous
10	Lake Winnipeg, MB		Nelson	50.40	97.15	normal
11	Lake Winnipegosis, MB		Nelson	52.45	100.15	normal
12	Lake of the Woods, ON		Winnipeg	49.18	94.44	normal
13	Lake Seul, ON		Winnipeg	50.30	92.02	normal
14	Lake Saganaga, MN		Winnipeg	48.15	90.52	HGR†
15a	La Grande Reservoir, QC	Hudson (East)	La Grande	55.00	78.30	large
15b						normal
16	Rupert River, QC		Rupert	51.00	78.50	anadromous
17	Povungnituk River, QC		Povungnituk	60.02	77.10	anadromous
18	Lake Nipigon, ON	Atlantic	St. Lawrence	49.45	88.30	normal
19	Lake Superior, ON		St. Lawrence	48.00	88.45	normal
20	Lake Ontario, ON		St. Lawrence	43.45	78.00	normal
21	Lake Opasatica, QC		Ottawa	48.05	79.15	normal
22	Kipawa Reservoir, QC		Ottawa	46.50	79.05	normal
23	Lac des Écorces, QC		Ottawa	46.32	75.25	spring spawner‡
24	Res. Poisson Blanc, QC		Ottawa	45.55	75.45	normal
25	White Partridge Lake, ON		Ottawa	45.50	78.06	LGR§
26	Lake Champlain, VT		Richelieu	45.03	73.09	normal
27	Three Lakes, QC	Ungava	Koksoak	58.00	68.00	normal

*Low total Gill Raker count (range 38–43). Putative *C. zenithicus* (M. Steinhilber, U. of Alberta).
†High total Gill Raker count (range 49–70). Putative *C. nipigon* (D. A. Etnier, U. of Tennessess).
‡Pariseau *et al*. (1983), Hénault & Fortin (1989).
§Low total Gill Raker count (range 33–37). Putative *C. zenithicus* (N. Mandrak, Youngstown St. University, OK).

were also reported by government and academic personnel providing samples (Table 1).

We employed SSCP (Orita *et al*. 1989; Sheffield *et al*. 1993) to detect sequence variation in a 375-bp fragment of the mitochondrial control (D-Loop) region. DNA was extracted using the classical phenol protocol (Sambrook *et al*. 1989) and a D-Loop segment was amplified using the H2/LN20 primers of Bernatchez & Danzmann (1993). This region of the D-loop excludes the 3′-end where length polymorphism has been detected in ciscoes (Reed *et al*. 1998). All amplifications were performed with a Perkin Elmer GeneAmp PCR System 9600 in a total reaction volume of 12.5 μL with 25–50 ng of genomic DNA, 800 nM of each primer, 75 μM of CTP, GTP, and TTP, 5 μM ATP, 0.2 μCi ^{35}S-ATP, 1.5 mM $MgCl_2$, and 0.25 unit of *Taq* polymerase. Polymerase chain reaction (PCR) cycles were as follow: 1 × (3 min at 95 °C), 32 × (1 min at 95 °C, 45 s at 45 °C, 1 min at 72 °C), 1 × (10 min at 72 °C). Formamide blue (5 μL) was added to the amplified samples which were denatured at 95 °C for 5 min and immediately put on ice prior to electrophoresis in a nondenaturating 5.5% acrylamide gel (1:49 bisacrylamide:acrylamide, i.e. 1% cross-linking) with 7% glycerol and 0.5X TBE. Migrations were performed at 15 W for 12–15 h at 4–10 °C, and conformation patterns were visualized by autoradiography following standard procedures (Sambrook *et al*. 1989). Reference conformant patterns were used during each run and all conformants were compared in at least one run.

The entire H2/LN20 fragment of all observed conformants was sequenced using at least two individuals from two distant locations when appropriate and possible (Table 2). Sequences were determined on an ABI 377 automated sequencer following PCR amplification in 75 μL volume (conditions as above) and agarose gel extraction using the Quiagen Quick Gel Extraction Kit protocol. Cycle sequencing

Table 2 Geographical distribution of *Coregonus artedi* H2/LN20 SSCP conformants and within population haplotype (*h*) and nucleotide (π) diversity based on sequence data (Fig. 2). Numbers of individuals sequenced are indicated in parentheses. Site codes are as in Table 1

	SSCP Conformant													
	Clade A										Diversity		Total	
Site	A	C	D	H	I	L	M	N	T	Clade B	N_i	N_h	*h*	π
1	20 (1)		1 (1)								20 (2)	2	0.10	0.0002
2a	24 (1)										24 (1)	1	0.16	0.0001
2b	16 (1)										16 (1)	1	0.16	0.0001
3	20										20	1	0	0
4	10				10 (2)						20 (2)	2	0.53	0.0009
5	19										19	1	0	0
6	19 (2)									5 (2)	24 (4)	2	0.41	0.0013
7a	10			1						9 (2)	20 (2)	3	0.57	0.0018
7b	9	2								10	21	3	0.65	0.0021
8	15			5 (2)							20 (2)	2	0.39	0.0003
9	12									8	20	2	0.51	0.0017
10	10									9	19	2	0.53	0.0017
11	20										20	1	0	0
12	12									8	20	2	0.51	0.0017
13	9								1 (1)	11	21 (1)	3	0.57	0.0019
14	19									1	20	2	0.10	0.0003
15a	4									15	19	2	0.35	0.0014
15b	6									15	21	2	0.43	0.0014
16	11(2)	1 (1)				1 (1)				9 (3)	22 (6)	4	0.72	0.0022
17	4									10	14	2	0.44	0.0012
18	14									6	20	2	0.44	0.0015
19	4	1					1 (1)			15	21 (1)	4	0.47	0.0016
20	13							1 (1)		3	17 (1)	3	0.47	0.0012
21	8									11	19	2	0.51	0.0017
22	17									3	20	2	0.27	0.0009
23	19										19	1	0	0
24	21									3	24	2	0.23	0.0008
25	19 (2)									5 (2)	24 (4)	2	0.41	0.0013
26	16									4	20	2	0.34	0.0011
27	20										20	1	0	0
Total	420 (9)	4 (1)	1 (1)	6 (2)	10 (2)	1 (1)	1 (1)	1 (1)	1 (1)	160 (9)	604 (27)	10	0.35	0.0015

