Molecular Ecology (2005) **14**, 1497–1509 doi: 10.1111/j.1365-294X.2005.02523.x

Allopatric origins of sympatric brook charr populations: colonization history and admixture

D. J. FRASER and L. BERNATCHEZ
Québec Océan, Département de Biologie, Université Laval, Sainte-Foy, Québec G1K 7P4

Abstract

Natural selection is presumed to be the driving force behind the occurrence of phenotypically and genetically divergent populations in sympatry within many north temperate freshwater fishes. If, however, these populations have different ancestral origins, history could also contribute to their divergence. We previously found evidence for the role of selection in the evolution of divergent outflow and inflow breeding populations of migratory brook charr (*Salvelinus fontinalis*) inhabiting postglacial Mistassini Lake (Québec, Canada). Here, we show that these populations do not have a common origin, through the use of admixture and spatial analyses with seven microsatellite loci. Divergent populations clustered into two different population groups when compared to samples from surrounding drainages, although inflow populations appeared to be more admixed between the two population groups than the outflow population. These results are noteworthy since outflow and inflow populations were monomorphic at mitochondrial DNA (338-bp sequence of the control region) and are only moderately differentiated (mean F_{ST} = 0.10). Colonization by two ancestral populations was also consistent with known outflow direction changes throughout lake formation. In addition to providing insight into how phenotypic divergence in sympatry may have been affected by the nature (i.e. timing and direction) of colonization of ancestral populations, our results also suggest that ancestral populations may have differed in their ability to colonize certain lake habitats.

Keywords: evolutionary history, mitochondrial DNA, postglacial dispersal and colonization, secondary contact, statistical phylogeography, sympatric divergence

Received 11 November 2004; revision accepted 26 January 2005

Introduction

Understanding how natural selection and history may interact to lead to evolutionary diversification is a fundamental goal of evolutionary biology (Travisano *et al*. 1995; Losos *et al*. 1998; Taylor & McPhail 2000; Langerhans & DeWitt 2004). The intraspecific population divergence observed within many freshwater fish species of northern temperate regions is useful for examining such interactions. On one hand, many populations within separate lakes exhibit parallel phenotypic and genetic differences, indicating an important role of divergent natural selection in their diversification since colonization of newly formed postglacial lake niches (7000–18 000 BP: Taylor & Bentzen 1993; Robinson & Wilson 1994; Schluter 1996; Taylor *et al*. 1996; Pigeon *et al*. 1997; Gislason *et al*. 1999; Lu & Bernatchez 1999; Taylor 1999; Rundle *et al*. 2000; Rogers *et al*. 2002; Saint-Laurent *et al*. 2003; Behrmann-Godel *et al*. 2004; Østbye *et al*. 2005; Rogers & Bernatchez 2005). On the other hand, dynamic changes in watershed connections during glacial advances and retreats also readily opened the way for multiple population lineages to come into secondary contact (Svärdson 1961; Bernatchez & Dodson 1990; Hynes *et al*. 1996; Taylor 1999; Taylor & McPhail 2000; Turgeon & Bernatchez 2001). Consequently, population differentiation may not be the sole product of divergent natural selection in sympatry. It may instead be partly caused by unique histories, presumably stemming from features of specific lineages and chance events (Harvey & Pagel 1991; McPhail 1993; Taylor & McPhail 2000).

Determining the ancestral origin of sympatrically occurring populations within individual lakes is therefore a critical first step to assess whether history could have played

Correspondence: Louis Bernatchez, Fax: 1-418-656-2043; E-mail: louis.bernatchez@bio.ulaval.ca

a role in their divergence. Indeed, history is less likely to be important in differentiation if divergent populations have a common origin. But if divergent populations have different origins, it is still difficult to know whether history has actually contributed to their divergence. Generally, obtaining clear evolutionary relationships of divergent populations with molecular data has been challenging. First, contemporary gene flow may erode pre-existing genetic differences between allopatrically derived populations within lakes. This can create the impression that divergent populations are more closely related than populations outside each lake (Taylor 1999; Lu *et al*. 2001; Hendry *et al*. 2002). Second, the detection of multiple colonization events within some lakes may be difficult given the potentially shallow divergence of colonizing lineages (Taylor & McPhail 2000; Tessier & Bernatchez 2000).

Three genetically distinct populations of brook charr (*Salvelinus fontinalis*) inhabit Mistassini Lake (Québec, Canada; Fig. 1). These populations are characterized by seasonal migrations between their feeding areas in the lake and breeding areas in rivers entering the lake (as well as within feeding areas) (Fraser *et al*. 2004; Fraser & Bernatchez 2005). Genetic differentiation is especially pronounced between an outflow-breeding population (Rupert) and two northeast inflow-breeding populations (Cheno and Pepeshquasati) (F_{ST} = 0.10; between inflow populations: F_{ST} = 0.02; Fraser *et al*. 2004). Relative to outflow charr, inflow charr have longer migrations between breeding and feeding areas (as well as within feeding areas), and marked spatial overlap within feeding areas. Inflow charr also have more fusiform bodies with longer posterior regions than outflow charr (Fraser & Bernatchez 2005). It appears that these differences have evolved in response to divergent natural selection in the environments that outflow and inflow charr utilize. Specifically, morphological characteristics of inflow charr are better suited for sustained swimming (Taylor & Foote 1991) needed for long migrations. Spatial segregation in feeding areas between outflow and inflow populations is also related to differential habitat selection and presumably reflects adaptive divergence in alimentation and/or predation regimes (Fraser & Bernatchez 2005). Outflow and inflow populations also differ in other traits related to divergent breeding environments, including breeding migration timing, breeding time, and age at maturity (Fraser *et al*. 2004; Fraser & Bernatchez 2005).

In a species-wide phylogeographical survey of mitochondrial DNA (mtDNA), Danzmann *et al*. (1998) reported that much of the northern distribution of *S. fontinalis* (including Ontario, Québec and eastern Canada) was colonized by fish sharing the same mtDNA haplotype '1' of assemblage 'B' (one of six assemblages). The authors provided evidence that haplotype 1 was common within two distinct glacial refugia (Atlantic and Mississippian), and based on the likely postglacial dispersal of *S. fontinalis* from these refugia, they proposed a possible contact zone

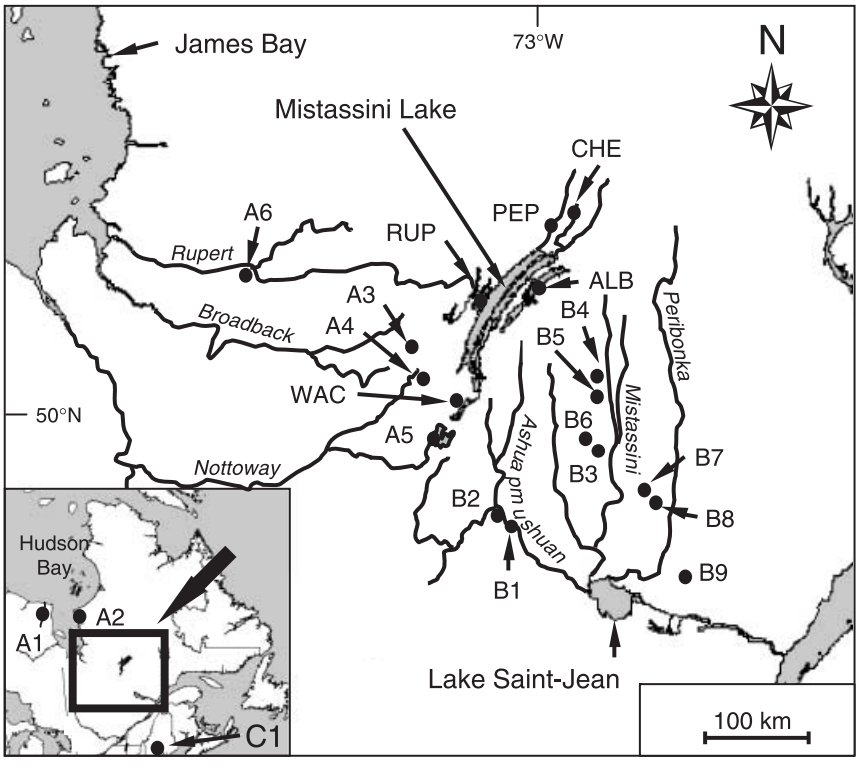

Fig. 1 Geographic locations of Mistassini Lake and *Salvelinus fontinalis* populations sampled from surrounding drainages between the Hudson/James Bay (A1–A6)–Lake Saint-Jean (B1–B9) region of central Québec, Canada, and Maine, USA (C1). Mistassini Lake populations: Rupert (RUP), Cheno (CHE), Pepeshquasati (PEP); Lake Waconichi, WAC; Lake Albanel, ALB.

in Ontario or Québec. Colonization of Québec from multiple refugia has also been suggested in other fish species (Legendre & Legendre 1984; Wilson & Hebert 1998; Turgeon & Bernatchez 2001). It is therefore noteworthy that striking directional changes in Mistassini Lake's outflow occurred from the south to west during postglacial formation (7000–8000 BP: Fig. 2; Bouchard 1980). This may have permitted *S. fontinalis* to colonize the lake from at least two routes, either by fish originating from two distinct glacial refugia possessing haplotype 1, or by different ancestral populations from within either refugium.

