

Phenotypic and genetic differentiation of two major phylogeographical lineages of arctic grayling *Thymallus arcticus* in the Lena River, and surrounding Arctic drainages

STEVEN WEISS^{1*}, IGOR KNIZHIN², ALEXANDER KIRILLOV³ and ELSA FROUFE^{4,5}

¹Karl-Franzens University Graz, Institute of Zoology, Universitätsplatz 2, A-8010 Graz, Austria

²Irkutsk State University, Biological Faculty, ul K. Marksa 1, Irkutsk 664003, Russia

³Institute of Applied Ecology of the North, AS RS (Ya), Yakutia, Russia

⁴CIBIO/UP, Campus Agrário de Vairão, 4480-661, Vairão, Portugal

⁵Faculdade de Ciências, Universidade do Porto, Praça Gomes Teixeira, 4000-002 Porto, Portugal

Received 25 November 2004; accepted for publication 3 October 2005

Two phylogeographical lineages of arctic grayling, *Thymallus arcticus*, in Siberia are extensively characterized based on both molecular genetic (mtDNA control region sequences) and phenotypic (12 meristic characters) data. One lineage, occurring in the delta region of the Lena River as well as all other Arctic draining rivers sampled, corresponds to the subspecific taxon *Thymallus arcticus pallasii*, whose type locality is the Kolyma River. This taxon is proposed to be a postglacial colonizer of the Lena delta. The second lineage occurs throughout the rest of the Lena basin and is proposed to have survived in a glacial refugium in the middle reaches of the Lena. These lineages form reciprocally monophyletic groups based on mtDNA sequences (net divergence 3.2%), a relationship that is concordant with phenotypic data, and thus reflects distinct taxa. The upper Lena taxon is given the preliminary name of *Thymallus arcticus lenensis*. Phylogenetic analysis, together with previously published data from North America, reveals that mtDNA sequences from North American populations group within the diverse clade corresponding to *T. a. pallasii* in Siberia. Despite the relatively close genetic relationship of most North American haplotypes with those in northern Siberia, inferences of fragmentation between the continents are supported, but bidirectional movements between the two continents are seen as likely. Despite inclusion in the clade representing *T. a. pallasii* in Siberia, the source of the relatively divergent Nahanni refuge haplotypes in North America is not resolved. Otherwise, inferences of postglacial expansion across several thousand kilometres are well supported within North America, but only smaller-scale colonization events among drainages are supported in Siberia. © 2006 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2006, 88, 511–525.

ADDITIONAL KEYWORDS: *lenensis* – morphology – mtDNA – *pallasii*.

INTRODUCTION

Like most salmonoid fishes, grayling *Thymallus* sp. are extremely popular for sport and commercial fishing and are an important ecological component of the temperate cold-water systems where they reside. Recent research has emphasized the complex within-basin diversity that grayling display, the extreme population substructure even within single rivers, and the broad-scale signatures that ice-age perturbations

have left on the genetic architecture of various lineages both within and between species (Koskinen *et al.*, 2000; Koskinen, Piironen & Primmer, 2001, 2002; Weiss *et al.*, 2002; Froufe *et al.*, 2003a, b; Knizhin *et al.*, 2004; Stamford & Taylor, 2004). Despite this attention and popularity, grayling systematics is far from resolved, especially for the paraphyletic Arctic grayling, *Thymallus arcticus* (Koskinen *et al.* 2002, Froufe, Knizhin & Weiss, 2005). Arctic grayling, is a polytypic species with an extremely broad distribution. It is currently found in the Ural mountain region where its range overlaps with European grayling *Thy-*

*Corresponding author. E-mail: steven.weiss@uni-graz.at

mallus thymallus, forming a hybrid zone (Zinoviev, 1980; Shubin & Zakharov, 1984); and exhibits a more or less continuous distribution across all of Siberia and into mostly Canadian North America; further range descriptions are provided in Scott & Crossman (1998) and Redenbach & Taylor (1999).

In the northern Arctic region, including the entire Lena basin, a subspecies (*Thymallus arcticus pallasii*) has been generally recognized, with its type locality found east of the Lena, in the Kolyma River, which drains into the east-Siberian sea (Novikov, 1966; Kirillov, 1972; Chereshnev *et al.*, 2001, 2002). The first molecular DNA sequences of Lena basin grayling (from the upper reaches of the basin) revealed a single highly distinct monophyletic lineage with respect to all other grayling lineages analysed (Koskinen *et al.*, 2002). A recent phylogenetic analysis across the entire genus (Froufe *et al.*, 2005) included several individuals from the mid or lower reaches of the Lena river, revealing haplotypes 3% divergent from those reported by Koskinen *et al.* (2002). In light of historical accounts of distinct lineages in the Lena basin characterized by alternate phenotypic and biological characters (Borisov, 1928; Tugarina & Knizhin, 1986; Makoedov, 1999), this result has raised the hypothesis of whether or not there is systematic-level correlation between the historically noted phenotypic variation and molecular variation in and around the Lena basin.

Recent investigations of Arctic grayling in North America demonstrate significant diversity, multiple refugia, and a hypothesized scenario of multiple colonization of the continent, potentially stemming from distinct Asian lineages (Stamford & Taylor, 2004). Because the Eurasian and North American Arctic are closely related biogeographically, the question arises as to the phylogeographical relation of grayling found in the Lena and surrounding basins to lineages found in North America.

The present study aimed to characterize lineages of Arctic grayling within the Lena basin (the seventh largest in the world) and several adjacent Arctic drainages using complimentary sets of phenotypic (meristic measurements) and genetic data. The use of phenotypic data strengthens inferences based on a single genetic (mtDNA) locus. The molecular results are phylogeographically integrated with existing data from North America, and an attempt is made to identify a colonization source in Siberia for extant North American lineages.

MATERIAL AND METHODS

SAMPLING

Grayling ($N = 400$) were collected by angling or nets in 1998–2003 from 23 locations in the Lena River basin

and adjacent Arctic systems. Some populations from the upper reaches of Lena have been genetically analysed previously (Koskinen *et al.*, 2002). A single population from the Big Hole River, Montana, USA, was added for comparative purposes. A small fin clip was preserved in 96% ethanol, standard length (mm) measured, and the whole specimen preserved in a 4% formalin solution for transport back to Irkutsk. Several individuals from the Baikal and Central Asian basins as well as North America (Koskinen *et al.*, 2002), along with several European grayling (Weiss *et al.*, 2002), were included in the phylogenetic analysis. The name and location of each sampling site and the source of previously published data are listed in Table 1 and shown in Figure 1.

MERISTIC ANALYSIS

Phenotypic characterization was based on 12 meristic characters: (1) pores along the lateral line; (2) unbranched dorsal rays; (3) branched dorsal rays; (4) total number of rays; (5) branched pectoral rays; (6) branched pelvic rays; (7) unbranched anal rays; (8) branched anal rays; (9) gill rakers on the first left branchial arch; (10) branchiostegal rays; (11) vertebrae (with central and haemal spines); and (12) pyloric caecae (Table 2). Counts were made on either fresh or fixed material following methodology described by Pravdin (1966) with the following modifications: vertebrae count does not include the urostyle; gill rakers and lateral line pore counts are all-inclusive aided by an 8–16 binocular; and the last branched ray in the dorsal and anal fins, which looks like two separate rays, is counted as one. A principal component analysis (PCA) was carried out on the correlation matrix (i.e. standardized data) of raw data because at least one variable (number of pores) had a considerably higher variance than the others (Johnson, 1998). Extracted factors were plotted against one another to evaluate the differences between the putative lineages in multivariate space. A canonical discriminant analysis (CDA) was used to quantify the discriminatory power of these factors in delineating these lineages. Prior probabilities of inclusion were considered equal for all groups. Both PCA and CDA were carried out using SPSS version 9.0 (SPSS Inc.).

