Molecular Ecology (1999) **8**, 23–35

Zoogeographical implications of variation in mitochondrial DNA of Arctic grayling (*Thymallus arcticus*)

Z. REDENBACH and E. B. TAYLOR
Department of Zoology, University of British Columbia, Vancouver, BC, Canada V6T 1Z4

Abstract

Most of the northern half of North America's freshwater fish survived the last glacial period in unfrozen refugia peripheral to the ice sheets. In our study, the question of which refugia Arctic grayling (*Thymallus arcticus*) inhabited during the Wisconsinan Ice Age, and how they subsequently dispersed to their present geographical range, was examined using mitochondrial DNA (mtDNA) analysis. mtDNA from 12 *T. arcticus* populations was analysed by direct sequencing and restriction fragment length polymorphism analysis (RFLP). Our data support the hypotheses that *T. arcticus* had a large refugial population in the Bering Refuge (shown by high mtDNA diversity in extant Alaskan populations) and that British Columbia was colonized from the Bering Refuge (shown by mtDNA haplotype similarities). Our data also show that a disjunct southern set of populations in Montana is significantly different from the northern grayling, in terms of restriction haplotype frequency and distinguishing sequence characteristics. Sequence results yielded an estimated divergence time of 370 000 years between the northern and Montana grayling haplotypes. We conclude that *T. arcticus* survived the Wisconsinan glaciation in at least two refugia: (i) the Bering Refuge north of the ice sheets; and (ii) either the Upper Missouri or the southwest Alberta Refuge, south of the ice sheet.

Keywords: mitochondrial DNA, postglacial dispersal, RFLP, *Thymallus arcticus*, zoogeography

Received 8 January 1998; revision received 5 June 1998; accepted 13 July 1998

Introduction

Over the last million years (Myr), as many as 20 glacial periods have sequentially covered most of Canada and parts of the northern United States with colossal ice sheets (Fig. 1) (Martinson *et al*. 1987). Flora and fauna persisted in ice-free refugia along the periphery of the glaciers and when the ice sheets periodically retreated, species recolonized the vast deglaciated areas (McPhail & Lindsey 1970).

Much zoogeographical and phylogeographical work has been carried out to determine which refugia were inhabited by particular species, and to deduce the postglacial dispersal routes and recolonization patterns that they used to arrive at their current distributions (Hocutt & Wiley 1986). Freshwater fish species, with their restriction to aquatic habitats, dispersed via routes that were often ephemeral in existence (such as pro-glacial lakes and headwater transfers) (Walters 1955; McPhail & Lindsey 1970). Of particular interest are fish species with disjunct distributions. These disjunct distributions can have many potential causes, such as a failure to exploit ephemeral dispersal routes or a subsequent extirpation from intervening areas, thus making these species appealing subjects of study.

The Arctic grayling, *Thymallus arcticus*, is a salt-water-intolerant, cold water, salmonid species whose distribution in North America has long puzzled zoogeographers (Fig. 1). The Arctic distribution is fairly simple, with populations in eastern Siberia, on St Lawrence Island in the Bering Strait, through Alaska and northern Canada to Hudson Bay. The southern limits of this distribution are the Peace and Athabasca rivers (in British Columbia and Alberta), which flow north into the Mackenzie River drainage basin. Two disjunct populations, however, are found over 850 and 1400 km south and east of the main

Correspondence: Zoë Redenbach, Department of Zoology, University of British Columbia, Vancouver, BC, Canada V6T 1Z4. Fax: +00 1-(604) 822-2416; E-mail: redenbac@zoology.ubc.ca

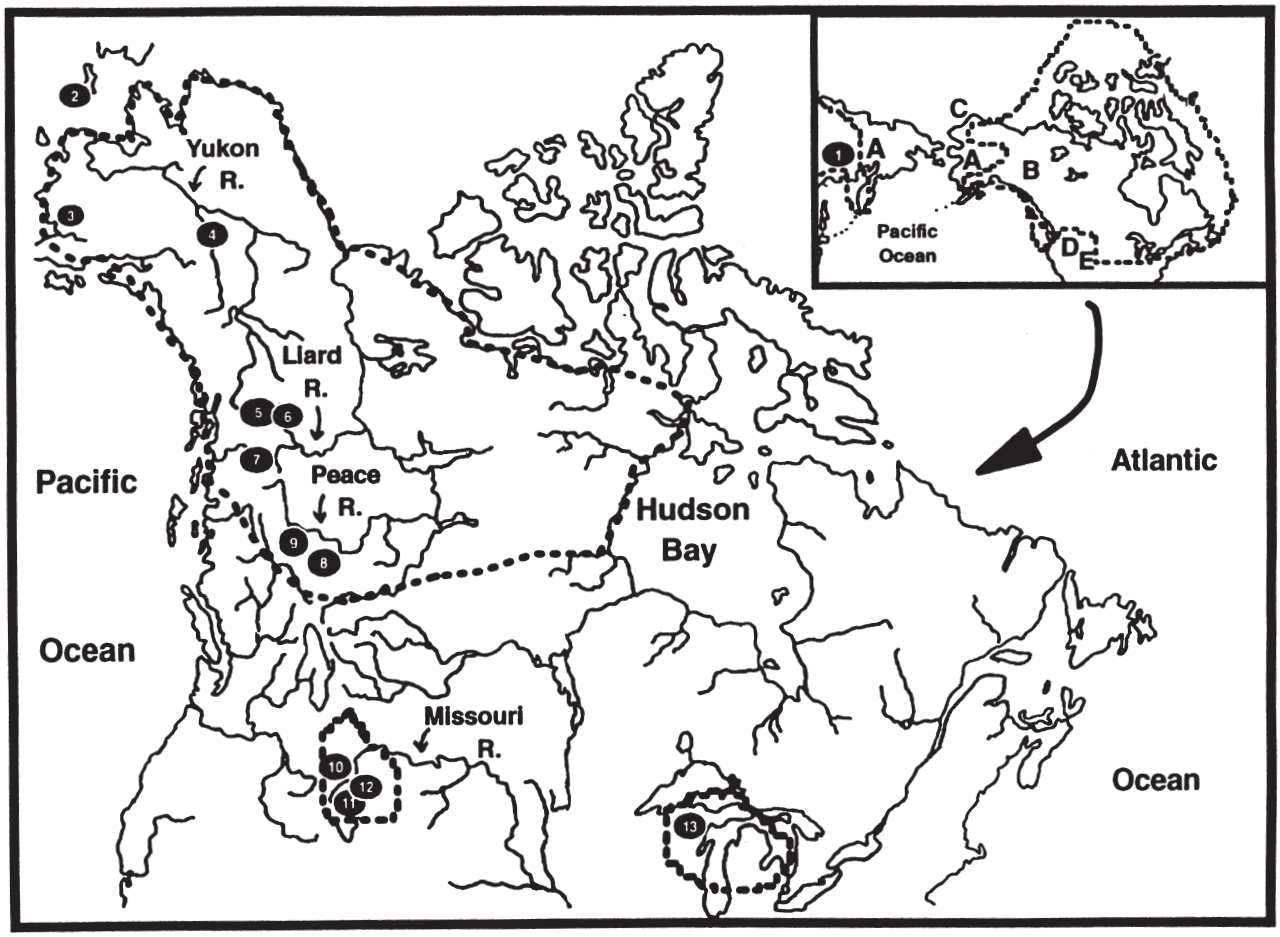

Fig. 1 Geographical distribution of *Thymallus arcticus*, sample sites, maximum extent of Wisconsinan glaciation and potential glacial refugia occupied by *T. arcticus* during the last glaciation. Dashed lines in the main map indicate the distribution of *T. arcticus* in North America. Sample sites are indicated by numbers: 1, Yana River, Russia; 2, St. Lawrence Island, AK; 3, unnamed stream in Bristol Bay drainage, AK; 4, Chena River, AK; 5, Plate Lake, BC; 6, Upper Tootsie Lake, BC; 7, Turnagain River, BC; 8, Burnt River, BC; 9, Table River, BC; 10, Big Hole River, MT; 11, Red Rock Lake, MT; 12, Madison River, MT; and 13, Otter River, MI. Inset: dashed lines indicate the maximum extent of Wisconsinan glaciation. Potential North American *T. arcticus* glacial refugia are indicated by letters: A, Bering Refuge; B, Nahanni Refuge; C, North Slope of the Brooks Range; D, southwest Alberta Refuge; and E, Missouri Refuge.