reactions were performed using each primer and performed twice to resolve single nucleotide ambiguities. For each of the H2/LN20 sequences documented, two additional mtDNA fragments were amplified: a portion of the ATPase 6 region using the L8558 and H9208 primers of Giuffra *et al*. (1994), and a portion of the ND1 region using the N1 (5′-GTAGGAAGCTCGTACTCA-3′) and N1S (5′-TGCAGCCG-CTATTAAGGGT-3′) primers of Doiron (2000). Amplifications (T_a = 48 °C) and gel extractions were performed as for H2/LN20; 494 bp of ATPase 6 and 355 bp of ND1 were sequenced using the H9208 and N1 primers, respectively. The composite (1224 bp) sequences were used to calculate genetic distances using Kimura 2-parameters with MEGA 1.02 (Kumar *et al*. 1993) and intrapopulation haplotype (*h*) and nucleotide diversity (π) were estimated with the DA module of REAP, version 4.0 (McElroy *et al*. 1992). The matrix of variable sites (Fig. 2) was used to visually construct a haplotype tree. We then estimated the probability of a parsimonious relationship between pairs of haplotypes differing by a given number of detected mutations (*j*) while sharing *m* identical nucleotides. This was evaluated using Templeton *et al*. (1992) algorithm and the program PROBPARS kindly provided by A. R. Templeton (Washington University, St-Louis, MI). This program first evaluates the probability that *one* site difference (*j* = 1) between two randomly drawn haplotypes is actually due to more than one mutation at that site (nonparsimonious state). Then, the limit of validity for inferring parsimonious relationships among haplotype pairs is evaluated by calculating the probability of non-parsimonious relationships for haplotype pairs that differ at *j* > 1 sites but that share *m* sites. This approach focuses on shared characters among haplotypes that differ minimally

Sequence Code (# . SSCP pattern)	Dloop (LN20)											ND1(N1)					ATPase 6 (H9208)								
								1	1	1	1	2				1	3		1	1	2	3	3	4	4
	2	8	8	8	9	9	1	2	6	9	7	2	8	9	2	3	3	1	5	9	1	8	1	6
	9	1	6	7	1	3	8	5	7	4	2	0	3	5	2	5	8	7	1	3	2	6	3	8	2
1. B	A	T	T	G	G	A	G	A	C	T	T	C	G	C	C	G	T	G	T	G	C	T	T	C	C
2. B	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	C	.	.	.	.	.	G	.	.
3. B	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	G	.	.
4. A	.	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.	.	.	G	T	.
5. A	.	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.	T	.	G	T	T
6. A	.	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	.	T	.	.	.	.	G	T	.
7. A	.	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.	.	C	G	T	.
8. A	G	.	.	.	.	.	.	.	A	.	.	T	.	.	T	.	.	.	.	.	.	.	G	T	.
9. A	.	.	.	.	.	.	.	.	A	C	.	T	.	.	.	A	.	.	.	.	.	.	G	T	.
10. C	.	.	.	.	.	.	A	G	A	.	.	T	.	.	.	.	.	.	.	.	.	.	G	T	.
11. D	G	.	.	.	.	.	.	.	A	.	C	T	.	.	.	.	.	.	.	.	.	.	G	T	.
12. H	.	.	.	.	.	G	.	.	A	.	.	T	.	.	.	.	.	.	.	.	.	.	G	T	.
13. I	.	C	.	.	.	.	.	.	A	.	.	T	A	.	.	.	.	.	.	.	.	.	G	T	.
14. L	.	.	.	C	.	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.	.	.	G	T	.
15. M	.	.	C	.	.	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.	.	.	G	T	.
16. N	.	.	.	.	A	.	.	.	A	.	.	T	.	.	.	.	.	.	C	.	.	.	G	T	.
17. T	G	.	C	.	.	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.	.	.	G	T	.

Fig. 2 Character matrix of variable nucleotides in the D-loop, ND1, and ATPase 6 mitochondrial segments (375, 355, and 494 nucleotides, respectively) of *Coregonus artedi*. Numbers refer to nucleotide positions in each segment (primer used) and correspond to the positions in sequences deposited in GenBank (Accession nos AF246932–34 for D-loop, ND1 and ATPase 6, respectively).

and has highest probabilities of confidence when differences are few. This approach contrasts with traditional phylogenetic tree bootstrapping analyses that attain higher statistical resolution when differences among compared units are many, a situation that is not likely for intraspecific comparisons of lineages of young expected ages (Templeton *et al*. 1992). For purposes of comparison with previous intraspecific molecular phylogeographic studies of other postglacial fish species, and in order to further support haplotype groupings suggested by the parsimonious mutational network, we also performed a parsimony/bootstrapping (1000 pseudoreplicates) analysis using the SEQBOOT, DNAPARS and CONSENSE modules of PHYLYP 3.5c (Felsenstein 1993).

In order to compare the geographical distributions of the two main groups of haplotypes (A and B, see Results), we first evaluated the null hypothesis of no association between genetic variation and geography. This was done by performing an X^2 test using Monte Carlo simulations (1000 iterations in the REAP package, McElroy *et al*. 1992) with the observed/expected occurrences of group B in each of the four main drainage areas of the sampled zone (Mackenzie, Hudson West, Hudson East, and Atlantic). We further compared the distribution of haplotypes by conducting a nested clade analysis (Templeton *et al*. 1995; Templeton 1998). Haplotypes (0-step clade) differing by one mutational step were grouped to form 1-step clade, and this procedure was repeated, using the 1-step clades, to form two 2-step clades. In order to evaluate the statistical significance of the observed clade distances [*Dc*, *Dn*, IT(*Dc*), and IT(*Dn*)], the expected distribution of each clade distance was established by 1000 random permutations of clade haplotypes across sites. Here, all composite sequences were used to define the nested clades, but those that could not be identified by SSCP were ignored for the calculation of clade distances because there was no information about their geographical distribution. For these analyses, the results reported by Bernatchez & Dodson (1990a) were also included (four populations with $n > 16$). We interpreted the nested clade analysis results by following the inference key of Templeton (1998) and by referring to the rationales of distance comparisons given in Templeton *et al*. (1995).

Results

Sequence relationships

SSCP analysis of 605 individuals revealed 10 conformants which corresponded to 12 H2/LN20 sequences and 17 composite sequences when ND1 and ATPase 6 partial sequences were considered (Fig. 2). The reference sequences (1.B) have been deposited in GenBank (Accession nos AF246932–34). A mutational network established on the basis of the 25 variable sites and defining the relationships among these sequences is presented in Fig. 3(A). The structure of the mutational network indicates that *Coregonus artedi* is composed of two closely related groups of haplotypes separated by three mutations, one in each of the sequenced segments.