MTDNA ANALYSIS

The complete mtDNA control region (CR) and a partial segment of both flanking tRNA genes were amplified in 45 individuals using primers (LRBT-25 and LRBT-1195) designed from the complete mtDNA sequence of rainbow trout *Oncorhynchus mykiss* (GenBank access. no. D83947) following the conditions described in Weiss *et al.* (2002). Polymerase chain

Table 1. Sample locations including river basin, geographical coordinates and the number of individuals scored for meristic and mtDNA characters

Map no.	Population	Form	Basin	Population code	No. of individuals			Longitude
					mtDNA	Meristics	Latitude	
19	Kalar River	<i>Thymallus arcticus</i> 'Lena'	Vitim→Lena→Arctic	Kal	–	11	56°16'	119°38'
18	Leprindocan Lake	<i>T. arcticus</i> 'Lena'	Vitim→Lena→Arctic	Lep	9	16	56°33'	117°29'
20	Davatchan River	<i>T. arcticus</i> 'Lena'	Vitim→Lena→Arctic	Dav	–	19	56°33'	117°29'
17	Yakichy Creek	<i>T. arcticus</i> 'Lena'	Baikal→Enisey→Arctic	Yak	3	–	56°05'	110°46'
21	Chitkanda Lake	<i>T. arcticus</i> 'Lena'	Olekema→Lena→Arctic	Chi	1	22	57°00'	119°33'
22	Lesha Lake	<i>T. arcticus</i> 'Lena'	Olekema→Lena→Arctic	Les	8	74	57°09'	119°37'
23	Olongdo River	<i>T. arcticus</i> 'Lena'	Olongdo→Lena→Arctic	Olo	6	81	57°10'	119°41'
16	Kutima River	<i>T. arcticus</i> 'Lena'	Kirenga→Lena→Arctic	Kut	2	28	57°02'	108°36'
15	Pilka River	<i>T. arcticus</i> 'Lena'	Lena→Arctic	Pil	3	20	60°00'	113°00'
14	Aldan River	<i>T. arcticus</i> 'Lena'	Lena→Arctic	Ald	–	11	64°43'	132°18'
13	Moloda River	<i>T. arcticus</i> 'Lena'	Lena→Arctic	Mol	2	–	69°43'	124°50'
3	Anabar River	<i>Thymallus arcticus pallasii</i>	Arctic	Ana	1	32	70°48'	113°13'
4	Olenek River	<i>T. arcticus pallasii</i>	Arctic	Ole	4	–	68°31'	112°00'
1	Lena Delta	<i>T. arcticus pallasii</i>	Lena→Arctic	Del	2	23	72°01'	127°21'
2	Kengdei River	<i>T. arcticus pallasii</i>	Lena→Arctic	Ken	4	21	71°48'	127°24'
6	Inderkey River	<i>T. arcticus pallasii</i>	Yana→Arctic	Ind	2	3	64°43'	132°33'
5	Tuostakh River	<i>T. arcticus pallasii</i>	Yana→Arctic	Tuo	–	10	67°09'	135°39'
7	Khatirniya River	<i>T. arcticus pallasii</i>	Yana→Arctic	Kha	3	30	65°36'	132°10'
11	Turakh Lake	<i>T. arcticus pallasii</i>	Indigirka→Arctic	Tur	2	–	64°15'	145°15'
10	Arga River	<i>T. arcticus pallasii</i>	Indigirka→Arctic	Arg	2	–	66°34'	143°00'
9	Syuryuntyakh River	<i>T. arcticus pallasii</i>	Indigirka→Arctic	Syu	4	16	67°04'	142°20'
8	Selenyakh River	<i>T. arcticus pallasii</i>	Indigirka→Arctic	Sel	3	–	68°01'	144°23'
12	Kolyma River	<i>T. arcticus pallasii</i>	Arctic	Kol	3	4	66°10'	151°03'
–	Sjamzhenga River	<i>T. arcticus pallasii</i>	Arctic	Sja	1	–	63°70'	46°26'
–	Big Hole River	<i>T. arcticus</i> 'North America'	–	Big	2	27	45°37'	113°27'
	Total				67	448		

Population code numbers correspond to those shown on the map (Fig. 1), and the three-letter code is used for haplotypes in the tree (Fig. 2). Samples are sorted by 'form' with 'Lena' referring to the phenotype occurring in the upper Lena basin, and the subspecies 'pallasii' occurring in the lowermost Lena as well as other arctic draining rivers sampled. Sequence data from populations marked in **bold** has been previously reported by Koskinen *et al.* (2002).

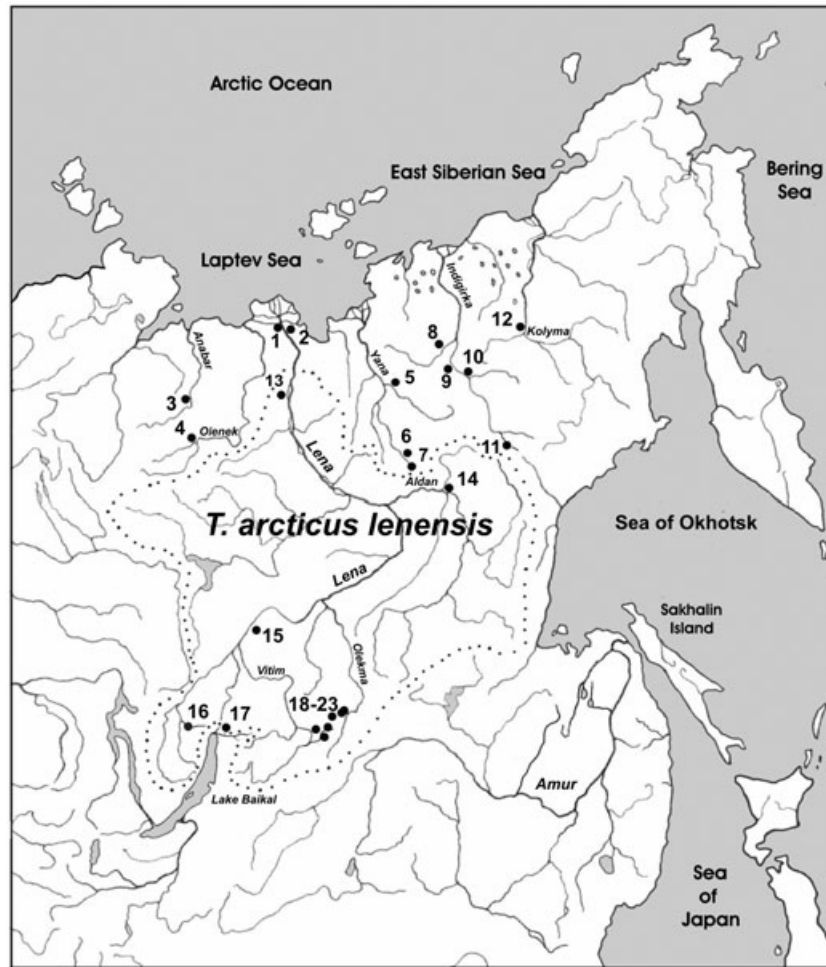


Figure 1. A map of the Siberian sample sites represented in Table 1. The major river basins are shown and, within the dotted line, the proposed range of the newly described taxon, *Thymallus arcticus lenensis*. Additional samples used in the meta analysis (nearly all North America) are found in Redenbach & Taylor (1999) and Stamford & Taylor (2004).

reactions (PCR) were carried out in 25 µl volumes. Each reaction contained 19 µl H₂O, 2.5 µl 10 µM Applied Biosystems Buffer II, 0.5 µl 10 mM of each primer, 1.5 µl 25 mM MgCl₂, 0.5 µl 10 mM dNTPs, 0.1 µl Applied Biosystems ATAQ DNA polymerase, and 0.5 µl 100 ng/µl DNA template. The PCR conditions were: initial denaturation at 94 °C for 3 min, followed by denaturation at 94 °C for 45 s, annealing at 55 °C for 45 s and extension at 72 °C for 45 s repeated for 30 cycles. Amplified DNA templates were cut from Agarose gels and purified using the NucleoSpin Extract PCR Purification Kit and 100 ng of purified PCR product was used in cycle sequencing reactions following ABI PRISM BigDye Terminator protocols. Sequences were visualized on an ABI-3100 capillary genotyper and aligned using the Sequence Navigator software. New mtDNA sequences generated in this study have been deposited in GenBank under Accession nos DQ683684–DQ683723.

