GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Nucleotide variation in the mitochondrial genome provides evidence for dual routes of postglacial recolonization and genetic recombination in the northeastern brook trout (*Salvelinus fontinalis*)

B.L. Pilgrim[1], R.C. Perry[2], J.L. Barron[1] and H.D. Marshall[1]

[1]Department of Biology, Memorial University of Newfoundland and Labrador, St. John's, Newfoundland and Labrador, Canada
[2]Department of Environment and Conservation, Wildlife Division, Government of Newfoundland and Labrador, Corner Brook, Newfoundland and Labrador, Canada

Corresponding author: B.L. Pilgrim
E-mail: brettneylp@mun.ca

Genet. Mol. Res. 11 (3): 3466-3481 (2012)
Received August 16, 2011
Accepted May 2, 2012
Published September 26, 2012
DOI http://dx.doi.org/10.4238/2012.September.26.2

ABSTRACT. Levels and patterns of mitochondrial DNA (mtDNA) variation were examined to investigate the population structure and possible routes of postglacial recolonization of the world's northernmost native populations of brook trout (*Salvelinus fontinalis*), which are found in Labrador, Canada. We analyzed the sequence diversity of a 1960-bp portion of the mitochondrial genome (NADH dehydrogenase 1 gene and part of cytochrome oxidase 1) of 126 fish from 32 lakes distributed throughout seven regions of northeastern Canada. These populations were found to have low levels of mtDNA diversity, a characteristic trait of populations at northern extremes, with significant structuring at the level of the watershed. Upon

comparison of northeastern brook trout sequences to the publicly available brook trout whole mitochondrial genome (GenBank AF154850), we infer that the GenBank sequence is from a fish whose mtDNA has recombined with that of Arctic charr (*S. alpinus*). The haplotype distribution provides evidence of two different postglacial founding groups contributing to present-day brook trout populations in the northernmost part of their range; the evolution of the majority of the haplotypes coincides with the timing of glacier retreat from Labrador. Our results exemplify the strong influence that historical processes such as glaciations have had on shaping the current genetic structure of northern species such as the brook trout.

Key words: Mitochondrial DNA; Genetic recombination; Brook trout; Postglacial recolonization; Glacial refugia; *Salvelinus fontinalis*

INTRODUCTION

Brook trout (*Salvelinus fontinalis* Mitchill, 1814) are ubiquitous across their native range of eastern North America (Power, 1980), existing as far north as Labrador, the mainland component of the province of Newfoundland and Labrador, Canada, and as far south as Georgia, USA. Populations have successfully adapted to a broad range of habitats, from small creeks to large rivers and lakes of various depths (Power, 1980), in which fish may exhibit sea-run or freshwater resident life histories. These salmonids therefore have intriguing ecologies and life histories. As a northern species in which population structure and distribution have been significantly affected by historical events such as glaciation, brook trout are also of interest from an evolutionary perspective.

The last glacial maximum, the Wisconsinan, which occurred ~18,000 years ago (Dyke and Prest, 1987), had a profound effect on the present-day distribution of genetic variation across the ranges of species, especially in populations at range extremes. As glacial ice advanced, northern species were restricted to southern unglaciated refugia. Relative to present-day northern populations, populations that persisted in isolation in southern refugia have had more time to accumulate variation (Hewitt, 1996; Bennett, 1997) while simultaneously diverging from populations in other refugia (Hewitt, 1996, 2004). The signature of these processes is evident in the distribution and genetic structure of contemporary populations. Northern regions typically have fewer species (Pielou, 1991), each with lower genetic diversity-a pattern referred to by Hewitt (1996) as "southern richness, northern purity". This pattern is exemplified by freshwater fish including brook trout; in northern, formerly glaciated regions, the number of species is relatively low, and populations within species display less genetic variation than that in species in southern locations (Danzmann et al., 1998).

Remarkably, certain brook trout populations at the northern limit of the range demonstrate complete introgression of mitochondrial DNA (mtDNA) from Arctic charr (*Salvelinus alpinus*). Originally documented by Bernatchez et al. (1995) in a brook trout population in Lake Alain, eastern Québec, this phenomenon has also been identified in allopatric populations of brook trout throughout the Rocheuse River subdrainage of the

Portneuf basin (Glémet et al., 1998). Because introgression in brook trout appears to be restricted to drainages where Arctic charr are absent, Bernatchez et al. (1995) have deduced that the timing of introgression was historical. A difference in the founding population sizes of brook trout and Arctic charr may have contributed to this event: the likelihood of hybridization is higher when one parental population is substantially smaller than the other and hence prone to outcrossing (Hubbs, 1955; Avise et al., 1988). Across the region of introgression, brook trout have a higher level of nuclear (allozyme) genetic diversity (McGlade, 1981) than that of Arctic charr (Kornfield et al., 1981; Anderson et al., 1983), consistent with a larger founding population size. Because the introgression event was likely historical, yet current brook trout populations retain Arctic charr mtDNA, the Arctic charr mtDNA genome may confer a selective advantage to these brook trout. Arctic charr mtDNA evolved in a cold environment; hence, the respiratory enzymes encoded by mtDNA may provide an advantage to brook trout at the northern end of their range in which conditions are much cooler (Glémet et al., 1998). The documentation of this event is a powerful illustration of how historical processes shape modern genetic structure.

MtDNA restriction fragment length polymorphism variation has been used to investigate large-scale patterns of postglacial recolonization of brook trout throughout their native range (Danzmann et al., 1998). Following the retreat of glacial ice from eastern North America after the Wisconsinan glacial maximum, watersheds are hypothesized to have been re-colonized by fish from three refugia: Mississippian, Atlantic, and Acadian. Fish with Mississippian origins dominate the southern Great Lakes region, whereas fish from the Atlantic refugium appear to have re-colonized the Lake Ontario watershed, the Atlantic coastal region, and to some extent, Ontario, Québec, and eastern Canada. Throughout eastern Canada a single haplotype, haplotype 1, postulated to have been present in both the Mississippian and the Atlantic refugia, is predominant. Contribution by the Acadian refugium is evident, however, from the presence of 10 private haplotypes and a lack of haplotype 2 fish.