The probability (*Pj*, Templeton *et al*. 1992) that the two groups of haplotypes differing by these three mutations have a parsimonious relationship is 0.998. The consensus parsimony cladogram presented in Fig. 3(B) clearly supported this grouping of haplotypes into clade A, which comprises several mutational derivatives of sequence 4.A,

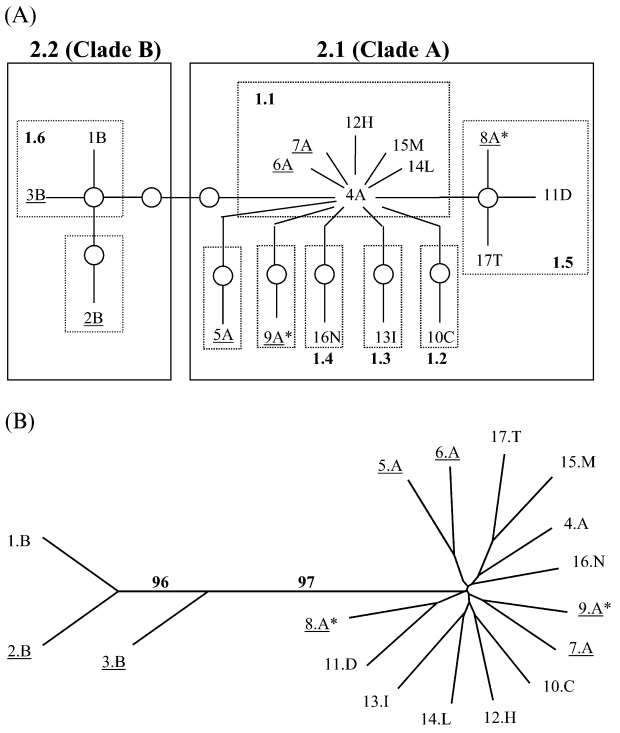

Fig. 3 (A) Haplotype network based on the mutational differences among composite mtDNA sequences of *Coregonus artedi* and showing the nesting clade design. Empty circles refer to unsampled haplotypes and sequence codes are those of Fig. 2. Haplotypes that were not detected by SSCP due to mutations located outside the D-loop segment are underlined, while * indicates D-loop mutations that were not revealed by SSCP. These haplotypes were used to establish the nested design but are not included in the nested clade analysis (see text). (B) Consensus maximum parsimony cladogram among sequences showing percent bootstrap values above 50.

and clade B, which is composed of only three haplotypes. Templeton nesting procedure resulted in the same two clades (A = 2.1, B = 2.2, Fig. 3A). The mean pairwise sequence divergence between these clades was estimated at 0.48%.

The D-loop mutation consists of a transversion at position 167 that is unambiguously discriminated by SSCP. Indeed, all B clade sequences corresponded to SSCP pattern B (167 = C) while the A clade sequences included all of the other SSCP patterns sharing in common an A at position 167 (i.e. conformants A, C, D, H, I, L-N, and T, Fig. 2 and Table 2). This D-loop mutation was also detected by Reed *et al*. (1998) and supported, with three other shared D-Loop mutations, the groupings of cisco sequences into two groups matching our clades A and B. More importantly, this mutation identified clades that perfectly corresponded to those defined by RFLP analysis of the whole mtDNA genome by Bernatchez & Dodson (1990a). Indeed, we re-analysed 36 individuals used in that study, and individuals of clades B and A defined by these authors consistently corresponded to our SSCP conformants/sequences of type

Table 3 Distribution of Clade B individuals among major drainages

Drainage	*n* sites	Total individual N (%)	B clade N_{obs} (%)	B clade N_{exp}
Mackenzie	5	120 (0.18)	0 (0.00)	35
Hudson (West)	9	211 (0.31)	62 (0.29)	62
Hudson (East)*	7	161 (0.24)	87 (0.54)	47
Atlantic	9	185 (0.27)	50 (0.27)	54
Total	30	677	199 (0.29)	

*Includes data from Bernatchez & Dodson (1990a) for samples with $n > 16$ (Eastmain R., La Grande R., Little Whale R., and Richmond Gulf).

A and B, respectively. One individual from Rupert River belonged to clade C of Bernatchez & Dodson (1990a) and clearly fell into our clade A (sequence 14.L, Figs 2 & 3). This result is consistent with the RFLP fragment patterns and sequence divergence matrix reported by these authors. Their clade C, which was composed of a few divergent haplotypes with many unique restriction sites, was most likely composed of mutational derivatives defined by undetected repeat polymorphisms now documented at the D-Loop 5′-end in *C. artedi* (Reed *et al*. 1998). Two other transitions, one in each of the ND1 and ATPase regions, further distinguished Clades A and B (Fig. 2).

Geographic distribution of C. artedi *clades*

The number of SSCP conformants per sampling site was low (1–4, mean 2.07) and corresponded to a low within population nucleotide diversity (mean $\pi = 0.0015$, range 0–0.0022). Conformants A and B were the only two that were widely distributed, A being present in all samples analysed, and B in 18 out of 27 samples (67%). All other conformants were mutational derivatives of type A and most were extremely rare in terms of frequency and distribution (Table 2); they all nested within clade 2.1 which is hereafter refer to as Clade A. There were no mutational derivatives of conformant B that could be detected by SSCP so that nested clade 2.2 will be referred to as Clade B.

Figure 1 shows the geographical distribution of *C. artedi* clades A and B, including the data of Bernatchez & Dodson (1990a). Both groups were widely distributed but contrasted in their frequency of occurrence among the four main drainage areas (Table 3, Monte Carlo X^2, $P < 0.0001$). Fish from the Mackenzie drainage were fixed for Clade A, while the proportion of Clade B individuals on the eastern side of Hudson Bay (54%) was nearly double those observed in the western Hudson Bay and Atlantic drainages (29% and 27%, respectively, Table 3). These significant differences persisted when the Mackenzie drainage was excluded from comparisons ($P = 0.012$).

0-setp (haplotype)			1-step			2-step			3-step	
No.	*Dc*	*Dn*	No.	*Dc*	*Dn*	No.	*Dc*	*Dn*	No.	*Dc*
A	1062^{L}	1062^{L}								
H	106^{S}	795								
L	0^{nc}	910								
M	0^{nc}	517								
I-T	1027^{L}	322								
1-2-3-4-9-10 No: F/IBD			**1.1**	1058^{L}	1059^{L}					
C	669	669	**1.2**	669	668					
I	0^{S}	0^{S}	**1.3**	0^{S}	1556^{L}					
N	0^{nc}	0	**1.4**	0^{S}	1556^{L}					
D	0^{nc}	908								
T	0^{nc}	987	**1.5**	947	969					
			I-T	654	−96					
			1-2-3-4-9-10 No: F/IBD			**2.1**	1067^{L}	1066^{L}		
B	700^{S}	700^{S}	**1.6**	700^{S}	700^{S}	**2.2**	700^{S}	773^{S}		
						I-T	367^{L}	293^{L}	**3.1**	983
						1-2 (inconclusive) or 1-2-3-4-9 No: Past fragmentation? (see text)				

Table 4 Results of the nested geographical analysis of *Coregonus artedi* mtDNA haplotypes detected by SSCP, following the nomenclature on Templeton (1998) inference key (F, fragmentation; IBD, isolation by distance). Superscripts indicate clade (*Dc*) and nested clade (*Dn*) distances that were significantly large (L) or small (S) at the 5% level (1000 permutations), or that could not be calculated due to single haplotype occurrence (nc). Similar information is given for average differences between interior (in grey) vs. tip clade distances (*I-T*) when tip/interior status is defined/relevant. The nested design is given in Fig. 3(A)