North American distribution: the Upper Missouri River (Montana) and former Michigan (extinct since 1936) populations, respectively (Scott & Crossman 1973).

These disjunct southern populations were initially thought to be so distinct from the main northern group of populations that they were given separate species designations: *T. tricolor* in Michigan (Cope 1865) and *T. montanus* in Montana (Milner 1874). Michigan and Montana grayling were later reduced to subspecific level (*T. arcticus tricolor* and *T. a. montanus*) and then to a single subspecies (*T. arcticus tricolor*, Walters 1955). Current views tend to agree that all North American grayling fall into a single species, requiring no subspecific designations (McPhail & Lindsey 1970; Northcote 1993).

Relationships between the three disjunct *T. arcticus* populations have been studied using traditional phylogenetic tools, such as osteology (Norden 1961), meristics and morphology (Hop & Gharrett 1989; Reed & McCann 1973), and allozyme studies (Lynch & Vyse 1979). These tools are applicable because isolated refugial populations evolve independently for thousands of years, allowing genetic drift, varying selection pressures and molecular evolution to create differences between refugial populations (Crossman & McAllister 1986). These genetic and morphological differences are reflected in current population distributions and can allow inferences about recolonization patterns. The grayling studies referred to above have shown that differences exist between populations, but have done little to determine the refugial origins and subsequent pattern of dispersal for *T. arcticus*.

Recent phylogenetic and zoogeographical studies have applied molecular techniques, particularly DNA amplifi-

cation, sequencing and restriction fragment length polymorphism (RFLP) analysis, to mitochondrial DNA (mtDNA) (e.g. Bermingham & Avise 1986; Meyer *et al*. 1990; Taylor *et al*. 1997). mtDNA is particularly useful for population level studies because it accumulates mutations very rapidly and has an effective population size that is only 25% of that of nuclear DNA, assuming mutation-drift equilibrium (Gyllensten & Wilson 1986). These characteristics mean that evolution over short time spans can be detected and that small, isolated populations will rapidly diverge in mtDNA haplotype owing to genetic drift.

The objectives of our study were to utilize RFLP and sequence analysis of mtDNA to study genetic differences between 13 widely distributed *T. arcticus* populations (Fig. 1). In particular, our aim was to determine whether the Montana grayling survived glaciation in a southern refuge and, if so, how long the two populations have been isolated, or whether the Montana population was founded by northern grayling with subsequent extirpation in the intervening areas of southern Canada.

Materials and methods

Sample locations

Ethanol-preserved adipose fin-clip samples were collected from 11 sample sites across most of the geographical range of *Thymallus arcticus* (Fig. 1). Sample sites ranged from Kamchatka in Russia, through Alaska and British Columbia, to the disjunct southern population in Montana. In Russia, the single sample site was the Yana River (sample site 1 in Fig. 1, $n = 4$), which drains into the northern Sea of Okhotsk (≈ 60°N, 150°E). The two Alaskan sample sites were an un-named stream draining into Bristol Bay (sample site 3, $n = 10$) and Chena River (sample site 4, $n = 20$), a central Alaskan tributary to the Yukon River. A second tributary to the Yukon River, Plate Lake (sample site 5, $n = 10$), located in northern British Columbia, is part of the extreme southern headwaters of the Yukon River. Other British Columbia sample sites were two drainage systems, the Peace and Liard rivers, which are in turn both tributaries to the Mackenzie River. On the Liard River, both sample sites were located upstream of the Liard Canyon, a barrier to upstream fish migration (Lindsey & McPhail 1986). Upper Tootsie Lake (sample site 6, $n = 10$) is located on the north fork of the Liard, while the Turnagain River (sample site 7, $n = 4$) is a tributary to the south fork. On the Peace River, the Table River (sample site 9, $n = 5$) is a tributary upstream of the Peace Canyon (current location the W. A. C. Bennett Dam), also a substantial barrier to upstream fish migration (Lindsey & McPhail 1986), while the Burnt River (sample site 8, $n = 19$) is a tributary downstream of the canyon. The Montana sample sites, Big Hole River (sample site 10, $n = 10$), Red Rock Lake (sample site 11, $n = 5$) and Madison River (sample site 12, $n = 10$), are headwater tributaries to the Missouri River, merging within about 100 km of each other, near Three Forks, Montana. Ethanol-preserved samples were also obtained for a single *T. thymallus* population (European grayling, from Dean Water, Scotland, UK; $n = 4$).

Formalin-preserved muscle tissue samples were obtained from an extant *T. arcticus* population on St Lawrence Island, located in the Bering Straight between mainland Alaska and Russia (sample site 2, $n = 1$) and from Otter River, Michigan, (sample site 13, $n = 2$), part of the Michigan *T. arcticus* population which has been extinct since 1936 (Scott & Crossman 1973).

DNA extraction

Ethanol-preserved tissue. DNA was extracted from 20 mg or less of adipose fin tissue by overnight digestion with pronase followed by phenol–chloroform extraction (as described in Taylor *et al*. 1996). DNA was precipitated in isopropanol, and the resulting DNA pellet was dried in a vacuum-condenser for 5 min, resuspended in 75–100 μL of TE buffer (pH 8.0), quantified by spectrophotometry and stored at – 20 °C.

Formalin-preserved tissue. DNA was extracted from 50 to 100 mg of formalin-preserved, ethanol-stored muscle tissue, using the method of Vachot & Monnerot (1996) (a modification of Kocher *et al*.'s (1989) technique). This method successfully extracted DNA from the St Lawrence Island sample, but yielded no useable DNA from the Michigan samples. A second extraction protocol, based on a 3-day glycine solution rinse and variation in PCR conditions (Shedlock *et al*. 1997), was attempted for the Michigan samples, but it also failed to yield useable DNA.

The poor results for the Michigan samples could be caused by the preservation technique originally used (such as a longer soak time in formalin), to the time the specimens spent unpreserved after capture (during which DNA would begin to degrade), or purely as a result of the age of the preserved tissue. The low final extracted DNA concentrations, consisting primarily of sheared DNA, required altered polymerase chain reaction (PCR) conditions (more cycles and a lower annealing temperature), which unfortunately allowed contaminating DNA to be amplified. Therefore, the Michigan samples could not be amplified in a condition sufficiently pure to allow sequencing and were not included in our study.

RFLP analysis

Restriction analysis was carried out on two PCR-ampli-

fied mtDNA fragments, the first containing the cytochrome *b* (*cyt-b*) gene and the noncoding D-loop (2.1 kb) and the second encompassing the NADH subunit 5 and 6 genes (2.5 kb). These mitochondrial regions were chosen because they often exhibit variability at the population level in fish (Park & Moran 1994). Primers for amplification were obtained from published studies of salmonid fishes: the *cyt-b*/D-loop fragment using primers HN20 and C-Glu described by Bernatchez & Osinov (1995) and the ND5/ND6 fragment using the complement of C-Glu and C-Leu3 described by Park *et al*. (1993). Fragments of this length typically cannot be amplified from formalin-preserved tissue, as formalin preservation degrades DNA such that shearing limits maximum fragment length to ≈ 500 bp. Therefore, RFLP analysis was limited to the single *T. thymallus* and 11 *T. arcticus* populations for which tissue samples had been preserved in ethanol.