Despite a comprehensive examination of brook trout mtDNA variation by Danzmann et al. (1998), the phylogeography of brook trout in Labrador, Canada, the northernmost part of the range, remains unknown. One possible scenario for the colonization of Labrador by brook trout is that fish from the Atlantic and Mississippian refugia dispersed through Québec and into Labrador, whereas other fish from the Atlantic refugium invaded coastal regions (Black et al., 1986). To determine the refugial origins and specific routes of postglacial recolonization of brook trout in the northern extreme of their natural range, we examined 1960-bp mtDNA sequences representing 126 fish from seven regions in eastern Canada, most extensively in Labrador. In addition to presenting the results of our phylogeographic analyses aimed at determining the ancestral origins of northeastern brook trout, we describe herein patterns of intraspecific haplotype and nucleotide diversity and discuss an unusual region of high variability in the brook trout genome.

MATERIAL AND METHODS

Sample collection

Brook trout samples were collected via gill netting by employees from the Wildlife Division of the Department of Environment and Conservation. Sampling took place in the

summer months from 2003 to 2008. Fin clippings were collected and stored in envelopes at -20°C. Six regions across eastern Canada were sampled - northern Labrador, west-central Labrador, southeastern Labrador, insular Newfoundland, Nova Scotia, and New Brunswick- including 35 lakes in 22 watersheds (Table 1; Figure 1). Representative Québec and Atikonak Lake (west-central Labrador) samples were provided by Dr. Louis Bernatchez (Université Laval) and Dr. Steven M. Carr (Memorial University of Newfoundland and Labrador), respectively. Regions in Labrador have been defined by Anderson (1985) based on geology, vegetation, and fish distributions, with watersheds in each region draining to a specific inlet or region of the coast.

Table 1. Sampled brook trout populations across eastern Canada, including population code, sample size (N), latitude and longitude.

Region	Watershed	Lake	N	Latitude	Longitude
Northern Labrador					
1	No Name	Shallow Lake	2	57°41'16"	63°21'56"
2	Saputit Brook	Saputit Lake	2	57°27'39"	62°36'26"
3	Kogaluk River	Cabot Lake	2	56°25'04"	63°38'21"
4	Anaktalik River	Anaktalik Lake	1	56°29'50"	62°51'38"
5	Iladlivik Brook	Walkabout Lake	2	56°19'39"	63°09'34"
6	No Name	Mistake Lake	1	56°25'01"	63°38'21"
7	Konrad Brook	Konrad Lake	2	56°13'22"	62°43'24"
West-Central Labrador					
8	Atikonak River	Atikonak Lake	17	52°40'00"	64°34'60"
9	Traverspine River	No Boat Pond	3	53°08'05"	60°38'05"
10	Traverspine River	The Right Lake	2	53°00'08"	60°45'54"
11	Kenamu River	Mercier Lake	1	52°55'05"	60°43'25"
12	Kenamu River	Brennan Lake	2	52°57'22"	60°15'25"
13	Kenamu River	Nikki's Pond	3	52°36'18"	60°25'55"
Southeast Labrador					
14	Eagle River	Fred's Lake	2	52°52'01"	59°43'39"
15	Eagle River	No Name Lake	3	52°40'58"	65°26'35"
16	Eagle River	NAP Pond	3	52°35'06"	59°04'44"
17	Eagle River	Nippard's Lake	3	53°05'10"	58°50'29"
18	Eagle River	Osprey Lake	3	52°44'39"	58°35'01"
19	Eagle River	Dead Dog Pond	2	52°45'52"	58°25'33"
20	St. Augustine River	St. Augustine	4	52°35'15"	59°18'21"
21	St. Augustine River	Bog Lake	3	52°31'45"	59°04'03"
22	Paradise River	Keith's Lake	3	52°59'23"	57°49'15"
23	Paradise River	Crooked Lake	3	53°20'28"	57°34'44"
24	Alexis River	Alexis Pond	2	52°31'59"	57°03'16"
25	Alexis River	Handkerchief Pond	2	52°31'04"	57°01'34"
26	Alexis River	Feeder Pond	3	52°32'10"	56°29'36"
27	Gilbert River	Tilt Pond	1	52°42'17"	56°18'39"
28	Gilbert River	Gilbert Lake	2	52°41'14"	56°11'57"
29	St. Lewis River	Curl's Pond	2	52°24'50"	56°00'52"
30	St. Mary's River	Mary's Harbour Big Pond	3	52°18'53"	56°01'24"
Newfoundland					
31	Middle Brook	Butt's Pond	15	48°49'22"	54°16'10"
32	Salmonier River	Little Gull Pond	18	47°15'34"	53°19'33"
Nova Scotia					
33	River Denys	Alder Brook	1	45°58'12"	64°35'10"
34	Salmon River	Farnham Brook	1	45°22'55"	63°16'14"
New Brunswick					
35	Saint John River	Moose Lake	1	46°51'54"	66°47'16"
36	Kennebecasis River	Walton Lake	2	45°36'39"	65°19'16"
Québec					
37	Rupert River	Mistassini Lake	4	50°59'49"	73°38'21"

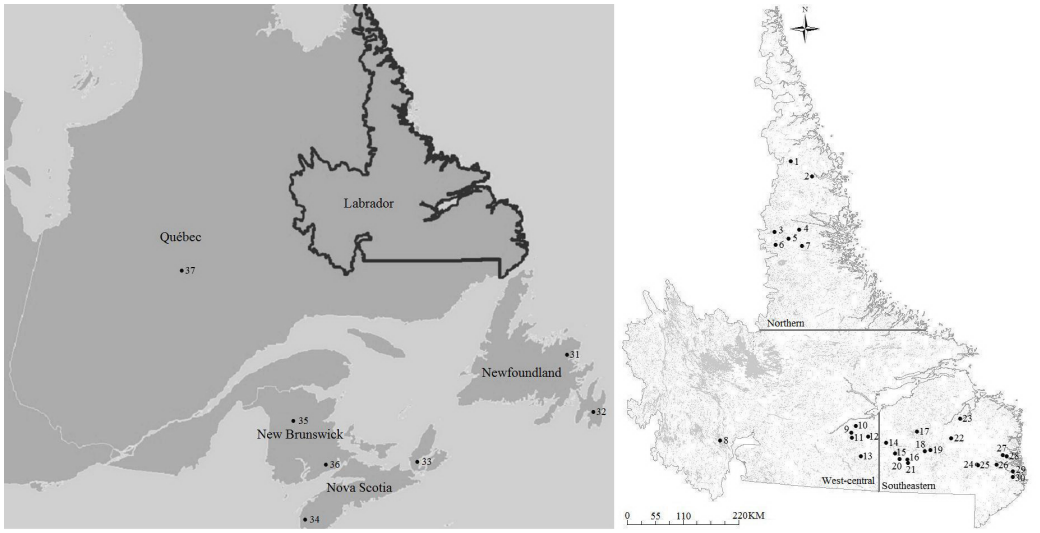

Figure 1. Lakes sampled for brook trout across eastern Canada. Population codes are given in Table 1.