The objective of this study was thus to determine whether outflow and inflow *S. fontinalis* populations in Mistassini Lake have a common origin or different origins. As expected based on previous studies, our initial screening of mtDNA (n = 75, 338 bp of the most variable segment of the control region, using single strand conformation polymorphism analysis; Sheffield *et al*. 1993) and verification with sequencing (n = 8, with at least two samples from each Mistassini Lake population) confirmed that divergent populations were monomorphic for haplotype 1 of assemblage B designated in Danzmann *et al*. (1998) (data not shown). Because of this limitation of mtDNA, we employed microsatellite DNA markers for discerning historical relationships of divergent populations. Indeed, recent studies have shown them to be effective for resolving phylogeographical structuring at potentially shallower evolutionary timescales or situations where gene flow still occurs between sympatric populations (e.g. Angers & Bernatchez 1998; Taylor & McPhail 2000; Lu *et al*. 2001; Koskinen *et al*. 2002). Alternative hypotheses of a common origin vs. different origins of divergent populations were principally tested using admixture analyses by clustering Mistassini Lake populations into population groups defined a posteriori from possible postglacial colonization routes surrounding the lake.

Materials and methods

Sampling scheme and microsatellite DNA polymorphism

Mistassini Lake is drained via the Rupert River into James Bay, Canada (Fig. 1). Genetic diversity of *Salvelinus fontinalis* in Mistassini Lake populations (Cheno, CHE; Pepeshquasati, PEP; Rupert, RUP) and two nearby lakes that empty into Mistassini Lake (Waconichi, WAC; Albanel, ALB) was compared with 16 surrounding samples predominantly from drainages in the James Bay (coded 'A') and Lake Saint-Jean (coded 'B') region (mean n = 23.6; Fig. 1, Table 1). Samples collected from these drainages represented all possible postglacial colonization routes for *S. fontinalis* into the Mistassini Lake region based on its postglacial formation (Bouchard 1980). Two of these samples were also from Hudson Bay, Canada (A1, A2) and one sample was from Maine, USA (C1), so that geographical outliers were

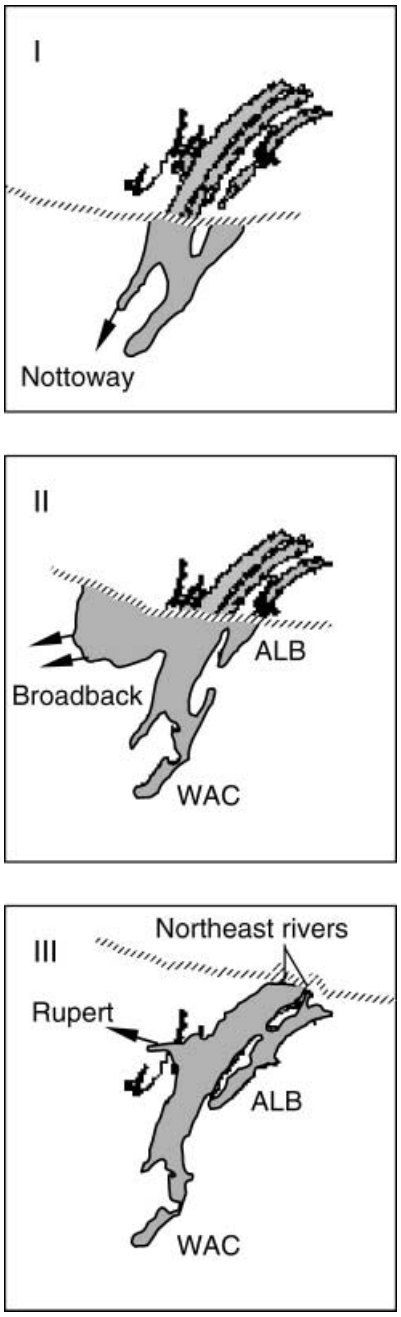

Fig. 2 Reconstruction of Mistassini Lake following the last Pleistocene glaciation (7000–8000 BP), following Bouchard (1980). Arrows represent the direction of discharge from Mistassini Lake, while striped lines indicate the location of the glacial ice margin at each successive stage. (I) A single lake initially covered Lake Waconichi (WAC) and parts of the Mistassini Lake basin, and the outflow was south-southwesterly from the southwest end of the lake. (II) Lake Albanel (ALB) was connected to the Mistassini basin, and the outflow shifted west-southwest via the Broadback River drainage. (III) Drainage then became westward via the Rupert River, Lake Waconichi became separated, Lakes Mistassini and Albanel were separate but interconnected, and eventually the ice margin arrived at the present day northeast rivers (Pepeshquasati and Cheno).

Table 1 Locations, sample sizes (n), and microsatellite diversity estimates within sample sites from drainages surrounding Mistassini Lake. Codes represent drainage regions where populations are found (James Bay, A, as well as Lake Waconichi, Mistassini Lake populations and Lake Albanel; Lake Saint-Jean, B; southern Maine, C1; see Fig. 1). Alleles per locus ($\hat{A}$), corrected allelic richness ($\hat{A}_C$, corrected to $n = 20$), expected heterozygosity (H_E) and observed heterozygosity (H_O) are shown as averages over the seven microsatellite loci used

River drainage	Sampling sites	Code	Latitude	Longitude	n	$\hat{A}$	$\hat{A}_C$	H_E	H_O
Sutton	Sutton R.	A1	55°15′N	83°45′W	24	5.4	5.2	0.57	0.62
Great Whale	Great Whale R.	A2	55°16′N	77°47′W	7	3.1	—	0.48	0.61
Broadback	L. Troilus	A3	50°55′N	74°30′W	10	3.7	—	0.50	0.54
Nottoway	Brock R.	A4	50°32′N	74°22′W	16	3.9	—	0.50	0.56
	L. Chibougamau	A5	49°50′N	74°15′W	12	3.6	—	0.43	0.33
Rupert	Rupert R. at Nemiscau	A6	51°26′N	76°43′W	25	4.6	4.5	0.57	0.61
	L. Waconichi	WAC	50°08′N	74°00′W	46	3.9	3.5	0.41	0.40
	Rupert R.	RUP	51°05′N	73°40′W	40	6.0	5.0	0.52	0.51
	Pepeshquasati R.	PEP	51°24′N	73°00′W	36	6.1	4.9	0.59	0.59
	Cheno R.	CHE	51°21′N	72°45′W	36	5.7	4.7	0.58	0.56
	L. Albanel	ALB	50°54′N	73°18′W	46	5.3	4.6	0.59	0.64
Ashuapmushuan	L. Agréable	B1	49°07′N	73°10′W	20	4.1	4.1	0.54	0.56
	L. Lorecu	B2	49°09′N	73°13′W	14	2.6	—	0.33	0.33
Mistassini	L. Baribal	B3	49°36′N	71°41′W	23	5.6	4.9	0.64	0.70
	L. Maupertuis	B4	50°26′N	71°45′W	20	4.1	4.1	0.48	0.49
	L. Mocassin	B5	50°21′N	71°44′W	20	3.4	3.4	0.49	0.63
	L. Paul-Horace Dumais	B6	49°46′N	72°07′W	20	3.0	3.0	0.28	0.31
Peribonka	L. de la Manne	B7	49°18′N	71°22′W	20	3.1	3.1	0.46	0.49
	L. de la Poule Folle	B8	49°17′N	71°18′W	20	3.6	3.6	0.55	0.57
Bethamisites	Ste. André Cr.	B9	48°27′N	70°54′W	20	4.4	4.4	0.56	0.58
Maine	Kennebago L.	C1	45°08′N	70°46′W	20	5.1	5.1	0.62	0.62