PCR reactions for these two fragments were run in a total volume of 40 μL with 800 μM of total dNTPs, 0.6 μM of each primer, 4 or 1.5 mM of $MgCl_2$ (for HN20/C-Glu and C-Leu3/C-Glu, respectively), 1 × GibcoBRL *Taq* DNA polymerase buffer, 2 units of *Taq* DNA polymerase and 1 μL of whole DNA. Cycling parameters were: 1 cycle of 2 min at 95 °C, 1 min at 55 °C and 1.5 min at 72 °C; 35 cycles of 1 min at 94 °C, 1 min at 55 °C and 1.5 min at 72 °C; followed by 1 cycle of 5 min at 72 °C.

The fragments were restricted using eight enzymes: *Alu*I, *Ban*I, *Hae*III, *Hin*fI, *Rsa*I, *Sau*3A1, *Sty*I and *Taq*I (New England Biolabs). Digests were performed according to the manufacturer's instructions, overnight, in a total volume of 20 μL. The digests were then run alongside molecular size standards on 1–1.5% agarose gels, stained with ethidium bromide and photographed under UV light. This type of visualization revealed any polymorphisms between individuals. These polymorphisms were then visualized more precisely by running digests of representative individuals on polyacrylamide gels, in order to yield better resolution of fragment length, particularly for smaller fragments.

DNA sequencing and analysis

Automated dideoxy termination sequencing was performed on three mtDNA fragments: a 213-bp fragment at the 3′ end of the D-loop, a 1406-bp fragment encompassing tRNA-Ile, NADH-1, tRNA-Leu and 16S rRNA, and a 876-bp subset of the 1406-bp fragment. The fragments were chosen for different purposes. The sequencing of 213 bp of the D-loop was intended to compare haplotypes from the different geographical areas, perhaps allowing a diagnostic characteristic to be identified for particular areas. Little variation was found via RFLP analysis and, although it is sometimes the case that RFLP will reveal more variation in fish than D-loop sequencing (e.g. Park *et al*. 1993; Bernatchez & Osinov 1995), it was hoped that additional variation could be found by sequencing (e.g. Thomas & Beckenbach 1989; Nielsen *et al*. 1994; M. McCusker & E. B. Taylor unpublished data), particularly as the two extremes of the noncoding D-loop region are typically very variable in fish (Park & Moran 1994).

Sequencing of the 1406-bp fragment was intended to give sequence divergence information for use in estimates of time since divergence. This would require a longer fragment, with both variable and conserved regions. We chose the *NADH-1* gene, which has been shown to be variable in bull trout, *Salvelinus confluentas* (Williams *et al*. 1995), and the adjacent, typically conserved, *tRNA* and *rRNA* genes. The 876-bp subfragment was sequenced in six individuals whose RFLP haplotypes were not represented by the 1406-bp sequence. For all three fragments sequenced, sample sizes were kept small and individuals were chosen to represent particular RFLP haplotypes.

The primers for amplification of the 213-bp fragment were S-Phe and P2, obtained from Nielsen *et al*. (1994). The 1406-bp fragment was a portion of a 2-kb fragment that was amplified using the primers t-Ile and t-Leu2 obtained from Park *et al*. (1993). Owing to the length of the 2-kb fragment, two internal primers had to be designed for sequencing. The first, named ND1C, was located ≈ 540 bp into the fragment from the t-Ile primer, in the same orientation, with sequence reading 5′-TGCAGC-CGCTATTAAGGGTTCG-3′. The second, named ND1D, was located ≈ 500 bp into the fragment from the t-Leu2 primer, in the same orientation, with sequence reading 5′-TCTGTGAAGTCAAATGGGGCAC-3′. The 1406 bp used for sequencing comprised the region between primers ND1C and t-Leu2, while the 876-bp fragment comprised the region between ND1C and ND1D.

PCR conditions for sequencing analysis were similar to those for RFLP analysis, with the following changes: 150 μL total volume, 4 mM $MgCl_2$, 5 units of *Taq* DNA polymerase, 5 μL of whole DNA and shorter extension times (72 °C for 1 min).

The initial 150-μL PCR volume was purified using the Qiagen Qiaquick PCR Purification Kit. The purified DNA was suspended in distilled, autoclaved water and quantified spectrophotometrically. Eighty to 120 ng of the purified DNA was amplified in a 20-μL termination PCR reaction, using 0.16 μM of primer, 0.05% DMSO and 8 μL of terminator premix from the Perkin-Elmer ABI Prism Dye Terminator Cycle Sequencing Ready Reaction Kit. The P2 primer was used for the 213-bp fragment, while both terminal primers (t-Leu2 and ND1C) and an internal primer (ND1D) were required for the 1406-bp fragment. Termination PCR conditions were: 1 cycle of 4 min at 94 °C; and 25 cycles of 0.5 min at 94 °C, 0.25 min at 55 °C and 4 min at 60 °C. The amplified, fluorescently labelled,

terminated DNA was purified using a Centrisep Spin Column and analysed with an ABI 377 automated sequencer.

Phylogenetic analysis

The DNA fragments produced by the restriction digests were used to infer restriction site data. All *T. arcticus* fragment pattern changes could be explained by a single restriction site gain or loss, while pattern changes between the two *Thymallus* species often required the inference of multiple site changes. A mutational network between RFLP haplotypes was constructed by eye. The inferred restriction site data were analysed using the REAP computer analysis package (McElroy *et al*. 1992). Estimates of evolutionary divergence between RFLP haplotypes were calculated as '*d*', the number of substitutions per nucleotide site (Nei & Tajima 1981) and weighted by enzyme class (Nei & Tajima 1983). Haplotype diversity was estimated according to Nei (1987), while nucleotide diversity and nucleotide divergence were estimated according to Nei & Tajima (1981).

The data from RFLP, the 213-bp sequence and the 876-bp sequence were then pooled into a single data set for phylogenetic analysis. One individual was sequenced to represent each RFLP haplotype, with the exceptions of RFLP haplotypes C (for which the DNA supply from the single individual was exhausted) and G (for which three individuals were sequenced, G1 from Montana, G2 from British Columbia and G3 from Alaska).

The sequence data was converted into binary format, coding the sequence variants by presence/absence and combined with RFLP-site data. The resulting 1184 characters were analysed using the PHYLIP computer analysis package (Felsenstein 1995). Data was subjected to bootstrapping (100 replicates, SEQBOOT) and analysed by a discrete character parsimony tree-building program (MIX, using the Wagner parsimony method and randomizing input order). The resulting trees were combined using majority rule consensus tree analyses (CONSENSE).

Sequence divergences for the 213-bp and 1406-bp fragments were calculated using the DNADIST program of PHYLIP, applying the Kimura 2-parameter distances model of nucleotide substitution (Kimura 1980).

Results

RFLP analysis

RFLP analysis with eight enzymes yielded 95 restriction sites, covering a total of 417 bp across about 4.6 kb of mtDNA. A total of 10 *Thymallus arcticus* restriction haplotypes were found, with sample sizes ranging from 1 to 60 individuals for each haplotype (Table 1). The two most common haplotypes, haplotypes A and G, represent the most common haplotypes of the Alaska/British Columbia and Montana grayling populations, respectively. *T. thymallus* and the Russian population of *T. arcticus* showed no haplotypes in common with North American *T. arcticus*, nor with each other (Table 1).