The nested clade analysis was only moderately informative at the 0- and 1-step levels due to the limited amount of polymorphism and the restricted distribution of several haplotypes/1-step clades (Table 4). At these levels, the similarity of most *Dc* and *Dn* values suggested short-distance movements (A, C, B, 1.1, 1.2, 1.5, 1.6), but the observed pattern of variation and the sampling scheme were inadequate to distinguish between two possible causes of gene flow restriction (fragmentation or isolation by distance). At the 2-step level, Clade A and B had significantly high and low *Dc* and *Dn* values, respectively. These results confirmed that Clade A was more widely distributed than Clade B. The interior/tip status are undefined at this level, but Clade A can arbitrarily be defined as the interior clade, since it is generally accepted that clades with larger distributions are older (Templeton *et al*. 1995). In this context, the significantly large 2-step IT-*Dc* value indicates that individuals of Clade A have dispersed farther than expected from the global centre of distribution while the dispersion of Clade B has been more restricted. If we arbitrarily accept hat the distribution range overlap between these clades is only partial, the inference key of Templeton *et al*. (1995) identifies past fragmentation as the most likely cause of gene flow restriction. Although this interpretation is subjective, it is consistent with the fact that haplotypes of each clade are connected to each other by a larger than average number of steps (5.8 vs. 1.9 for interclade and overall number of differences, respectively).

Genetic vs. phenotypic polymorphisms

There were no obvious associations between the presence or co-occurrence of the two mitochondrial groups and the presence of particular *C. artedi* eco/morphotypes in sympatry or allopatry. Sympatric morphotypes were fixed for a clade A haplotype in Barrow Lake (site 2) (Table 2), whereas both clades were equally represented in each of the sympatric morphotypes of Cormorant Lake (site 7) and La Grande Reservoir (site 15) (Table 2, Fisher exact tests: $P = 1.0$ and 0.72, respectively), and haplotypes of the B type dominated in the latter case. Allopatric populations with eco/morphotypes were more often fixed for Clade A but some also contained haplotypes of Clade B. Populations characterized by unusual size distribution [large: Great Slave L. (site 1), Utikuma L. (site 4), and Pasqua L. (site 8); dwarf: Peerless L. (site 3)] were associated only with Clade A, and the only two cases where mutational derivatives of the dominant haplotype A were abundant were populations with very large-sized individuals [Utikuma L. (site 4), and Pasqua L. (site 8)]. Spring-spawners of Lac des Écorces (site 23) were also fixed for Clade A haplotype. However, both clades were present in populations of lake cisco displaying particularly low or high total gill-raker number [White Partridge L (site 25) and Saganaga L. (site 14), respectively].

Discussion

This phylogeographic survey of *Coregonus artedi* throughout

its entire range of distribution confirmed the existence of the two main clades identified by Bernatchez & Dodson (1990a), and our estimate of divergence between these two clades (0.48%) closely matched their estimate based on mtDNA RFLP polymorphisms for populations of the James/Hudson Bay area (0.52%). The method employed to establish the global geographical distribution (SSCP) of each clade was slightly homoplasic, as two iso-conformants of clade A corresponded to sequence variants. However, this did not affect our ability to identify the D-loop mutation defining these clades, and their respective geographical distribution could be established.

The geographical distributions of these clades were extensively overlapping, but we could reject the null hypothesis of no association between genetic (haplotype) variation and geography due to the absence of Clade B from the Mackenzie drainage, and the increased occurrence of that group in the north-eastern Hudson Bay drainage. Moreover, the nested clade analysis revealed that the extent of their distribution were significantly contrasted. This geographical pattern of genetic variation cannot readily be used to identify the refugial origin(s) of *C. artedi*. Refugial areas are generally identified either by a centre of increased genetic diversity (Cann *et al*. 1987) which was not revealed in our data. Alternatively, higher frequency of occurrence of a distinct lineage in a zone can also be associated with the proximity of that zone to a refuge. Here, the increased occurrence of Clade A in the Mackenzie drainage would thus suggest its dispersal from a Beringian refuge, but the global distribution of *C. artedi* stands in sharp contrast with that of all other fish species having dispersed from Beringia (Hocutt & Wiley 1986). On the other hand, dispersal of Clade B could have proceeded from eastern Hudson Bay, but there certainly were no refugial areas in this zone (Hocutt & Wiley 1986). Alternatively, higher frequency of occurrence of a lineage in a given area can result from an earlier colonization and/or privileged access to that zone. Our results are more compatible with the latter hypothesis and challenge the current hypothesis of a single Mississippian refuge for *C. artedi*.

The hypothesis of a single Mississippi refuge is based on the fact that the extant distribution of lake cisco does not overlap with any other recognized or putative refuge, namely the Beringian and Atlantic refuges whose position flanked the north-western and south-eastern distribution limits of *C. artedi*. Individuals bearing haplotypes belonging to groups A and B may both have survived in and dispersed only from the Mississippi refuge. However, differences in dispersal success and relative demography must be invoked to account for the significant difference in geographical distribution of each haplotype group among drainages. If both clades were equally represented in the Mississippian refuge, the probability of finding only clade A haplotypes in the five sampled sites of the Mackenzie drainage as a result of random drift is only 0.03 ($P = 0.5^5$). This probability increases if Clade A outnumbered Clade B in the refuge, but it is then difficult to explain its reduced occurrence in drainages *closer* to its origin of dispersal because the western Hudson Bay drainage was necessarily on the itinerary of fish dispersing from the Mississippi towards the Mackenzie drainage. The increased occurrence of group B in eastern Hudson Bay is even more problematic and requires selective arguments to account for a putative dispersal success bias. To our knowledge, young mitochondrial clades have never been associated with selective differences in any other organisms, and particularly not in fishes.

Alternatively, differences in distribution of each clade could have resulted from survival in and dispersal from distinct refuges. Here, we contend that the level of divergence between these clades, their contrasted geographical distribution, as well as the global distribution of *C. artedi* in NA together support the Mississippian and Atlantic origins of Clade A and B, respectively. First, the level of divergence between *C. artedi* clades corresponds to the level of divergence between refugial races in other nearctic fish species with distributional ranges mostly restricted to formerly glaciated areas (reviewed in Bernatchez & Wilson 1998). Moreover, the estimated level of divergence between Clades A and B was lower than what has been documented for comparisons involving Beringian refugial races in *C. clupeaformis* (1.15%, Bernatchez & Dodson 1991), *Salvelinus alpinus* (1.32%, Wilson *et al*. 1996), and *S. namaycush* (0.88%, Wilson & Hebert 1998) and resembled the level of divergence among eastern refugial groups of Mississippian, Atlantic, or Acadian origins in *Stizostedion vitreum* (0.47%, Billington & Hebert 1988), *C. clupeaformis* (0.40%, Bernatchez & Dodson 1991), *Osmerus mordax* (0.7%, Baby *et al*. 1991; 0.78%, Taylor & Bentzen 1993), and *S. namaycush* (0.42%, Wilson & Hebert 1998).