included to expand the scale at which relationships among outside samples could be interpreted. Only samples with no histories of human enhancement were chosen for analyses. There are currently no dams within the principal river drainages of our sampling sites in Québec which could have potential effects on gene flow. Although Mistassini Lake, WAC and ALB were initially interconnected after glacial recession (Bouchard 1980; Fig. 2), contemporary movement of *S. fontinalis* is limited between the three lakes due to the presence of barrier falls (Flick 1977; Bouchard 1980). Genomic DNA extraction from tissue stored in 95% ethanol followed Olsen *et al*. (1996). Microsatellite DNA polymorphism in all samples was quantified by amplifying seven microsatellite loci (*SfoB52*, *SfoC86*, *SfoC88*, *SfoC113*, *SfoC129*, *SfoD75*, *SfoD91*; T.L. King, US Geological Survey, unpublished). Polymerase chain reaction (PCR) conditions and electrophoresis followed that of Fraser *et al*. (2004). Only samples from the three Mistassini Lake populations were genotyped previously in Fraser *et al*. (2004).

Genetic diversity and genetic differentiation

Genetic diversity within each sample at microsatellite loci was quantified with standard descriptive statistics [alleles per locus (A), observed (H_O) and expected (H_E) heteroygosities] and analysed by verifying Hardy–Weinberg equilibrium (HWE) expectations of genotypic frequencies (across all loci in each population and at each locus using GENEPOP 3.3; Raymond & Rousset 1995). Within-sample tests for genotypic disequilibrium between all loci pairs were also performed using GENEPOP. Mean allelic richness per locus ($\hat{A}$) was corrected ($\hat{A}_C$) based on $n = 20$ to increase power for detecting differences in the numbers of alleles per locus (A), using the rarefaction method of Goudet (2001).

Testing between differing ancestral histories was supplemented by comparing global variance in allelic identity (F-statistics, e.g. F_{ST}, θ_{ST}) and allelic size (R-statistics, e.g. R_{ST}) measures of genetic differentiation between all 16 samples outside Mistassini Lake, WAC and ALB. Here, we wanted to determine the relative impact of drift (θ_{ST}) vs. mutation (R_{ST}) on genetic differentiation in the region for subsequent analyses of molecular variance (AMOVAS) in the succeeding discussion. Following Hardy *et al*. (2003), global pairwise θ_{ST} and R_{ST} estimates were first computed among samples. Using the software SPAGEDI 1.1 (Hardy & Vekemans 2002), different allele sizes at each locus were randomly permuted among allelic states (2000 permutations) to provide a simulated distribution of R_{ST} values (ρR_{ST}), with 95% confidence intervals (CI) covering the 50th and 1950th ordered values of ρR_{ST}. Under the null hypothesis $R_{ST} = \theta_{ST}$, differentiation is caused mainly by drift and the observed R_{ST} should not exceed the 95% CI of ρR_{ST} values.

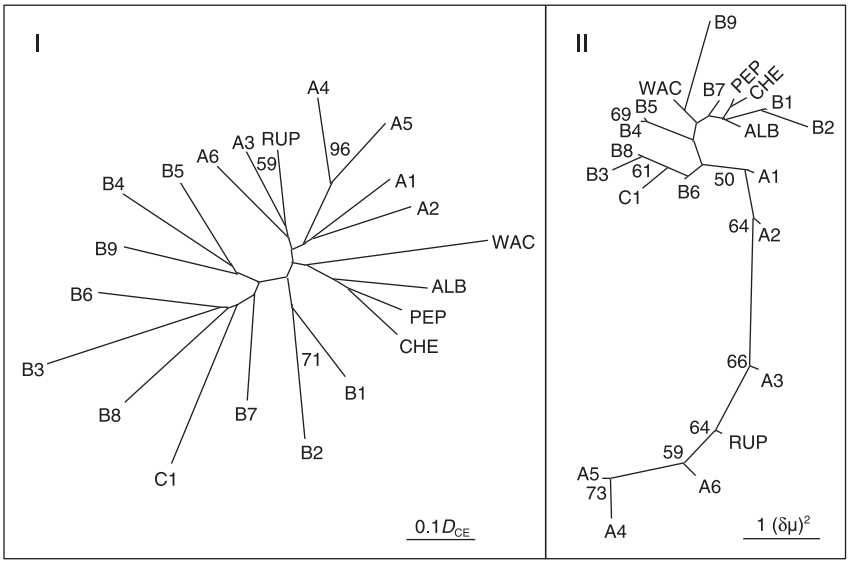

Fig. 3 D_{CE} (I) and $(\delta\mu)^2$ (II) distance-based relationships (with unrooted neighbour-joining trees) illustrating genetic relationships between divergent populations of *Salvelinus fontinalis* from Mistassini Lake (outflow: RUP; inflows: CHE, PEP) and samples from drainages of Hudson/James Bay (A1–A6), Lake Saint-Jean (B1–B9), and Maine (C1). Trees were constructed using POPULATIONS 1.2.14 (Langella 2001). Consistency of tree topology was assessed by bootstrapping over loci with replacement and 2000 replicates, and trees were visualized using TREEVIEW (Page 1996). Numbers at branch points represent bootstrap resampling percentages of ≥ 50%.

Under the alternative hypothesis $R_{ST} > \theta_{ST}$, stepwise-mutation model (SMM)-like mutations have contributed to differentiation and the observed R_{ST} should exceed the 95% CI of ρR_{ST} values (see Hardy *et al*. 2003), thereby supporting a mutational component to differentiation.

Testing ancestral origins

Initial analyses involved the estimation of distance-based phylogenetic relationships among all 21 samples to determine whether outflow and inflow populations clustered together or more closely with outside samples. This was performed using an unrooted neighbour-joining (NJ) clustering analysis with one drift-based estimator, Cavalli-Sforza & Edwards's (1967) chord distance, D_{CE}, and one mutation-based estimator, Goldstein *et al*.'s (1995) $(\delta\mu)^2$. Similar approaches have been adopted in previous investigations of molecular ecology of other north temperate freshwater fishes (e.g. Taylor & McPhail 2000; Tessier & Bernatchez 2000). In both D_{CE} and $(\delta\mu)^2$ topologies, outflow and inflow populations from Mistassini Lake did not cluster together, providing a first indication that they did not have a common origin (Fig. 3). However, tree topologies had relatively low bootstrap support for major nodes (and nodes separating outflow and inflow populations) (Fig. 3). More emphasis should be placed on $(\delta\mu)^2$ since genetic differentiation in the region had a mutational component because of allelic size variance (see Results). Nevertheless, the weak bootstrap support put into question whether Mistassini Lake populations represented 'pure' lineages or varying degrees of admixed lineages, especially given that gene flow occurs between outflow and inflow populations (Fraser *et al*. 2004). This necessitated admixture analyses to provide further support that outflow and inflow populations had different origins, and involved (i) clustering the surrounding 16 samples into groups, and (ii) quantifying admixture proportions to evaluate the likelihood that Mistassini Lake populations, WAC and ALB, originated from each of these groups. Different origins of divergent populations in Mistassini Lake would be further supported if they clustered into different surrounding population groups with good clustering confidence.