The mutational network (Fig. 2) shows that the common Montana haplotype, haplotype G, is most closely related to the *T. thymallus* haplotype, and therefore is probably most similar to an ancestral *T. arcticus* haplotype. The

Table 1 Composite restriction enzyme haplotypes for 11 *Thymallus arcticus* populations and a single *T. thymallus* population

Haplotype	Composite fragment pattern*	Sample sites											
		Russia	Alaska		British Columbia					Montana			
		1	3	4	5	6	7	8	9	10	11	12
A	a a a a a a a a		6	11	10	9	1	18	5			
B	a a a b a a a a		3									
C	a a a a b a a a		1									
D	a a a a a a b a					1						
E	a a a c a a a a			7								
F	a a b c a a a a						3					
G	b a a a c a a a			2				1		10	5	9
H	b a a a c b a a											1
I	c a c d c a a a	3										
J	c a d d c a a a	1										
K	d a e e d c c b				*Thymallus thymallus* (Scotland)							

*The composite fragment pattern consists of the fragment patterns produced by eight enzymes: *Alu*I, *Ban*I, *Hae*III, *Hin*fI, *Rsa*I, *Sau*3AI, *Sty*I and *Taq*I, respectively. Each individual restriction enzyme pattern is a sum of the patterns produced by that enzyme on the *cyt-b*/D-loop and ND5/ND6 PCR-amplified fragments. Information on the inferred cut sites and fragment patterns can be obtained from the authors.

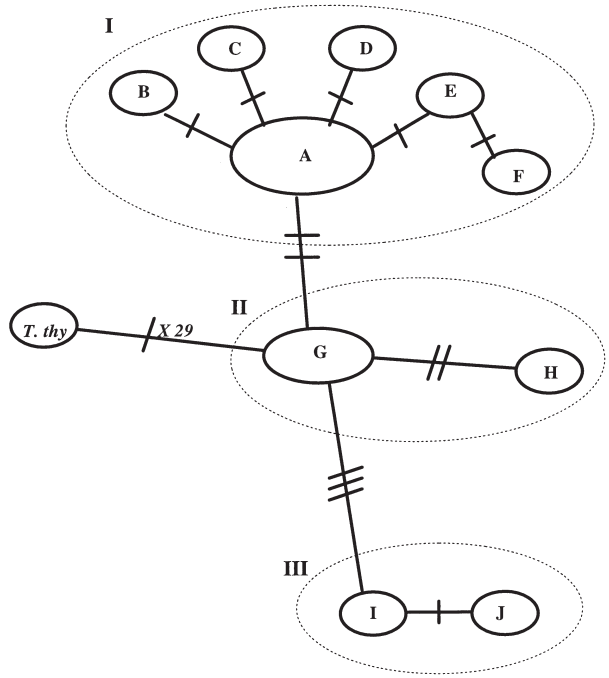

Fig. 2 Parsimony network showing the mutational relationships between the single *Thymallus thymallus* and 10 *T. arcticus* restriction haplotypes described in Table 1. Each slash represents a change in a restriction site (loss or gain) and large solid ellipses represent the common haplotypes ($n > 25$). Dashed lines represent the boundaries of three haplotype groups, whose geographical distribution is approximately restricted to three geographical areas: I, Alaska/British Columbia; II, Montana, III, Russia.

Russian haplotypes (I and J) and the main North American group (haplotypes A to F) are separated from haplotype G by three and two restriction site differences, respectively. These separations suggest that the 10 *T. arcticus* haplotypes consist of three groups, Groups I and III defined by inclusion (separated from other haplotypes by two and three restriction site changes, respectively) and Group II defined by exclusion (it falls between Groups I and III). These three groups correspond approximately to three geographical areas: Group I represents Alaska and British Columbia, Group III represents Russia and Group II represents Montana (with ≈ 11% of individuals characterized by G haplotypes from Alaska/British Columbia).

The restriction site data were analysed by both parsimony (as discrete character data, using the presence (1) or absence (0) of restriction sites) and distance analysis methods (number of substitutions per site, *d*, Nei & Miller 1990). The trees produced by these different methods (data not shown) were all quite similar in overall topology, resembling the mutational network (Fig. 2). Variation between trees occurred primarily within Group I and in the placement of the Montana haplotypes (haplotypes G and H) relative to Groups I and III. The variable position of the Montana haplotypes in the trees is explained by the ancestral, central nature of haplotype G relative to Groups I and III.

Calculation of haplotype and nucleotide diversity within each of the sampled drainage systems revealed that most of the intrapopulation variation in North America is found in the Liard River and the two Alaskan drainage systems (Table 2). Furthermore, these drainage systems each have different variant haplotypes: haplotypes E and G in the Chena River, at frequencies of 0.37 and 0.11, respectively, Haplotypes B and C in the Bristol Bay drainage, at frequencies of 0.3 and 0.1, respectively, and haplotype F in the Liard River, at a frequency of 0.21. The Peace and Missouri rivers had the lowest haplotype and nucleotide diversities.

The nucleotide divergence between pairs of drainage basins was calculated from the genetic distances between haplotypes and the haplotype frequency distributions for each drainage basin (Table 3). These values reveal which drainage basins are most closely related in terms of mitochondrial haplotypes. The pairwise nucleotide divergence between each of Bristol Bay, Yukon River, Liard River and Peace River show that the most closely related basins are the Yukon and Liard rivers, while the greatest divergence is seen between the Bristol Bay drainage and both the Yukon and Liard rivers. Divergence between the Missouri River and any of the four northern drainage basins is an order of magnitude greater than between the most divergent pair of northern basins (Table 3).

Sequencing analysis

Seven *T. arcticus* sequence haplotypes were found for the 213-bp D-loop fragment (Fig. 3). The Russian *T. arcticus*

Geographical area	Drainage system	*n*	Haplotype diversity	Nucleotide diversity
Russia	Yana River	4	0.4286 ± 0.16870	0.000823
Alaska	Bristol Bay	10	0.5684 ± 0.08629	0.001147
	Yukon River	29	0.4588 ± 0.06104	0.001097
British Columbia	Liard River	14	0.4550 ± 0.09412	0.001527
	Peace River	24	0.0816 ± 0.05293	0.000286
Montana	Missouri River	25	0.0784 ± 0.05104	0.000270

Table 2 Haplotype and nucleotide diversity of *T. arcticus* by drainage system, calculated with the DA program of REAP (McElroy *et al*. 1992)

Table 3 Nucleotide divergence among pairs of North American drainage basins ($\times 10^{-4}$), calculated with the DA program of REAP (McElroy *et al*. 1992)

	Liard River	Yukon River	Bristol Bay	Missouri River
Peace River	1.157	0.741	1.162	31.43
Liard River		0.327	2.394	35.64
Yukon River			2.174	30.70
Bristol Bay				35.37

haplotypes (I and J) were most closely related to *T. thymallus*, owing to a diagnostic C→A transversion at position 80. The specimen from St Lawrence Island (SLI) had the common D-loop haplotype A, which was also found in eight individuals throughout Alaska and British Columbia, representing four different RFLP haplotypes. Three RFLP haplotype G individuals, from different areas, were sequenced and yielded different haplotypes from the 213-bp and 876-bp sequences. They have been labelled G1 (from Montana), G2 (from British Columbia) and G3 (from Alaska).

The most salient result obtained from the 213-bp sequences was the observation that individuals from Montana had a T insertion at position 15 (haplotypes G1 and H). Although sample sizes were small, this insertion was only found once in the northern population, in a British Columbian individual whose restriction haplotype was the common 'Montana' RFLP haplotype G (haplotype G2). An Alaskan individual with RFLP haplotype G (haplotype G3) did not have this T insertion. Haplotype G1/H/G2, with the insertion, was no more divergent from the common sequence haplotype A/D/E/F/SLI than were either of the other variant sequence haplotypes found in the Alaskan/British Columbian populations (haplotypes F and G3).

