Molecular Ecology (2004) **13**, 1533–1549 doi: 10.1111/j.1365-294X.2004.02174.x

Phylogeographical lineages of Arctic grayling (*Thymallus arcticus*) in North America: divergence, origins and affinities with Eurasian *Thymallus*

M.D. STAMFORD*† and E.B. TAYLOR*

**Department of Zoology and Native Fish Research Group, University of British Columbia, 6270 University Blvd., Vancouver, British Columbia, Canada, V6T 1Z4; †RR#3 COMP, Gibsons, British Columbia, Canada*

Abstract

The number and location of Arctic glacial refugia utilized by taxa during the Pleistocene are continuing uncertainties in Holarctic phylogeography. Arctic grayling (*Thymallus arcticus*) are widely distributed in freshwaters from the eastern side of Hudson Bay (Canada) west to central Asia. We studied mitochondrial DNA (mtDNA) and microsatellite DNA variation in North American *T. arcticus* to test for genetic signatures of survival in, and postglacial dispersal from, multiple glacial refugia, and to assess their evolutionary affinities with Eurasian *Thymallus*. In samples from 32 localities, we resolved 12 mtDNA haplotypes belonging to three assemblages that differed from each other in sequence by between 0.75 and 2.13%: a 'South Beringia' lineage found from western Alaska to northern British Columbia, Canada; a 'North Beringia' lineage found on the north slope of Alaska, the lower Mackenzie River, and to eastern Saskatchewan; and a 'Nahanni' lineage confined to the Nahanni River area of the upper Mackenzie River drainage. Sequence analysis of a portion of the control region indicated monophyly of all North American *T. arcticus* and their probable origin from eastern Siberian *T. arcticus* at least 3 Mya. Arctic grayling sampled from 25 localities displayed low allelic diversity and expected heterozygosity (H_E) across five microsatellite loci (means of 2.1 alleles and 0.27 H_E, respectively) and there were declines in these measures of genetic diversity with distance eastward from the lower Yukon River Valley. Assemblages defined by mtDNA divergences were less apparent at microsatellite loci, but again the Nahanni lineage was the most distinctive. Analysis of molecular variance indicated that between 24% (microsatellite DNA) and 81% (mtDNA) of the variance was attributable to differences among South Beringia, North Beringia and Nahanni lineages. Our data suggest that extant North American Arctic grayling are more diverse phylogeographically than previously suspected and that they consist of at least three major lineages that originated in distinct Pleistocene glacial refugia. *T. arcticus* probably originated and dispersed from Eurasia to North America in the late to mid-Pliocene, but our data also suggest more recent (mid–late Pleistocene) interactions between lineages across Beringia.

Keywords: Holarctic phylogeography, microsatellites, mtDNA, postglacial dispersal, *Thymallus*

Received 2 September 2003; revision received 15 January 2004; accepted 9 February 2004

Introduction

Dramatic climate change throughout the Pleistocene epoch has profoundly shaped biodiversity in north temperate and Arctic ecosystems (Pielou 1991; Hewitt 2000). Survival in these habitats over the past 2.5 Myr depended on the ability of species to tolerate climatic oscillations between frigid glaciation and warm interglacial periods that roughly followed a 90 000-year cycle (Pielou 1991). Freshwater faunas were particularly vulnerable to such changes because land and high salinity of the oceans restricted their ability to disperse between watersheds, and phylogeographical analyses have revealed strong signatures of such perturbations (e.g. reviewed in Bernatchez & Wilson 1998). For instance,

Correspondence: E. B. Taylor. Tel.: +1 604 822 9152; Fax: +1 604 822 2416; E-mail: etaylor@zoology.ubc.ca

one generality apparent among north temperate species is that the greatest genetic diversity tends to occur in regions located in or close to putative glacial refugia with a progressive decline at more distant localities (e.g. Sage & Wolff 1986; Merilä *et al*. 1996; Costello *et al*. 2003). This observation probably reflects sequential bottlenecks and founder events during postglacial range expansion that took place behind the receding glaciers of the last glaciation that ended between 10 and 12 000 years ago (Hewitt 1996; Ibrahim *et al*. 1996). Further, studies of phylogeography have revealed that repeated fluctuations in a species' range combined with long periods of isolation have resulted in the evolution of intraspecific phylogenetic lineages whose geographical distributions are strongly associated with the locations of putative Pleistocene refugia (reviewed in Avise 2000; Hewitt 2000).

Despite these general effects of Pleistocene glaciations, the extent and timing of ice coverage combined with variation in the vagility of individual species have generated distinct phylogeographical signatures found within particular regions and some strong contrasts between northern glaciated and southern nonglaciated regions for Nearctic and Palearctic freshwater fishes (Bernatchez & Wilson 1998). For instance, phylogeographical analysis of the lake whitefish species complex (*Coregonus clupeaformis* and *C. laveratus*) suggested that greater diversity exists in Eurasian lineages compared with those in North America which is consistent with the idea that some eastern Palearctic lineages were impacted less severely by Pleistocene glaciations (e.g. Bernatchez *et al*. 1989; Bernatchez & Dodson 1994; Brunhoff *et al*. 2003). For fishes with a contemporary Holarctic distribution, the smaller degree of ice cover in Eurasia may have provided a reservoir of biodiversity from which lineages dispersed through Siberia and Beringia into northwestern North America during Pleistocene interglacial periods (Lindsey & McPhail 1986; Skopets 1991; Bernatchez & Dodson 1994; see also Fedorov *et al*. 2003 for small mammals). Curiously, however, dispersal of Eurasian lineages of *C. clupeaformis* and of the burbot, *Lota lota*, was limited to Arctic watersheds in northwestern North America (e.g. Beringia, Bernatchez & Dodson 1994; van Houdt *et al*. 2003). Similarly, Fedorov & Stenseth (2002) and Fedorov *et al*. (2003) reported limited dispersal of Beringian lineages of lemmings (*Dicrostonyx* and *Lemmus*) to previously glaciated regions of North America. Clearly, further studies of phylogeography of cold-adapted taxa in North America are needed to identify and better define the influence of Arctic refugia on extant phylogeographical patterns.

Arctic grayling (*Thymallus arcticus*, Salmonidae) are distributed largely in mainland Arctic drainages from the west coast of Hudson Bay west to northern British Columbia, Alaska, and eastern and central Russia. At one time there were also two disjunct southern populations in North America: in the upper Missouri River system in Montana and in tributaries of Lake Michigan, but the latter has now been extirpated. Redenbach & Taylor (1999) examined mitochondrial DNA (mtDNA) variation in Arctic grayling sampled from much of North America and concluded that Beringia was the major Pleistocene refuge for the species from which it dispersed postglacially. Arctic grayling in Montana, however, are genetically distinct from grayling in the northern part of the species' range (Lynch & Vyse 1979; Redenbach & Taylor 1999), which suggests that they may represent descendants of fish that survived the Pleistocene glaciations in the Great Plains (upper Missouri River) glacial refuge (Lindsey & McPhail 1986). In addition, morphological evidence suggests that there are two distinct lineages in northwestern North America; north and south of the Brooks Range in the Alaskan portion of Beringia (McCart & Pepper 1971; Reed 1973). There is also evidence that points to an inland glacial refuge (Nahanni Refuge) that harboured freshwater fish in the Northwest Territories (Lindsey *et al*. 1981; Foote *et al*. 1992). Finally, now extirpated *T. arcticus* from Michigan suggest survival in the upper Mississippi River valley. There are, therefore, potentially five geographical areas that may have served as refugia for *T. arcticus*: two in Beringia (north and south of the Brooks Range in Alaska), and one each in the Nahanni, Missouri and upper Mississippi river valleys. Currently, however, there is only enough evidence to suggest the existence of two refugia in North America for *T. arcticus*: a more general Beringia refuge and an upper Missouri River refuge (Redenbach & Taylor 1999).

Previous work on other fishes has provided evidence of a Eurasian origin for northwestern North American lineages (e.g. Bernatchez & Dodson 1994; van Houdt *et al*. 2003). Recent phylogeographical work on far eastern Asian *T. arcticus* has placed variation within the species into a broader context of Eurasian *Thymallus* (Koskinen *et al*. 2002a), but the only information on the interrelationships between Eursian *Thymallus* and North American *T. arcticus* consists of comparsions among very few localities or haplotypes (Redenbach & Taylor 1999; Koskinen *et al*. 2002a). Because most of the taxonomic diversity within *Thymallus* resides in Asia (Schoffmann 2000; Koskinen *et al*. 2002a), the recent data for Asian *Thymallus*, provides an opportunity to test for ancestral affinities of North American lineages with one or more of several candidate areas in northeast Asia.

This study extends earlier analysis of the phylogeography of Arctic grayling in North America to better investigate the evolutionary history of a species whose distribution spans Beringia, an important biogeographical 'crossroads' between Eurasian and North American taxa (Sher 1999). Specifically, we assayed populations for mtDNA and microsatellite DNA variation to test for distinct phylogenetic lineages and their association with areas proposed as refugia in North America, particularly in the Nahanni River area, where evidence for a refuge is