DNA extraction, polymerase chain reaction (PCR) amplification, and sequencing

Genomic DNA was extracted from ~25 mm^2 of caudal fin tissue using a Qiagen QIAamp DNA Mini Kit (Qiagen Inc., Mississauga, Canada) following the tissue protocol. Two gene regions [nicotinamide adenine dinucleotide (NADH) dehydrogenase 1 (ND1) and cytochrome oxidase 1 (CO1)] from the mitochondrial genome were amplified using three sets of primers as follows: ND1-1F: 5'-AAGTGGCAGAGCCCGGTAATTG-3'; ND1-1R: 5'-AATGGTCCTCCAGCGTATTCTAC-3'; ND1-2F: 5'-CGTAGCTCAAGAAAG CATCTGAC-3'; ND1-2R: 5'-TATTTGGGGTATGGGCCCGAAAG-3'; CO1-F: 5'-TCAACC AACCACAAAGACATTGGCAC-3'; CO1-R: 5'-TAGACCTCTGGGTGGCCAAAGAATC A-3'. All primers were tailed by attaching one of the following M13 sequences to the 5'-end: 5'CAGGAAACAGCTATGAC-3' (forward primers) or 5'-GTAAAACGACGGCCAGT-3' (reverse primers). PCRs (25 µL) containing 2.5 µL 10X buffer (Qiagen Inc.), 0.5 µL 10 mM dinucleotide triphosphates (New England Biolabs, Pickering, Canada), 10 µM forward and reverse primers (1 µL each; Operon, Huntsville, USA), 0.2 µL 5 U/µL HotStar *Taq* Plus DNA Polymerase (Qiagen Inc.), and 2 µL DNA template (2-570 ng) were subjected to the following thermal profile on an Eppendorf Mastercycler EP Gradient S (Hamburg, Germany): 95°C for 5 min followed by 40 cycles at 93°C for 30 s, 55°C for 30 s, and 72°C for 2 min, and a final extension of 72°C for 5 min. Successfully amplified PCR products were purified using a multi-well filter plate (Pall Life Sciences, Port Washington, USA) according to manufacturer instructions. Sequencing reactions were carried out in 20-µL reactions composed of 0.5 µL BigDye® Terminator v3.1 Cycle Sequencing Ready Reaction mix (Applied Biosystems Inc., Carlsbad, USA), 2 µL 1.6 pmoL/µL forward or reverse M13 primer, 5 µL 5X sequencing buffer, and 2 µL DNA template. The following PCR profile was used: 96°C for 6 min and 25 cycles at 96°C for 10 s followed by 5 s at 50°C and 4 min at 60°C. Sequencing reaction products were purified via ethanol precipitation and resuspended in 10 µL Hi-Di Formamide

(Applied Biosystems Inc.). Samples were electrophoresed in an Applied Biosystems 3730 DNA Analyzer using the Sequencing Analysis v. 5.2 Software at the Genomics and Proteomics Facility, Core Research Equipment and Instrument Training Network, Memorial University of Newfoundland.

Data analysis

Sequences were edited and aligned through comparison to the homologous ND1 and CO1 genes from the complete mitochondrial genome of *S. fontinalis* available from GenBank (AF154850) using Sequencher v4.8 (Gene Codes, Ann Arbor, USA), and contiguous sequences for each individual were assembled. Numbers of variable sites, transitions, transversions, synonymous and nonsynonymous substitutions, and average pairwise sequence diversity were determined for each gene and the concatenated sequence using MEGA 4.0 (Tamura et al., 2007). Individual haplotypes were identified using DnaSP v.5 (Librado and Rozas, 2009) in combination with visual inspection to account for gap polymorphisms.

A hierarchical analysis of molecular variance (AMOVA) was carried out in Arlequin v. 3.11 (Excoffier et al., 2005) to investigate the partitioning of genetic variation in Newfoundland and Labrador. Brook trout were grouped in two ways: 1) according to lake, then to watershed; 2) according to watershed, then to region. Estimates of the variance components were quantified using conventional *F*-statistics over 1000 bootstrap replicates. Arlequin was also used to calculate pairwise F_{ST} values among watersheds in Newfoundland and Labrador based on conventional *F*-statistics (F_{ST}). The pairwise F_{ST} matrix was used to construct a neighbour-joining dendrogram of watersheds in MEGA 4.0 (Tamura et al., 2007). Differentiation between groups of watersheds revealed by the dendrogram was quantified with F_{ST} using Arlequin.

The Sneato v.2.0 program (http://www.xmission.com/~wooding/Sneato/index.html) was used to construct a haplotype network, and the minimum spanning network option under AMOVA in Arlequin 3.11 (Excoffier et al., 2005) was used to find alternate connections. A neighbour-joining phylogeny of haplotypes was constructed from a maximum composite likelihood distance matrix using pairwise deletion and tested with 1000 bootstraps in MEGA 4.0. Haplotype 13 was excluded from the phylogeny owing to its high level of missing data.

Dates were inferred at the nodes of major clusters using the divergence time and linearized tree options in MEGA 4.0. We used an evolutionary rate of 3.018%/million years calculated from the sequence divergence between brook trout from this study and Arctic charr reference sequence AF154851 (overall mean distance = 0.06036) and the time of divergence (~1 million years) between brook trout and Arctic charr reported by Brunner et al. (2001).

To investigate the possibility of demographic expansion of brook trout populations into eastern Canada, Tajima (1989) *D* and a pairwise mismatch distribution among all sequences were calculated using DnaSP v.5.0. This analysis was repeated on a subset of the data containing all representatives of the most common haplotype and its derivatives.

RESULTS

mtDNA diversity of brook trout in eastern Canada

Contiguous 1960-bp sequences containing a 49-bp segment of transfer RNA (tRNA)-

Leu 1, the complete ND1 gene, the complete tRNA-Ile and tRNA-Gln, an 18-bp segment of the tRNA-Met, and a 655-bp segment of the CO1 gene were assembled for 126 brook trout. Sequences were submitted to GenBank and assigned the accession Nos. JF979036-JF979060. A total of 22 variable sites were identified (Table 2); one in tRNA-Leu 1, 10 in the ND1 gene, one in tRNA-Ile, and the remaining 10 in the CO1 gene. Of the 22 substitutions, 18 were transitions and four were transversions. All protein-coding gene substitutions were synonymous with the exception of two, both found in the ND1 gene (Val ↔ Ile at position 350, and Val ↔ Met at position 988; see Table 2). Average pairwise sequence diversity among all individuals was 0.00087 in the ND1 gene and 0.00080 in the CO1 gene; overall pairwise sequence diversity was 0.00094 (see Table 2).