For phylogenetic reconstruction both maximum parsimony (MP) and Neighbour-joining (NJ) methodologies were used in the program PAUP Version 4.0b10 (Swofford, 2001). The following previously published CR sequences were added to the analysis: 23 from the upper Lena basin; four from the Baikal/Enisey basin, and two European grayling *Thymallus thymallus* (Weiss *et al.*, 2002), making a total of 74 sequences for the final phylogenetic reconstruction. For MP, insertions or deletions (indels) were included as a fifth character. A heuristic search (ten replicates) with TBR branchswapping was employed to find the most parsimonious trees. For NJ, a Kimura two parameter substitution model was chosen. Support for the resulting nodes in the trees were obtained with 1000 bootstrap replicates.

A second phylogenetic analysis used partial CR (754 bp) in order to include sequences in Stamford & Taylor (2004), primarily stemming from North

Table 2. Number of individuals, means (SDs) for standard length (mm) and the 12 meristic characters used to characterize the two major population groups (mtDNA clades) in Siberia (*Thymallus arcticus lenensis* and *Thymallus arcticus pallasii*), in addition to the sample from North America

Character	Population group			Mann–Whitney <i>U</i> -test maximum \uparrow in mean rank (chi squared)	<i>P</i>
	<i>Thymallus arcticus lenensis</i> (Siberia) (N = 282)	<i>Thymallus arcticus pallasii</i> (Siberia) (N = 139)	<i>Thymallus arcticus</i> (North America) (N = 27)		
SL	191.7 (40.3)	239.5 (62.1)	258.2 (24.6)		
(a) Lateral line pores	92.0 (4.5)	89.1 (5.1)	91.8 (4.0)	238 (43.0)	< 0.001
(b) Unbranched dorsal rays	8.0 (0.8)	9.6 (0.9)	11.7 (1.1)	315 (176.9)	< 0.001
(c) Branched dorsal rays	12.9 (0.9)	13.7 (1.0)	10.9 (0.9)	265 (53.7)	< 0.001
(d) Total number dorsal rays	20.9 (1.0)	23.3 (1.0)	22.6 (1.1)	334 (228.7)	< 0.001
(e) Branched pectoral rays	14.3 (0.7)	15.2 (0.1)	14.9 (0.9)	284 (88.9)	< 0.001
(f) Branched pelvic rays	9.0 (0.3)	9.3 (0.5)	9.3 (0.7)	248 (43.2)	< 0.001
(g) Unbranched anal rays	4.1 (0.5)	4.1 (0.4)	5.3 (0.7)	212 (0.2)	= 0.680
(h) Branched anal rays	9.0 (0.6)	9.1 (0.6)	9.8 (0.9)	219 (1.3)	= 0.247
(i) Gill rakers	18.8 (1.1)	19.1 (1.3)	17.8 (1.2)	217 (3.9)	= 0.046
(j) Branchiostegal rays	9.1 (0.6)	9.0 (0.6)	8.4 (0.6)	217 (3.8)	= 0.052
(k) Vertebrae	55.7 (1.3)	54.8 (1.5)	57.2 (0.8)	231 (42.7)	< 0.001
(l) Pyloric caeca	15.2 (1.9)	19.3 (2.8)	19.4 (2.7)	303 (147.5)	< 0.001

The Mann–Whitney *U*-test is made between the two Siberian groups only.

American populations. For this analysis, MP and NJ approaches were repeated as above.

The between-group variation in haplotypes (corrected for within-group variation) was calculated using the net nucleotide divergence (*Da*) between groups (Kimura two parameter model), and the mean pairwise divergence (uncorrected '*p*' distances) was calculated within clades or populations using Mega Version 2.1 (Kumar *et al.*, 2001). To further characterize the differences between the major groups of haplotypes for the Siberian lineages studied, we calculated nucleotide diversity (Nei & Li, 1979), as well as the pairwise mismatch distribution (Rogers & Harpending, 1992) using the program DnaSP version 3.51 (Rozas & Rozas, 1999).

PHYLOGEOGRAPHICAL ANALYSIS

To evaluate a suite of historical and contemporary processes responsible for the spread of genetic variation across northern Siberia and North America a nested clade phylogeographical analysis (NCPA) was performed (Templeton, 2004). We additionally sought to address the hypothesis posed by Stamford & Taylor (2004) concerning multiple colonizations of North America from Asian source populations. Thus, we combined our data with that of Stamford & Taylor (2004), allowing analysis over a broad range of longitudes from the Ural mountains, across Siberia, and into North America as far east as Saskatchewan.

First, a haplotype network was constructed using the 95% parsimony criterion (Templeton, Crandall & Sing, 1992a) with gaps counted as mutational events (i.e. treated as a fifth state) using the TCS 1.13 program (Clement, Posada Crandall, 2000). The network was transformed into a nested series of clades (Templeton, Boerwinkle & Sing, 1987; Templeton, Crandall & Sing, 1992b) and the association of these clades with geographical distance was tested with a permutational contingency analysis using the program GeoDis 2.2. The significant genetic–geographical associations were interpreted following the November 2003 update of an inference key available at <http://darwin.uvigo.es/software/geodis.html>.

RESULTS

MTDNA

The alignment for 73 CR sequences across *Thymallus* included 1110 bp (including 87 bp of tRNA proline and 11 bp of tRNA phenylalanine genes). There were 112 variable sites of which 94 were parsimony informative, and 18 insertion/deletions. The shortest MP tree was 193 steps [confidence interval (CI) = 0.720; relative incidence (RI) = 0.963] and depicts two highly distinct monophyletic clades (100% bootstrap support),

one representing all Lena basin locations except for the delta (A), and the other representing the Lena delta and all other Arctic draining rivers, as well as locations from North America (B) (Fig. 2, NJ tree shown). A third monophyletic clade (99% bootstrap support) represented the four individuals from the Baikal/Enisey basin. There was bootstrap support for several subclades within clade B because haplotypes from the most eastern drainages (Indigirka and Kolyma) grouped together with those from North America, in a clade distinct from the remainder of clade B, with the North American samples from the Missouri River drainage showing some bootstrap support for a basal position within this subclade. Within this remaining portion of clade B, the most eastern populations (Anabar, Olenek, and the Ural sample) also grouped together with moderate bootstrap support. The NJ tree (Fig. 2) was essentially identical to the MP tree.

For the two major lineages (A and B), there were 17 haplotypes among 34 individuals for Lena River populations, and 21 haplotypes among 33 individuals for the Lena Delta/Arctic group. The net mean divergence between these clades was 3.2%, and the maximum pairwise divergence was 4.3%. The Lena Delta/Arctic group (*T. a. pallasii*) was more diverse ($\pi = 0.00690$) than the Lena River grayling ($\pi = 0.00174$), whose pairwise mismatch distribution was unimodal compared to a multimodal distribution for the Lena Delta/Arctic group.

For the second (partial CR) phylogenetic analysis, 14 sequences from Stamford & Taylor (2004) were added. For this data set (754 bp), there were 79 variable sites of which 68 were parsimony informative. The shortest tree had 149 steps (CI = 0.685, RI = 0.9623), and all newly added sequences fell into clade B, which retained high bootstrap support (NJ = 99%, MP = 97%). Therefore, only clade B is depicted (Fig. 3), showing that most of the North American sequences group most closely with our samples from the eastern-most drainages of Siberia (Indigirka and Kolyma). Two sequences (Nah1 and Mac2), however, representing the Nahanni Refuge (Stamford & Taylor, 2004) remain distinct from all other sequences in clade B. The mean divergence between these sequences and the rest of the clade is 1.6%.