To initiate the clustering of surrounding 16 samples into groups, we ran the Bayesian clustering method of STRUCTURE (Pritchard *et al*. 2000) under a model assuming admixture, correlated allele frequencies between K population groups [burn-in 50 000 replications, 100 000 Markov chain Monte Carlo (MCMC) replicates], and no prior information on sample location for individuals. Runs were started at $K = 2$ and proceeded to $K = 13$. Runs were replicated three times at each K to confirm consistency of log-likelihood probabilities. At each run (i.e. $K = 2$, 3, 4, etc.), we assigned each sample to the K where their individuals had the greatest posterior probability (highest mean individual admixture proportion q) of originating (e.g. for $K = 3$, samples could be assigned to one of three groups). As no unanimous criterion exists for defining groups of populations (Dupanloup *et al*. 2002), we defined groups as those that were maximally differentiated from one another in genetic terms (*sensu* Dupanloup *et al*. 2002). We thus performed analyses of molecular variance (AMOVA; Schneider *et al*. 2000) at each K to determine at which K the proportion of between-group genetic variance was maximized while minimizing the within-group proportion of genetic variance. We did not specifically use the spatial analysis of molecular variance (SAMOVA) of Dupanloup *et al*. (2002) because it incorporates geographical coordinates between samples for variance component calculations, which do not necessarily reflect watershed connections in

our sampled region. We stopped at $K = 13$ because it was clear after this point for AMOVAS based on either allelic identity differences (F_{ST}) or allelic size variance (R_{ST}) that the between-group component of variance had continued to decline after reaching maximum peaks at $K = 9$ (F_{ST}) or $K = 8$ (R_{ST}) while the within-group component of variance began to increase again (see Results). We complemented this approach by assessing trends in log-likelihood probabilities {ln[$P(X \mid K)$]} outputted in STRUCTURE for each model of K (i.e. $K = 2, 3, 4$, etc.).

All admixture methods make generalizations about a set of evolutionary processes that would otherwise be difficult to address quantitatively (Chakraborty 1986; Bertorelle & Excoffier 1998; Dupanloup *et al*. 2004). They assume that allele frequencies within admixed populations and their source groups are accurate and have remained unchanged since admixture events, and that all potentially contributing source groups have been sampled and are of constant size (Chakraborty 1986; Dupanloup *et al*. 2004). Thus, these methods also assume that the effects of genetic drift or gene flow since the admixture event have been negligible (Dupanloup *et al*. 2004), conditions that may or may not be met among populations of *S. fontinalis* in this study. As a consequence, we employed three different approaches for estimating admixture proportions in CHE/PEP/RUP/WAC/ALB. First, we ran STRUCTURE for each population separately with individuals considered as unknowns to cluster their individuals into *a posteriori*-defined major groups, and ultimately to determine their Bayesian posterior probabilities (mean of individual admixture proportions q) of originating in each group. We then used ADMIX 2.0 (Dupanloup & Bertorelle 2001) to obtain conventional admixture proportions based on allele frequencies in the defined major groups (*sensu* Chakraborty 1986). These two approaches did not take into account information regarding the degree of molecular divergence between alleles in the defined major groups. Consequently, we also ran ADMIX to account for such information (in addition to allele frequency differences) by computing the admixture coefficient (mY) of Bertorelle & Excoffier (1998) for CHE/PEP/RUP/WAC/ALB. Under this model, molecular divergence was estimated from the average squared difference in allele sizes, and microsatellites were assumed to follow an SMM (Bertorelle & Excoffier 1998).

Results

Microsatellite DNA diversity

The number of significant ($P < 0.05$) genotypic disequilibrium tests between microsatellite loci pairs was lower than expected by chance (13 of 441), supporting the conclusion that all seven loci used were independent. Only two samples deviated from HWE expectations (A4, global heterozygote deficit; B5, global heterozygote excess), and only five of 154 individual locus tests in samples displayed significant departures from HWE ($P < 0.05$), with no consistent patterns. Mean H_E and A among samples were 0.51 (range 0.28–0.64) and 4.2 (range 2.6–6.1) (Table 1), while A was higher across all samples than within samples (range 5–18; mean 10) (Fig. 4). No differences in $\hat{A}_C$ were detected among those samples with a minimum $n = 20$ (one-way ANOVA; $F_{15,96} = 1.37$, $P = 0.18$) (Table 1).

Table 2 Mean single locus and multilocus pairwise estimates rounded to three decimal places of R_{ST}, θ_{ST} and ρR_{ST} (95% distribution of central ρR_{ST} values in parentheses) between 16 samples collected from outside of Lakes Mistassini/Waconichi/Albanel following 2000 allele permutations (Hardy *et al*. 2003)

Locus	θ_{ST}	R_{ST}	ρR_{ST} (95% range)
SfoB52	0.307	0.164	0.265 (0.067–0.445)
SfoC86	0.417	0.636*	0.397 (0.163–0.636)
SfoC88	0.189	0.180	0.171 (0.085–0.248)
SfoC113	0.295	0.258	0.284 (0.124–0.489)
SfoC129	0.302	0.376*	0.276 (0.104–0.371)
SfoD75	0.235	0.194	0.225 (0.086–0.389)
SfoD91	0.207	0.329*	0.203 (0.081–0.325)
Multilocus	0.277	0.331*	0.253 (0.167–0.330)

*Significance values at the $P = 0.05$ level.

Genetic differentiation

Global multilocus pairwise R_{ST} estimates among the surrounding 16 samples were larger than the 95% range of ρR_{ST} values ($P = 0.007$). The same trend was observed when calculations were done using all 21 samples (data not shown). Overall, this supported that $R_{ST} > \theta_{ST}$ in the surrounding region, indicating a mutational component to genetic differentiation based on allelic size variance (Table 2). Namely, at individual loci with $R_{ST} > \theta_{ST}$, alleles *SfoC86**101 and *SfoD91**220–240 were common in samples from most Lake Saint-Jean drainages (B1–B9) and C1, whereas *SfoC86**116–119 and *SfoD91**256–268 were common to samples from Hudson/James Bay drainages (A1–A6: Fig. 4). Relative to the 16 surrounding samples, similar trends were observed at the same allelic suites for *SfoD91* at the level of RUP/WAC vs. CHE/PEP/ALB, but allele *SfoC86**101 was infrequent in all populations relative to alleles *SfoC86**116–119 (Fig. 4). The presence of both 'Hudson/James Bay' and 'Lake Saint-Jean drainages/C1' alleles in Mistassini Lake was also demonstrated at other loci: there was a strong presence of allele *SfoB52**219 shared between CHE/PEP/WAC/ALB and samples from Lake Saint-Jean drainages (B1–B9)/C1, and the presence of alleles *SfoD75**200–208 shared between CHE/PEP/RUP/WAC/ALB and certain samples from Hudson/James Bay (A3–A6) drainages (Fig. 4).

Table 3 Summary of results of combined Bayesian-clustering (STRUCTURE) for inferring k populations in the surrounding 16 samples outside Mistassini/Waconichi/Albanel, including untransformed log-likelihood probabilities ln[$P(X \mid K)$] and corresponding analyses of molecular variance (AMOVA) based on allelic identity (F_{ST}) and allelic size variance (R_{ST}) differences at each K. All variance components for each K were significant at the $P = 0.05$ level and are expressed as proportions. Bold values represent the models of k where ln[$P(X \mid K$] was greatest (lowest negative values), and where the within-group component of variance (F_{SC}/R_{SC}) was minimized and the between-group component of variance (F_{CT}/R_{CT}) maximized