Table 2. Number of variable sites, transitions (Ts), transversions (Tv), synonymous (dS) and nonsynonymous (dN) substitutions, and average pairwise diversity (π) obtained within each protein-coding region characterized, and over all brook trout sampled in eastern Canada.

Gene	No. of variable sites	Ts	Tv	dS	dN	π
ND1	12	10	2	9	2	0.00087
CO1	10	8	2	10	0	0.00080
Overall	22	18	4	19	2	0.00094

Unusual pattern of variability in the brook trout mtDNA genome

While comparing our sequences to GenBank sequence AF154850, the whole mtDNA genome of a brook trout from a population in Québec (Doiron et al., 2002), we discovered an unusual region of high diversity. Eleven third-position base changes within a 220-bp region of the ND1 gene were discovered in which all the fish we sequenced differed from the reference sequence (Figure 2). The majority (nine) of these positions were transitions, but two were transversions. Only a single transition was nonsynonymous, at position 1084, resulting in the conversion of Val ↔ Ile (a conservative amino acid change). Interestingly, the 11 variable sites in the 220-bp region in the reference sequence perfectly matched a sequence in Arctic charr (GenBank AF154851), whereas the nucleotides at those positions in the fish in this study matched the sequence of a brook trout from a naturalised British Columbia population (GenBank AF126000; Taylor et al., 1999; see Figure 2).

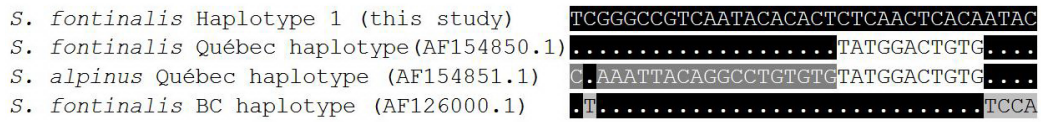

Figure 2. Alignment of a portion of the ND1 gene in four *Salvelinus* species, with conserved portions of sequence highlighted in the same shade. When considering the 220-bp region of high diversity between our haplotype 1 brook trout and the reference brook trout sequence from GenBank (AF154850), the sequence is conserved between the reference brook trout and Arctic charr.

Population structure of brook trout in eastern Canada

To investigate the distribution of genetic diversity in Newfoundland and Labrador brook

trout populations, two separate hierarchical AMOVAs were performed. First, brook trout grouped according to lake and subsequently into corresponding watershed partitioned most of the variation among watersheds (64.98%; F_{CT} = 0.6499, P < 0.00001), followed by among lakes within watersheds (30.22%; F_{SC} = 0.8629, P < 0.00001). The least amount of variation was detected within lakes (4.80%; F_{ST} = 0.9520, P < 0.00001). Next, fish were grouped by watershed, then into regions. Most of the variation was detected within watersheds (71.30%; F_{ST} = 0.2870, P < 0.00001), followed by among watersheds within regions (28.78%; F_{SC} = 0.2876, P < 0.00001). The amount of variation partitioned among regions (-0.07%) was not significantly different from zero (F_{CT} = -0.0007, P > 0.1).

Regional distribution of haplotypes and population differentiation

The 22 variable sites identified among 126 brook trout define 13 haplotypes, which we numbered according to frequency (Table 3). The geographic distribution of these haplotypes was investigated to identify patterns of heterogeneity that could correspond to distinctive post-glacial founding lineages (Figure 3). Haplotypes 1 and 2 (accounting for 61.1 and 18.2% of all fish, respectively) are clearly heterogeneously distributed (P = 0; Fisher exact test). Haplotype 1 is the most common haplotype in Newfoundland and Labrador. Haplotype 2 is found in every region of Labrador except the easternmost part of southeastern Labrador. Haplotype 1 is present in Québec, haplotype 2 in New Brunswick, and neither in Nova Scotia. Newfoundland is characterized by a high frequency of region-specific (and watershed-specific) haplotype 3 and the presence of a distinctive haplotype 4, also seen in the westernmost watershed in west-central Labrador but nowhere else. Québec and New Brunswick share haplotype 5. All regions except northern Labrador contain unique (singleton) or at least watershed-specific haplotypes. Nova Scotia contains only two unique haplotypes.

Table 3. Haplotype frequency (N) and distribution among *Salvelinus fontinalis* populations in eastern Canada.

Haplotype	N	Regions	Variable sites
1	77	Northern Labrador, West-Central Labrador, Southeast Labrador, Newfoundland, Québec	AACGGGGACA GATATTCCAT GG
2	23	Northern Labrador, West-Central Labrador, Southeast Labrador, New Brunswick	C.......T. ....
3	7	Newfoundland	.....AA... .... --
4	5	West-Central Labrador, Newfoundland	CGTA...C.G ..CCC...GC A.
5	4	New Brunswick, Québec	C......... ....
6	2	Southeast Labrador	.... .A
7	2	West-Central Labrador	.... .G........ ..
8	1	Newfoundland	....C.... ..
9	1	New Brunswick	C......... .....C.... ..
10	1	Nova Scotia	C......... A......... ..
11	1	Southeast Labrador	-....-.... ......A... ..
12	1	Southeast Labrador	-....-.... ....T.. ..
13	1	Nova Scotia	....A..... ..-------- --

Haplotypes are numbered according to frequency. Variable sites are with respect to haplotype 1.

Regional heterogeneity in haplotype distribution was supported by the neighbour-joining dendrogram of pairwise F_{ST} values among watersheds in Newfoundland and Labrador (Figure 4). Two well-differentiated clusters (F_{ST} = 0.4464, P < 0.00001) are apparent, one containing some northern, west-central, and southeastern Labrador populations and the other comprised of watersheds from all regions of Newfoundland and Labrador.