MERISTICS

There were statistically significant differences ($P < 0.001$) between the two lineages at eight of the 12 meristics characters measured, although ranges overlapped for each character considering minimum and maximum values (Table 2). The most obvious differences were seen in the number of dorsal rays and pyloric caecae. The North American population (Big

Hole River) showed markedly larger differences to Siberian populations in almost all characters. With PCA, four factors with eigenvalues of over 1 were extracted and these explained cumulatively 56% of the variance in the data set. Plotting the first two factors against each other revealed a relatively clear trend in delineating the three geographical groups, despite only accounting for 36% of the total variance in the data (Fig. 4). Using three of the four PCA factors in the CDA (based on stepwise inclusion) allowed correct assignment of 94% of the individuals from the *T. a. pallasii* group, 95% of the individuals from the Lena grayling group, and 100% of the individuals from the Montana population. The mean individual probability of group inclusion for both the *pallasii* and Lena grayling group was approximately 92%.

COMPOSITE PHYLOGEOGRAPHICAL ANALYSIS

The 95% parsimony network represents all haplotypes from clade B except Nah1 and Mac2 (Stamford & Taylor, 2004), which fall outside the parsimony criterion (Fig. 5). The location and frequency of each haplotype included in the analysis is listed in the Appendix. The network spanned a maximum of 13 mutations and revealed several ambiguities (e.g. clades 1–2 and 2–7). The nested design included 33 clades across four nesting levels including the total cladogram. Of these, 15 exhibited both geographical and genetic variation and thus were included in the GeoDis input file to test for significant associations. A total of six clades revealed statistically significant associations between genetic and geographical variation and Table 3 lists the inferences drawn based on the GeoDis output and application of the inference key.

At the first nesting level (one step, or tip clades), reflecting events of the most recent past, clade 1–12 included the most frequent and widespread haplotypes in North America with haplotype 20 supporting an inference of contiguous range expansion (CRE). At the second nesting level, clade 2–1 contains six haplotypes distributed between the lower Lena and two Yana basin populations, whereby CRE could be inferred from the Lena to the Yana dependant on accepting an adequate sampling scale (the Yana is the next large river basin east of the Lena). Clade 2–6, containing one haplotype from the Big Hole River in Montana (haplotype 25), and two haplotypes from the very distant Beaufort Sea and Yukon River region (haplotypes 23 and 24) leads to the result of insufficient evidence to discriminate between long-distance movement and the combined effects of gradual movement during past range expansion and fragmentation (see Discussion). Clade 2–7 was highly significant and contained one-step clades from North America as well as

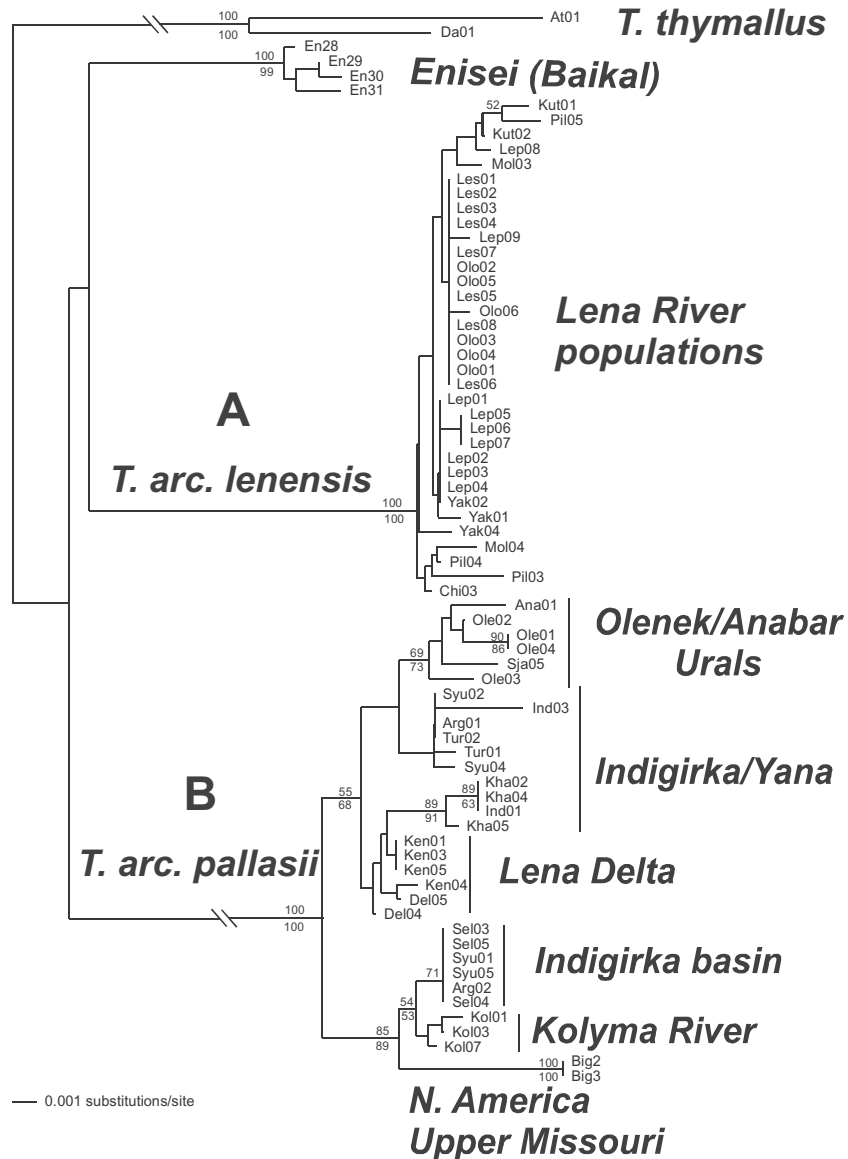


Figure 2. A neighbour-joining (NJ)-tree (Kimura two-parameter) of the complete mtDNA control region (1023 bp) and 98 bp of flanking tRNA. NJ bootstrap values (1000 replicates) are shown above nodes and maximum parsimony values below the nodes. Clade A represents the newly described taxon *Thymallus arcticus lenensis*.

the Yana (two basins), Indigirka, and Kolyma basin leading to a signal of allopatric fragmentation if the ocean is considered as the barrier (see total cladogram inference).

At the third nesting level, clade 3-1 gave a clear signal of CRE stemming from the occurrence of a tip haplotype (haplotype 7) in the upper Yana River related to Indigirka basin haplotypes. Surprisingly, this was the only clear inference among the third nesting level. For the total cladogram, the inference leads to a small ambiguity (as for clade 2-7) in considering what has,

or has not been sampled between North American locations and those in Siberia (Question 16). If considering the genetic distance between clades 3-3 and 3-1, and the lack of sampling in far north-eastern Siberia, one could draw the conclusion that the sampling scheme is inadequate to discriminate between fragmentation and isolation by distance (IBD). However, most of the geographical distance between North American and Siberian locations consists of ocean, meaning it is not grayling habitat, and thus allopatric fragmentation is inferred.

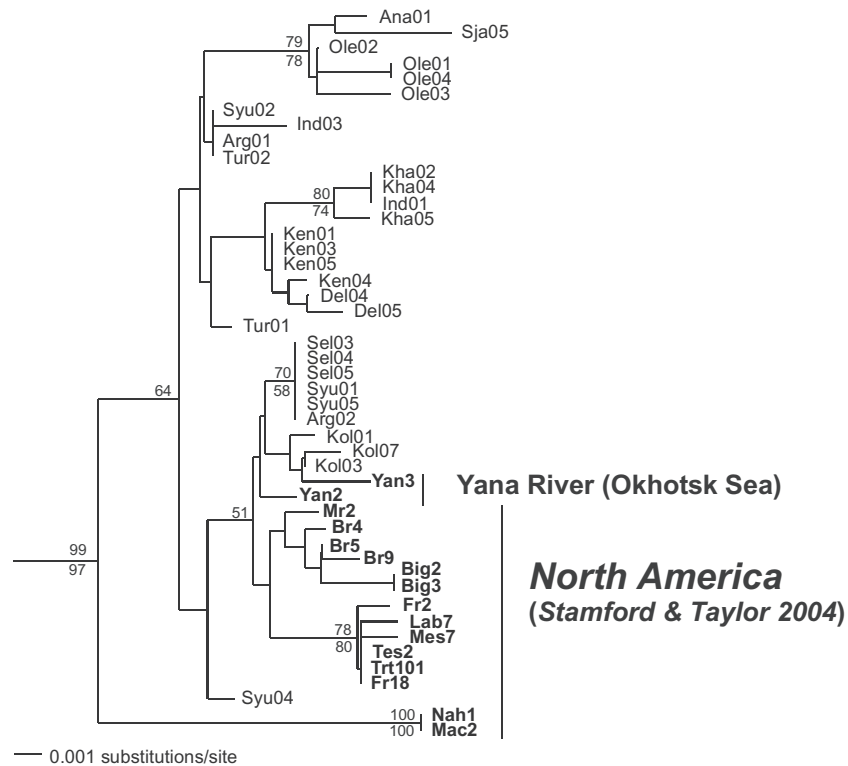


Figure 3. A neighbour-joining (NJ)-tree (Kimura two-parameter) 754 bp of the control region, analysed to incorporate sequences described by Stamford & Taylor (2004) (shown in bold). Only clade B is shown because all additional samples belong to it. Bootstrap values have been recalculated for the entire tree, as shown in Figure 2.