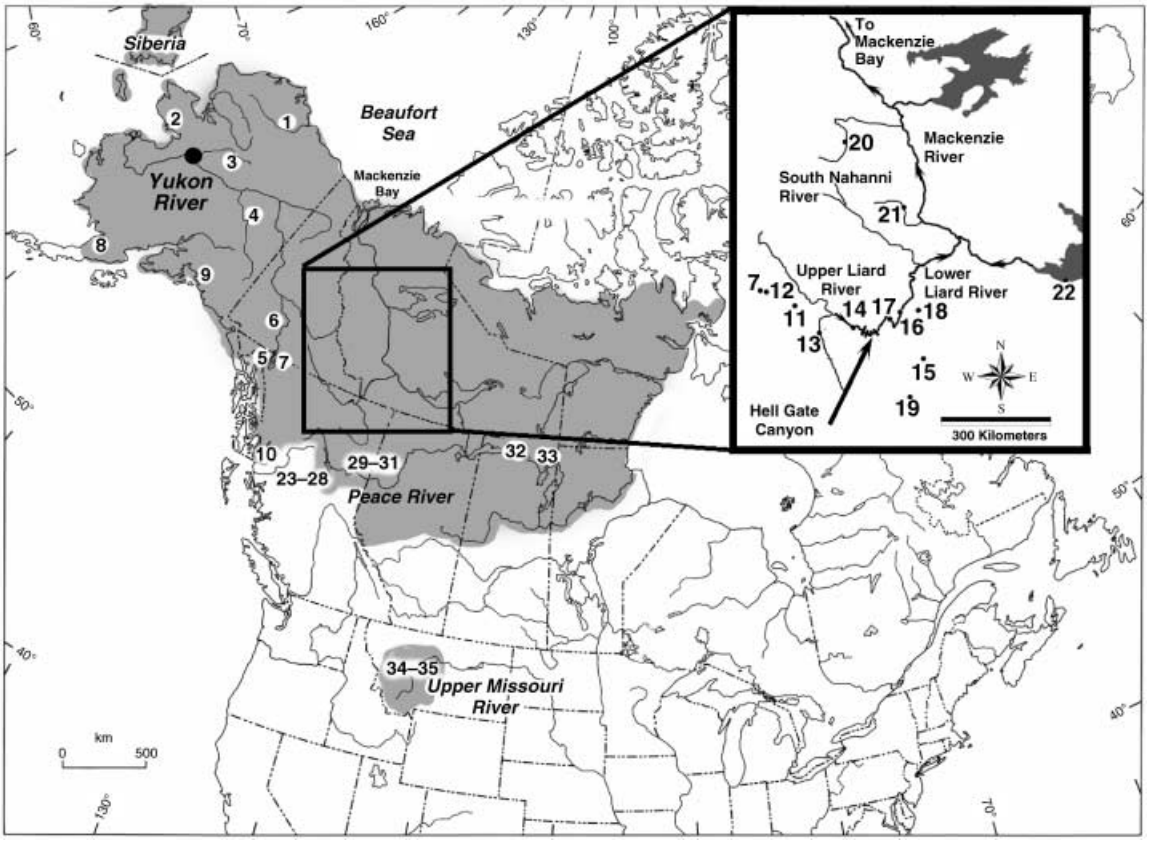

Fig. 1 Geographic distribution of Arctic grayling sample localities. Number codes refer to place names in Table 1. Shaded area shows the geographical distribution of *Thymallus arcticus* in North America and a small portion of Siberia (after Scott & Crossman 1973). Inset shows location of sample areas within the Liard/Mackenzie rivers. The black dot on the lower Yukon River indicates the location from which pairwise geographical distances were recorded for all localities.

controversial. In addition, we assess relationships among *trans*-Beringian *Thymallus* to infer the phylogenetic and temporal origin of North American grayling from Eurasian lineages.

Materials and methods

Sample collection

Arctic grayling tissue samples were collected for mtDNA analysis from 32 localities distributed throughout the species' range in North America (Table 1, Fig. 1). For microsatellite analysis, samples came from 25 localities (Table 1, Fig. 1). Samples consisted of fin tissue that was preserved in 95% ethanol in the field.

Mitochondrial DNA analysis

Total genomic DNA was extracted from the tissue samples using either phenol–chloroform extraction (Taggart *et al*. 1992) or the PUREGENE (Gentra Systems) extraction kit. The DNA precipitate was resuspended in TE buffer then stored at –20 °C. The polymerase chain reaction (PCR) was used to amplify two adjacent fragments of mtDNA: a 2100 bp fragment that included the cytochrome *b* gene and the noncoding control region, and a 2500 bp fragment encompassing genes 5 and 6 in the NADH subunit (ND 5/6). Primers and PCR conditions for both regions are described in Redenbach & Taylor (1999).

Restriction fragment length polymorphisms (RFLPs) in mtDNA were resolved with 15 restriction enzymes (New England Biolabs): *Alu*I, *Ava*II, *Ban*I, *Bfa*I, *Bst*UI, *Dde*I, *Dpn*II, *Hae*III, *Hha*I, *Hin*cII, *Hin*fI, *Msp*I, *Rsa*I, *Sty*I and *Taq*I. Each enzyme was incubated with a mixture of both PCR fragments, the products were then run alongside a 1 kb BRL size standard on 2% agarose gel and stained with ethidium bromide. The two PCR products were digested and electrophoresed once separately for each unique restriction site sequence (fragment pattern) for all enzymes to accurately determine the presence or absence of restriction sites. Some geographical samples were diagnosed into major mtDNA lineages (see below) by incubating their mtDNAs with restriction enzymes diagnostic for each lineage (e.g. *Hin*fI and *Msp*I see Table 2).

Table 1 Sample localities, longitude, latitude, sample sizes, haplotype diversity and nucleotide diversity of Arctic grayling mtDNA. Sample sites, sample sizes and mean (SD) number of alleles, and observed and expected heterozygosity across five microsatellite loci for Arctic grayling. The diversity measures were derived from transformed allele frequencies (see text) and sample localities are listed from the northwestern (Beringia at the top) to the southeastern (Montana) part of the species' range in North America (see Fig. 1). Number preceding each locality designates locations in Figs 1 and 3. The Becharof Lake and Copper River mtDNA samples were assayed for major mtDNA grouping only so no diversity measures were calculated

Population	Longitude	Latitude	mtDNA *N*	Haplotype diversity	Nucleotide diversity (×100)	Microsatellite *N*	N_a	H_O	H_E
Beaufort Sea									
1. Sagavonirktok River	147°50′ W	70°13′ N	17	0.5348	0.0874	20	2.6 (0.065)	0.370 (0.207)	0.323 (0.228)
Seward Peninsula									
2. Niukluk River	163°25′ W	64°49′ N	21	0.5110	0.3543	30	3.0 (0.043)	0.452 (0.258)	0.420 (0.247)
Lower Yukon River									
3. Koyukuk River	150°21′ W	67°17′ N	—	—	—	24	2.6 (0.054)	0.346 (0.302)	0.339 (0.291)
4. Chena River	147°52′ W	64°49′ N	20	0.5795	0.1627	20	2.8 (0.065)	0.363 (0.253)	0.370 (0.297)
Upper Yukon River									
5. Atlin River	133°49′ W	59°36′ N	—	—	—	20	1.4 (0.129)	0.136 (0.173)	0.133 (0.207)
6. Teslin River	134°53′ W	61°53′ N	15	0.4598	0.0548	15	2.2 (0.086)	0.323 (0.225)	0.280 (0.184)
7. Plate Lake—	130°45′ W	59°54′ N	5	0.0000	0.0000	—	—	—	—
Pacific Coast									
8. Becharof Lake	156°21′ W	57°55′ N	10	—	—	24	1.8 (0.054)	0.152 (0.259)	0.137 (0.199)
9. Copper River	145°28′ W	62°18′ N	18	—	—	24	2.4 (0.054)	0.284 (0.259)	0.229 (0.199)
10. Stikine River	131°51′ W	56°39′ N	10	0.0000	0.0000	—	—	—	—
Upper Liard River									
11. Blue River	128°58′ W	59°42′ N	19	0.3983	0.0942	19	2.4 (0.068)	0.329 (0.213)	0.295 (0.215)
12. Upper Tootsie Lake	130°27′ W	59°54′ N	5	0.0000	0.0000	—	—	—	—
13. Turnagain River	127°36′ W	59°09′ N	4	—	—	—	—	—	—
14. Trout River	126°01′ W	59°24′ N	3	—	—	—	—	—	—
Lower Liard River									
15. Muskwa River	122°33′ W	58°50′ N	6	0.0000	0.0000	19	2.4 (0.068)	0.285 (0.308)	0.332 (0.284)
16. LaBiche River	123°52′ W	59°55′ N	15	0.6805	1.1057	18	2.2 (0.072)	0.294 (0.274)	0.300 (0.268)
17. Beaver River	124°19′ W	59°43′ N	13	0.1477	0.3141	17	2.6 (0.076)	0.445 (0.181)	0.346 (0.110)
18. Petitot River	122°59′ W	60°00′ N	2	—	—	—	—	—	
19. Minnaker River	123°01′ W	57°52′ N	2	—	—	—	—	—	—
Mackenzie River									
20. Keel River	127°54′ W	63°51′ N	4	—	—	—	—	—	—
21. North Nahanni River	123°53′ W	63°43′ N	4	—	—	—	—	—	—
22. Great Slave Lake	115°45′ W	60°51′ N	16	0.1210	0.0144	16	2.8 (0.081)	0.378 (0.252)	0.296 (0.265)
Upper Peace River									
23. Table River	122°18′ W	54°42′ N	20	0.0000	0.0000	60	1.4 (0.022)	0.184 (0.252)	0.197 (0.270)
24. Anzac River	122°32′ W	54°48′ N	20	0.0000	0.0000	60	1.6 (0.022)	0.190 (0.259)	0.187 (0.252)
25. Nation River	123°31′ W	55°30′ N	20	0.0000	0.0000	49	1.4 (0.026)	0.200 (0.274)	0.184 (0.250)

Table 1 *Continued*

Population	Longitude	Latitude	mtDNA *N*	Haplotype diversity	Nucleotide diversity (x100)	Microsatellite N	*Na*	H_O	H_E
26. Mesilinka River	124°28′ W	56°08′ N	20	0.0974	0.0114	36	1.6 (0.036)	0.207 (0.290)	0.229 (0.308)
27. Ingenika River	125°02′ W	56°44′ N	20	0.0000	0.0000	23	1.6 (0.056)	0.199 (0.285)	0.209 (0.292)
28. Finlay River	124°57′ W	56°54′ N	—	—	—	29	1.4 (0.046)	0.106 (0.199)	0.111 (0.210)
Lower Peace River									
29. Burnt River	121°44′ W	55°20′ N	24	0.1560	0.1078	34	1.6 (0.038)	0.145 (0.211)	0.148 (0.210)
30. Beatton River	120°23′ W	56°06′ N	14	0.0000	0.0000	14	1.6 (0.092)	0.192 (0.265)	0.209 (0.293)
31. Halfway River	121°27′ W	56°13′ N	—	—	—	18	1.8 (0.072)	0.148 (0.235)	0.178 (0.289)
Saskatchewan									
32. Black Lake	103°57′ W	58°56′ N	10	0.0000	0.0000	20	1.4 (0.065)	0.272 (0.255)	0.210 (0.227)
33. Wollaston Lake	101°11′ W	58°20′ N	20	0.0000	0.0000	20	2.0 (0.065)	0.326 (0.239)	0.281 (0.209)
Upper Missouri River									
34. Madison River	111°281′ W	45°15′ N	2	—	—	—	—	—	—
35. Big Hole River	112°15′ W	45°25′ N	1	—	—	—	—	—	—

We also sequenced ~790 bp of the control region for each of the RFLP haplotypes that we resolved as well as those resolved by Redenbach & Taylor (1999). The sequence analysis was conducted to examine the relationships of North American *Thymallus arcticus* haplotypes relative to *Thymallus* haplotypes in Siberia resolved by Koskinen *et al*. (2002a) and Froufe *et al*. (2003). The control region sequences were obtained by first amplifying a ~1.2 kb fragment using primers and conditions outlined in Froufe *et al*. (2003). The fragment was purified using Qiagen columns, ethanol precipitated, air-dried and cycle sequenced using the LRBT-25 primer on an ABI automated sequencer using BigDye Terminator methods.