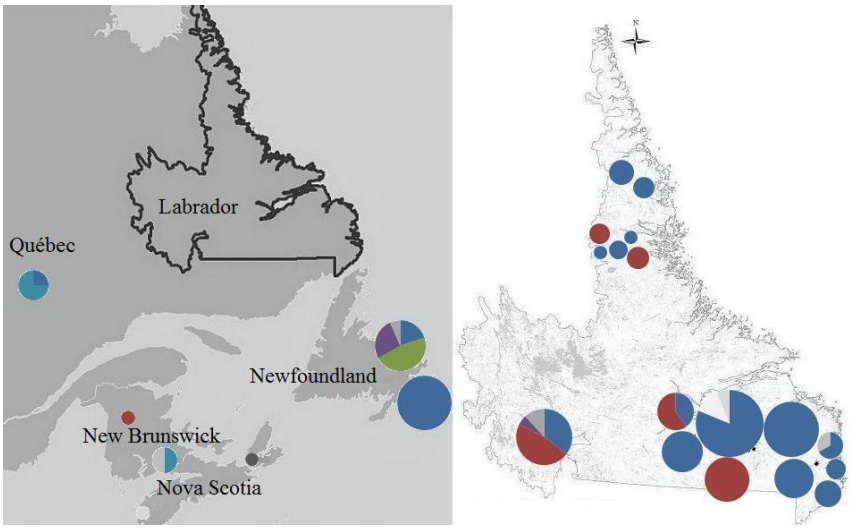

Figure 3. Geographical distribution of the 13 haplotypes among 126 brook trout in eastern Canada within each watershed. Watershed specific haplotypes are shaded whereas shared haplotypes appear in color. Pie charts are sized according to relative sample size.

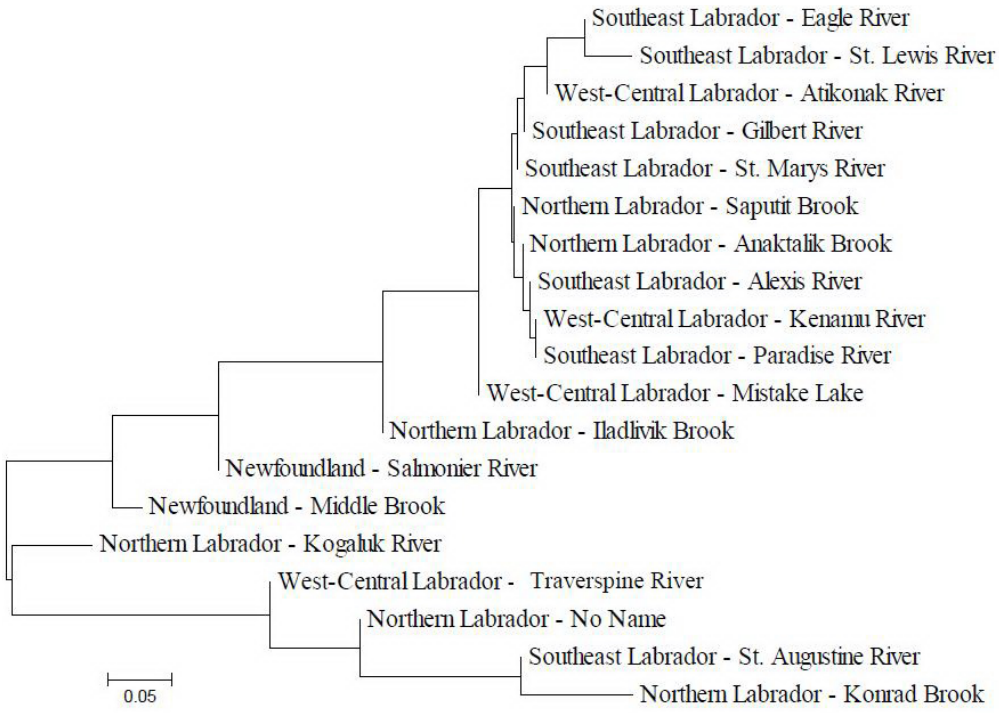

Figure 4. Neighbour-joining dendrogram of watersheds throughout Newfoundland and Labrador inferred from pairwise F_{ST} values among watersheds, based on conventional *F*-statistics.

Relationships among haplotypes

The haplotype network depicted in Figure 5A demonstrates that the majority of haplotypes (3, 5, 6, 7, 8, 11, 12, 13) are connected to haplotype 1, the most frequently observed haplotype, through only 1 or 2 bp changes. Haplotype 9 is connected to haplotype 1 by two changes through haplotype 8, whereas haplotype 10 is connected by two changes through haplotype 5. Haplotype 2, the second most frequent haplotype, is also connected to haplotype 1 by 2 bp changes through haplotype 5. A notable exception to this pattern is haplotype 4; this haplotype differs from haplotype 1 by 12 bp through haplotype 5. Also notable is that only haplotype 1, the most frequent haplotype, has a number of derivative haplotypes; haplotype 2, the second most frequent haplotype, does not. Assuming that midpoint rooting identifies the most ancestral haplotype, the neighbour-joining phylogeny of haplotypes (Figure 5B) indicates that haplotype 4 is ancestral and that haplotypes 2, 5, and 10 are basal to a clade including haplotypes 1, 3, 6, 7, 8, 9, 11, and 12, making it the most recently evolved clade.

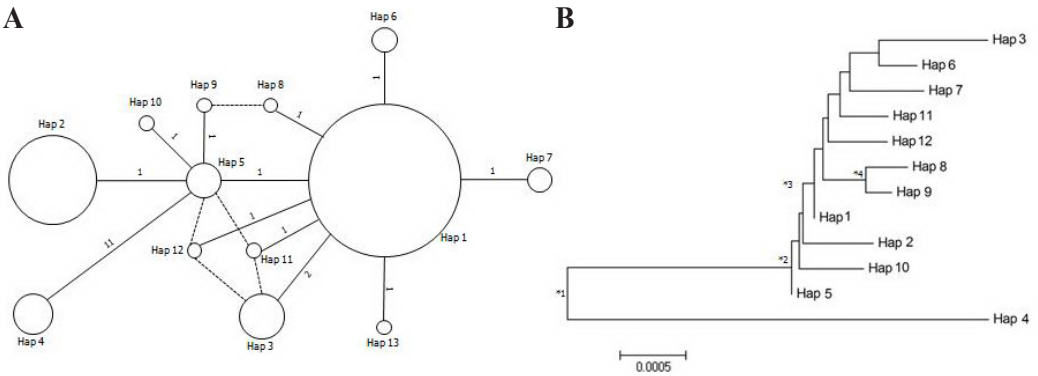

Figure 5. A. Haplotype network of the ND1 gene and partial CO1 gene of brook trout in eastern Canada. Haplotypes are numbered and nodes are sized according to relative frequency found in Table 3. Numbers on the connections indicate the number of mutational steps between haplotypes. Dashed lines indicate alternate connections. **B.** Neighbour-joining dendrogram relating 12 of the mtDNA haplotypes of brook trout in eastern Canada, based on sequence analysis of the ND1 gene and a portion of the CO1 gene. The dendrogram was constructed from a maximum composite likelihood distance matrix using pairwise deletion. Numbered asterisks refer to the following period of time since divergence: *1 = 92,980 years ago, *2 = 12,280 years ago, *3 = 10,260 years ago, *4 = 8,500 years ago.