Six individuals were sequenced for the 1406-bp fragment spanning the 16S *rRNA*, *tRNA-Leu*, *NADH-1* and *tRNA-Ile* genes (Fig. 4). These individuals represented RFLP haplotypes A, E, G (G1 from Montana and G3 from Alaska), I and *T. thymallus*, and yielded six different sequence haplotypes. Over the 1406 bp, 93 variable sites were found, of which 86 were located within the 978 bp of the NADH-1 gene. Twenty-five of the variable sites were among *T. arcticus* individuals and 10 were among North American individuals.

Pairwise estimates of evolutionary divergence were calculated between haplotypes within each of the four sources of data (RFLP and the 213-bp, 1406-bp and 876-bp sequences). These divergence estimates were sorted into five categories: between *T. thymallus* and *T. arcticus*, between Russian and North American *T. arcticus*, between Montana and Alaskan/British Columbian *T. arcticus*, and within Alaskan/British Columbian or Montana *T. arcticus* (Table 4). In each case, the divergence was greatest between haplotypes from different species, followed by the divergence between haplotypes from different continents. The divergences calculated between North American geographical areas often could not be distinguished from those within North American geographical areas. For any given geographical comparison, except within Montana, the greatest evolutionary divergences were those calculated using the 213-bp sequence, located at the hypervariable 3′ end of the D-loop.

Composite data analysis

A Wagner parsimony tree of combined RFLP and sequence data divided the composite mitochondrial haplotypes into three groups (Fig. 5). The first bifurcation within the *T. arcticus* clade separated the North American and Russian haplotypes, while the second bifurcation separated the North American haplotypes into groups representing the two disjunct approximate geographical areas of the grayling range. These three groups corresponded to the three groups described in the RFLP mutational network (Fig. 2) and were labelled similarly: Group I is Alaska/British Columbia, Group II is Montana and Group III is Russia. These groups were all supported with bootstrap values greater than 50%, which is at least moderate support given that bootstrap values ≥ 70% generally correspond to ≥ 95% probability that the corresponding clade is real (Hillis & Bull 1993).

There are two haplotypes that were interesting exceptions from the conclusions described above. Haplotype G3, belonging to an Alaska individual, wavered between Groups I and II and was placed in Group II with only 20% bootstrap support. Haplotype G2, however, belonging to a British Columbian individual, was securely placed in

RFLP Haplotype	n	Sequence Position 1 1 1 1 1 1 1 1 1 1 1 4 4 4 5 8 8 9 3 4 5 5 5 6 6 7 8 9 5 4 5 6 0 0 4 3 9 0 4 7 8 4 5 1 4 5
A, B, D, E, SLI	9	- C T T G A T C C T C T A G G T A G
F	1	A T
G1, G2, H	5	T
G3	1	A
I	1	C
J	1	A A C
K	2	T A C C T C T C C A C A

Fig. 3 Variable sequence positions for a 213-bp sequence located at the 3′ end of the mitochondrial D-loop. Sequences are listed according to the RFLP haplotype characterizing the individuals sequenced. The five individuals characterized by RFLP haplotype A are from sample sites 3, 4, 8 and 9. SLI indicates a formalin-preserved individual from St Lawrence Island (sample site 2), for which a RFLP haplotype could not be determined. The complete sequences have been submitted to GenBank under Accession nos AF036368–AF036375.

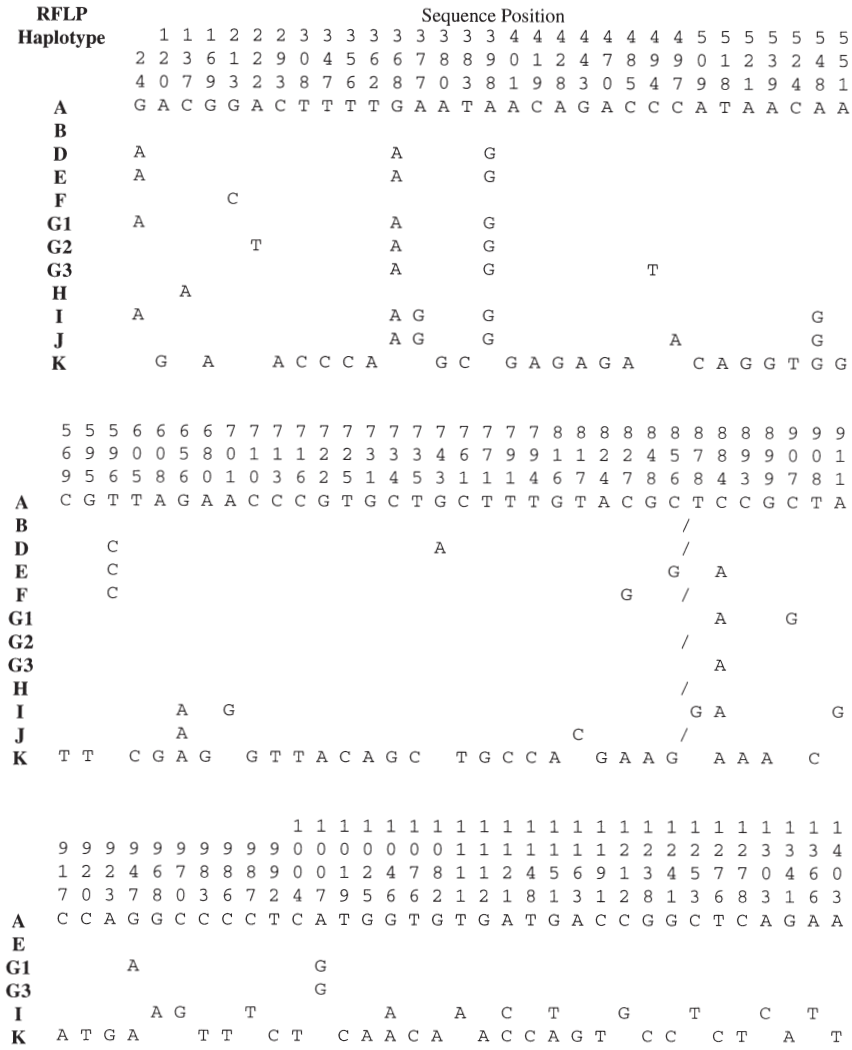

Fig. 4 Variable sequence positions for a 1406-bp sequence encompassing part of the *16S rRNA* (bp 1–199), *Leu-tRNA* (bp 200–275), *NADH-1* (bp 276–1247) and *Ile-tRNA* (bp 1248–1319) genes. Sequences are listed according to the RFLP haplotype characterizing the individuals sequenced and are in the plus (+) strand orientation. Sequences A, E, G1, G3, I and K contain the full 1406-bp, while sequences B, D, F, G2, H and J contain only the first 876 bp (with termination indicated by backslash). The 1406-bp and 876-bp sequences have been submitted to GenBank under Accession nos AF036376–AF036381 and AF076905-AF076910 respectively.

the Montana group, Group II.

Discussion

Russian haplotypes were shown to differ from North American haplotypes by genetic distances ranging from 0.5 to 2.4% sequence divergence, depending on which North American haplotype and which class of molecular data was used (Table 4). If a calibration of 1% divergence per Myr is applied to this range (Smith 1992; but see below), Russian and North American haplotypes could have diverged 500 000–2 400 000 years ago. This time estimate implies that separation occurred long before the Bering Land Bridge was submerged by rising sea levels 14 000 years ago (McPhail & Lindsey 1970). The Yana River *Thymallus arcticus*, from the southern extreme of the grayling range in Asia, are therefore unlikely to be derived from Bering refugial origins.

North Pacific threespine stickleback (*Gasterosteus aculeatus*), which are also likely to have survived in the Bering Refuge, show divergent groupings within the northwestern Pacific (Higuchi & Goto 1996). It is therefore not surprising that *T. arcticus* should show a similar divergence. In fact, variation in morphological and life-history characteristics of Arctic grayling in Asia have resulted in putative subspecific designations: Chukotka Peninsula and Alaskan grayling are included in the subspecies *T.a. signifer* (Skopets 1991) while Kamchatkan grayling, whose range encompasses the Yana River, are designated *T.a. mertensi* (Skopets & Prokop'yev 1990).