Microsatellite DNA analysis

Tests for polymorphism using published primer sets and optimization of PCR conditions were carried out using ^{32}P end-labelled primers, then running the PCR products on 6% polyacrilamide gels. General conditions for PCR and gel electrophoresis can be found in Costello *et al*. (2003). Initial tests for polymorphism were carried out on 10 Arctic grayling samples that were selected from a wide geographical range (e.g. lower Liard River, Yukon River and upper Peace River). A total of 16 loci were tested, only three of which resolved polymorphisms with PCR: *One 8*, *Ots 100* and *BFRO004* (Scribner *et al*. 1996; Nelson & Beacham 1999; Snoj *et al*. 1999). Annealing temperatures for these loci were 60, 60 and 63 °C, respectively. Two polymorphic dinucleotide microsatellite loci from an Arctic grayling partial genomic library were also employed: *Tar 1* (forward primer: 5′-ACATATCATTCTTAGCATATC-3′; reverse primer 5′-CAAAATAGTAATTGAAATGC-3′, 50 °C) and *Tar 8* (forward primer: 5′-GAAGTGGTGGATGTGAATTT-3′; reverse primer: 5′-GCCATGCATTTTTACCTCTA-3′, 54 °C, E.B. Taylor, unpublished data). Our objective for the microsatellite analysis was to assess the presence of the same broad geographical groupings revealed by our more extensive mtDNA analyses.

Data analysis

The mtDNA restriction site data was converted to binary (1, 0) character matrix format that was imported into the Restriction Enzyme Analysis Package (REAP, McElroy *et al*. 1992). This was used to calculate the number of nucleotide substitutions per site, *d* (Nei & Tajima 1981; Nei & Miller 1990), weighted by the class of restriction enzymes used (Nei & Tajima 1983). Haplotype diversity and nucleotide diversity within populations and nucleotide divergence between all population pairs were also determined using the DA program from REAP employing algorithms described by Nei & Tajima (1981) and Nei (1987).

Cluster analysis among haplotypes and populations was carried out using PHYLIP Version 3.57c (Felsenstein 1995).

Table 2 Geographic distribution of 12 mtDNA composite haplotypes resolved with 15 restriction enzymes. The haplotypes (1–12) are AAAAAAAAAAAAAAA, AAAAAAAAAAABAAA, AAAABAAAAAAAAAA, AAABBAAAAAAAAAA, AACABAAAAAAAAAA, AAABCAAAAAAAAAA, BAAAABAAAABCBAA, BAAAABAAAABABAAA, BAAAABAAAABEBAA, BAAAAAAAAAABABAA, BABCDCBABACDCAA and BABCECBABACDCAA resolved using enzymes *Alu*I, *Ban*I, *Hae*III, *Hin*fI, *Rsa*I, *Dpn*II, *Sty*I, *Taq*I, *Ava*II, *Bfa*I, *Bst*UI, *Dde*I, *Hha*I, *Hinc*II and *Msp*I. 'A', 'B' and 'C' refer to the major haplotype lineages (South Beringia, North Beringia and Nahanni, respectively) shown in Fig. 2. '?' indicates samples that were assigned to major lineage only, using restriction enzymes diagnostic for each lineage, and for which specific haplotype definitions are unknown. Major geographical regions or drainage basins are italicized, major tributaries are then arranged within each region or basin. Numbers before each locality designate location in Figs 1 and 3. Locations 3, 5, 28 and 31 had no mtDNA data and, consequently, are not listed below

	Composite haplotype													
	A							B					C	
Geographic locality	1	2	3	4	5	6	?	7	8	9	10	?	11	12
Beaufort Sea														
1. Sagavonirktok River									3	11	3			
Seward Peninsula														
2. Niukluk River							10					11		
Lower Yukon River														
4. Chena River	10		7						2					
Upper Yukon River														
6. Teslin River	10		5											
7. Plate Lake			5											
Pacific Coast														
8. Becharof Lake							10							
9. Copper River							18							
10. Stikine River				10										
Upper Liard River														
11. Blue River	14			5										
12. Upper Tootsie River	5													
13. Turnagain River	1			3										
14. Trout River				2	1									
Lower Liard River														
15. Muskwa River													6	
16. Beaver River				1									12	
17. LaBiche River	2			6		1							6	
18. Petitot River													2	
19. Minnaker River				1									1	
Mackenzie River														
20. Keel River							4							
21. North Nahanni River													4	
22. Great Slave Lake													15	1
Upper Peace River														
23. Table River	20													
24. Anzac River	20													
25. Nation River	20													
26. Mesilinka River	19	1												
27. Ingenika River	20													
Lower Peace River														
29. Burnt River	22							2						
30. Beatton River	14													
Saskatchewan														
32. Black Lake								10						
33. Wollaston Lake								20						
Upper Missouri River														
34. Madison River								2						
35. Big Hole River								1						

The restriction site matrix was imported into PHYLIP and replicated 1000 times using SEQBOOT, then a most parsimonious tree was built from these replicate data sets using Wagner MIX parsimony and the CONSENSE programs in PHYLIP. Nucleotide divergence estimates among haplotypes and nucleotide divergence estimates from all population pairs were used to generate neighbour-joining dendrograms (NJ; Saitou & Nei 1987). Control region sequences were aligned using BIOEDIT (Hall 1999; available at: http://www.mbio.ncsu.edu/BioEdit/bioedit.html) with gaps treated as single characters. Sequences obtained were compared with 19 sequences representing haplotypes from Siberian *T. arcticus*, *T. grubii* and *T. brevirostris* obtained by Koskinen *et al*. (2002a) or Froufe *et al*. (2003) (Accession nos. AY168358, AY168365, AY168368, AY168370, AY168378, AY168379, AY168384, AY168386, AY168389, AY168397, AY168398, AY168400–402, AY168405, AY246414, AY246420, AY246423–424). These sequences were selected to represent a range of divergences within each of the major clades resolved by Koskinen *et al*. (2002a) and Froufe *et al*. (2003). We calculated measures of sequence divergence (Kimura 2-parameter distance) and clustered these estimates using NJ. We also subjected the sequence matrices to parsimony and maximum-likelihood analyses using routines in PAUP (Version 4b10, Swofford 2002). Parsimony analysis incorporated heuristic searches with 10 replicates and TBR branch swapping. Likelihood analyses employed the Hasegawa–Kishino–Yano model (Hasegawa *et al*. 1985) and heuristic searches (10 replicates) were used to find the most likely tree. The resulting tree topologies were identical among all three tree-building algorithms and therefore only the NJ tree is presented. All sequences obtained in this study have been deposited in GenBank under Accession nos. AY528426-439.

Descriptive statistics of microsatellite loci included expected and observed heterozygosity, average number of alleles per locus, and Hardy–Weinberg equilibrium exact tests, which were calculated using GENEPOP Version 3.1 (Raymond & Rousset 1995). We used the CONTML algorithm of PHYLIP to derive and cluster maximum-likelihood estimates of population interrelationships. Partitioning of the genetic variance of both mtDNA and microsatellite loci among populations and among distinct geographical regions was analysed with analysis of molecular variance (AMOVA; Excoffier *et al*. 1992). Mitochondrial DNA RFLP haplotype and microsatellite DNA allele frequencies were imported into ARLEQUIN (Schneider *et al*. 1997) for these analyses. Specifically, we arranged populations into putative refugial units suggested by the geographical distribution of mtDNA haplotype groups and assessed the levels of among group variance compared with other geographical groupings that reflected contemporary watershed divides (e.g. east vs. west of the continental divide). Straight-line geographical distances of all populations from the lower Yukon River (at the confluence of the Koyukuk and Yukon rivers) were determined from topographic maps. These distances were correlated with measures of genetic diversity within populations (haplotype diversity, mean number of alleles, etc.). If Beringia indeed served as a major refuge for Arctic grayling, we expected to observe a reduction in genetic diversity in populations located progressively farther from the North American centre of this refuge, i.e. a negative correlation between measures of genetic diversity and geographical distance. We tested this relationship for Beringia only because it had the greatest number of localities sampled and was probably the largest refuge.

Results

Mitochondrial DNA diversity and phylogenetic relationships among haplotypes

Restriction enzyme analysis of Arctic grayling mtDNA surveyed 118 restriction sites and 446 bp, representing 9.6% of the ND 5/6, cytochrome b/D-loop fragment. We resolved 12 composite haplotypes that differed from each other by an average pairwise nucleotide divergence of 0.92% (range: 0.1–2.3%, restriction site matrix available from E.B. Taylor).

The topologies from maximum parsimony and neighbour-joining clustering of sequence divergence estimates among the 12 haplotypes each resolved three groupings (Fig. 2). One group ('A' in Fig. 2) consisted of haplotypes 1–6. Another group ('B') consisted of haplotypes 7–10 and was distinguished from group A with 95% bootstrap support. A third group ('C') consisted of haplotypes 11–12 and was resolved with 100% bootstrap support (Fig. 2).

Average nucleotide divergence among the haplotype groups was greatest between groups A and C (2.1%, range

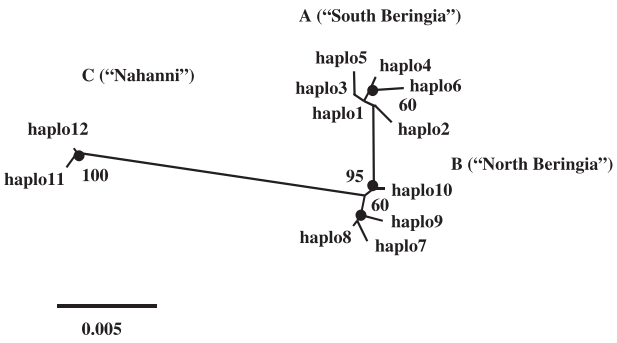

Fig. 2 Neighbour-joining tree of interrelationships among *Thymallus* mtDNA haplotypes inferred from estimates of pairwise nucleotide divergence (*d*) among 12 mtDNA RFLP haplotypes from North American Arctic grayling resolved with 15 restriction enzymes. Numbers at branch points represent the bootstrap support levels from 1000 pseudoreplicate analyses. Only mean bootstrap values > 50% are shown.