Based on the evolutionary rate of 3.018%/million years, the most ancestral haplotype, haplotype 4, evolved 92,980 years ago. The most basal clade (haplotypes 2, 5, and 10) arose 12,280 years ago, whereas the clade including haplotypes 1, 3, 6, 7, 11, and 12 dates to 10,260 years ago. The most recently evolved clade including haplotypes 8 and 9 is approximately 8 500 years old.

Demographic history

We measured Tajima (1989) *D* to investigate the possibility of demographic expan-

sion of brook trout populations in Labrador. Tajima's *D* for all brook trout across eastern Canada measured -1.33 but was not significantly negative ($P > 0.10$). To investigate the possibility of demographic expansion associated with haplotype 1 and its derivatives, Tajima's *D* was recalculated for fish represented by these haplotypes alone and measured -1.30, remaining non-significantly negative ($P > 0.10$).

DISCUSSION

We investigated the population genetic structure of brook trout in the northernmost extreme of their natural range by sequencing two mtDNA genes in 126 fish from seven regions in northeastern Canada. Our goals were to describe levels of genetic diversity and use patterns of variation within and among populations to make inferences about the postglacial recolonization and demographic history of brook trout in this region. This study expands earlier research by Danzmann et al. (1998) to complete the characterization of brook trout mtDNA diversity throughout their entire continental range. The fine scale at which we investigated the structure allows a detailed analysis of the refugial origins of the northernmost populations of brook trout. This scale may be important in the broader context of conservation genetics, as populations at northern extremes are often characterised by lower diversity and are more vulnerable to stochastic events (Anderson et al., 2011).

Low genetic variability characterizes brook trout populations in eastern Canada

Pairwise sequence divergences among brook trout populations in eastern Canada are higher in the ND1 gene (0.00087) than in the CO1 gene (0.00080) but are very low overall (0.00094). Regardless, two nonsynonymous changes occur in the ND1 gene, one of which (Val ↔Met) involves the exchange of dissimilar amino acid side chains. Low genetic diversity in northeastern brook trout populations was also illustrated by the haplotype distribution; two haplotypes (1 and 2) account for 78.6% of fish, whereas 76.2% of fish are represented by haplotype 1 and its derivatives.

The level of diversity we observed is low compared to that found in other studies of brook trout and other North American freshwater fish. For example, in a brook trout population in Ontario, nine haplotypes were identified among 33 individuals (300 bp of control region; Bernatchez and Danzmann, 1993), compared to 13 haplotypes among 126 individuals in our study. In the Ontario population, average pairwise sequence diversity measured 0.01, an order of magnitude higher than that we observed in brook trout throughout eastern Canada. Taylor et al. (1999) characterized mtDNA variation in bull trout (*Salvelinus confluentus*) populations in northwestern North America and reported a pairwise sequence diversity of 0.079, almost two orders of magnitude higher than the diversity we observed in northeastern brook trout. Arctic charr populations across their entire range (including northern regions such as alpine Europe, Norway, and Alaska) were also much more diverse (average pairwise sequence diversity = 0.0203; Brunner et al., 2001) than those found in our study of brook trout. The pairwise sequence diversity of lake whitefish (*Coregonus clupeaformis*), which have a northern distribution from Alaska to Labrador, was 0.007 (Bernatchez and Dodson, 1991) - again much higher than that of northeastern brook trout.

Low diversity is characteristic of northern populations owing to the bottleneck effect

during population expansion. Hence, low mtDNA diversity was not an unexpected finding. Notably, levels of mtDNA diversity found in northern brook trout are much lower than those observed in other northern-distributed freshwater fish such as Arctic charr and whitefish. Both of these species likely colonized their current northern distribution from a northern refugium (Bernatchez and Dodson, 1991; Brunner et al., 2001), however, brook trout populations are hypothesized to have originated from southern refugia (Danzmann et al., 1998). As a result, northern brook trout may have experienced a larger loss of alleles owing to bottlenecking during expansion from more distant source populations than those of northern Arctic charr and whitefish populations which are in closer geographic proximity to their founding populations.

Unusual region of high diversity in the ND1 gene of brook trout

A comparison of the brook trout sequences in this study with the mitochondrial genome of brook trout in GenBank (AF154850) revealed a 220-bp region in the ND1 gene in which there were 11 differences between the brook trout in this study and the reference sequence. This difference corresponds to 5% sequence diversity, a very high level compared with the average pairwise sequence diversity reported intra-specifically in brook trout (the present study and Bernatchez and Danzmann, 1993). The nucleotide variants in our fish match a GenBank entry of brook trout from a naturalised population in British Columbia, Canada (GenBank AF126000; the only brook trout mtDNA sequence available in GenBank other than the whole mtDNA genome sequence), whereas the nucleotide variants in the whole mtDNA genome sequence (GenBank AF154850) match an Arctic charr entry (GenBank AF154851) exactly. Except for this 220-bp region, the consensus brook trout mtDNA sequence obtained in this study is identical to the whole mtDNA genome over the 1960 bp sequenced. Complete replacement of brook trout mtDNA with Arctic charr mtDNA has been reported in brook trout populations in eastern Québec (Bernatchez et al., 1995; Doiron et al., 2002). Therefore, a reasonable explanation for this phenomenon is that the whole mtDNA genome sequence represents a brook trout with mtDNA that has recombined with Arctic charr mtDNA.

Evidence of mtDNA recombination has been documented in a number of other species, including vertebrates (Kraytsberg et al., 2004; Ujvari et al., 2007), invertebrates (Lunt and Hyman, 1997; Ladoukakis and Zouros, 2001), plants (Städler and Delph, 2002; Jaramillo-Correa and Bousquet, 2005), and fungi (Saville et al., 1998). The most commonly suggested mechanism for mtDNA recombination is by means of paternal leakage of mitochondria during fertilization (Lunt and Hyman, 1997; Ladoukakis and Zouros, 2001; Städler and Delph, 2002; Jaramillo-Correa and Bousquet, 2005) which could be likely between brook trout and Arctic charr because the two species interbreed. Paternal leakage is more likely to occur in hybridization zones because reproductive barriers are not always entirely developed within these regions, and the mechanism that eliminates male mitochondria frequently may not be fully functional (Wagner et al., 1991; Jaramillo-Correa and Bousquet, 2005). As previously mentioned, the instance documented in this study occurs in a region (Québec) in which hybridization of the two species has been documented (Bernatchez et al., 1995; Doiron et al., 2002). This increases the likelihood that reproductive barriers of the two species were not fully developed at the time, thus increasing the chances of paternal leakage and therefore mtDNA recombination.