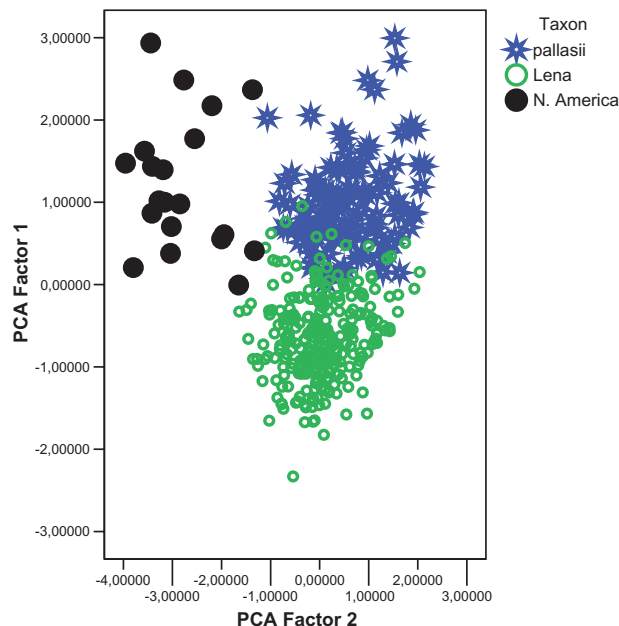


Figure 4. Scatterplot of the first two principal components from the principal component analysis of 13 meristic characters. Individuals scored from clades A and B are coded separately, as well as individuals from the Big Hole River population in North America.

DISCUSSION

PHYLOGENETIC ANALYSIS

Expanding molecular based investigations of *Thymallus* in Eurasia continue to reveal extensive genetic variation, multiple lineages within major river basins, and phylogenetic relations that do not correspond to the existing systematic scheme (Koskinen *et al.*, 2000, 2002; Sušnik, Snoj & Dovc, 2001; Weiss *et al.*, 2002; Froufe *et al.*, 2003b, 2005), and additionally lend support to a scenario of large-scale, palaeo-hydrological dynamics in Siberia that have been previously under appreciated in terms of its affects on the distribution of freshwater organisms (Froufe *et al.*, 2003a). The present study focuses on grayling lineages that have been given some historic attention in the Russian literature from a phenotypic perspective, but have lacked geographically comprehensive description or any attempt at molecular genetic verification. Our phylogenetic analysis clearly supports two distinct lineages within the Lena basin (clades A and B; Fig. 2), with a net mean divergence that reflects a split reaching into the mid-Pliocene epoch (3.2% corresponding to 3.2 Myr, based on the 1% calibration in Koskinen *et al.*, 2002). It is additionally clear that all samples derived from other Arctic draining rivers in Siberia, as

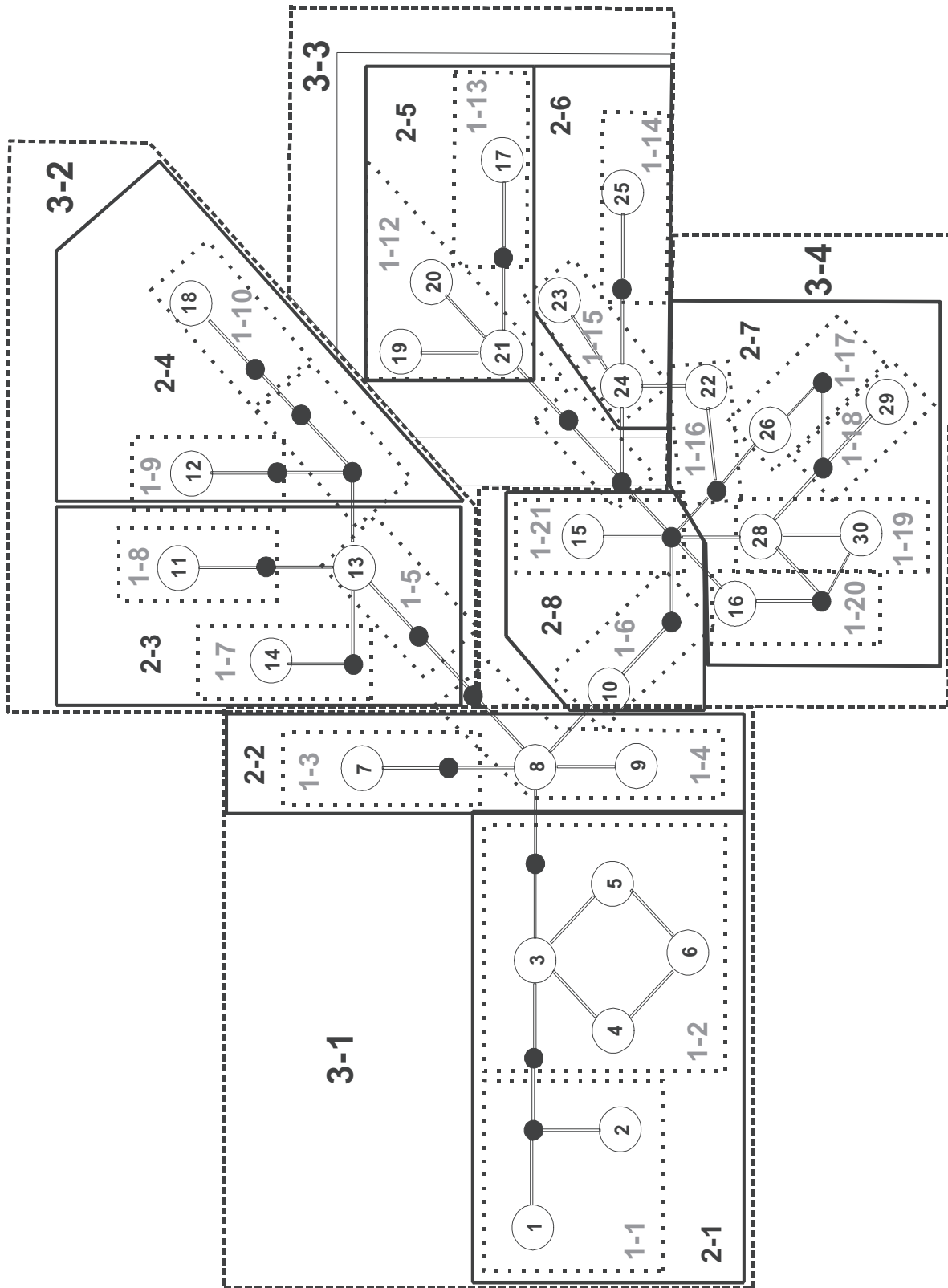


Figure 5. The 95% parsimony network and the nested design used in the nested clade phylogeographical analysis. All haplotypes are included from clade A (Fig. 3) excluding Nah1 and Mac2.