1.7–2.3%) and between groups B and C (1.6%, range 1.5–1.7%). There were 21 restriction site changes between groups A and C and 14 changes between groups B and C. Groups B and C had similar recognition sites with *Alu*I, *Rsa*I, *Bst*UI, *Dde*I and *Hha*I. By contrast, although groups A and C show site changes with the same restriction enzymes (*Dpn*II, *Hae*III and *Hin*fI), they involve different site changes within the same recognition sequence, suggesting that these groups diverged independently. Groups A and B differed from each other the least (0.75%, range 0.45–0.96%) and variation within lineages A, B and C averaged 0.25, 0.17 and 0.12%, respectively.

Mitochondrial RFLP haplotype distribution and differentiation among assemblages

Group A haplotypes (1–6) were distributed from northern Alaska, along the north Pacific Coast to the Peace River in central British Columbia (Table 2, Fig. 3). Haplotype 1 was by far the most common (183 individuals, 49% of the whole data set) and was widely distributed from the lower Yukon River in western Alaska to the lower Peace River and lower Liard River in central and northern British Columbia (Table 2, Fig. 3). The remaining haplotypes in group A were generally more locally distributed in interior Alaska and northern British Columbia. Given the distribution of these haplotypes in the southern and central areas of Beringia, we have designated group A haplotypes as the 'South Beringia' lineage.

Group B (haplotypes 7–10) ranged most widely, from northern Alaska into the lower Mackenzie River, south to the upper Missouri River, and as far east as Wollaston Lake, Saskatchewan (Table 2, Fig. 3). Haplotypes 8, 9 and 10 were predominant in the north and haplotype 7 was the only one found in the extreme eastern and southern portions of the study area. Using diagnostic enzymes, 11 individuals from the Niukluk River in the Seward Peninsula were identified as belonging to group B (Table 2). There were no group B haplotypes found in other populations located between the northern and southeastern extremes

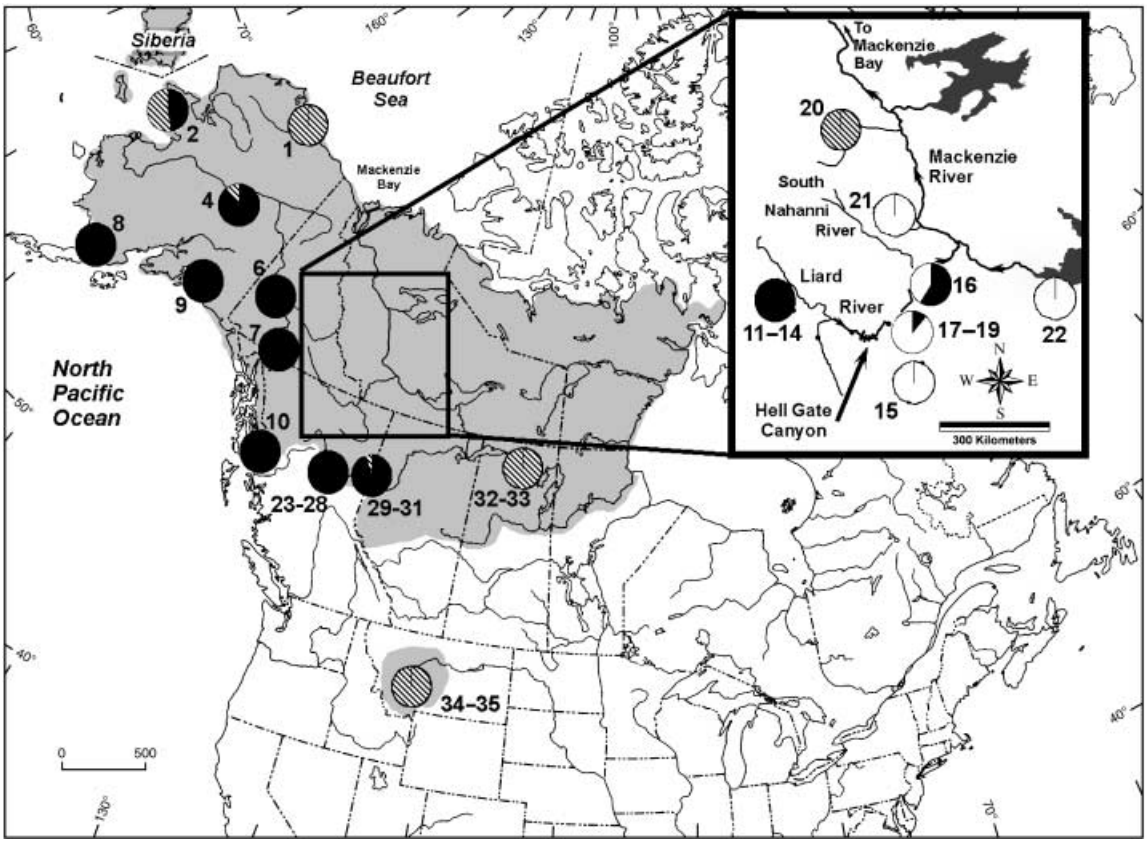

Fig. 3 Distribution and relative abundance of the three major mtDNA lineages of Arctic grayling at each sample locality (identified by the numbers beside the pies; see Table 1). Pie shading represents proportional abundance of each of the three mtDNA lineages found at each locality: black = lineage A; hatched = lineage B, and white = lineage C. Lineages refer to those shown in Fig. 2. Shaded area shows the current range of the species in North America and includes a small portion of Siberia. Inset shows location of sample areas within the Liard/Nahanni rivers.

of the geographical range of this group (e.g. Great Slave Lake or lower Liard River, Fig. 3). Although these haplotypes are widely distributed, the greatest haplotype diversity within this group was found on the north slope of Beringia (Beaufort Sea, Table 1, Fig. 3). Consequently, we have designated group B haplotypes as the 'North Beringia' lineage.

Group C consisted of only two haplotypes (11 and 12) and their geographical range did not extend beyond the Mackenzie River drainage (including the Liard, Nahanni rivers and Great Slave Lake, Table 2, Fig. 3). Given the close association with the Nahanni River (a tributary of the lower Liard River), we have designated group C haplotypes as the 'Nahanni' lineage. Nahanni lineage Arctic grayling therefore occupy a geographical location near the middle section of the Mackenzie River basin and intermediate to the disjunct distribution of North Beringia lineage Arctic grayling,

Populations in the lower Liard River had both group A and group C haplotypes and populations in the Seward Peninsula, lower Yukon River and lower Peace River had both group A and group B haplotypes (Table 2, Fig. 3). The greatest level of mtDNA lineage mixing occurred in the Seward Peninsula (Niukluk River) and lower Liard River (LaBiche, Beaver and Minnaker rivers) and where group A haplotypes were found in the same populations with group B and group C haplotypes, respectively. Although they were rare, group B haplotypes were also found in the lower Yukon River (Chena River) and the lower Peace River (Burnt River) where group A haplotypes typically predominated. There were, however, no localities where both group B and group C haplotypes occurred together in one population.

When we grouped our population samples of haplotypes into geographical assemblages that corresponded to major watersheds, three major population assemblages were resolved: a South Beringia group consisting of populations from the Peace River north and west to the Yukon River, upper Liard River and the Pacific coast, a North Beringia group consisting of populations from the Beaufort Sea and Seward Peninsula and populations east to Saskatchewan and south to Montana, and a Nahanni group consisting of populations from the lower Liard River and Mackenzie River (Fig. 4).

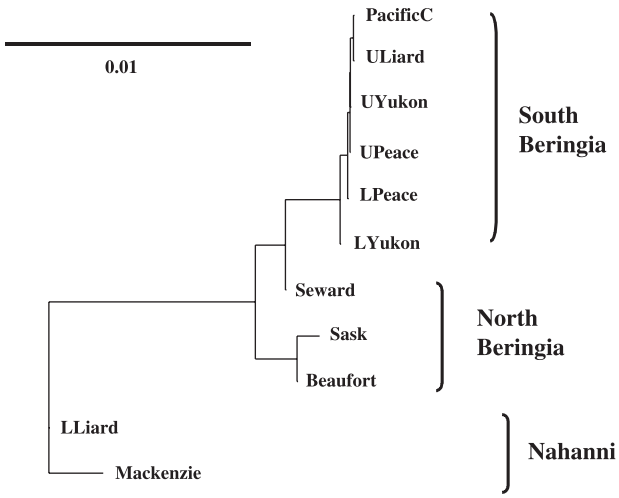

Fig. 4 UPGMA cladogram of mtDNA nucleotide divergence estimates among regions. Population groups were suggested from clustering of net mtDNA divergences among component populations denoted in Table 1. Beaufort = Beaufort Sea, Mackenzie = Mackenzie River, LLiard = lower Liard River, ULiard = upper Liard River, LPeace = lower Peace River, LYukon = lower Yukon River, UYukon = upper Yukon River, PacificC = Pacific Coast, Sask = Saskatchewan, Seward = Seward Peninsula.

Mitochondrial DNA sequence interrelationships

We obtained control region sequences (788 bp) from *T. arcticus* representing the 12 RFLP haplotypes resolved in our study and the two Siberian haplotypes resolved by Redenbach & Taylor (1999). We analysed 13 of these haplotypes (two of the original RFLP haplotypes had identical control region sequences) with 19 *Thymallus* haplotypes from eastern Asia and a small portion of North America (Koskinen *et al*. 2002a; Froufe *et al*. 2003). Sequence divergence among *T. arcticus* haplotypes ranged from 0.53 to 6.5% (matrix available from EBT) consistent with earlier work on a much smaller set of partial control region and ND1 sequences (Redenbach & Taylor 1999). Our inferred phylogeny of a partial set of samples of *T. arcticus* was consistent with the more extensive previous analysis of Koskinen *et al*. (2002a). When rooted with the *T. thymallus* haplotypes, two major clades were resolved (Fig. 5). The most basal clade was that consisting of *T. grubii*, whereas the other clade consisted of *T. brevirostris* from Central Asia and Asian *T. arcticus* from the Lake Baikal basin, the Lena River Basin, Yana River Basin (Sea of Okhotsk), and all North American *T. arcticus* (Fig. 5). Within the *T. arcticus/brevirostris* clade, however, all North American *T. arcticus* formed a strongly supported subclade (100%) that was sister to an Asian *T. arcticus/brevirostris* subclade and was the most derived relative to *Thymallus* from Asia (Fig. 5). Within the North American *T. arcticus*, the haplotypes consituting the 'Nahanni' RFLP lineage (C) were clearly distinct from all other haplotypes in North America (100% support). The other haplotypes segregated in the sequence-based phylogeny as in the RFLP phylogeny. Lineage A (haplotypes 1–6) clustered together with weak support (63%) and were the most derived haplotypes. Lineage B formed a loosely defined lineage intermediate between lineage A and lineage C (Fig. 5). The North American lineage B also clustered with the two Asian haplotypes sampled from the Yana River Basin, tributary to the Sea of Okhotsk (Redenbach & Taylor 1999).