STRUCTURE/AMOVA K	ln[P(X | K)]	F_{SC}	F_{CT}	R_{SC}	R_{CT}	Inferred groups
2	−5509	0.124	0.185	0.230	0.146	(A1–A6; B1–B2) (B3–B9; C1)
3	−5256	0.113	0.188	0.162	0.201	(A1–A6; B1–B2) (B3, B6, B8, C1) (B4, B5, B7, B9)
4	−5084	0.121	0.174	0.127	0.239	(A1–A6) (B1–B2) (B3, B6, B8, C1) (B4, B5, B7, B9)
5	−4948	0.116	0.178	0.152	0.202	(A1–A6) (B1–B2)(B6–B8) (B3, C1) (B4, B5, B9)
6	−4778	0.143	0.148	0.135	0.217	(A1–A6) (B1–B2) (B3) (B4, B5) (B6, C1) (B7–B9)
7	−4691	0.124	0.164	0.106	0.241	(A1–A3, A6) (A4–A5) (B1–B2) (B3) (B4–B5, B9) (B6, C1)(B7–B8)
8	−4487	0.103	0.183	**0.093**	**0.251**	(A1–A3, A6) (A4–A5) (B1–B2) (B3) (B4–B5) (B6, B8) (B7, B9) (C1)
9	−4570	**0.091**	**0.195**	0.102	0.240	Same as $K = 8$ but (B4) (B5, B9) (B7)
10	−4454	0.097	0.185	0.109	0.228	Same as $K = 8$ but (A1, A6) (A2,A3) (B4) (B5, B9) (B7)
11	**−4422**	0.093	0.192	0.109	0.227	Same as $K = 8$ but (A1, A6) (A2,A3) (B4)(B5) (B7)(B9)
12	−5072	0.100	0.174	0.140	0.193	Same as $K = 11$ but (B6) (B8)
13	−5012	0.106	0.172	0.141	0.191	Same as $K = 11$ but (B1) (B2) (B6) (B8)

Testing ancestral histories

Analyses for defining groups among the surrounding 16 samples revealed that in AMOVAs based on allelic size variance (R_{ST}), the within-group component of variance declined while the between-group component of variance increased from $K = 2$–8 (Table 3). Opposite trends were observed from $K = 9$–13 (Table 3). Maximally differentiated groups (where the within-group component of variance was minimized and the between-group component of variance maximized) were therefore realized at $K = 8$. The same analyses using AMOVAs based on allelic identity differences (F_{ST}) were slightly more variable, but suggested maximally differentiated groups at $K = 9$. These differences between R_{ST}- and F_{ST}-based AMOVAs are not surprising given their emphasis on different evolutionary forces (mutation vs. drift), but we place more emphasis on AMOVAs based on allelic size differences because of the mutational component to genetic differentiation detected in the 16 samples. Untransformed log-likelihood probabilities for models of $K = 2$–13 in STRUCTURE were highest (i.e. with the lowest negative values) between $K = 8$–11 with a bimodal distribution (Table 3). In such instances of weak discrimination between competing K, the suggested conservative approach is to assume that the first mode of the distribution (i.e. $K = 8$) is the most likely K (see Pritchard *et al*. 2000) (Table 3). At $K = 8$, individuals were also assigned to the group in which they had been clustered (where their admixture proportion q was greatest) at a mean rate of 90.4% (range 84.8–100%), and assignment success did not improve at competing K (e.g. $K = 9$: 90.4%; $K = 10$: 90.8%; $K = 11$: 88.4%; data not shown). Together, these results justified the use of eight *a posteriori*-defined population groups in the surrounding region as the most conservative approach to estimate admixture proportions for Mistassini Lake populations, WAC and ALB.

Admixture proportion estimates were consistent with the hypothesis of different origins for divergent populations in Mistassini Lake. Results from all three approaches supported (i) segregation of outflow (RUP) and inflow (CHE, PEP) populations into different groups; (ii) clustering of RUP with a group composed of samples collected west of Mistassini Lake draining into Hudson/James Bay (A1–A3, A6); and (iii) clustering of CHE and PEP principally into a group south-southwest of the lake comprised of samples from the Ashuapmushuan drainage (B1–B2) that drain contemporarily into Lake Saint-Jean (Table 4). Admixture proportions in these respective groups were particularly high (0.80–0.96) for RUP and more varied in CHE and PEP depending on the approach employed (0.53–0.88; Table 4). Admixture proportions for ALB followed similar trends as CHE and PEP (Table 4). Admixture proportions for WAC were the most variable of any population, clustering into different groups depending on the approach, but there was nevertheless a clustering signal for the same two groups as all other populations (A1–A3, A6 vs. B1–B2; Table 4: I, II). A closer inspection of allele frequency and molecular-based admixture proportions using ADMIX revealed a number of admixture values that were either > 1 for WAC (Table 4: III) or less than 0 for other populations (Table 4: II, III). Slightly negative values (up to −0.15) are expected when a group is not a plausible source/parental population (Dupanloup *et al*. 2004). Larger negative values, admixture proportions > 1, or large standard deviations may occur when more complex exchanges have occurred than simple admixture events, such as pronounced genetic

Table 4 Admixture proportions of Mistassini Lake populations (RUP, PEP, CHE), Lake Waconichi (WAC), and Lake Albanel (ALB) for originating from the eight groups defined in the surrounding region, based on three approaches: (I) Bayesian clustering using STRUCTURE (Pritchard *et al*. 2000); (II) allele frequencies using ADMIX (Dupanloup *et al*. 2002); (III) accounting for information regarding the degree of molecular divergence between alleles (*mY*) using ADMIX. Admixture proportions using STRUCTURE are the mean of individual admixture proportions (Bayesian-based posterior probabilities *q*) within populations, and weighted averages across loci using ADMIX. Standard deviations are denoted in parentheses. Bold values indicate the inferred group with the highest admixture proportion using each approach

		Inferred group							
		A1–A3, A6	A4–A5	B1–B2	B3	B4–B5	B6, B8	B7, B9	C1
I	WAC	**0.92 (0.09)**	0.01 (0.02)	0.03 (0.05)	< 0.01	< 0.01	< 0.01	0.02 (0.05)	< 0.01
	RUP	**0.80 (0.21)**	0.03 (0.04)	0.04 (0.07)	< 0.01	0.04 (0.06)	0.02 (0.04)	0.06 (0.04)	< 0.01
	CHE	0.04 (0.08)	0.02 (0.02)	**0.88 (0.13)**	< 0.01	0.01 (0.01)	0.01 (0.02)	0.03 (0.07)	< 0.01
	PEP	0.12 (0.15)	0.07 (0.08)	**0.57 (0.27)**	0.03 (0.02)	0.04 (0.03)	0.04 (0.03)	0.13 (0.21)	< 0.01
	ALB	0.05 (0.06)	0.05 (0.06)	**0.78 (0.21)**	0.01 (0.03)	0.04 (0.09)	0.03 (0.06)	0.04 (0.09)	< 0.01
II	WAC	**0.42 (0.33)**	−0.23 (0.16)	**0.43 (0.11)**	−0.15 (0.20)	0.24 (0.13)	−0.17 (0.15)	0.27 (0.19)	0.19 (0.09)
	RUP	**0.92 (0.13)**	−0.14 (0.08)	0.29 (0.08)	−0.09 (0.05)	0.07 (0.08)	−0.04 (0.07)	−0.01 (0.09)	0.00 (0.07)
	CHE	0.27 (0.39)	−0.15 (0.21)	**0.65 (0.29)**	−0.15 (0.12)	−0.05 (0.33)	0.23 (0.20)	0.13 (0.08)	0.07 (0.11)
	PEP	0.37 (0.34)	−0.05 (0.17)	**0.53 (0.11)**	−0.09 (0.10)	−0.14 (0.12)	0.13 (0.14)	0.17 (0.20)	0.08 (0.07)
	ALB	0.33 (0.57)	−0.04 (0.28)	**0.64 (0.19)**	−0.21 (0.16)	0.18 (0.20)	0.17 (0.24)	−0.15 (0.31)	−0.02 (0.13)
III	WAC	0.27 (0.29)	−0.17 (0.29)	0.34 (0.42)	−1.04 (0.68)	1.01 (0.72)	−0.52 (0.55)	−0.35 (1.48)	**1.47 (1.68)**
	RUP	**0.96 (0.16)**	−0.25 (0.08)	0.35 (0.07)	−0.04 (0.05)	0.11 (0.13)	−0.04 (0.07)	−0.10 (0.14)	0.01 (0.10)
	CHE	0.16 (0.14)	0.01 (0.12)	**0.79 (0.22)**	−0.26 (0.25)	0.33 (0.28)	−0.35 (0.25)	0.03 (0.40)	0.29 (0.63)
	PEP	0.10 (0.11)	0.05 (0.09)	**0.73 (0.12)**	−0.21 (0.12)	−0.24 (0.21)	0.02 (0.17)	0.18 (0.22)	0.37 (0.53)
	ALB	0.21 (0.12)	0.11 (0.10)	**0.94 (0.12)**	−0.22 (0.13)	0.48 (0.29)	−0.47 (0.15)	−0.05 (0.18)	0.00 (0.23)

drift or additional gene flow (deviations of the admixture model assumed in ADMIX; Dupanloup *et al*. 2004).