Second, contrasts in the distribution of each clade are compatible with different times and routes of postglacial dispersal into interior NA freshwater habitats. On one hand, differences in time of dispersal are suggested by the fact that only individuals of Clade A were documented in the Mackenzie drainage. The presence of Clade A in the upper Mackenzie watershed indicates that dispersal into this area may have occurred while Lake Agassiz offered a connection between the Mississippian and Mackenzie watersheds. This opportunity may have been present around 11 kya BP (McPhail & Lindsey 1970; Lindsey & McPhail 1986) but most likely allowed effective fish dispersal between 9.9 and 9.5 kya BP when proglacial Lake Agassiz discharged into Lake McConnell via the Clearwater River (Dyke & Prest 1987; Fisher & Smith 1994; Smith 1994; Rempel & Smith 1998). Conversely, the absence of Clade B from the Mackenzie watershed and its more restricted distribution both suggest that it arrived in the western upper reaches of extant Hudson drainage after this connection

Table 5 Distribution and proposed refugial origin of freshwater fish species present on the western and/or eastern sides of Hudson Bay. (Adapted from Tables 3.1 and 3.2 of Crossman & McAllister 1986)

	Distribution near Hudson Bay		Proposed refgium*				
Species	West	East	Mp	Ms	B	A	W
Ichthyomyzon unicuspis	X		+				
Acipenser fulvescens	X		+	?			
Hiodon alosoides	X		+	+			
Thymallus arcticus	X		?		+		
Notropis atherinoides	X		+	+			
N. heterolepsis	X		+				
N. hudsonius	X		+	?			
N. volucellus	X		+				
Phoxinus eos	X		+	+			
P. neogaeus	X		+	+			
Pimephales promelas	X		+	+			
Semotilus margarita	X		+	+			
Moxostoma macrolepidotum	X		+	?			
Percopsis omiscomaycus	X		+	?	?		
Culaea inconstans	X		+	+			
Etheostoma exile	X		+	?			
E. nigrum	X		+				
Perca flavescens	X		+			+	
Percina caprodes	X		+				
P. shumardi	X		+				
Stizostedion canadense	X		+	+			
Aplodinotus grunniens	X		+	+			
Cottus ricei	X		+	?			
***Coregonus artedi* (this study)**	**X**	**X**	⊕			○	
Coregonus clupedaformis†	X	X	⊕		⊕	⊕	
Prosopium cylindraceum	X	X	+		+		
Salmo salar		X				+	
Salvelinus alpinus‡	X	X			⊕	⊕	
Salvelinus fontinalis§	X	X	○			⊕	
Salvelinus namaycush¶	X	X	⊕	○	⊕	⊘	
Esox lucius	X	X	+		+		
Couesius plumbeus	X	X	+	+	+	?	
Rhinichthys cataractae	X	X	+	+		+	
Catostomus catostomus	X	X	+	+	+		
Semotilus atromaculatus		X	+			?	
C. commersoni	X	X	+	+		?	
Lota lota	X	X	+	+	+		
*Gasterosteus aculeatus***	X	X				⊕	+
Pungitius pungitius	X	X	+		+	?	
Stizostedion vitreum	X	X	+	+		+	
Cottus bairdi	X	X	+	+		?	?
C. cognatus	X	X		+	+		?
Myoxocephalus quadricornis	X	X				+	

*Mp, Mississippi; Ms, Misouri; B, Beringia; A, Atlantic; W, Western. +/?, inferred/putative from distribution data; ○, molecular data. †Bernatchez & Dodson (1991); ‡Wilson *et al*. (1996); §Danzmann *et al*. (1998); ¶Wilson & Hebert (1998); **J.J. Dodson, Laval University, unpublished data.

was severed. On the other hand, differences in dispersal routes are suggested by the ubiquity of Clade A and the polarized distribution of Clade B. The presence of Clade A over the entire range indicates that it has taken advantage of the extensive dispersal opportunities offered by the interconnected proglacial lakes that were formed along the margin of the Laurentian ice sheet during the initial phase of retreat (12–8 kya BP, Dyke & Prest 1987). In contrast, the

distribution of Clade B suggests that it took advantage of preferential dispersal avenues toward eastern Hudson Bay but that it also expanded its distribution via the proglacial lakes (see below).

Third, the global pattern of distribution of *C. artedi* is not compatible with that of fish species having dispersed only from the Mississippian (and Missourian) refuge. A close comparison of species distribution and inferred postglacial history of fish currently present around Hudson Bay reveals this discrepancy (Table 5). It is first worth noting that the number of freshwater species present on the eastern side of Hudson Bay is remarkably lower than on the western side (20 species in Quebec; 31 and 39 in Manitoba and Ontario, respectively, Crossman & McAllister 1986). All of the fish species limited to the western side of Hudson Bay share a Mississippian origin (along with a Missourian origin when their range greatly extends south-westwardly). More importantly, their distribution rarely suggests an Atlantic or Beringian refuge. The only exceptions to this pattern are *Thymallus arcticus*, whose distribution clearly indicates dispersal from Beringia alone, *Percopsis omiscomaycus*, whose distribution may also imply a Beringian origin, and *Perca flavescens*. The proposed postglacial history of fish species present on the eastern side of Hudson Bay is in sharp contrast. Indeed, 19 of these 20 species exhibit patterns of distribution that are suggestive of an Atlantic and/or Beringian refuge. The only exception to this pattern is *C. artedi*: unlike any other species presumably of solely Mississippian origin, lake cisco is widely distributed on the eastern side of Hudson Bay, and *C. artedi* is the only species of that region for which an Atlantic or Beringian refuge has not been suspected (but see Mandrak & Crossman 1992). The existence of *C. artedi* having colonized the eastern side of Hudson Bay from another refuge is highly compatible with the increased occurrence of Clade B documented in this area by the present study (Table 3, Hudson East vs. Hudson West: $P = 0.006$).

Several refugial origins are possible for eastern Hudson Bay fishes. Molecular phylogeography of *S. alpinus* revealed that fish of Beringian origin could disperse via the Arctic Ocean to eastern Hudson Bay (Belcher Is.) (Wilson *et al*. 1996). However, the absence of fossils and extant populations of *C. artedi* in Beringia argues against the existence of this refuge for lake cisco, and the decreasing westward occurrence of Clade B is incompatible with its dispersal from a western refuge with subsequent extinction. A Mississippian contribution such as that documented for *C. clupeaformis* (Bernatchez & Dodson 1991) cannot be excluded, but it can hardly explain the pattern of increased frequency of clade B away from that refuge. The Atlantic origin of Clade B appears more plausible. Previous molecular data have revealed a contribution of the Atlantic refuge to Hudson/Ungava Bay for *C. clupeaformis* (Bernatchez & Dodson 1991) and *S. fontinalis* (Danzmann *et al*. 1998), and, as stated above, the distribution of many fish species present on eastern Hudson Bay suggests dispersal from the Atlantic region (Crossman & McAllister 1986; Morin & Dodson 1986).