T. arcticus further north in Russia, particularly on the Chukotka Peninsula (the eastern extreme of the Bering Refuge), would be more likely than Yana River grayling to share haplotypes with Alaskan populations. Support for this conjecture is found on St Lawrence Island, a remnant of the submerged land bridge (McPhail & Lindsey 1970). A formalin-preserved St Lawrence Island specimen yielded the D-loop sequence haplotype common in Alaskan and

Table 4 Range of genetic distances between mitochondrial haplotypes from different geographical areas*

Comparison of haplotypes:	RFLP (*cyt-b*, D-loop, *ND5/6*)	213-bp sequence (3′ end of D-loop)	1406-bp sequence (*ND1*, *tRNA*, *rRNA*)	876-bp sequence (*ND1*, *tRNA*, *rRNA*)
Between *Thymallus* spp.	5.2–6.0	5.4–7.0	5.7–5.8	5.4–5.7
Russian vs. N. American	0.5–1.4	0.5–2.4	1.4–1.5	0.5–0.9
Montana vs. Alaska/BC.	0–0.7	0–1.4	0.3–0.5	0.2–0.6
Within Alaska/BC	0.2–0.5	0.5–1.4	0.3–0.5	0.2–0.7
Within Montana	0.3	0	–	0.2

*The range of genetic distances between mitochondrial haplotypes from different geographical areas were calculated using four sources of molecular data: (i) RFLP analysis of *cyt-b*, D-loop and *ND5/ND6*; (ii) the 213-bp sequence at the 3′ end of the D-loop; (iii) the 1406-bp sequence spanning *NADH-1*, *tRNA-Ile*, *tRNA-Leu* and part of the *16S rRNA* genes; and (iv) a 876-bp subset of the 1406-bp sequence. Exact distances between particular sequence haplotypes are shown in Figs 3 and 4.

British Columbian populations (Fig. 3). As the land bridge was criss-crossed with a network of freshwater systems though which fish could disperse, it is possible that the large Bering Refuge had a relatively homogenous population of *T. arcticus*, particularly as the species is known to make long yearly migrations between feeding, overwintering and spawning habitats (Northcote 1993).

Our measures of haplotype and nucleotide diversity (Table 2) support the hypothesis that the Alaskan population is descended from a large refugial population (McPhail & Lindsey 1970), as there is high mitochondrial diversity within the two Alaskan drainage systems sampled, compared with populations in British Columbia and Montana. A higher level of diversity would be expected in the *in situ* descendants of a large refugial population (Alaska) than in populations founded elsewhere from that large population (British Columbia) or in populations with historically small population sizes (Montana), owing to the likely elimination of genetic variation by founder effects and population bottlenecks (Nei *et al*. 1975; Avise 1994; Jaarola & Tegelström 1996). In fact, similar high levels of nucleotide diversity have been found in Alaskan populations of lake trout, *Salvelinus namaycush* (Wilson & Hebert 1998), and lake whitefish, *Coregonus clupeaformis* (Bernatchez & Dodson 1991).

The British Columbian populations of *T. arcticus* are located in previously glaciated regions of Canada and therefore must have colonized from one or more glacial refugia. The low levels of nucleotide diversity observed in the Peace River support the conclusion that the Peace was founded from a single refuge, while the similarity between Peace River and Alaska RFLP haplotypes (Group I in Figs 2 and 3) indicates that the primary, if not only, source of Peace River colonizers was the Bering Refuge. The Liard River, with its high nucleotide and haplotype diversities, is discussed below with respect to a putative Nahanni Refuge.

The formalin-preserved samples from the extinct Michigan population did not yield any DNA suitable for analysis, preventing estimation of Michigan haplotype divergence. We are therefore unable to either support or refute the hypothesis that *T. arcticus* survived glaciation in the Mississippi Refuge (McPhail & Lindsey 1970), a hypothesis supported by both distributional (i.e. potential refuge located near or overlapping current range; Crossman & McAllister 1986) and fossil (> 730 000-year-old *T. arcticus* fossils in Indiana, Miller *et al*. 1993) evidence.

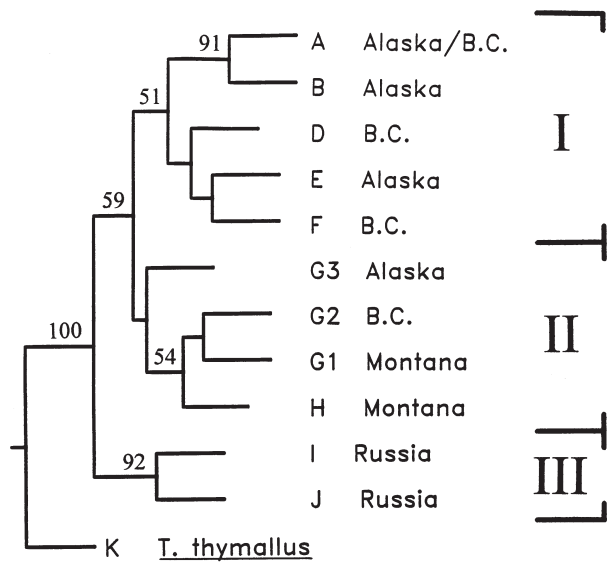

Fig. 5 Bootstrapped MIX Wagner parsimony phenogram clustering the composite character divergence haplotypes (01) of the single *Thymallus thymallus* and 12 *T. arcticus* composite mitochondrial haplotypes. G1, G2 and G3 indicate individuals with restriction haplotype G from Montana, British Columbia and Alaska, respectively. Bootstrap values are indicated for nodes with ≥ 50% support. Each branch is labelled with the RFLP haplotype and the area in which the haplotype was found. Haplotypes have been separated into three groups (as in Fig. 2), each of which has a distribution approximately restricted to three geographical areas: I, Alaska/British Columbia; II, Montana; and III, Russia.

Our data may help to clarify how long the Montana grayling have been separated from the northern grayling populations, and the circumstances surrounding that separation. Both the RFLP and sequence results showed that the Montana grayling haplotypes are distinct from the northern assemblage of mtDNA haplotypes. In terms of the RFLP analysis, the difference between the two disjunct groups of grayling is one of haplotype frequency. The Montana samples were almost fixed for RFLP haplotype G (one individual had a variant, haplotype H), but haplotype G was found at frequencies of only 0.07 and 0.04 in Alaska and British Columbia, respectively. This extreme difference in haplotype frequency shows that the Montana grayling are isolated from the northern grayling. It does not, however, reveal how long the populations have been separated.

A particular difficulty with the RFLP analysis arises from the observation that haplotype G appears to be ancestral both to the Russian and British Columbian/Alaskan haplotypes (Fig. 2). This implies that haplotype G did not originate in the Montana population (unless by convergence to the ancestral state), but rather was already present in the *T. arcticus* species in general. This ancestral nature, combined with the fact that haplotype G is still found in the main northern population, strongly suggests that Montana grayling have also had a Bering origin and probably achieved fixation of haplotype G (and its derivative, haplotype H) by genetic drift and/or founder effects. Montana grayling may have been founded by the ancestral haplotype alone and/or have lost any other founding haplotypes owing to genetic drift. Therefore, the only conclusion that can be drawn from the RFLP analysis is that the Montana population is isolated from the northern population and has been so long enough for genetic drift to have reduced levels of haplotype diversity in the small population.