Discussion

Secondary contact of at least two ancestral groups; congruence with geological history

Our genetic data confirmed that divergent outflow (Rupert) and inflow (Cheno, Pepeshquasati) breeding populations of *Salvelinus fontinalis* in Mistassini Lake do not have a common origin. Divergent populations clustered into two different population groups in surrounding drainages based on three diverse admixture analyses of microsatellite DNA. Clustering was neither weak nor random (i.e. there was consistent segregation into two particular groups). It is also unlikely that non-neutrality at some microsatellite loci (Kashi & Soller 1999) has led to closer genetic relationships of either outflow or inflow populations with outside samples. Specifically, charr from different samples have dissimilar life histories suitable for occupying physiologically variable environments and presumably experience different selective regimes (e.g. A1, A2 are anadromous; Mistassini Lake populations are lake migratory; A6, B9 are non-migratory). Furthermore, our results do not support that an accumulation of many new microsatellite mutations in sympatry has occurred since postglacial colonization (maximum 7000–8000 BP or 1750–2000 generations). Indeed, the most parsimonious explanation for the prevalence of the same microsatellite alleles between many samples collected from physically isolated drainages, coupled with the low allelic diversity observed, is that part of the genetic divergence observed between divergent populations in Mistassini Lake precedes the formation of the lake.

Several possibilities could account for the contrasting signals we observed between the strong allelic size variance subdivision at microsatellite DNA and mtDNA monomorphism in Mistassini Lake populations. Based on the work of Danzmann *et al*. (1998), divergent populations sharing haplotype 1 of mtDNA assemblage 'B' could have originated from two distinct glacial refugia or from different ancestral populations within either refugium. Additionally, because some degree of gene flow occurs between outflow and inflow populations (Fraser *et al*. 2004), divergent populations may have had different 'B' assemblage haplotypes in allopatry, with mtDNA introgression and hybridization upon secondary contact. Neither these possibilities, nor whether stochastic or selective loss of certain haplotypes led to monomorphism of haplotype 1 in Mistassini Lake and other northern regions, can be ascertained with our data. However, they would not affect our interpretations regarding different origins. Although we tentatively favour colonization from two glacial refugia sharing 'B' haplotypes, either scenario implies that divergence times between ancestral populations were relatively recent (Danzmann *et al*. 1998), but long enough for substantial mutational events at microsatellites to occur. Both Bernatchez & Dodson (1991) and Danzmann *et al*. (1998) have argued

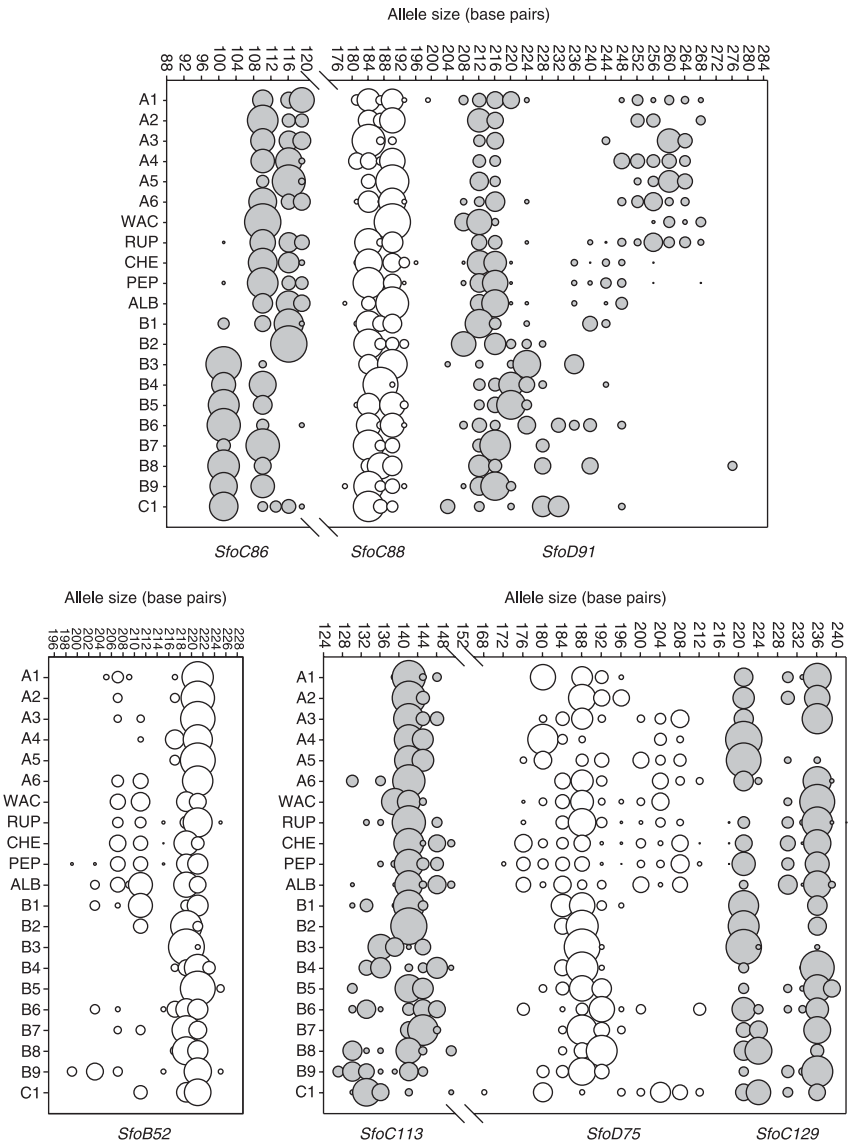

Fig. 4 Allele frequency (bubble area) and size distributions (in base pairs) for seven microsatellite loci in *Salvelinus fontinalis* populations: Mistassini Lake (RUP, CHE, PEP); Lake Waconichi, WAC; Lake Albanel, ALB; Hudson/James Bay (A1–A6); Lake Saint-Jean (B1–B9); Maine (C1).

that historical isolation of recently derived haplotype mtDNA assemblages in eastern North American freshwater fishes (including 'B' in *S. fontinalis*) likely resulted from the Wisconsian glacier advance (100 000 BP, peak 18 000 BP; Denton & Hugues 1981).

The scenario that at least two ancestral groups colonized Mistassini Lake is also supported by a sequence of known outflow direction changes during the lake's formation (7000–8000 BP: Fig. 2; Bouchard 1980). An initial south-southwest discharge would have permitted a first colonization from this direction and is supported by the genetic clustering of inflow populations (Cheno, Pepeshquasati) and Lake Albanel with Ashuapmushuan River drainage samples. The subsequent shift to a west outflow direction (a process of several hundred to a thousand years; A. Bouchard, personal communication) would have facilitated dispersal to the lake by a second ancestral group from this direction and is consistent with the close genetic affinities of the outflow population (Rupert) to samples west of Mistassini Lake. Presumably at this same time, the first ancestral group evolved to inhabit available habitat in the southwestern portion of the expanding lake basin, thereby permitting it to later colonize inflows (northeast rivers) and Lake Albanel once glaciers had receded (Fig. 2II, III). In agreement with this scenario, inflow populations were more similar to Lake Albanel than to the outflow population, despite a contemporary barrier (waterfall) to gene flow between the two lakes (Flick 1977; Bouchard 1980). The clustering of Lake Waconichi into the same group as the outflow population (Bayesian approach) is one discrepancy to our colonization scenario, and it was perhaps allelic fixation at two of seven microsatellite loci employed

in this lake (*SfoC86* and *SfoC88*) that contributed to this pattern (and to the overall discrepancies between different admixture models in this population).