Brook trout and Arctic charr hybridization has also been reported in northern Labrador (Hammar et al., 1991), in which the hybrids are physically most similar to brook trout but

can be distinguished by protein electrophoresis. It is interesting that none of the brook trout in our study appear to contain any Arctic charr mtDNA even though the habitats of the two species overlap in our study area, which is located in a northern environment.

Doiron et al. (2002) has suggested that the introgression event could be advantageous for brook trout in northern regions because these fish would have mitochondrial respiratory enzymes encoded by Arctic charr mtDNA that evolved in a cold environment. In support of this theory, they found evidence that some regions of the mtDNA genome, especially in the ND2 and ND5 genes, are under selection, but none of those regions occur in ND1. We found only one conservative amino acid substitution (Val ↔ Ile) in the 220-bp region of high diversity, also suggesting no influence of selection in the ND1 gene.

The possibility that the 220-bp region of high diversity represents a pseudogene is poorly supported, as the nucleotide variants show the typical pattern of mtDNA variation; they do not introduce any frameshifts or stop codons and, with the exception of two transversions, are third-position transitions. Although we cannot rule out laboratory error, we found no evidence for this occurrence, as none of the combinations of primers used by Doiron et al. (2002) to amplify mtDNA sequences would produce PCR amplicons specifically of the size of the region of unusual diversity (220 bp).

Differentiation among Newfoundland and Labrador brook trout populations

Despite low levels of variation within populations of brook trout in eastern Canada, significant structure was revealed among watersheds by the hierarchical AMOVA. This pattern was also observed by Danzmann et al. (1998), who found that ~40% of variation among units consisting of drainages over a broad geographical scale. Danzmann et al. (1998) attributed this observation to heterogeneity in the founding populations of each watershed. Hence, during postglacial recolonization each watershed was founded primarily by a single or small number of haplotypes that differ from each other by chance; subsequent drift and lack of gene exchange among watersheds has maintained the heterogeneity. Differentiation at the level of the watershed is consistent with the generally potamodromous life history of these fish in Labrador based on tagging data (Robert Perry, personal communication); even if fish migrate to the sea, they remain near the mouth of the river. Significant differentiation also occurred among lakes but to a lesser degree; because many watersheds were characterized by a single haplotype, the resolution is simply not present in our data to make strong inferences about this phenomenon. Although no regional component to the mtDNA differentiation was obvious, phylogeographic structuring at this level may have been shaped by postglacial recolonization history.

Inferences about postglacial recolonization history

The regional distribution of haplotypes suggests genetic contributions from at least two refugial sources to the postglacial recolonization of Labrador and environs by brook trout. The absence of haplotype 2 fish and the presence of regionally restricted haplotypes in watersheds in the southeastern-most part of Labrador and insular Newfoundland suggest that these regions may have an founding source different than that of the rest of Labrador. Western Labrador (part of the designated region of west-central Labrador) does contain haplotype 2 fish but is also characterized by the presence of regionally restricted haplotypes suggesting a mix-

ture of the two refugial sources in this watershed. West Labrador and Newfoundland are also characterized by the presence of unusual ancestral haplotype 4 fish, providing further support for shared ancestry in these two areas. The hypothesis of two refugial sources is statistically supported by the neighbour-joining dendrogram, which reveals two significantly differentiated groups, including one that contains watersheds in the northern, west-central, and part of southeastern Labrador, corresponding to the presence of haplotype 2.

Previous authors have postulated that brook trout from the Atlantic and Mississippian refugia migrated through Québec to recolonize Labrador, whereas the Atlantic refugium alone was the source for fish invading coastal regions (Black et al., 1986). Conversely, Danzmann et al. (1998) have suggested that northeastern brook trout populations may have been founded by fish from an Acadian refugium. Evidence from our study supports an intermediate hypothesis: one group of fish, possibly originating from the Atlantic or Mississippian refugia, may have colonized northern, west-central, and part of southeastern Labrador, and a second founding group, possibly from an Acadian refugium, invaded Newfoundland and the southeastern-most part of Labrador, with some contribution to west-central Labrador watersheds. Danzmann et al. (1998) have attributed the presence of a number of restricted haplotypes plus the absence of their haplotype 2 fish in eastern populations to contributions from an alternate refugium, the Acadian, which is parallel to the absence of our haplotype 2 fish in the southeastern-most part of Labrador and insular Newfoundland, combined with the presence of a number of restricted haplotypes in these regions. The presence of our haplotypes 1 and 2 and restricted haplotypes in other Atlantic Canadian regions (Québec, Nova Scotia, and New Brunswick) also support the contribution of both Atlantic/Mississippian and Acadian refugia.

Analysis of mtDNA reveals that the evolution of the majority of the haplotypes coincides with the timing of recolonization of brook trout into Labrador, reported to have occurred anywhere from 13,000 to 5000 years ago (Lamb, 1980; Dyke and Prest, 1987). The most frequent haplotype, haplotype 1 (along with haplotypes 3, 6, 7, 8, 9, 11, and 12 in this clade) originated approximately 10,260 years ago, whereas the clade composed of haplotypes 2, 5, and 10 evolved 12,280 years ago. The only exception is haplotype 4 which is ancestral and evolved approximately 92,980 years ago. Although the evolution of these haplotypes coincides with glacial retreat from Labrador, no strong evidence couples recolonization with population expansion. The star-like network of haplotype 1 and its derivatives, as well as the magnitude of Tajima's *D* (for all haplotypes = -1.33; for haplotype 1 and its derivatives = -1.30), are indicative of population expansion; however, Tajima's *D* was not significant. The levels of variation in this portion of the mtDNA genome may possibly be too low to allow significance.

CONCLUSIONS

Very low mtDNA diversity but significant structure at the watershed level characterize populations of brook trout in northeastern Canada. Our study provides evidence of two founding populations in Newfoundland and Labrador: one that may have recolonized the north, west-central, and part of southeastern Labrador from some combination of Atlantic and Mississippian refugium, and another Acadian source, that prevails in insular Newfoundland and the southeastern-most part of Labrador. The evolution of the majority of the haplotypes coincides with the timing of glacial retreat from Labrador. We suggest that the mtDNA genome sequence available in GenBank (AF154850) may represent a brook trout with mtDNA that has

recombined with Arctic charr mtDNA. The evolutionary patterns of mtDNA variation characterized herein demonstrate the strong influence that historical processes have had on shaping the modern population genetic structure of brook trout.