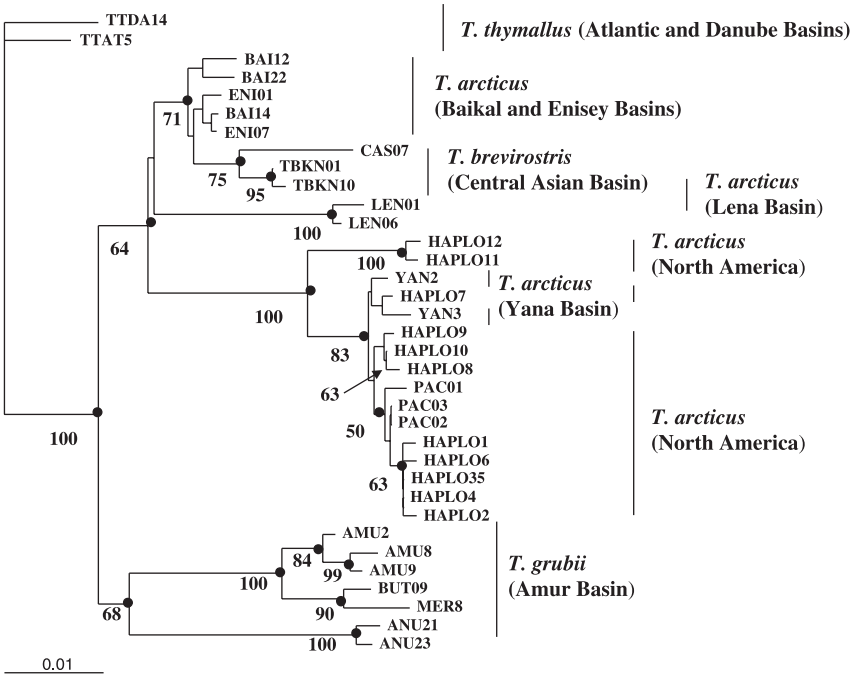

Fig. 5 Neighbour-joining tree of interrelationships among *Thymallus* mtDNA haplotypes inferred from pairwise Kimura 2-parameter distances from partial control region sequences (788 bp) from the 12 RFLP haplotypes resolved in the current study, two *T. arcticus* haplotypes from the Yana River Basin (Sea of Okhotsk) (Redenbach & Taylor 1999), two *T. thymallus* haplotypes and 19 haplotypes of *Thymallus* obtained from Koskinen *et al*. (2002a) and Froufe *et al*. (2003). Numbers at branch points represent the mean of bootstrap support levels from distance (1000 replicates), parsimony (1000 replicates) and maximum-likelihood (100 replicates) analyses. Only mean bootstrap values > 50% are shown. ENI = Enisey River Basin, Bai = Lake Baikal Basin, Len = Lena River Basin, Cas = Central Asian Basin, Amu = Amur River Basin, Anu = Annui River Basin, But = Buta River Basin, Mer = Merek River Basin, KN = Khökh Nuur, YAN = Yana River Basin, PAC = eastern Pacific Ocean Basin, TTAT and TTDA = *T. thymallus* haplotypes from the Atlantic Ocean and Danube River basins, respectively (see Koskinen *et al*. 2002a and Froufe *et al*. 2003).

Microsatellite variation within samples

There were several deviations from Hardy–Weinberg equilibrium both at *Tar 1* and *Ots 100* that took the form of heterozygote deficiences. In addition, there were several cases of apparently failed PCR amplifications for individuals that were scored at other loci. These observations suggested the presence of null alleles at these loci and, consequently, we estimated null allele frequencies using maximum likelihood (Dempster *et al*. 1977) as implemented in GENEPOP. Arctic grayling displayed low levels of intrapopulation variation; the number of alleles and expected heterozygosity averaged across loci and populations were 2.1 and 0.266, respectively (Table 1; allele frequency matrix available from EBT).

Genetic diversity (mean number of alleles and expected heterozygosity) declined with straight-line geographical distance from the lower Yukon River in Alaska (Fig. 6). Among all populations there was a significant negative correlation between geographical distance and the mean number of alleles per locus ($r = -0.424$, $P = 0.028$). The correlation coefficient was greater ($r = -0.701$, $P < 0.0005$), however, when South Beringia populations (containing mtDNA lineage A) located in the southern interior of Alaska, Yukon and British Columbia were considered alone. This was due to removing genetically diverse populations from the North Beringia and Nahanni lineages that contained mtDNA lineages B and C, respectively, that could obscure the signature of dispersal from the Yukon River valley which is dominated by mtDNA lineage A (Fig. 3). Although expected heterozygosity also declined with geographical distance from Beringia, the correlation coefficient was not significant among all sampled populations ($r = -0.123$, $P > 0.125$), but it was larger and significant among South Beringia populations ($r = -0.639$, $P = 0.0103$).

Microsatellite divergence among assemblages

We grouped our population samples into the same geographical assemblages as suggested by the mtDNA analyses (Fig. 4). Genetic divergence at microsatellite loci among these assemblages only partially corresponded with the evolutionary history inferred from the geographical locations of lineages of mtDNA (Fig. 7). The most divergent populations measured with microsatellites contained mtDNA group C and were located in the lower Liard River and Great Slave Lake (Fig. 7). Other than the general

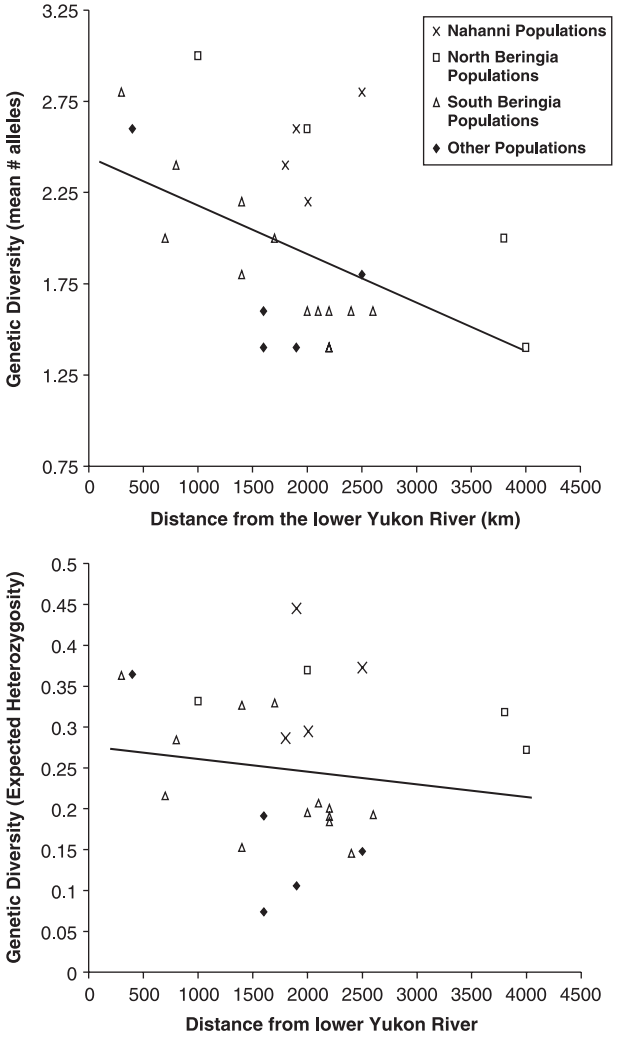

Fig. 6 Change in microsatellite genetic diversity within populations vs. their geographical distance from the lower Yukon River, at the confluence with the Koyukuk River (see Fig. 1). The three population groups (Nahanni, North Beringia and South Beringia) were defined by the mixture of mtDNA lineages that they contained. (a) Mean number of alleles per locus vs. geographical distance (km); (b) Expected heterozygosity vs. geographical distance (km).

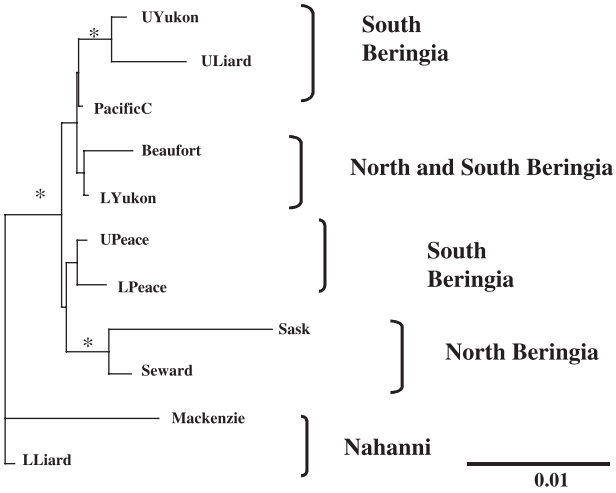

Fig. 7 Phylogram derived from maximum-likelihood analysis of microsatellite allele frequencies among populations. The mtDNA lineage that is known for each population and the geographical location of the populations within each group are also shown and component sample localities are shown in Table 1. Asterisks denote significant ($P < 0.05$) branch lengths. Beaufort = Beaufort Sea, Mackenzie = Mackenzie River, LLiard = lower Liard River, ULiard = upper Liard River, LPeace = lower Peace River, LYukon = lower Yukon River, UYukon = upper Yukon River, PacificC = Pacific Coast, Sask = Saskatchewan, Seward = Seward Peninsula.

distinctiveness of these Nahanni lineage populations, those from the mtDNA South and North Beringia lineages, although tending to cluster in lineage-specific pairs were interspersed with one another (Fig. 7). There was also one fixed allelic difference across the five loci that supported the distinction of the Nahanni lineage. The One8 179 base pairs allele was found only in Great Slave Lake, and in the LaBiche, Muskwa and Beaver rivers.