The sequence results were more informative regarding the Montana grayling. A single bp insertion in the D-loop region was found in the four Montana grayling sequenced and in only a single northern individual, a British Columbian fish characterized by RFLP haplotype G (labelled G2). This insertion was not found elsewhere, not even in an Alaskan individual characterized by RFLP haplotype G (labelled G3). The limited distribution of the insertion implies that this characteristic evolved in the Montana population, or the founders of the Montana population, but that the Montana and northern populations have been in contact since, enabling G2 to be present in British Columbia. It is worth noting that the 213-bp sequence haplotype G1/G2 is no more divergent from the common Alaskan/British Columbian sequence haplotype than are the other variant haplotypes present in Alaskan and British Columbian populations (haplotypes F and G3).

This 213-bp sequence, containing only a single mutation, was able to yield only a very approximate estimate of time since haplotype divergence, owing to the stochastic nature of the molecular clock (Wilson *et al*. 1977). To reduce this source of error, a longer fragment, encompassing both variable (*NADH-1* gene) and conserved (*tRNA* and *rRNA* genes) regions of the mitochondrial genome, was sequenced. When the two sequences were pooled, it was found that G1 and G3 (the Alaskan haplotype G without the T-insertion) were 0.37% divergent. This value implies 370 000 years since haplotype divergence, using the calibration of 1% sequence divergence per Myr, commonly used in fish studies (Smith 1992). However, this calibration is contentious, as alternative estimates of 2%/Myr (e.g. Becker *et al*. 1988; Billington *et al*. 1990; Bernatchez & Dodson 1991) are also used, and recent work has suggested that poikilothermic vertebrates have rates six to seven times slower than the 2% calibration, which was originally derived for homeotherms (Thomas & Beckenbach 1989; Avise *et al*. 1992; Martin & Palumbi 1993).

However approximate, this time estimate certainly implies that the haplotypes characteristic of the Montana grayling, or their founder population, diverged from the northern haplotypes prior to the Wisconsinan glaciation (which only represents about 70 000 years of separation). Furthermore, the populations have been in contact since that time, as the T-insertion is present in southern British Columbia. The sample location of G2 implies that British Columbia was colonized from both the Bering and a southern refuge after the Wisconsinan, although it is also possible that contact was made prior to the Wisconsinan, and the G2 individual arrived in British Columbia via the Bering.

There are two refugial hypotheses that satisfy both this timescale and the locality requirements of a refuge which could found the Montana population. The first hypothesis is that *T. arcticus* survived in the Upper Missouri system, part of the Missouri Refuge (Fig. 1). This hypothesis has been quite generally accepted for *T. arcticus* (McPhail & Lindsey 1970; Northcote 1993), primarily because one of the main types of evidence for refugial origin is distributional: when the current range of a species overlaps a past refugial area, it is assumed that the species survived in that refuge (Crossman & McAllister 1986).

The second refugial hypothesis is that *T. arcticus* survived the glaciation in a refuge postulated to have existed in southwest Alberta (Crossman & McAllister 1986). Geological evidence has suggested that part of southwest Alberta was unglaciated during the Wisconsinan glaciation (Lindsey & McPhail 1986). Both mitochondrial RFLP analysis of lake trout (*Salvelinus namaycush*, Wilson & Hebert 1998) and allozyme analysis of rainbow trout (*Oncorhynchus mykiss*, Carl *et al*. 1994) have suggested that

this unglaciated area may have harboured a refuge for freshwater fish. The presence of *T. arcticus* in this refuge has been suggested by 23 000–31 000-year-old fossil *T. arcticus* scales found in January Cave, in southwest Alberta (Burns 1991). If *T. arcticus* did survive in this refuge, it may have colonized Montana as proglacial lakes formed along the margins of the retreating ice sheets (Lindsey & McPhail 1986).

Given these two hypotheses, it is possible that *T. arcticus* survived in either or both of these southern refugia. Mitochondrial haplotypes cannot be used to determine which of these hypotheses are correct, because neither refuge has produced a second, independent population of grayling with which to compare haplotypes. Furthermore, the very low mtDNA haplotype diversity seen in the Montana population could be the result of a single refugial origin (i.e. either the southwest Alberta or the Upper Missouri Refuge), or it could simply be caused by genetic drift removing variation from small populations at the periphery of the *T. arcticus* range. It is not possible to tell whether the G2 haplotype migrated to British Columbia after the retreat of the ice sheets, or whether some contact previous to the Wisconsinan allowed the T-insertion to enter a northern refugial population and colonize British Columbia from the north.

It is also possible that *T. arcticus* may have had a second northern refuge distinct from the Bering Refuge. The first such possible refuge is a putative Nahanni Refuge, thought to have been present in the shifting corridor between the Laurentide and Cordilleran ice sheets in the Yukon Territory (Lindsey & McPhail 1986). Mitochondrial RFLP analysis of lake trout, *Salvelinus namaycush* (Wilson & Hebert 1998), and lake whitefish, *Coregonus clupeaformis* (Bernatchez & Dodson 1991) has suggested that these two species survived in a freshwater Nahanni Refuge. As these two coldwater species have life histories similar to *T. arcticus*, it is likely that grayling would have been able to survive in a Nahanni Refuge.

If grayling did survive in a Nahanni Refuge, the most likely place to find supporting evidence would be in the Upper Liard River above the Liard Canyon (McPhail, personal communication). Fish colonizing from the Nahanni Refuge would reach the Liard River sooner than those colonizing from the Bering Refuge, and so would be the first to establish populations. More importantly, these early colonizers would be more likely to use ephemeral postglacial lakes to bypass the Liard Canyon, which is currently a barrier to upstream fish migration (Lindsey & McPhail 1986). Our samples from the Upper Liard River (sites 6 and 7) have two variant haplotypes (haplotypes D and F), which were not observed in Alaska. These haplotypes greatly increased haplotype and nucleotide diversity in the Liard River compared with the Peace River in central British Columbia, as would be expected if the Liard River population had arisen from two refugial populations. While this data cannot confirm the presence of grayling in the Nahanni Refuge, it certainly is consistent with the possibility.

The second putative northern refuge for grayling was located on the north slope of the Brooks Range in Alaska, which was partially isolated from the main Bering Refuge during the Wisconsinan (Lindsey & McPhail 1986). McCart & Pepper (1971) suggested from lateral line scale counts that Alaskan grayling fell into two groups, with low counts north of the Brooks Range and along the Bering coast to Bristol Bay and with high counts in the Yukon Basin. Both populations were thought to have contributed to grayling recolonization of previously glaciated regions of northern Canada (McCart & Pepper 1971). While the difference in RFLP haplotypes between sample sites 3 and 4 (representing low and high-count grayling groups, respectively; McCart & Pepper 1971) is striking, further sampling in Alaska would be required to rigorously test for the existence or location of a putative second northern refuge for *T. arcticus* in western North America.

Acknowledgements

We would like to thank Derek Louie, who extracted the Peace and Missouri samples and completed the initial PCR and restriction trials, J. D. McPhail for helpful comments, and the many people who helped with sample collection around the world: Brian Blackman (British Columbia Ministry of Environment, Land and Parks, Prince George), John Wenberg (University of Washington), Pat Dwyer (U.S. Fish and Wildlife Service, Bozeman, Montana), Gerald Smith (University of Michigan), Jim Reynolds (University of Alaska, Fairbanks) and Ross Gardiner (Freshwater Fisheries Laboratory, Scottish Office of Agriculture, Environment, and Fisheries Department). This research was funded by a grant from NSERC (5-81696) awarded to E.B.T.

References

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

Avise JC, Bowen BW, Lamb T, Meylan AB, Bermingham E (1992) Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in the Testudines. *Molecular Biology of Evolution*, **9**, 457–473.

Becker II, Grant WS, Kirby R, Robb FT (1988) Evolutionary divergence between sympatric species of southern African Hakes, *Merluccius capensis* and *M. paradoxus*. II. Restriction enzyme analysis of mitochondrial DNA. *Heredity*, **61**, 21–30.