Table 3. Summary results of the nested clade phylogeographical analysis

Haplotype or clades tested	Chi square; <i>P</i>	Inference chain
Haplotypes in clade 1-12	152; 0.009	1N; 2N; 11Y; 12N = contiguous range expansion
One step clades in clade 2-1	10; 0.008	1N; 2N; 11Y; 12N = contiguous range expansion
One step clades in clade 2-6	10; 0.045	1N; 2N; 11Y; 12Y; 13Y; 21N = inconclusive (see text)
One step clades in clade 2-7	231; <0.001	1Y; 19N = AF (see text)
Two step clades in clade 3-1	12.7; 0.013	1N; 2N; 11Y; 12N = contiguous range expansion
Three step clades in total	869; 0.001	1N; 2Y; 3Y; 5Y; 15Y; 16Y + N = AF + IBD

The chi-squared statistic and *P*-value are shown for the permutational contingency analysis performed in GeoDis. For each test, the haplotypes or clades involved is listed, and the results of the dichotomous inference chain (Yes/No questions). AF, allopatric fragmentation; IDD, isolation by distance.

well as all North American samples screened by Stamford & Taylor (2004) belong to only one of these lineages, to which the samples from Kolyma River (type locality of *T. a. pallasii*) also belongs (clade B). Furthermore, excluding the Nahanni refuge haplotypes (Nah1 and Mac2), the North American haplotypes all fall within a subclade of Clade B, consisting of Indigirka and Kolyma river haplotypes, the most eastern sites sampled in our study.

The inclusion of the relatively divergent (~1.5%) Nahanni refuge haplotypes (Stamford & Taylor, 2004) does not challenge this perspective because clade B is still highly supported (97–98% bootstrap values; Fig. 3). Although we cannot presently suggest that North American populations should be considered *T. a. pallasii*, the mtDNA clade B is quite widespread, containing populations as far west as the western slopes of the Urals (haplotype Sja05) and as far east as Saskatchewan, North America (haplotype Mr2). Haplotype Sja05 corresponds to the same individual first sequenced for the NADH-1 gene by Koskinen *et al.* (2000), shown to be highly divergent from all other European samples, and haplotype Mr2 corresponds to the restriction fragment length polymorphism haplotype 7 (Stamford & Taylor, 2004) found in two populations in Saskatchewan, as well as the Big Hole drainage in the upper Missouri River in Montana (their sample). Considering the hypothesis of multiple colonization of North America from source populations in Siberia, our data cannot exclude the possibility that the Nahanni haplotypes stem from a separate colonization event preceding that generating the spread of genetic variation elsewhere in North America. However, because the first split in clade B is between the Nah1 and Mac2 haplotypes and all others in the clade, one cannot also exclude the possibility that Siberian clade B haplotypes stem from a 'return' colonization from North America. Such a route is additionally supported because the Missouri drainage haplotypes of North America are basal to Kolyma and Indigirka drainage haplotypes in Siberia. A return, or even

multiple bilateral movements between North American and Siberian lineages is further supported by the lack of monophyly for North American lineages (Fig. 3). This analysis sheds little light on the historical development of the Nahanni refuge haplotypes described by Stamford & Taylor (2004), except that their divergence clearly predates the postglacial (or perhaps several glacial cycles) spread of the remaining variation in North America as well as the delta regions of major north-flowing Siberian river systems.

All samples from the Lena basin upstream of the Delta region exhibit a haplotype belonging to the highly supported clade A. The clade is far less diverse than clade B, although samples sites span over 2000 km of drainage area. This clade further contains several haplotypes carried by individuals captured outside of the Lena Basin, in a small north Baikal tributary, Yakchiy Spring (Yak-pop 17; Fig. 1). All previously investigated individuals from the Baikal basin (Koskinen *et al.*, 2002; Froufe *et al.*, 2005; S. Weiss, unpubl. data) belong to a diverse but shallow clade first shown by Koskinen *et al.* (2002) and represented in Figure 2 by the 'Eni' haplotypes. The Yakchiy system is small, and flows over a steep gradient from a plateau just north of the Lake Baikal down into the lake. Upstream migration of grayling from Baikal is clearly impossible, and we hypothesize that the plateau region containing several small lakes where the fish were sampled represents a postglacial connection with the nearby headwaters of the Lena basin.

PHENOTYPIC ANALYSIS

There were highly significant differences in the mean values for eight of the 12 meristic characters but multivariate analysis (CDA) was necessary for adequate diagnosis of the two Siberian lineages. Although North American populations genetically group with clade A, our one population sample is inadequate to assess the generality of phenotypes there, but we assume that these meristic characters will vary significantly over

such a vast area with heterogeneous environmental conditions. Nonetheless, it is noteworthy that the Missouri drainage population differs substantially from all other Siberian grayling operational taxonomic units (OTUs) thus far sampled. For example, both the number of vertebrae as well as dorsal rays exceeds (nonoverlapping mean and SD) that of all three OTUs described by Froufe *et al.* (2003b) ($N = 207$), as well as three additional OTUs of *T. arcticus* from the Baikal basin ($N = 310$, S. Weiss *et al.*, unpubl. data). The Big Hole population also exhibited the highest mean number of anal rays and the lowest mean number of branchiostegal rays among the aforementioned OTUs. Such disparate counts support our ability to distinguish the Big Hole population from the two Siberian taxa with 100% assignment using CDA although correct assignment between the two Siberian taxa was also quite high (94–95%). Because Siberian samples come from a variety of localities, both lakes and rivers, we assume that the within-taxon phenotypic similarities are based on shared ancestry (whether adaptive or not) and not environmental conditions.

Although sample sites for clade B extend into the middle and lower reaches of the Lena River, we have not yet found a zone where the potential for reproductive isolation could be assessed. The full congruency of phenotypic and mitochondrial data and the lack of any site with a mixture of the mtDNA clades lends confidence to the distinction of these two lineages. In this framework, we have thus chosen to provide a preliminary name for the undescribed Lena River taxon (*Thymallus arcticus lenensis*), which emphasizes its apparent geographical limitation to the Lena basin, distinguishes it from *T. arcticus pallasii*, and holds to the tradition (for now) of subspecific designation for distinct units of *T. arcticus* in Siberia. We are currently collecting systematic phenotypic and genetic data for all potential OTUs within the genus, as well as comprehensive geographical sampling to uncover potentially cryptic OTUs. Thus, a more formal effort of systematic revision of the *T. arcticus* complex as well as other *Thymallus* taxa must await a more comprehensive meta-analysis.

PHYLOGEOGRAPHY

A strength of NCPA is that it limits speculative inferences based on limited sample sizes or geographical coverage. In our application of NCPA to the broadly distributed clade B (Urals to Saskatchewan), several limitations in sample size, geographical resolution and geographical coverage were noted. Whereby 311 individuals from the frequency based study of Stamford & Taylor (2004) could be incorporated into our joint analysis, there were only eight distinct haplotypes in the network (because Nahanni haplotypes lie

beyond the 95% parsimony network) and, in our own analysis, several geographically distant populations were only represented by one or a few individuals. The more limited sampling in Siberia prevented any inference, for example, at the first nesting level, which included clades 1-5 and 1-7 through 1-11, representing haplotypes in the most western range of our sampling, including Sja05 found west of the Urals. These haplotypes are positioned in presumably derived positions (as tip clades) suggestive of range expansion but, with only one haplotype per clade, no test for geographical \times genetic correlation can be made. Likewise, for clades 1-1 and 1-2, which contained variation, there was little within-clade geographical variation. Thus, only in North America, with the denser geographical sampling, was an inference of contiguous range expansion supported across a region where the Cordilleran ice sheet must have severely limited available grayling habitat during the Wisconsinian glaciation, an inference essentially described by Stamford & Taylor (2004). At the next nesting level, another inference across the region is drawn, involving locations at the extreme north and south of the former extent of the Cordilleran ice sheet. However, some ambiguity arises when attempting to precise this inference (i.e. long distance movement vs. gradual movement). Within clade 2-6, the sample location of clade 1-14 (Big Hole River) is separated from the rest of grayling's distribution by an area that does not contain the species, but there is also a large sampled area between clade 1-14 and clade 1-15, which contains haplotypes from distance locations in Alaska. This complication (the species both existing and being absent over large regions between areas representing the clades involved in the inference) is not adequately covered in the key, but it is difficult to imagine that the historical process responsible for this spread of genetic variation was simple (i.e. long distance movements) and thus the more complex scenario of the combined effects of movement during past range expansion and fragmentation (fragmentation is an absolute) appears likely. A similar interpretive dilemma is seen for clade 2-7 whereby an inference of allopatric fragmentation is drawn based on the clades existing on separate continents, but noting that extreme north-east Siberia has not been adequately sampled; the inference involving the total cladogram is likewise affected. Within Siberia, the CRE inference for clade 2-1 (Lena to Yana) and clade 3-1 (Yana to Indigirka) suggests a stepwise refuge and expansion pattern from north central Siberian systems (perhaps both to the east and west considering tip positions of haplotypes in the west, although only eastern expansion is supported in this analysis).