Several routes were available for dispersal from an Atlantic refuge to eastern James/Hudson Bay. The main dispersal avenue was via Lake Barlow-Ojibway (10.1–8.0 kya BP), either when it abruptly discharged into the Tyrell Sea at approximately 8.0 or 8.5 kya BP (Veillette 1994; Barber *et al*. 1999, respectively), or via its connection with a network of inland waterways located in the Quebec peninsula west of the Otish Mountains (Legendre & Legendre 1984). Lake Barlow-Ojibway was first accessible from the Atlantic refuge via its Ottawa River outlet and the Hudson River/Lake Vermont connection during the Champlain Sea inundation (11.8–9.7 kya BP). Fossils of *C. artedi* have been found in deposits associated with the Champlain Sea (McAllister *et al*. 1988), supporting the dispersal of this species from an Atlantic refugium (Mandrak & Crossman 1992). Access to Lake Barlow-Ojibway was also possible later when Lake Algonquin drained into the Ottawa River (10.5–9.5 kya BP) via the Fossmill-Petawawa outlet (Underhill 1986; Mandrak & Crossman 1992). This dispersal pathway was available from the Atlantic refuge following dispersal into the Great Lakes via the Susquehanna/Mohawk/Hudson outlets, as well as from the Mississippi refuge, and may have contributed to the intermixing of fish of proposed Mississippian (Clade A) and Atlantic (Clade B) origins in the Great Lakes and Ottawa River watersheds. However, the dispersal via the Champlain Sea provided an early access to Lake Barlow-Ojibway, and may have favoured the earlier establishment of an Atlantic race in eastern Hudson Bay area. The anadromous life cycle of Hudson/James Bay lake cisco populations indicates a tolerance to saline water that is compatible with facilitated dispersal via the Champlain Sea.

In conclusion, our results clearly identify the presence of two distinct lineages of *C. artedi* that display significantly different and polarized distributions within the species range. These results cast doubt on the current hypothesis of a single Mississippian refuge and are better compatible with the existence of a Mississippian as well as an Atlantic race. Following survival in these distinct refugial areas, each race initially took advantage of privileged dispersal routes, and subsequently met with the other race over an extensive area of secondary contact. However, the latter hypothesis necessarily implies an historical range fragmentation that our nested clade analysis could not objectively support. Signal of past fragmentation can only persist (be detected) if gene flow between allopatric isolates has been restricted (biologically or geographically) when they were brought back into contact following range expansion events (Templeton *et al*. 1995). Therefore, this method cannot distinguish between primary or secondary contacts in

cases of extensive distribution overlap of two clades as was observed in *C. artedi*. Finally, the absence of extant lake cisco populations in the maritime provinces of Canada and most of New England is counterintuitive with the recognition of an Atlantic race for that species. Zoogeographic regions have been legitimately delineated with respect to the concordance among *current* distributional limits of several species, but absence of extant distributional evidence is not necessarily evidence of previous absence. One possibility is that the postglacial climate change have made the more southern habitats unsuitable for lake cisco (Mandrak & Crossman 1992). As for most coregonids, *C. artedi* is clearly a cold water species. However, the spawning season of lake cisco is delayed relative to that of whitefish when they co-occur, suggesting that their requirement for very cold water may be more pronounced. Lake cisco *may* have once been present in the area associated with an Atlantic refuge, but ecological explanations for its current absence can only rely on indirect evidence and remain speculative. Stronger evidence for distinct dispersal origins and colonization patterns may be obtained by looking at nuclear and more variable markers suitable for genetic analysis using the coalescence theory. This approach uses the historical (age) information contained in gene frequencies and has the power of identifying the direction of colonization patterns in young nonequilibrium systems (e.g. Slatkin 1993).

Contact between refugial races and origin of polymorphism in C. artedi

The distribution of proposed refugial groups A and B is very extensive, and contact among these lineages are proposed over the entire Atlantic and Hudson drainages. Secondary contacts among lineages do not appear to be a necessary condition for the evolution of divergent polymorphism in *C. artedi*. Indeed, the sympatric occurrence of both mtDNA lineages is not associated with that of any specific eco/morphotypes examined in this study. Moreover, divergent morphotypes are observed in areas where a single race prevails. Thus, this situation may seem to contrast with what has been observed in the closely related *C. clupeaformis*, where ecotypes are reported mostly from areas corresponding to contact zones among refugial races (Bernatchez & Dodson 1991; Bernatchez *et al*. 1996; Pigeon *et al*. 1997; but see Vuorinen *et al*. 1993). However, it is worth mentioning that eco/morphotypes are reported from all parts of *C. artedi* range (Clarke 1973; Scott & Crossman 1973), and that our documenting of a very wide contact zone between refugial lineages is in agreement with other indications that secondary contacts can favour the emergence of phenotypic diversity. Nevertheless, such contacts appear neither necessary nor sufficient, and suitable environmental conditions for the coexistence of sympatric forms are certainly required, as exemplified by the co-occurrence of sympatric morphotypes belonging to a single clade as well as that of distinct lineages without any phenotypic divergence in both of these coregonid complexes.

Comparative phylogeography of North American postglacial fishes

The global phylogeographic structure of *C. artedi* revealed by this study conforms to what has been documented for the regional NA freshwater fish fauna of formerly glaciated areas (Bernatchez & Wilson 1998). As expected, it is composed of young clades (0.5% sequence divergence), each dominated by a single and widely dispersed haplotype. Comparisons with other postglacial fish phylogeographic patterns reveal both similarities and contrasts. Similarly to lake whitefish and lake trout, dispersal was most extensive in areas comprised within the maximum extent of proglacial lakes. Relative to these species, the postglacial dispersal of *C. artedi* has apparently been even more closely linked to the existence of the proglacial lakes, as indicated by its western limit of distribution which does not extend far beyond the maximal western extent of Lakes Agassiz, McConnell and Peace. In contrast to what has been found for lake whitefish, *C. artedi* presents a pattern of almost total overlap in refugial group distribution, whereas each of the *C. clupeaformis* glacial races display ranges that are largely distinct (Bernatchez & Dodson 1991). Unlike for lake whitefish, the proposed Atlantic race of *C. artedi* dispersed well into the Great Lakes basin and in this regard is more similar to lake trout (Wilson & Hebert 1998). The planktivorous nature of lake cisco, as well as its cold-water preference may have better enabled it to closely follow the retreating ice margin in this area, and in so doing allowed it to better take advantage of the early westward dispersal opportunities. Finally, tolerance to low salt concentrations appears to have allowed a dispersal pattern similar to species such as *S. fontinalis* via the Champlain Sea into the Ottawa River basin for an early colonization of eastern Hudson Bay. In this regard, it is particularly interesting to note that only euryhaline species of Atlantic origin have extended their range far into the Great Lakes watershed and beyond the Ottawa River system into the James/Hudson Bay area (Underhill 1986). Altogether, the vast distribution of each clade once again confirms the necessity of considering historical conditions and species potential to cope with them in order to interpret the current distributional patterns of north temperate aquatic fauna.