ACKNOWLEDGMENTS

The authors are grateful to the numerous conservation officers of the Department of Natural Resources and staff of the Wildlife Division, Newfoundland and Labrador Department of Environment and Conservation who were involved in the collection of samples. Research supported by a grant from the Wildlife Division of the Newfoundland and Labrador Department of Environment and Conservation, and a Natural Sciences and Engineering Research Council Discovery Grant to H.D. Marshall. We also gratefully acknowledge the assistance of Dr. Elizabeth Perry and Dr. Ed Yaskowiak at the Genomics and Proteomics Facility, Core Research Equipment and Instrument Training, Memorial University of Newfoundland.

REFERENCES

Anderson L, Ryman N and Ståhl G (1983). Protein loci in the Arctic charr, *Salvelinus alpinus* L.: electrophoretic expression and genetic variability patterns. *J. Fish. Biol.* 23: 75-94.

Anderson LW, Lydersen C, Frie AK, Rosing-Asvid A, et al. (2011). A population on the edge: genetic diversity and population structure of the world's northernmost harbour seals (*Phoca vitulina*). *Biol. J. Linn. Soc.* 102: 420-439.

Anderson TC (1985). The Rivers of Labrador. *Can. Spec. Publ. Fish. Aquat. Sci.* 81: 389.

Avise JC, Ball RM and Arnold J (1988). Current versus historical population sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. *Mol. Biol. Evol.* 5: 331-344.

Bennett KD (1997). Evolution and Ecology: The Pace of Life. Cambridge University Press, Cambridge.

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

Bernatchez L, Glémet H, Wilson CC and Danzmann R (1995). Introgression and fixation of Arctic charr (*Salvelinus alpinus*) mitochondrial genome in an allopatric population of brook trout (*Salvelinus fontinalis*). *Can. J. Fish. Aquat. Sci.* 52: 179-185.

Black GA, Dempson JB and Bruce WJ (1986). Distribution and postglacial dispersal of freshwater fishes of Labrador. *Can. J. Zool.* 64: 21-31.

Brunner PC, Douglas MR, Osinov A, Wilson CC, et al. (2001). Holarctic phylogeography of Arctic charr (*Salvelinus alpinus* L.) inferred from mitochondrial DNA sequences. *Evolution* 55: 573-586.

Danzmann RG, Morgan RPM, Jones MW, Bernatchez L, et al. (1998). A major sextet of mitochondrial DNA phylogenetic assemblages extant in eastern North American brook trout (*Salvelinus fontinalis*): distribution and postglacial dispersal patterns. *Can. J. Zool.* 76: 1300-1318.

Doiron S, Bernatchez L and Blier PU (2002). A comparative mitogenomic analysis of the potential adaptive value of Arctic charr mtDNA introgression in brook charr populations (*Salvelinus fontinalis* Mitchill). *Mol. Biol. Evol.* 19: 1902-1909.

Dyke AS and Prest VK (1987). Late wisconsinan and holocene history of the laurentide ice sheet. *Geog. Phys. Quat.* 41: 237-263.

Excoffier L, Laval G and Schneider S (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1: 47-50.

Glémet H, Blier P and Bernatchez L (1998). Geographical extent of Arctic char (*Salvelinus alpinus*) mtDNA introgression in brook char populations (*S. fontinalis*) from eastern Québec, Canada. *Mol. Ecol.* 7: 1655-1662.

Hammar J, Dempson JB and Verspoor E (1991). Natural hybridization between Arctic charr (*Salvelinus alpinus*) and brook trout (*Salvelinus fontinalis*): evidence from northern Labrador. *Can. J. Fish. Aquat. Sci.* 48: 1437-1445.

Hewitt GM (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.*

Lond. 58: 247-276.
Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 359: 183-195.
Hubbs CL (1955). Hybridization between fish species in nature. *Syst. Biol.* 4: 1-20.
Jaramillo-Correa JP and Bousquet J (2005). Mitochondrial genome recombination in the zone of contact between two hybridizing conifers. *Genetics* 171: 1951-1962.
Kornfield I, Beland KF, Moring JR and Kircheis FW (1981). Genetic similarity among endemic Arctic char (*Salvelinus alpinus*) and implications for their management. *Can. J. Fish. Aquat. Sci.* 38: 32-39.
Kraytsberg Y, Schwartz M, Brown TA, Ebralidse K, et al. (2004). Recombination of human mitochondrial DNA. *Science* 304: 981.
Ladoukakis ED and Zouros E (2001). Direct evidence for homologous recombination in mussel (*Mytilus galloprovincialis*) mitochondrial DNA. *Mol. Biol. Evol.* 18: 1168-1175.
Lamb HF (1980). Late Quaternary Vegetational history of southeastern Labrador. *Arct. Antarct. Alp. Res.* 12: 117-135.
Librado P and Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.
Lunt DH and Hyman BC (1997). Animal mitochondrial DNA recombination. *Nature* 387: 247.
McGlade JM (1981). Genotypic and Phenotypic Variation in Brook Trout *Salvelinus Fontinalis* (Mitchill). PhD thesis. Department of Zoology, University of Guelph, Guelph.
Pielou EC (1991). After the Ice Age: the Return of Life to Glaciated North America. University of Chicago Press, Chicago.
Power G (1980). The Brook Charr, *Salvelinus Fontinalis*. In: Charrs: Salmonid Fishes of the Genus *Salvelinus* (Balon EK, ed.). Dr. W. Junk Publishers, The Hague, 141-203.
Saville BJ, Kohli Y and Anderson JB (1998). mtDNA recombination in a natural population. *Proc. Natl. Acad. Sci. U. S. A.* 95: 1331-1335.
Städler T and Delph LF (2002). Ancient mitochondrial haplotypes and evidence for intragenic recombination in a gynodioecious plant. *Proc. Natl. Acad. Sci. U. S. A.* 99: 11730-11735.
Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.
Tamura K, Dudley J, Nei M and Kumar S (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596-1599.
Taylor EB, Pollard S and Louie D (1999). Mitochondrial DNA variation in bull trout (*Salvelinus confluentus*) from northwestern North America: implications for zoogeography and conservation. *Mol. Ecol.* 8: 1155-1170.
Ujvari B, Dowton M and Madsen T (2007). Mitochondrial DNA recombination in a free-ranging Australian lizard. *Biol. Lett.* 3: 189-192.
Wagner DB, Dong J, Carlson MR and Yanchuk AD (1991). Paternal leakage of mitochondrial DNA in *Pinus*. *Theor. Appl. Genet.* 82: 510-514.