Hierarchical analysis of genetic diversity

Relative to other hierarchies tested, partitioning of genetic variance among putative North Beringia, South Beringia and Nahanni refugial groups accounted for the greatest proportion of the total mtDNA variation found among regions; 81% of the molecular variance occurred among regions, whereas 19% was found among populations within regions and within populations (Table 3). By contrast, when populations were pooled to reflect a partition across the continental divide, the between group component of mtDNA variation, although significant, was greatly reduced (34%). Similarly, when samples were pooled into Pacific and Arctic drainage groups, between group variance was effectively 0% (Table 3).

Analysis of genetic variance carried out for microsatellite allele frequency data produced similar results to that with mtDNA, although among region differences were substantially less (Table 4). A significant amount of variation was resolved among the three putative refugial groups, although the among region component was marginally higher when the Saskatchewan populations were treated as a separate, and fourth, group even though they contained the same mtDNA lineage (mtDNA B) as other North Beringia populations located in Alaska (24 and 32%, respectively, Table 4). By contrast, only 18 and 2.8%, respectively, of the variation was attributable to that among regions when the samples were arranged as east vs. west of the continental divide or between Arctic and Pacific drainage basins (Table 4).

Table 3 Analysis of mtDNA variance (AMOVA) among populations grouped into geographical hierarchies on either side of the Rocky Mountain divide (east vs. west) and putative origins from glacial refugia. NB = North Beringia, SB = South Beringia, Nah = Nahanni, SE = Saskatchewan and Montana populations

Geographic hierarchy	Variance component	*F* statistic	Percent variation	*P*
NB vs. SB vs. Nah	Among regions	0.81	80.8	0.000
	Among populations within regions	0.52	10. 0	0.000
	Within	0.91	9.2	0.000
NB vs. SB vs. Nah vs. SE	Among regions	0.78	77.8	0.000
	Among populations within regions	0.47	10.3	0.000
	Within populations	0.88	11.8	0.000
East vs. West	Among regions	0.34	34.4	0.001
	Among populations within regions	0.81	53.3	0.000
	Within populations	0.87	12.2	0.001
Pacific vs. Arctic	Among regions	–0.05	–5.3	0.47
	Among populations within regions	0.85	89.9	0.000
	Within populations	0.83	15.4	0.001

Table 4 Analysis of microsatellite variance (AMOVA) among populations grouped into geographical hierarchies on either side of the Rocky Mountain divide (east vs. west) and putative origins from glacial refugia. NB = North Beringia, SB = South Beringia, Nah = Nahanni, SE = Saskatchewan and Montana populations

Geographic hierarchy	Variance component	*F* statistic	Per cent variation	*P*
NB vs. SB vs. Nah	Among regions	0.24	24.3	0.000
	Among populations within regions	0.34	24.7	0.000
	Within populations	0.49	51.0	0.000
NB vs. SB vs. Nah vs. SE	Among regions	0.32	32.1	0.000
	Among populations within regions	0.28	19.3	0.000
	Within populations	0.51	48.6	0.000
East vs. West	Among regions	0.18	18.3	0.000
	Among populations within regions	0.35	28.7	0.000
	Within populations	0.47	53.0	0.000
Pacific vs. Arctic	Among regions	0.03	2.8	0.080
	Among populations within regions	0.44	41.3	0.000
	Within populations	0.42	55.9	0.000

Discussion

Multiple glacial refugia for Arctic grayling in North America

The precise number and locations of Arctic refugia and their influence on extant patterns of phylogeographical variation in a diversity of taxa are continuing uncertainties (e.g. Abbott & Brochmann 2003; Fedorov *et al*. 2003); however, our data have helped to clarify a number of issues. For instance, we have provided support for three mtDNA assemblages of North American Arctic grayling partitioned across geographical areas suggested to serve as refugia based on independent data: the North Beringia lineage in the Brooks Range Refuge, the South Beringia lineage in the Yukon River Valley Refuge and the Nahanni lineage in the Nahanni River Refuge. Furthermore, regional partitioning of genetic variation across independent data sets (mitochondrial and microsatellite) and relatively high genetic variation in samples from near putative refugia support the hypothesis that Arctic grayling survived in multiple refugia during the Pleistocene glaciations. The observation of higher correlations between distance from the Yukon River and measures of genetic diversity when populations were separated into the three putative refugial groups also supports the idea that current diversity in North American *Thymallus arcticus* is the result of survival in and dispersal from multiple glacial refugia in the Pleistocene.

The decline of genetic diversity with distance from the lower Yukon River suggests that postglacial dispersal of

Arctic grayling from South Beringia into south and eastern regions of North America occurred as the glaciers retreated during the Holocene, and is consistent with theoretical expectations (Nei *et al*. 1975; Hewitt 1996; Ibrahim *et al*. 1996) and empirical data for a variety of organisms (e.g. Sage & Wolf 1986; Reiss *et al*. 1999; Costello *et al*. 2003; Taylor *et al*. 2003; Weider & Hobæk 2003). Our data therefore indicate a pronounced influence of Beringian populations on establishment of, and genetic variation in, populations located in previously glaciated areas located to the east and south. This result, although observed in other taxa including fish (cf. McPhail 1963; Wilson & Hebert 1998; Reiss *et al*. 1999) apparently contrasts with the situation in other species where dispersal of Beringian lineages into North America has been more limited, perhaps by the presence of closely related taxa in adjacent areas in North America (e.g. Bernatchez & Dodson 1994; Fedorov *et al*. 2003).

McCart & Pepper (1971) described so-called 'large'- and 'small'-scale phenotypes of *T. arcticus* distributed north and south of the Brooks Range in Alaska. Although we did not study mtDNA variation from the same populations, our documentation of mtDNA haplotype groups A and B, that also roughly coincide with a division north and south of the Brooks Range is consistent with the idea that Arctic grayling survived glaciation in two isolated regions in Beringia (i.e. North and South Beringia). There is geological evidence that north and south slope Brooks Range freshwater drainages in Beringia were isolated from each other for at least 130 000 years, until at least the end of the Sangamon interglacial (Lindsey & McPhail 1986).

Unique sequence haplotypes in the upper Missouri River reported by Redenbach & Taylor (1999), and the presence of grayling fossils dated to between 22 000 and 50 000 years BP in southern Alberta (Burns 1991), suggest that Arctic grayling also survived at least the most recent glaciation in a Great Plains or Missouri glacial refuge. A similar scenario has been hypothesized for lake trout, *Salvelinus namaycush* (Wilson & Hebert 1998) and burbot (van Houdt *et al*. 2003). Our data, however, do not constitute strong support, in the form of divergent clades that might signal isolation during multiple glacial advances, for persistence of *T. arcticus* in an upper Missouri refuge. By contrast, our expanded RFLP analysis resolved shared haplotypes between Montana and the lower Peace River (haplotype 7) and the sequence analysis indicated strong similarities between the Missouri area haplotype and haplotypes in Asia from the Yana River basin (Sea of Okhotsk). Further, the haplotype representative of Montana area grayling clustered within haplotype group B and was similar (average of 0.6% sequence divergence) to all other North American haplotypes with the exception of high divergence from Nahanni haplotypes (1.2–3.8% divergence). These observations, coupled with high genetic diversity within populations carrying group B mtDNA along the Arctic coast and apparently low mtDNA diversity in Montana (Redenbach & Taylor 1999) suggest that these extant Montana were founded by North Beringia grayling. The Montana grayling appear phenotypically intermediate, however, between the large- and small-scale forms of *T. arcticus* (McCart & Pepper 1971) and the greatest proportion of genetic variance at microsatellite loci occurred among groups when Saskatchewan populations (that share haplotypes at high frequency with Montana grayling) were treated as a separate group. These subtle differences suggest that although these southern and eastern populations of Arctic grayling probably owe their ultimate origin to dispersal from North Beringia during early Pleistocene glaciations, they may have originated most recently by survival in and dispersal from southern Great Plains refugia during the last glaciation (cf. Redenbach & Taylor 1999). Recent data for taxa as diverse as cisco, *Coregonus artedii* (Turgeon & Bernatchez 2001) reindeer, *Rangifer tarandus* (Flagstaff & Roed 2003), and black spruce, *Picea mariana* (Gamache *et al*. 2003) also suggest survival in and extensive northward dispersal of cold-adapted taxa from refugia south of the North American ice sheets.

The existing geological (Ford 1976; Bodaly & Lindsey 1977) and genealogical evidence (Lindsey & McPhail 1986; Bernatchez & Dodson 1991; Foote *et al*. 1992; Wilson & Hebert 1998) it's equivocal in terms of is support for a North American refuge in the Nahanni River area. Two distinct regions located between the Cordilleran and Laurentian ice sheets (just west of the Mackenzie River drainage) where glacial lakes persisted during Pleistocene glaciations may have served as refugia: a southern most refuge consisting of a large northward draining glacial lake that filled the lowlands of the South Nahanni River and North Nahanni River has been called the Nahanni Refuge (Ford 1976), and a more northern refuge encompassing glacial lakes that flooded portions of the Peel and Porcupine drainages that flowed west into the Yukon River watershed (Bodaly & Lindsey 1977; Lindsey & McPhail 1986). In our study, the apparent localization of the highly divergent haplotype group C around the Nahanni River Valley supports the existence of the more southern, Nahanni Refuge. For instance, all of the Arctic grayling populations that contain mtDNA lineage C are locally distributed around the putative Nahanni Refuge and the relatively high genetic diversity among the populations in this area suggests few founder events associated with postglacial dispersal from other areas. A sample located further downstream in the Mackenzie River, closer to the Peel River contained only mtDNA lineage B haplotypes (Keel River, Table 2). In addition, Arctic grayling with the large-scale phenotype that was associated with mtDNA group B (North Beringia) are found in the Peel River north of the Nahanni River (McCart & Pepper 1971; Bodaly & Lindsey 1977). Both these observations suggest a North Beringia lineage origin for grayling found between the Nahanni River Valley and the Arctic Ocean rather than the

existence of a second, more northern refuge in the Mackenzie River area.

Nahanni lineages have also been reported in lake whitefish (*Coregonus clupeaformis*, Foote *et al*. 1992) and lake trout (Wilson & Hebert 1998), but they have wider geographical ranges than Nahanni Arctic grayling that may have been promoted by postglacial dispersal through large proglacial lakes Peace and Agassiz (Crossman & McAllister 1986; Lindsey & McPhail 1986; Dyke & Prest 1987; Pielou 1991). Arctic grayling, however, are most often found in streams, especially in northwestern North America near the limits to their geographical range (Armstrong 1986; Northcote 1995). This greater dependence on fluvial habitat may therefore have limited their potential for postglacial dispersal through glacial lakes.