Acknowledgements

This work was supported by a research grant from the Natural Sciences and Engineering Research Council (NSERC, Canada) to LB and postgraduate scholarships from NSERC and the Fonds pour

la Formation des Chercheurs et l'Aide à la Recherche (FCAR, Québec) to JT. We sincerely thank all those who very kindly provided samples of lake cisco for this study. In Canada, they were, in Quebec: M. Legault, M. Hénault, and B. Hardy (Société Fauneet Parcs, FAPAQ), and P. May (Makivik); in Ontario: R. Salmon, L. Mohr, J. Casselman, W. McCallum, T. Mosindy, B. Allan, and F. Hicks (Ontario Ministry of Natural Resources); in Manitoba: I. Hagenson, D. Leroux, and K. Campbell (Manitoba Natural Resources), and R. Baker (North/South Consultants); in Saskatchewan: M. Walker and G. Mitchell (Saskatchewan Environmental Resources Management); in Alberta: D. Brown (Alberta Environmental Protection), M. Steinhilber (University of Alberta); in NWT: G. Low (Department of Fisheries and Oceans). In the USA, samples were obtained from T. Todd, and G. Fleisher (National Biological Service); B. Chipman (U.S. Fish and Wildlife); R. Phillips and K. Reed (University of Milwaukee); D. A. Etnier (University of Tennessee). An earlier version of this manuscript was substantially improved by the constructive comments of Eric Taylor and an anonymous reviewer. JT also sincerely thanks Lucie Papillon for her valuable laboratory assistance for methodological optimization and J.E. Stacy for the basic Excel sheets first used to calculate clade distances. Contribution to the research program of GIROQ.

References

Angers B, Bernatchez L (1998) Combined use of SMM and non-SMM method to infer fine structure and evolutionary history of closely related brook charr (*Salvelinus fontinalis*, Salmonidae) populations from microsatellites. *Molecular Biology and Evolution*, **15**, 143–159.

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York.

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

Barber D, Dyke A, Hillaire-Marcel C *et al.* (1999) Forcing of the cold event of 8,200 years ago by catastrophic drainage of Laurentide lakes. *Nature*, **400**, 344–348.

Bernatchez L (1997) Mitochondrial DNA analysis confirms the existence of two glacial races of rainbow smelt *Osmerus mordax* and their reproductive isolation in the St Lawrence River estuary (Québec, Canada). *Molecular Ecology*, **6**, 73–83.

Bernatchez L, Danzmann RG (1993) Congruence in control-region sequence and restriction-site variation in mitochondrial DNA of brook charr (*Salvelinus fontinalis* Mitchill). *Molecular Biology and Evolution*, **10**, 1002–1014.

Bernatchez L, Dodson JJ (1990a) Mitochondrial DNA variation among anadromous populations of cisco (*Coregonus artedii*) as revealed by restriction analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, **47**, 533–543.

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

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

Bernatchez L, Dodson JJ (1994) Phylogenetic relationships among Palearctic and Nearctic whitefish (*Coregonus* sp.) populations as revealed by mitochondrial DNA variation. *Canadian Journal of Fisheries and Aquatic Sciences*, **51**, 240–251.

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

Bernatchez L, Wilson CC (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*, **7**, 431–452.

Billington N, Barrette RJ, Hebert PDN (1992) Management implications of mitochondrial DNA variation in walleye stocks. *North American Journal of Fisheries Management*, **12**, 276–284.

Billington N, Hebert PDN (1988) Mitochondrial DNA variation in Great Lakes walleye (*Stizostedion vitreum*) populations. *Canadian Journal of Fisheries and Aquatic Sciences*, **45**, 643–654.

Bodaly RA, Vuorinen J, Ward RD, Luczynski M, Reist JD (1991) Genetic comparisons of New and Old World coregonid fishes. *Journal of Fish Biology*, **38**, 37–51.

Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. *Nature*, **325**, 31–36.

Clarke RM (1973) *The systematics of ciscoes (Coregonidae) in central Canada*. PhD Thesis. University of Manitoba, Winnipeg, Canada.

Crossman EJ, McAllister DE (1986) Zoogeography of freshwater fishes of the Hudson Bay drainage, Ungava Bay and the Arctic Archipelago. In: *Zoogeography of North American Freshwater Fishes* (eds Hocutt CH, Wiley EO), pp. 53–104. John Wiley and Sons, New York.

Danzmann RG, Morgan RPI, Jones MW, Bernatchez L, Ihssen PE (1998) A major sextet of mitochondrial DNA phylogenetic assemblages extant in eastern North American brook charr (*Salvelinus fontinalis*): distribution and post-glacial dispersal patterns. *Canadian Journal of Zoology*, **76**, 1300–1318.

Doiron S (2000) *Analyse comparative de la séquence complète du génome mitochondrial chez l'Omble de fontaine* (Salvelinus fontinalis) *et l'Omble chevalier* (S. alpinus). MSc Thesis. Université Laval, Québec, Canada.

Dyke AS, Prest VK (1987) Late Wisconsinan and Holocene history of the Laurentide Ice sheet. *Géographie Physique et Quaternaire*, **41**, 237–263.

Felsenstein J (1993) *PHYLIP: Phylogeny inference package, 3.5c*. University of Washington, Seattle, WA.

Fisher TG, Smith DG (1994) Glacial Lake Agassiz: its northwestern maximum extent and outlet in Saskatchewan (Emerson Phase). *Quaternary Science Reviews*, **13**, 845–858.

Giuffra E, Bernatchez L, Guyomard R (1994) Mitochondrial control region and protein coding genes sequence variation among phenotypic forms of brown trout *Salmo trutta* from northern Italy. *Molecular Ecology*, **3**, 161–1171.

Grewe PM, Hebert PDN (1988) Mitochondrial DNA diversity among broodstocks of the lake trout, *Salvelinus namaycush*. *Canadian Journal of Fisheries and Aquatic Sciences*, **45**, 2114–2122.

Hénault M, Fortin R (1989) Comparison of meristic and morphometric characters among spring- and fall-spawning ecotypes of cisco (*Coregonus artedii*) in southern Quebec, Canada. *Canadian Journal of Fisheries and Aquatic Sciences*, **46**, 166–173.

Hewitt GM (1996) Some genetic consequence of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hocutt CH, Jenkins RE, Stauffer JR (1986) Zoogeography of fishes of the central Appalachians and central Atlantic Coastal Plain.

In: *Zoogeography of North American Freshwater Fishes* (eds Hocutt CH, Wiley EO), pp. 161–212. John Wiley and Sons, New York.

Hocutt CH, Wiley EO, eds (1986) *The Zoogeography of North American Freshwater Fishes*. John Wiley and Sons, New York.

Kumar S, Tamura K, Nei M (1993) *MEGA: Molecular Evolutionary Genetics Analysis, 1.01*. The Pennsylvania State University, PA.