Bermingham E, Avise JC (1986) Molecular zoogeography of freshwater fishes in the southeastern United States. *Genetics*, **113**, 939–965.

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

Bernatchez L, Osinov A (1995) Genetic diversity of trout (genus *Salmo*) from its most eastern native range based on mitochondrial DNA and nuclear gene variation. *Molecular Ecology*, **4**, 285–297.

Billington N, Hebert PDN, Ward RD (1990) Allozyme and mitochondrial DNA variation among three species of *Stizostedion* (Percidae): phylogenetic and zoogeographical implications. *Canadian Journal of Fisheries and Aquatic Sciences*, **47**, 1093–1102.

Burns JA (1991) Mid-Wisconsinan vertebrates and their environment from January Cave, Alberta, Canada. *Quaternary Research*, **35**, 130–143.

Carl LM, Hunt C, Ihssen PE (1994) Rainbow trout of the Athabasca River, Alberta: a unique population. *Transactions of the American Fisheries Society*, **123**, 129–140.

Cope ED (1865) Partial catalogue of the cold-blooded vertebrata of Michigan. *Proceedings of the Academy of Natural Science*, **2**, 78–88.

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

Felsenstein J (1995) PHYLIP (Phylogeny Inference Package) Version 3.57c. Department of Genetics, University of Washington, Seattle, WA.

Gyllensten U, Wilson AC (1986) Mitochondrial DNA of salmonids: inter- and intraspecific variability detected with restriction enzymes. In: *Population Genetics and Fishery Management* (eds Ryman N, Utter F), pp. 301–317. University of Washington Press, Seattle.

Higuchi M, Goto A (1996) Genetic evidence supporting the existence of two distinct species in the genus *Gasterosteus* around Japan. *Environmental Biology of Fishes*, **47**, 1–16.

Hillis DM, Bull JJ (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology*, **42**, 182–192.

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

Hop H, Gharrett AJ (1989) Genetic relationships of Arctic grayling in the Koyokuk and Tanana rivers, Alaska. *Transactions of the American Fisheries Society*, **118**, 290– 295.

Jaarola M, Tegelström H (1996) Mitochondrial DNA variation in the field vole, *Microtus agrestis*: regional population structure and colonization history. *Evolution*, **50**, 2073–2085.

Kimura M (1980) A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.

Kocher TD, Thomas WK, Meyer A *et al*. (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Acadamy of Sciences of the USA*, **86**, 6196–6200.

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

Lynch JC, Vyse ER (1979) Genetic variability and divergence in grayling, *Thymallus arcticus*. *Genetics*, **92**, 263–278.

Martin AP, Palumbi SR (1993) Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Acadamy of Sciences of the USA*, **90**, 4087–4091.

Martinson DG, Pisias NG, Hays JD, Imbrie J, Moore TC Jr, Shackleton NJ (1987) Age dating and the orbital theory of the ice ages: development of a high-resolution 0–300 000-year chronostratigraphy. *Quaternary Research*, **27**, 1–29.

McCart P, Pepper VA (1971) Geographic variation in the lateral line scale counts of the Arctic grayling, *Thymallus arcticus*. *Journal of the Fisheries Research Board of Canada*, **28**, 749–754.

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

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

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

Miller BB, Palmer DF, McCoy WD, Smith AJ, Colburn ML (1993) A pre-Illinoian Pleistocene fossil assemblage from near Connersville, Southeastern Indiana. *Quaternary Research*, **40**, 254–261.

Milner JW (1874) Notes on the grayling of North America (*Thymallus tricolor*). Report of the U.S. Fish Commission, Vol. 2 for 1872 and 1873, pp. 729–742.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variation in populations. *Evolution*, **29**, 1–10.

Nei M, Miller JC (1990) A simple method for estimating average number of nucleotide substitutions within and between populations from restriction data. *Genetics*, **125**, 873–879.

Nei M, Tajima F (1981) DNA polymorphism detectable by restriction endonucleases. *Genetics*, **97**, 145–163.

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

Nielson JL, Gan C, Thomas WK (1994) Differences in genetic diversity for mitochondrial DNA between hatchery and wild populations of *Oncorhynchus*. *Canadian Journal of Fisheries and Aquatic Sciences*, **51** (Suppl. 1), 290–297.

Norden CR (1961) Comparative osteology of representative salmonid fishes, with particular reference to the grayling (*Thymallus arcticus*) and its phylogeny. *Journal of the Fisheries Research Board of Canada*, **18**, 679–791.

Northcote TG (1993) *A Review of Management and Enhancement Options for the Arctic grayling (*Thymallus arcticus*) with Special Reference to the Williston Reservoir Watershed in British Columbia*. B.C. Ministry of Environment, Land and Parks, Fisheries Management Report no. 101.

Park LK, Brainard MA, Dightman DA, Winans GA (1993) Low levels of intraspecific variation in the mitochondrial DNA of chum salmon (*Oncorhynchus keta*). *Molecular Marine Biology and Biotechnology*, **2**, 362–370.

Park LK, Moran P (1994) Developments in molecular genetic techniques in fisheries. *Reviews in Fish Biology and Fisheries*, **4**, 272–299.

Reed RJ, McCann JA (1973) Analyses of some meristic and morphometric data from the Arctic grayling, *Thymallus arcticus*, in Alaska. *Copeia*, **1973**, 819–822.

Scott WB, Crossman EJ (1973) *Freshwater Fishes of Canada*. Bulletin of the Fisheries Research Board of Canada, no. 184, Ottawa.

Shedlock AM, Haygood MG, Pietsch TW, Bentzen P (1997) Enhanced DNA extraction and PCR amplification of mitochondrial genes from formalin-fixed museum specimens. *Biotechniques*, **22**, 394–400.

Skopets MB (1991) Biological features of a subspecies of the Arctic grayling in northeast Asia. II. The Alaskan grayling, *Thymallus arcticus signifer*. *Journal of Ichthyology*, **31**, 87–102.

Skopets MB, Prokop'yev NM (1990) Biological characteristics of subspecies of the Arctic grayling in northeastern Asia. I. The Kamchatkan grayling, *Thymallus arcticus mertensi*. *Journal of Ichthyology*, **30**, 43–58.

Smith GR (1992) Introgression in fishes: significance for paleontology, cladistics and evolutionary rates. *Systematic Biology*, **41**, 41–57.

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.

Taylor EB, Harvey S, Pollard S, Volpe J (1997) Postglacial genetic differentiation of reproductive ecotypes of kokanee *Oncorhynchus nerka* in Okanagan Lake, British Columbia. *Molecular Ecology*, **6**, 213–216.

Thomas WK, Beckenbach AT (1989) Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanisms of substitution. *Journal of Molecular Evolution*, **29**, 233–245.

Vachot A-M, Monnerot M (1996) Extraction, amplification and sequencing of DNA from formaldehyde-fixed specimens. *Ancient Biomolecules*, **1**, 3–16.

Walters V (1955) Fishes of Western Arctic America and Eastern Arctic Siberia. *Bulletin of the American Museum of Natural History*, **106**, 257–368.

Williams RN, Evans RP, Shiozawa DK (1995) Mitochondrial DNA diversity in bull trout in the Upper Columbia River basin. pp. 283–297. In: Mackay WC, Brewin MK, Monita M (eds.) Friends of the bull trout Conference Proceedings. Bull trout Task Force (Alberta, Canada). Trout Unlimited Canada, Calgary, Canada.

Wilson AC, Carlson SS, White TJ (1977) Biochemical evolution. *Annual Review of Biochemistry*, **46**, 573–639.

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.

Molecular work and analysis was carried out by Zoë Redenbach for the Honours Thesis of a BSc degree. Her primary interest is molecular, morphological and ecological evolution. The study was initiated and supported by Eric Taylor, whose interests focus on the evolution, ecology and conservation of fishes.