Although we believe that our conclusions regarding different ancestral affinities for outflow and inflow populations are conservative, it could be argued that a group that was polymorphic for a wide range of alleles colonized Mistassini Lake and that genetic drift under limited gene flow made them appear to have different origins by chance. However, it is very unlikely that stochastic genetic drift would be responsible for creating similarity with other populations from outside the lake. On the other hand, ongoing gene flow within divergent population groups (i.e. between inflow populations) might have given the impression that they originated from the same ancestor when they did not. For instance, it could be that Pepeshquasati and Rupert originated from the same ancestral group but that subsequent gene flow between Cheno and Pepeshquasati made these latter two appear more similar. Clearly, the extent of admixture was different among inflow populations. Furthermore, both admixture analyses and allele frequency distributions at each of the seven loci suggested overall that the inflow populations were likely more admixed between the divergent lineages in the lake than the outflow population, the latter appearing to represent a more 'pure' lineage from a Hudson/James Bay ancestor. Consequently, one must be cognizant of the assumptions underlying the admixture analyses employed (see Materials and methods). Additional investigation into admixture proportions within inflow populations especially, using larger numbers of loci and greater sample sizes from outside populations, would thus be beneficial.

The potential role of history in divergence

Because *S. fontinalis* populations in Mistassini Lake have different origins, history may have contributed to their divergence. Where secondary contacts are known in other divergent north temperate fish fauna, the general view is that much of the differentiation (particularly phenotypic) has likely occurred from selective processes in sympatry (Taylor 1999; Bernatchez 2004). For instance, equilibrium conditions for phenotypic traits have probably been reached since postglacial colonization (Hendry *et al*. 2002; Bell 2004). Furthermore, ancestral populations appear to have been initially more similar in phenotype upon secondary contact, with differences in ecology or morphology arising or at least being magnified in sympatry (Thompson *et al*. 1997; Taylor 1999; Lu *et al*. 2001).

From this standpoint, one question in Mistassini Lake remains. Given that the outflow was available for colonization well before inflows, why did fish from the first colonizing ancestral group not colonize the outflow (as these fish were obviously capable of colonizing several other areas)? It may be that inflows and outflows are divergent enough environments that colonization and adaptation of the outflow by inflow ancestors was difficult. For instance, outflows and inflows typically differ in thermal and flow regimes (Carmack *et al*. 1979; Burger *et al*. 1997). Alternatively, inflow ancestors may in fact have colonized the outflow, but were out-competed by outflow ancestors once these fish colonized the region. Either case would suggest that outflow and inflow ancestors had at least some pre-existing features that permitted them to successfully colonize their respective environments in Mistassini Lake. In parallel, intraspecific differences in colonization potential for occupying different environments are known in another north temperate fish species. Sockeye salmon (*Oncorhynchus nerka*) use lakes for spawning and are commonly characterized by outflow, inflow, and shoreline-spawning forms (Burger *et al*. 1997, 2000). In a previously unoccupied lake, Burger *et al*. (2000) showed that following introductions of external outflow, inflow, and shoreline-spawning populations, it was predominantly shoreline and inflow forms that led to the successful colonization and populating of shoreline and inflow breeding areas, respectively.

The different littoral zone feeding habitats of outflow and inflow populations were also strongly associated with their ancestors' colonization directions. Inflow populations dominated the distinctive middle island chain and south shore, while the outflow population was more restricted to the area west and southwest of its river mouth (Fraser & Bernatchez 2005). Thus, even if outflow and inflow ancestors were largely identical upon colonizing Mistassini Lake, the mode of divergence in sympatry (i.e. habitat use, short vs. long migrations) may have been affected by the nature of colonization (i.e. direction and timing). Indeed, the formation of new habitats within postglacial lakes likely initiated episodes of directional selection for colonizing fish (McPhail 1993; Schluter 1996). Assuming this was the case in Mistassini Lake, the response to selection acting on habitat use and foraging behaviour within initial areas colonized by first colonizers (inflow ancestors) may have provided fitness advantages that facilitated their colonization of these same habitats as further lake expansion occurred in a northeast direction, while limiting access for later colonizers (outflow ancestors).

Conclusions

Our study yields two significant findings. First, based on microsatellite DNA, our results demonstrate that divergent outflow and inflow populations of *Salvelinus fontinalis* in Mistassini Lake do not have a common origin, despite being monomorphic at mtDNA. This serves as a cautionary reminder that monomorphism at mtDNA is insufficient evidence to infer a common ancestry and thus the origin of divergent populations found in specific

systems (i.e. lakes) (see also: Taylor 1999; Lu *et al*. 2001). Second, our results suggest that differential colonization potential of ancestral populations and/or the nature of colonization (i.e. direction and timing) could have affected the mode of divergence in sympatry. Thus, although Mistassini Lake populations share common elements with other north temperate freshwater fish species with respect to their environmental regimes and evolution (broadly speaking: alternative niche use resulting from natural selection), history may have also promoted unique ecological interactions.

Acknowledgments

The diverse comments and thoughtful suggestions of A. Calvert, V. Castric, M. Hansen, J. Turgeon, D. Veliz, G. Wallis and five anonymous reviewers greatly improved the article. We also thank A. Bouchard for insights into the glacial history of central Québec. We are indebted to R. Danzmann (University of Guelph), S. Gravel, P. Ouellet (Société de la faune et des parcs du Québec, FAPAQ, Jonquière and Chibougamau), N. LeClerc, R and R. Proulx (Société de l'établissement des activités du plein air du Québec, SÉPAQ, Chibougamau), M. Prince (Cree Nation of Mistissini), N. Saganash, R. Dion (Grand Council of the Crees of Québec), R. White (Squaretail Lodge), V. Castric (Université de Lille), Pourvoirie La Jeanoisse, Pourvoirie Aventures Nipissi, for contributing samples. B. Angers and R. Saint-Laurent provided advice on mtDNA processing and we thank K. Allard for laboratory assistance. This work was supported by a field research grant from the Eeyou Namess Corporation (Montréal) to LB and DJF, a research grant from Science and Engineering Research Canada (NSERC, Genomic program) to LB, and NSERC Postgraduate and Northern Research Scholarships to DJF.

References

Angers B, Bernatchez L (1998) Combined use of SMM and non-SMM methods 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.

Behrmann-Godel J, Gerlach G, Eckmann R (2004) Postglacial colonization shows evidence for sympatric population splitting of Eurasian perch (*Perca fluviatilis* L.) in Lake Constance. *Molecular Ecology*, **13**, 491–497.

Bell MA (2004) Twelve years of contemporary evolution in a three-spine stickleback population. *Evolution*, **58**, 814–824.

Bernatchez L (2004) Ecological theory of adaptive radiation: an empirical assessment from coregonine fishes (Salmoniformes). In: *Evolution Illuminated: Salmonids and their Relatives* (eds Hendry AP, Stearns SC), pp. 175–207. Oxford University Press, New York.

Bernatchez L, Dodson JJ (1990) Allopatric origin of sympatric populations of 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.

Bertorelle G, Excoffier L (1998) Inferring admixture proportions from molecular data. *Molecular Biology and Evolution*, **15**, 1298–1311.

Bouchard A (1980) *Late Quaternary geology, Temiscamie region, central Québec*. Doctoral Thesis, McGill University, Montréal, Canada.

Burger CV, Scribner KT, Spearman WJ, Swanton CO, Campton DE (2000) Genetic contribution of three introduced life history forms of sockeye salmon to colonization of Frazer Lake, Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, **57**, 2096–2111.

Burger CV, Spearman WJ, Cronin MA (1997) Genetic differentiation of sockeye salmon subpopulations from a geologically young Alaskan lake system. *Transactions of the American Fisheries Society*, **126**, 926–938.

Carmack EC, Gray CBJ, Pahro CH, Daley RJ (1979) Importance of lake–river interaction on seasonal patterns in the general circulation of Kamloops Lake, British Columbia. *Limnology and Oceanography*, **24**, 634–644.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution*, **32**, 550–570.

Chakraborty R (1986) Gene admixture in human populations: models and predictions. *Yearbook of Physical Anthropology*, **29**, 1–43.

Danzmann RG, Morgan RP, Jones MW, Bernatchez L, Isshen PE (1998) A major sextet of mitochondrial DNA phylogenetic assemblages in eastern North American brook trout (*Salvelinus fontinalis*): distribution and postglacial dispersal patterns. *Canadian Journal of Zoology*, **76**, 1300–1318.