Origins of North American Arctic grayling

That Eurasia contains the greatest taxonomic diversity within the genus *Thymallus* suggests Eurasian ancestry for North American *T. arcticus* and our sequence data are consistent with this scenario. North American *T. arcticus* were the most derived lineage in the phylogenetic tree and appear to be related to the Asian *T. arcticus/brevirostris* lineage. More specifically, our inclusion of novel sequence data from a population from the northern Sea of Okhotsk (Yana River basin) and the tight clustering of these haplotypes with North American grayling clearly suggests that, with the possible exception of Nahanni lineage grayling, the latter were founded by Asian *T. arcticus* from a Pacific basin ancestral lineage rather than by a lineage from the Central Asian or Arctic basins. Galbreath & Cook (2004) recently also showed close genetic affinities, implying recent divergence, between Beringian populations of tundra volves (*Microtus oeconomus*) from eastern Siberia, Alaska and northwestern Canada. Stronger differentiation of central Beringian *T. arcticus* in eastern Siberia and North America from lineages farther west and south in Asia may stem from greater isolation promoted by extensive glacier formation in the Kolyma Highlands near the western border of Beringia ~85 000–55 000 years ago (Galbreath & Cook 2004). Our data therefore highlight the importance of obtaining sequence data from *T. arcticus* from the extreme eastern portions of Siberia or Kamchatka boarding the Gulf of Anadyr and Bering Sea as they are the most likely source of Eurasian colonists across Beringia. Affinities of *T. arcticus* from these areas with those from more western areas in Eurasia is a significant gap in our knowledge of *Thymallus* evolutionary history.

Although there were multiple opportunities for faunal exchange throughout the Pleistocene (Lindsey & McPhail 1986), Arctic grayling probably first crossed the Bering land bridge from Siberia earlier in the mid-late Pliocene (3–5 Mya, Makoedov 1987). For instance, assuming a divergence rate of ~1%/Myr (based on calibration for the entire control region sequences by Koskinen *et al*. 2002a and references therein) suggests that Nahanni lineage grayling diverged from other North American grayling in the late Pliocene to early Pleistocene because they were 2–3% divergent from all other North American *T. arcticus*. Further, with the exception of the Yana River basin *T. arcticus*, all other Eurasian *T. arcticus* were 3–5% divergent in sequence from North American grayling suggesting a pre-Pleistocene divergence and initial colonization of North America. Certainly, multiple opportunities for freshwater dispersal across Beringia from west to east were available during the Pliocene before the Bering Seaway first opened 3–5 Mya (Marincovich & Gladenkov 1999). By contrast, the closer relationship between Yana River haplotypes (Sea of Okhotsk) and most North American *T. arcticus* (0.5–0.9% divergence) indicates more recent divergence and dispersal events from Asia to North America also occurred during the mid-late Pleistocene. Interestingly, although North American burbot are also thought to have had an origin in Eurasia, dispersal of ancestral lineages into the North American portion of Beringia is thought to have occurred only during more recent times (760 000–1 000 000 years ago) relative to *T. arcticus* (van Houdt *et al*. 2003).

Comparative divergence among Holarctic grayling lineages

Notwithstanding the distinctiveness of Nahanni grayling, levels of divergence among the North American Arctic grayling haplotype groups were substantially lower than divergences found among three European grayling (*T. thymallus*) haplotype groups, which ranged between 3.1 and 4.6% for the same mtDNA region (D-loop/Cyt *b*, ND 5/6; Koskinen *et al*. 2000). There were also substantially more haplotypes found across the European *T. thymallus* populations surveyed (27 haplotypes from 540 individuals, among 27 populations; Koskinen *et al*. 2000) relative to North American *T. arcticus* populations in this study (12 haplotypes from 320 individuals, among 32 populations). Similarly, Koskinen *et al*. (2002a) and Froufe *et al*. (2003) reported pairwise control region sequence divergences of 2.3–2.8 and 4.6% for *T. arcticus* from the Lena River and Lake Baikal basins and *T. grubii* from the Amur Basin, respectively. Comparisons among *Thymallus* for the one locus in common (*BFRO004*) in our study and that of Koskinen *et al*. (2002b) for *T. thymallus* showed that the total number of alleles observed across all populations in both species was similar (5–6). Nevertheless, comparing variation across the five non-*T. thymallus*-derived loci (of 17 total) from Koskinen *et al*. (2002b) with the five from our study as a conservative approach, the total number of alleles across loci observed ranged from 7 to 28 and averaged 14 for *T. thymallus* (Koskinen *et al*. 2002b) vs. 3–10 and an average of

6.4 in our, geographically more widespread, study. Altogether, these observations suggest that the total variation in mtDNA and at microsatellite loci is relatively lower in North American *T. arcticus* than in Eurasian species of *Thymallus*.

Greater diversity and divergence among phylogeographical groups of Eurasian *Thymallus* imply that the various taxa may have suffered fewer historical extinctions and persisted in relatively more stable, larger and/or more glacial refugia than North American *T. arcticus*. The extent of ice cover in Eurasia was considerably less than in North America throughout the Pleistocene (Anderson & Borns 1997) and large proglacial lakes in Russia (Mangerud *et al*. 2001) presumably imposed fewer founder events and bottlenecks on Eurasian *Thymallus* during postglacial range expansion. A progressive decline of genetic diversity with distance from putative refugia is a theoretically and empirically well-documented consequence of range contraction during glacial advances followed by postglacial range expansion (Sage & Wolff 1986; Hewitt 1996; Bernatchez & Wilson 1998; Costello *et al*. 2003). Further evidence for more pronounced effects of Pleistocene glaciations on North American *T. arcticus* therefore comes from the genetic signature for large-scale range expansion, which was prominent for *T. arcticus* in North America (Fig. 6), whereas there was no such trend for *T. thymallus* in Europe (Koskinen *et al*. 2000; Koskinen *et al*. 2002b). Similarly, higher levels of diversity in European populations of the lake whitefish species complex (*Coregonus clupeaformis* and *C. lavaretus*) were also attributed to a smaller extent and lesser impact of Eurasian ice sheets relative to events in North America (Bernatchez *et al*. 1989; Bernatchez & Dodson 1994; reviewed in Bernatchez *et al*. 1999). The lake whitefish species complex, however, had shallow divergence among clades (ranged from 0.4 to 1.2%; Bernatchez *et al*. 1999) relative to clades within North American *T. arcticus*. Furthermore, the lake whitefish mtDNA clades were distributed from northwestern Europe, across Eurasia to northeastern North America (Bernatchez & Dodson 1994), whereas all five grayling species are located in distinct regions in North America and Siberia (*T. arcticus*), Mongolia (*T. nigrescens*, *T. grubii*, *T. brevirostris*) and Europe (*T. thymallus*) (Schoffmann 2000). Locally distributed clades and greater divergence within and among *T. thymallus* and *T. arcticus* suggest that Pleistocene range expansion for *Thymallus* species was more limited relative to *Coregonus*.

Alternatively, for more eastern Asian populations of *Thymallus*, glacial and hydrological perturbations were perhaps more extreme (e.g. Grosswald 1998) and resulted in well-documented genetic signatures of pronounced bottlenecks and subsequent expansion during glacial times (e.g. Koskinen *et al*. 2002a). Notwithstanding the existence of such dramatic events in Eurasia, the remaining higher levels of diversity in eastern Eurasian *Thymallus* relative to *T. arcticus* in North America may stem from the longer history of *T. arcticus* in Eurasia and the relatively recent origin of North American *T. arcticus*.

Conclusions

Our data are a contribution to understanding the complex nature of glacial refugia in Arctic regions of North America as well as to the evolutionary history of a single species. These data strongly suggest that Arctic regions of northwestern North America contain more than a single refugium and we have provided perhaps the strongest geneological evidence for the existence of a refuge in the Nahanni River area. It is, however, possible that additional Arctic refugia exist in areas farther east as suggested for lemmings and *Daphnia* (Fedorov & Stenseth 2002; Weider & Hobæk 2003). In addition, our analysis has demonstrated that there is a greater degree of phylogenetic diversity in North American *T. arcticus* than was previously appreciated (e.g. Redenbach & Taylor 1999). Furthermore, the combined use of mtDNA and microsatellite DNA data has permitted us to resolve genetic signatures both of long-term isolation in distinct refugia and postglacial expansion into previously glaciated regions. At a minimum, the divergences among lineages should be recognized in broad-based programs to conserve the evolutionary legacy of the species (e.g. in Committee for the Status of Endangered Wildlife in Canada [COSEWIC] species status assessments). Specifically, the apparently narrow distribution for the highly distinctive Nahanni grayling suggests that it might be at inherent higher conservation risk. Our data also contribute to an increased understanding of Beringia and adjacent high latitude environments as important sources of Holarctic biodiversity. Lastly, Siberian and northwestern North American biota have clearly interacted over time within Beringia (e.g. Pielou 1991; Elias *et al*. 2000; Galbreath & Cook 2004). Owing to their rich phylogeographical structure (this study; Redenbach & Taylor 1999; Koskinen *et al*. 2002a) and a distribution throughout Beringia, Arctic grayling present an excellent model system to resolve the details of this interaction for freshwater organisms.

Acknowledgements

The Peace-Williston Fish and Wildlife Compensation Program provided funding for this project. We thank Brian Blackman for support and helpful discussion throughout the study. Many samples were kindly provided by individuals associated with the following groups: Peace Williston Fish and Wildlife Compensation Program, BC Ministry of Environment Lands and Parks, Alaska Department of Fish and Game, Camp Grayling Fly Fishing Lodge, Wollaston Lake Lodge, Fred Gundersen collected samples from North Nahanni and Keel rivers, George Low (of Canadian Department of Fisheries and Oceans collected samples from Great Slave Lake. This paper benefited from critical reviews by L. Bernatchez and anonymous reviewers.

References

Abbott RJ, Brochmann C (2003) History and evolution of the arctic flora: in the footsteps of Eric Hultén. *Molecular Ecology*, **12**, 299–313.

Anderson BG, Borns HW Jr (1997) *The Ice Age World. An Introduction to Quaternary History and Research with an Emphasis on North America and Europe During the Last 2.5 Million Years*. Scandinavian University Press, Oslo.

Armstrong RH (1986) *A Review of Arctic Grayling Studies in Alaska, 1952–82*. Alaska Cooperative Fisheries Research Unit, University of Alaska.

Avise JC (2000) *Phylogeography*. Harvard University Press, Cambridge, MA.