These inferences can be put into the framework of a brief palaeo-hydrological synopsis of Siberia. One,

albeit controversial perspective on Siberian Pleistocene glaciation depicts a series of ice sheets prevailing along the polar continental shelves and coastal Pacific lowlands, which variously blocked the flow of north flowing rivers and created a series of proglacial lakes (Grosswald, 1998). Such lakes are well documented for the Enisey and Ob systems (Yamskikh Anatoly, 1998). Furthermore, interior mountain regions (e.g. Trans-Bakalian) were glaciated perhaps above 1000–1200 m, and thus most if not all of the upper Lena basin was glaciated (Grosswald, 1998). Thus, north flowing rivers such as the Lena would have been isolated from far east Siberian systems (the Gobi–Amur system) as well as the central Asian region, which remains a closed basin to this day. However, even under this model (which depicts more extensive glaciation as alternative minimum models), the southern extent of marine ice sheets never reached the continental divide, and thus potential refugia for cold adapted organisms must have existed in central and east Siberia, north of the interior mountain systems. *Thymallus a. lenensis* must have survived in a refuge located near the present middle reaches of the Lena basin, and its postglacial expansion did not involve leaving the basin. *Thymallus a. pallasii*, in contrast, must have entered the lower reaches of the Lena, after marine ice sheets retreated. Interestingly, *T. a. pallasii* exists throughout the Olenek, Yana, Indigirka, and Kolyma systems, but not beyond the delta region of the Lena, although there is no physical barrier preventing upstream migration. The other systems, however, lack a second grayling taxon. Thus, as suggested for the largely allopatric distribution of brown trout *Salmo trutta* lineages in Europe (Bernatchez, 2001), ecological and/or genetic factors may be limiting the expansion of *T. a. pallasii* into a system previously occupied by another closely related taxon.

CONCLUSIONS

The comprehensive description of grayling lineages in the Lena and surrounding drainages falls in line with previous *Thymallus* studies, which continue to present rich genetic diversity, complex phylogeographical structure with respect to river drainages, and, for salmonid fishes, a remarkable concordance between phenotypic and genetic data. The present study is the third relatively extensive investigation to reveal no evidence of interbreeding among grayling lineages (Koskinen *et al.*, 2002; Froufe *et al.*, 2003b). Although we hypothesize that certain phenotypic characters, namely fin and body coloration (data not shown), may help drive reproductive isolation, and thus the good phenotypic/genetic correlation, this must be tested. Furthermore, it will require meta-analysis involving more populations within taxa to determine if the mer-

istic or other phenotypic differences among genetically delineated lineages that we find in this study, those of Froufe *et al.* (2003b) and Knizhin *et al.* (2004) as well as data from the Lake Baikal (S. Weiss, unpubl. data), correspond to evolutionary (phyletic) trends, or represent local divergence in response to a combination of environmental selection and divergent selection in response to re-enforcement for reproductive isolation. Answering this question will both serve further understanding of the evolutionary dynamics of grayling as well as salmonid fishes at the basic level, but also aid in our long-term goal of systematic revision of the genus, which must include both genetic and phenotypic diagnostic characters.

ACKNOWLEDGEMENTS

We thank C. Primmer for use of the Sja5 (Ural) sample; W. Dwyer, K. Kostow, and D. DeHart for providing the Big Hole river sample, E. Taylor for haplotype names enabling phylogeographical analysis of North American data, and D. J. Harris for comments and suggestions. Financial support for this study was provided by the Portuguese Ministry of Science and Technology, Fundação para a Ciência e Tecnologia grant POCTI/BSE/33364/99 to P. Alexandrino and S. Weiss, as well as a PhD grant to E. Froufe (SFRH/BD/11377/2002); Russian Ministry of Education (grant A03-2.12-265); and Russian program 'Universities of Russia' grant no. UR.07.01.009.

REFERENCES

- Bernatchez L. 2001.** The evolutionary history of the brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analysis of mitochondrial DNA variation. *Evolution* **33**: 351–379.
- Borisov PG. 1928.** *Ribi Reki Leni (fishes of the Lena River)*. Trudi Komissii AN SSSR po izucheniyu Yakutskoy; ASSR, Vol. IX, 181.
- Chereshnev IA, Shestakov AV, Skopetz MB. 2001.** *Opre-delitel' presnovodnikh rib severo-vostoka Rossii (guide to freshwater fish of Northeast Russia)*. Vladivostok: Dalnauka.
- Chereshnev IA, Volobuev VV, Shestakov AV, Frolov SV. 2002.** *Salmonoid fishes in Russian North-East*. Vladivostok: Dalnauka.
- Clement M, Posada D, Crandall KA. 2000.** TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**: 1657–1660.
- Froufe E, Alekseyev S, Knizhin I, Alexandrino P, Weiss S. 2003a.** Comparative phylogeography of salmonid fishes (Salmonidae) reveals late to post-Pleistocene exchange between three now-disjunct rivers basins in Siberia. *Diversity and Distributions* **9**: 269–282.
- Froufe E, Knizhin I, Koskinen MT, Primmer CR, Weiss S. 2003b.** Identification of reproductively isolated lineages of Amur grayling (*Thymallus grubii* Dybowski 1869): concor-