Denton GH, Hugues TJ (1981) *The Last Great Ice Sheets*. Wiley, Toronto, Canada.

Dupanloup I, Bertorelle G (2001) Inferring admixture proportions from molecular data: extension to any number of parental populations. *Molecular Biology and Evolution*, **18**, 672–675.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, **11**, 2571–2581.

Dupanloup I, Bertorelle G, Chikhi L, Barbujani G (2004) Estimating the impact of prehistoric admixture on the genome of Europeans. *Molecular Biology and Evolution*, **21**, 1361–1372.

Flick WA (1977) Some observations, age, growth, food habits and vulnerability of large brook trout (*Salvelinus fontinalis*) from four Canadian lakes. *Naturaliste Canadienne*, **104**, 353–359.

Fraser DJ, Bernatchez L (2005) Adaptive migratory divergence among sympatric brook charr populations. *Evolution*, **59**, in press.

Fraser DJ, Lippé C, Bernatchez L (2004) Consequences of unequal population size, asymmetric gene flow, and sex-biased dispersal on population structure in brook charr (*Salvelinus fontinalis*). *Molecular Ecology*, **13**, 67–80.

Gislason D, Ferguson MM, Skulason S, Snorrason SS (1999) Rapid and coupled phenotypic and genetic divergence in Icelandic Arctic charr (*Salvelinus alpinus*). *Canadian Journal of Fisheries and Aquatic Sciences*, **56**, 2229–2224.

Goldstein DB, Linares AR, Cavalli-Sforza LL, Feldman MW (1995) Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Science of the United States of America*, **92**, 6725–6727.

Goudet J (2001) FSTAT, *a program to estimate and test gene diversities and fixation indices, Version 2.9.3*. Institut d'Écologie, Université de Lausanne, Switzerland.

Hardy OJ, Vekemans X (2002) SPAGEDI: a versatile computer program to analyze spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Hardy OJ, Charbonnel N, Freville H, Heuertz M (2003) Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics*, **163**, 1467–1482.

Harvey PH, Pagel MD (1991) *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford.

Hendry AP, Taylor EB, McPhail JD (2002) Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the Misty system. *Evolution*, **56**, 1199–1216.

Hynes RA, Ferguson A, McCann MA (1996) Variation in mitochondrial DNA and postglacial colonization of northwestern Europe by brown trout. *Journal of Fish Biology*, **48**, 54–67.

Kashi Y, Soller M (1999) Functional roles of microsatellites and minisatellites. In: *Microsatellites: Evolution and Applications* (eds Goldstein DB, Schlötterer C), pp. 10–23. Oxford University Press, Oxford.

Koskinen MT, Nilsson J, Veselov A, Potutkin AG, Ranta E, Primmer CR (2002) Microsatellite data resolve phylogeographic patterns in European grayling, *Thymallus thymallus*, Salmonidae. *Heredity*, **88**, 391–401.

Langella O (2001) POPULATIONS *1.2.24. Population genetic structure (individual or population distances, phylogenetic trees)*. Available at http://www.pge.cnrs-gif.fr/bioinfo/populations/.

Langerhans RB, DeWitt TJ (2004) Shared and unique features of evolutionary diversification. *American Naturalist*, **164**, 335–349.

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

Losos JB, Jackman TR, Larson A, De Queiroz K, Rodriguez-Schettino L (1998) Contingency and determinism in replicated adaptive radiations of island lizards. *Science*, **279**, 2115–2118.

Lu G, Bernatchez L (1999) Correlated trophic specialization and genetic divergence in sympatric lake whitefish ecotypes (*Coregonus clupeaformis*): support for the ecological speciation hypothesis. *Evolution*, **53**, 1491–1505.

Lu G, Basley DJ, Bernatchez L (2001) Contrasting patterns of mitochondrial DNA and microsatellite introgressive hybridization between lineages of lake whitefish (*Coregonus clupeaformis*): relevance for speciation. *Molecular Ecology*, **10**, 965–985.

McPhail JD (1993) Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): origins of the species pairs. *Canadian Journal of Zoology*, **71**, 515–523.

Olsen JB, Wenburg JK, Bentzen P (1996) Semiautomated multilocus genotyping of Pacific salmon (*Oncorhynchus* spp.) using microsatellites. *Molecular Marine Biology and Biotechnology*, **5**, 259–272.

Østbye K, Naesje TF, Bernatchez L, Sandlund OT, Hindar K (2005) Morphological divergence and origin of sympatric populations of European whitefish (*Coregonus lavaretus*, L.). *Journal of Evolutionary Biology*, in press.

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences*, **12**, 357–358.

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

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Robinson BW, Wilson DS (1994) Character displacement and release in fishes: a neglected literature. *American Naturalist*, **144**, 596–627.

Rogers SM, Bernatchez L (2005) Integrating QTL mapping and genomic scans towards the characterization of candidate loci under parallel directional selection in the lake whitefish (*Coregonus clupeaformis*). *Molecular Ecology*, **14**, 351, http://www.blackwell-synergy.com/links/doi/10.1111/j.1365-294X.2004.02396.x/full/.

Rogers SM, Gagnon V, Bernatchez L (2002) Genetically based phenotype association for swimming behavior in lake whitefish ecotypes (*Coregonus clupeaformis* Mitchill). *Evolution*, **56**, 2322–2329.

Rundle H, Nagel L, Boughman J, Schluter D (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science*, **287**, 306–308.

Saint-Laurent R, Legault M, Bernatchez L (2003) Divergent selection maintains differentiation despite high gene flow between sympatric rainbow smelt (*Osmerus mordax*, Mitchill). *Molecular Ecology*, **12**, 315–330.

Schluter D (1996) Ecological speciation in postglacial fishes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 351, 807–814.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN, *Version 2.000: a software for population genetics data analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

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.

Svärdson G (1961) Young sibling fish species in northwestern Europe. In: *Vertebrate Speciation* (ed. Blair WF), pp. 498–513. University of Texas Press, Austin.

Taylor EB (1999) Species pairs of north temperate freshwater fishes: taxonomy, evolution, and conservation. *Reviews in Fish Biology and Fisheries*, **9**, 299–324.

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, Foote CJ (1991) Critical swimming velocities of juvenile sockeye salmon and kokanee, the anadromous and non-anadromous forms of *Oncorhynchus nerka* (Walbaum). *Journal of Fish Biology*, **38**, 407–419.

Taylor EB, McPhail JD (2000) Historical contingency and ecological determinism interact to prime speciation in sticklebacks, *Gasterosteus*. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **267**, 2375–2384.

Taylor EB, Foote CJ, Wood CC (1996) Molecular genetic evidence for parallel life history evolution within a pacific salmon (sockeye salmon and kokanee, *Oncorhynchus nerka*). *Evolution*, **50**, 401–416.

Tessier N, Bernatchez L (2000) A genetic assessment of single versus double origin landlocked Atlantic salmon (*Salmo salar*) from Lake Saint-John, Quebec, Canada. *Canadian Journal of Fisheries and Aquatic Sciences*, **57**, 797–804.

Thompson CE, Taylor EB, McPhail JD (1997) Parallel evolution of lake–stream pairs of threespine stickleback (*Gastersteus aculeatus*) inferred from mitochondrial DNA variation. *Evolution*, **51**, 1955–1965.

Travisano M, Mongold JA, Bennett AF, Lenski RE (1995) Experimental tests of the roles of adaptation, chance, and history of evolution. *Science*, **267**, 87–90.

Turgeon J, Bernatchez L (2001) Mitochondrial DNA phylogeography of lake cisco (*Coregonus artedi*): evidence supporting extensive secondary contacts between two glacial races. *Molecular Ecology*, **10**, 987–2001.

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.

This study is part of D.J. Fraser's PhD thesis on the evolutionary biology and conservation of *Salvelinus fontinalis* in central Québec, Canada, a collaborative project in the Bernatchez laboratory between the Cree Nation of Mistissini, Cree Regional Authority of Québec and Université Laval. The authors' career choices are likely based on a deterministic desire to understand the nature of things, but also historically contingent upon past exposure to aquatic environments.