Bernatchez L, Chouinard A, Lu G (1999) Integrating molecular genetics and ecology in studies of adaptive radiation: whitefish, *Coregonus* sp., as a case study. *Biological Journal of the Linnean Society of*, **68**, 173–194.

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

Bernatchez L, Dodson JJ (1994) Phylogenetic relationships among Palearctic and Nearctic whitefish (*Coregonus* sp.) populations as revealed by mitochondrial DNA variation. *Canadian Journal of Fisheries and Aquatic Sciences*, **51** (Suppl. 1), 240–251.

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

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

Bodaly RA, Lindsey CC (1977) Pleistocene watershed exchanges and the fish fauna of the Peel River Basin, Yukon Territory. *Journal of the Fisheries Research Board of Canada*, **34**, 388–395.

Brunhoff C, Galbreath KE, Fedorov VB, Cook JA, Jaarola M (2003) Holarctic phylogeography of the root vole (*Microtus oeconomus*): implications for late quaternary biogeography of high latitudes. *Molecular Ecology*, **12**, 957–968.

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

Costello AB, Down TE, Pollard SM, Pacas CJ, Taylor EB (2003) The influence of history and contemporary stream hydrology on the evolution of genetic diversity within species: an examination of microsatellite DNA variation in bull trout, *Salvelinus confluentus* (Pisces: Salmonidae). *Evolution*, **57**, 328–344.

Crossman EJ, McAllister (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. Wiley, New York.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorothm. *Journal of the Royal Statistical Society, Series B*, **39**, 1–38.

Dyke ASVK, Prest VK (1987) Late Wisconsisnan and Holocene history of the Laurentide ice sheet. *Geological Physics of the Quaternary*, **41**, 237–263.

Elias SA, Berman A, Afimov A (2000) Late Pleistocene beetle faunas of Beringia: where east meets west. *Journal of Biogeography*, **27**, 1349–1363.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from the metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Fedorov VB, Goropashnaya AV, Jaarola M, Cook JA (2003) Phylogeography of lemmings (*Lemmus*): no evidence for postglacial colonization of Arctic from Beringian refugium. *Molecular Ecology*, **12**, 725–731.

Fedorov VB, Stenseth NC (2002) Multiple glacial refugia in the North American arctic: inference from phylogeography of the collared lemming (*Dicrostonyx groenlandicus*) *Proceedings of the Royal Society of London B*, **268**, 809–814.

Felsenstein J (1995) PHYLIP *(Phylogeny Inference Package), Version 3.5c*. Distributed by the author. Department of Genetics, University of Washington, Seattle.

Flagstaff Ø, Røed KH (2003) Refugial origins of reindeer (*Rangifer tarandus* L). inferred from mitochondrial DNA sequences. *Evolution*, **57**, 658–670.

Foote CJ, Clayton JW, Lindsey CC, Bodaly RA (1992) Evolution of lake whitefish (*Coregonus clupeaformis*) in North America during the Pleistocene. Evidence for a nahanni glacial refuge race in the northern Cordilleran region. *Canadian Journal of Fisheries and Aquqtic Sciences*, **49**, 760–768.

Ford DC (1976) Evidence of multiple glaciations in South Nahanni National Park, Mackenzie Mountains, Northwest Territories. *Canadian Journal of Earth Sciences*, **13**, 1433–1445.

Froufe E, Knizhin I, Koskinen MT, Primmer CR, Weiss S (2003) Identification of reproductively isolated lineages of Amur grayling (*Thymallus grubii* Dybowski 1869): concordance between phenotypic and genetic variation. *Molecular Ecology*, **12**, 2345–2355.

Galbreath KE, Cook JA (2004) Genetic consequences of Pleistocene glaciations for the tundra vole (*Microtus oeconomus*) in Beringia. *Molecular Ecology*, **13**, 135–148.

Gamache I, Jaramillo-Correa JP, Payette S, Bousquet J (2003) Diverging patterns of mitochondrial and nuclear DNA diversity in subarctic black spruce: imprint of a founder effect associated with postglacial colonization. *Molecular Ecology*, **12**, 891–902.

Grosswald MG (1998) New approaches to the ice age paleohydrology of northern Eurasia. In *Paleohydrology and Environmental Change* (eds Benito G, Baker VR, Gregory KJ), pp. 199–214. Wiley, Chichester.

Hall TA (1999) BIOEDIT Version 5.0.9: a user-friendly biological sequence alignment editor and analysis for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Hasegawa M, Kishino H, Yano T (1985) Dating the human-age split by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

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

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Houdt Van JK, Hellemens B, Volckaert FAM (2003) Phylogenetic relationships among Palearctic and Nearctic burbot (*Lota lota*): Pleistocene extinctions and recolonization. *Journal of Biogeography*, **29**, 599–612.

Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, **77**, 282–291.

Koskinen MT, Knizhin I, Primmer CR, Schoetterer C, Weiss S (2002a) Mitochondrial and nuclear DNA phylogeography of *Thymallus* spp. (grayling) provides evidence of ice-age mediated environmental perturbations in the world's oldest body of fresh water, Lake Baikal. *Molecular Ecology*, **11**, 2599–2611.

Koskinen MT, Nilsson J, Veselov AJ, Potutkin AJ, Ranta E, Primmer CR (2002b) Microsatellite data resolve phylogeographic

patterns in European grayling, *Thymallus thymallus* Salmonidae. *Heredity*, **88**, 391–401.

Koskinen MT, Ranta E, Piironen J *et al*. (2000) Genetic lineages and postglacial colonization of grayling (*Thymallus thymallus*, Salmonidae) in Europe, as revealed by mitochondrial DNA analyses. *Molecular Ecology*, **9**, 1609–1624.

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. Wiley, New York.

Lindsey CC, Patalas K, Bodaly RA, Archibald CP (1981) *Glaciation and the Physical, Chemical and Biological Limnology of Yukon Lakes*. Canadian Technical Report of Fisheries and Aquatic Sciences, no. 966.

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

Makoedov AN (1987) Interpopulational differences and distribution of grayling, *Thymallus*: study of colour variation of dorsal fin. *Journal of Ichthyology*, **28**, 24–30.

Mangerud J, Astakhov V, Jakobsson M, Svendsen JI (2001) Huge ice-age lakes in Russia. *Journal of Quaternary Science*, **16**, 773–777.

Marincovich L, Gladenkov AY (1999) Evidence for an early opening of the Bering Strait. *Nature*, **397**, 149–151.

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 (1963) Geographic variation in North American ninespine sticklebacks, *Pungitius pungitius*. *Journal of the Fisheries Research Board of Canada*, **20**, 27–44.

Merilä J, Bjorklund M, Baker AJ (1996) Genetic structure and gradual northward decline of genetic variability in the greenfinch (*Carduelis chloris*). *Evolution*, **50**, 2548–2557.

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**, 145–167.

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

Nelson RJ, Beacham TD (1999) Isolation and cross species amplification of microsatellite DNA useful for study of Pacific salmon. *Animal Genetics*, **30**, 228–229.

Northcote TG (1995) Comparative biology and manangement of Arctic and European grayling (Salmonidae, *Thymallus*). *Reviews in Fish Biology and Fisheries*, **5**, 141–194.

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

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

Redenbach Z, Taylor EB (1999) Zoogeographical implications of variation in mitochondrial DNA of Arctic grayling (*Thymallus arcticus*). *Molecular Ecology*, **8**, 23–35.

Reed RJ (1973) Analysis of some meristic and morphometric data from the Arctic grayling, *Thymallus arcticus*, in Alaska. *Copeia*, **4**, 819–822.

Reiss RA, Ashworth AC, Schwert DP (1999) Molecular genetic evidence for post-Pleistocene divergence of populations of the arctic-alpine ground beetle *Amara alpina* (Paykul) (Coloeptera: Carabidae). *Journal of Biogeography*, **26**, 785–794.

Sage RD, Wolff JO (1986) Pleistocene glaciations, fluctuating ranges, and low genetic variability in a large mammal (*Ovis dalli*). *Evolution*, **40**, 1092–1095.

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406–425.

Schneider S, Kueffer J-M, Roessli D, Excoffier L (1997) ARLEQUIN, *Version 1: An Exploratory Population Genetics Software Environment*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Schoffmann J (2000) The grayling species (Thymallinae) of three different catchment areas of Mongolia. The *Grayling Society Journal*, **Spring 2000**.

Scott WB, Crossman EJ (1973) Freshwater fishes of Canada. *Fisheries Research Board of Canada Bulletin*, **184**.

Scribner KT, Gust J, Fields RL (1996) Isolation and characterization of novel salmon microsatellite loci: cross-species amplification and population genetic applications. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**, 833–841.

Sher A (1999) Traffic lights at the Beringian crossroads. *Nature*, **397**, 103–104.

Skopets MB (1991) Biological features of a subspecies of the Arctic grayling in northeast Asia. 2. The Alaskan grayling, *Thymallus arcticus signifer*. *Journal of Ichthyology*, **31**, 87–102.

Snoj A, Sušnik S, Pohar J, Dovc P (1999) The first microsatellite marker (BFRO 004) for grayling, informative for its Adriatic population. *Animal Genetics*, **30**, 74–75.

Swofford DL (2002) PAUP*. *Phylogenetic Analysis Using Parsimony (*and other methods). Version 4*. Sinauer Associates, Sunderland MA, USA.

Taggart JB, Hynes RA, Prodohl PA, Ferguson A (1992) A simplified protocol for routine total. DNA isolation from salmonids. *Journal of Fish Biology*, **40**, 963–965.

Taylor EB, Stamford MD, Baxter JS (2003) Population subdivision in westslope cutthroat trout (*Oncorhynchus clarki lewisi*) at the northern periphery of its range: evolutionary inferences and conservation implications. *Molecular Ecology*, **12**, 2609–2622.

Turgeon J, Bernatchez L (2001) Clinal variation at microsatellite loci reveals historical secondary intergradation between glacial races of *Coregonus artedi* (Teleostei: Coregoninae). *Evolution*, **55**, 2274–2286.

Weider LJ, Hobæk A (2003) Glacial refugia, haplotype distributions, and clonal richness of the *Daphnia pulex* complex in arctic Canada. *Molecular Ecology*, **12**, 463–473.

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

This research formed the basis for a MSc thesis completed by M Stamford at University of British Columbia stimulated by an interest in understanding the basic biology of taxa affected by hydroelectric development in the Peace River area of British Columbia. M Stamford has interests in the ecology, evolution, and conservation of aquatic taxa. This research illustrates E Taylor's interests in the intersection of evolutionary history and conservation of aquatic faunas.