Lee DS, Gilbert CR, Hocutt CH *et al*. (1980) *Atlas of North American freshwater fishes*. North Carolina Biological Survey, Publishers no. 1980–12. Raleigh, NC.

Legendre P, Legendre V (1984) Postglacial dispersal of freshwater fishes in the Québec peninsula. *Canadian Journal of Fisheries and Aquatic Sciences*, **41**, 1781–1802.

Lindsey CC, McPhail JD (1986) Zoogeography of fishes of the Yukon and Mackenzie Basins. In: *The Zoogeography of North American Freshwater Fishes* (eds Hocutt CH, Wiley EO), pp. 639–674. John Wiley and Sons, New York.

Mandrak NE, Crossman EJ (1992) Postglacial dispersal of freshwater fishes into Ontario. *Canadian Journal of Zoology*, **70**, 2247–2259.

McAllister DE, Harington CR, Cumbaa SL, Renaud CB (1988) Paleoenvironmental and biogeographic analyses of fossil fishes in peri-Champlain Sea deposits in eastern Canada. *Geological Association of Canada Special Paper*, **35**, 241–258.

McAllister DE, Platania SP, Schueler FW, Baldwin ME, Lee DS (1986) Ichthyofaunal patterns on a geographic grid. In: *Zoogeography of North American Freshwater Fishes* (eds Hocutt CH, Wiley EO), pp. 17–52. John Wiley and Sons, New York.

McElroy DP, Moran P, Bermingham E, Kornfield I (1992) REAP: an integrated environment for the manipulation and phylogenetic analysis of restriction data. *Journal of Heredity*, **83**, 157–158.

McPhail JD, Lindsey CC (1970) Freshwater Fishes of Northwestern Canada and Alaska. *Bulletin of the Fisheries Research Board of Canada*, no. 173.

Morin R, Dodson JJ (1986) The ecology of fishes in James Bay, Hudson Bay and Hudson Strait. In: *Canadian Inland Seas* (ed. Martini IP), pp. 293–325. Elsevier, Amsterdam.

Orita M, Iwahana H, Kanazawa H, Sckiya T (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using polymerase chain reaction. *Genomics*, **5**, 874–879.

Pariseau R, Dumont P, Migneault J-G (1983) Découverte, dans le sud-ouest du Québec, d'une population de cisco de lac, *Coregonus artedii*, frayant au printemps. *Canadian Journal of Zoology*, **61**, 2365–2368.

Pielou EC (1991) *After the Ice Age: the Return of Life to Glaciated North America*. University of Chicago Press, Chicago.

Pigeon D, Chouinard A, Bernatchez L (1997) Multiple modes of speciation involved in the parallel evolution of sympatric dwarf and normal morphotypes of lake whitefish (*Coregonus clupeaformis*, Salmonidae). *Evolution*, **51**, 196–205.

Reed KM, Dorschener TN, Phillips RB (1998) Sequence analysis of the mitochondrial DNA control region of ciscoes (genus: *Coregonus*): taxonomic implications for the Great Lakes species flock. *Molecular Ecology*, **7**, 1091–1096.

Rempel LL, Smith DG (1998) Postglacial fish dispersal from the Mississippi refuge to Mackenzie River basin. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 893–899.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, New York.

Schmidt RE (1986) Zoogeography of the Northern Appalachians. In: *Zoogeography of North American Freshwater Fishes* (eds Hocutt CH, Wiley OE), pp. 137–159. John Wiley and Sons, New York.

Scott WB, Crossman EJ (1973) Freshwater Fishes of Canada. *Bulletin of the Fisheries Research Board of Canada*, **184**, 1–1092.

Sheffield VC, Beck JS, Kwitek AE, Sanstrom DW, Stone EM (1993) The sensitivity of Single-Strand Conformation Polymorphism analysis for the detection of single base substitutions. *Genomics*, **16**, 325–332.

Slatkin M (1993) Isolation by distance equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.

Smith DG (1994) Glacial Lake McConnell: paleogeography, age, duration, and associated river deltas, MacKenzie River Basin, Western Canada. *Quaternary Science Reviews*, **13**, 829–844.

Smith GR, Todd TN (1984) Evolution of species flocks of fishes in north temperate lakes. In: *Evolution of Fish Species Flocks* (eds Echelle AA, Kornfield I), pp. 45–68. University of Maine at Orono Press, Orono.

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

Taylor EB, McPhail JD (1999) Evolutionary history of an adaptive radiation in species pairs of threespine sticklebacks (*Gasterosteus*): insights from mitochondrial DNA. *Biological Journal of the Linnean Society*, **66**, 271–291.

Taylor E, Pollard S, Louie D (1999) Mitochondrial DNA variation in bull trout (*Salvelinus confluentus*) from northwestern North America: implications for zoogeography and conservation. *Molecular Ecology*, **8**, 1155–1170.

Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381–397.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.

Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics*, **140**, 767–782.

Turgeon J, Estoup A, Bernatchez L (1999) Species flock in the North American Great Lakes: molecular ecology of Lake Nipigon ciscoes (Teleostei: Coregonidae: *Coregonus*). *Evolution*, **53**, 1857–1871.

Turner TF, Trexler JC, Kuhn DN, Robinson HW (1996) Life-history variation and comparative phylogeography of darters (Pisces: Percidae) from the North American Central Highlands. *Evolution*, **50**, 2023–2036.

Underhill JC (1986) The fish fauna of the Laurentian Great Lakes, the St. Lawrence lowlands, Newfoundland and Labrador. In: *Zoogeography of North American Freshwater Fishes* (eds Hocutt CH, Wiley EO), pp. 105–136. John Wiley and Sons, New York.

Veillette JJ (1994) Evolution and paleohydrology of glacial Lakes Barlow and Ojibway. *Quaternary Science Reviews*, **13**, 945–971.

Vuorinen JA, Bodaly RA, Reist JD, Bernatchez L, Dodson JJ (1993) Genetic and morphological differentiation between dwarf and normal size forms of lake whitefish (*Coregonus clupeaformis*) in Como Lake, Ontario. *Canadian Journal of Fisheries and Aquatic Sciences*, **50**, 210–216.

Wilson CC, Hebert PDN (1998) Phylogeography and postglacial dispersal of lake trout (*Salvelinus namaycush*) in North America. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 1010–1024.

Wilson CC, Hebert PDN, Reist JD, Dempson JB (1996) Phylogeography and postglacial dispersal of arctic charr *Salvelinus alpinus* in North America. *Molecular Ecology*, **5**, 187–197.

Zink RM (1996) Comparative phylogeography in North American birds. *Evolution*, **50**, 308–317.

This paper is part of JT's doctoral research on the evolutionary biology of ciscoes in North America. She is currently conducting postdoctoral research on the genetic signature of variable demography during rapid colonization process. LB's research programme focuses on the understanding of patterns and processes of molecular and organismal evolution, and its relevance to conservation.