- dance between phenotypic and genetic variation. *Molecular Ecology* **12**: 2345–2355.
- Froufe E, Knizhin I, Weiss S. 2005.** Phylogenetic analysis of the genus *Thymallus* (grayling) based on mtDNA control region and ATPase 6 genes, with inferences on control region constraints and broad-scale Eurasian phylogeography. *Molecular Phylogenetics and Evolution* **34**: 106–117.
- Grosswald MG. 1998.** New approach to the ice age paleohydrology of Northern Eurasia. In: Benito G, Baker VR, Gregory KJ, eds. *Paleohydrology and environmental change*. Chichester: Wiley and Sons, 199–214.
- Johnson DE. 1998.** *Applied Multivariate Methods for Data Analysis*. Pacific Grove, California: Duxbury Press.
- Kirillov FN. 1972.** *Ribi Yakutii Nauka (fishes of the Yakutiya)*. Moscow: Nauka.
- Knizhin I, Weiss S, Antonov AL, Froufe E. 2004.** Morphological and genetic diversity of Amur graylings (*Thymallus*; Thymallidae). *Journal of Ichthyology* **44**: 59–76.
- Koskinen MT, Knizhin I, Primmer CR, Schlötterer C, Weiss S. 2002.** Mitochondrial and nuclear DNA phylogeography of *Thymallus* spp. (grayling) provides evidence of ice-age mediated environmental perturbations in the world's oldest body of freshwater, Lake Baikal. *Molecular Ecology* **11**: 2599–2611.
- Koskinen MT, Piironen J, Primmer CR. 2001.** Interpopulation genetic divergence in European grayling (*Thymallus thymallus*, Salmonidae) at a microgeographic scale: implications for conservation. *Conservation Genetics* **2**: 133–143.
- Koskinen MT, Ranta E, Piironen J, Veselov A, Titov S, Haugen TO, Nilsson J, Carlstein M, Primmer CR. 2000.** Genetic lineages and postglacial colonization of the grayling (*Thymallus thymallus*, Salmonidae) in Europe, as revealed by mitochondrial DNA analyses. *Molecular Ecology* **9**: 1609–1624.
- Kumar S, Tamura K, Jakobsen IB, Nei M. 2001.** mega2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**: 1244–1245.
- Makoedov AN. 1999.** *Rodstvenniye otnosheniya khariusov Sibiri i Dalnego Vostoka*. Moscow: UMK Psikhologiya.
- Nei M, Li WH. 1979.** Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America* **76**: 5269–5273.
- Novikov AS. 1966.** *Ribi Reki Kolimi (fishes of the Kolyma River)*. Moscow: Nauka.
- Pravdin IF. 1966.** *Rukovodstvo po izucheniyu rib* (Manual for fish investigation). Moscow: Pishevaya promoshlennost. [In Russian.]
- Redenbach Z, Taylor EB. 1999.** Zoogeographical implications of variation in mitochondrial DNA of Arctic grayling (*Thymallus arcticus*). *Molecular Ecology* **8**: 23–25.
- Rogers AR, Harpending HC. 1992.** Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**: 552–569.
- Rozas J, Rozas R. 1999.** Dnasp, version 3.1. An integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15**: 174–175.
- Scott WB, Crossman EJ. 1998.** *Freshwater fishes of Canada*, 5th edn. Oakville: Galf House Publications, Ltd.
- Shubin P, Zakharov A. 1984.** Hybridization between European grayling, *Thymallus thymallus*, and Arctic grayling, *Thymallus arcticus*, in the contact zone of the species. *Journal of Ichthyology* **4**: 159–163.
- Stamford MD, Taylor EB. 2004.** Phylogeographic lineages of Arctic grayling (*Thymallus arcticus*) in North America: divergence, origins and affinities with Eurasian *Thymallus*. *Molecular Ecology* **13**: 1533–1549.
- Sušnik S, Snoj A, Dovec P. 2001.** Evolutionary distinctness of grayling (*Thymallus thymallus*) inhabiting the Adriatic river system, as based on mtDNA variation. *Biological Journal of the Linnean Society* **74**: 375–385.
- Swofford DL. 2001.** *PAUP* ver 4.0.b3a. Phylogenetic Analysis Using Parsimony and Other Methods*. Sunderland, MA: Sinauer Associates.
- Templeton A. 2004.** Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology* **13**: 789–809.
- Templeton AR, Boerwinkle E, Sing CF. 1987.** A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics* **117**: 343–351.
- Templeton AR, Crandall KA, Sing CF. 1992a.** A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**: 619–633.
- Templeton AR, Crandall KA, Sing CF. 1992b.** A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**: 619–633.
- Tugarina PY, Knizhin IB. 1986.** K voprosu o morfologicheskoi neodnorodnosti rib roda *Thymallus* vodoyomov Yakutii Tesisi dokladov XI Vsesoyuznogo simposiuma 'Biologicheskie problemi Severa'. *Ikhtiologiya, gidrobiologiya, gidrokhimiya, entomologiya i parazitologiya*. Vip. 4. Yakutsk: YaF SO AN SSSR, 64–65.
- Weiss S, Persat H, Eppe R, Schlötterer C, Uiblein F. 2002.** Complex patterns of colonization and refugia revealed for European grayling *Thymallus thymallus*, based on complete sequencing of the mtDNA CR. *Molecular Ecology* **11**: 1393–1407.
- Yamskikh Anatoly F. 1998.** Late Pleistocene and Holocene Siberian river valley geomorphogenesis as a result of palaeogeographical cyclic changes. In: Benito G, Baker VR, Gregory KJ, eds. *Paleohydrology and Environmental Change*. Chichester: Wiley and Sons, 111–124.
- Zinoviev EA. 1980.** *Parallelizm Izmenchivosti U Evropeiskogo I Sibirskogo Khariusov // Lososevidniye Rybi*. Leningrad: ZIN AN SSSR, 69–80.

APPENDIX

Haplotype frequencies by populations for all individuals included in the nested clade phylogeographical analysis (NCPA). Shown are the geographical coordinates for each population, the single digit population code (1-13) for populations from this manuscript, and a two digit code (1-1) for populations from Stamford & Taylor (2004) whereby the second digit corresponds to their code. The last population (Yana) stems from data reported by Redenbach & Taylor (1999). Haplotype numbers appearing in Figure 5 (NCPA) are given, as are numbers taken from Stamford & Taylor (2004) (1–10), whereby haplotypes 3–5 collapse into one haplotype. Haplotype names correspond to those appearing in the tree of clade B (Fig. 3), whereby some codes correspond to multiple sequences (haplotypes 1, 3, 8, 14, 16), but only one code is given.

				Stam. & Tay_no												
				Hap_code_Fig3												
				Kha2	Kha5	Ken1	Ken4	Del4	Del5	Ind3	Suy2	Tur1	Suy6	Ole3		
Population				NCPA_Hap_no.	1	2	3	4	5	6	7	8	9	10	11	
Latitude	Longitude	No.	Name													
72.01	127.21	1	Lena delta	0	0	0	0	1	1	0	0	0	0	0	0	
71.48	127.24	2	Kengdei River	0	0	3	1	0	0	0	0	0	0	0	0	
70.48	113.13	3	Anabar River	0	0	0	0	0	0	0	0	0	0	0	0	
68.31	112.0	4	Olenek River	0	0	0	0	0	0	0	0	0	0	0	1	
67.09	135.39	6	Inderkey River	1	0	0	0	0	0	1	0	0	0	0	0	
64.43	132.33	7	Khatiriniya R.	2	1	0	0	0	0	0	0	0	0	0	0	
65.36	132.1	8	Selenyakh R.	0	0	0	0	0	0	0	0	0	0	0	0	
68.01	144.23	9	Syuryuntyakh R.	0	0	0	0	0	0	0	1	0	1	0	0	
67.04	142.2	10	Arga River	0	0	0	0	0	0	0	1	0	0	0	0	
66.34	143.0	11	Turakh Lake	0	0	0	0	0	0	0	1	1	0	0	0	
64.15	145.15	12	Kolyma River	0	0	0	0	0	0	0	0	0	0	0	0	
66.1	151.03	13	Sjamzhenga	0	0	0	0	0	0	0	0	0	0	0	0	
70.13	-147.5	1_1	Sagavonirktok	0	0	0	0	0	0	0	0	0	0	0	0	
64.49	-145.52	1_4	Chena River	0	0	0	0	0	0	0	0	0	0	0	0	
61.53	-134.53	1_6	Teslin River	0	0	0	0	0	0	0	0	0	0	0	0	
59.54	-130.35	1_7	Plate River	0	0	0	0	0	0	0	0	0	0	0	0	
56.39	-131.51	1_10	Stikine River	0	0	0	0	0	0	0	0	0	0	0	0	
59.42	-128.58	1_11	Blue River	0	0	0	0	0	0	0	0	0	0	0	0	
59.54	-130.27	1_12	Upper Tootsie R.	0	0	0	0	0	0	0	0	0	0	0	0	
59.09	-127.36	1_13	Turnagain R.	0	0	0	0	0	0	0	0	0	0	0	0	
59.24	-126.01	1_14	Trout River	0	0	0	0	0	0	0	0	0	0	0	0	
59.55	-123.52	1_16	Beaver River	0	0	0	0	0	0	0	0	0	0	0	0	
59.43	-124.19	1_17	LaBiche River	0	0	0	0	0	0	0	0	0	0	0	0	
57.52	-123.01	1_19	Minnaker River	0	0	0	0	0	0	0	0	0	0	0	0	
54.42	-122.18	1_23	Table River	0	0	0	0	0	0	0	0	0	0	0	0	
54.48	-122.32	1_24	Anzac River	0	0	0	0	0	0	0	0	0	0	0	0	
55.3	-122.31	1_25	Nation River	0	0	0	0	0	0	0	0	0	0	0	0	
56.08	-124.28	1_26	Mesilinka River	0	0	0	0	0	0	0	0	0	0	0	0	
56.44	-125.02	1_27	Ingenika River	0	0	0	0	0	0	0	0	0	0	0	0	
55.2	-121.44	1_29	Burnt River	0	0	0	0	0	0	0	0	0	0	0	0	
56.06	-120.23	1_30	Beatton River	0	0	0	0	0	0	0	0	0	0	0	0	
58.56	-103.57	1_32	Black Lake	0	0	0	0	0	0	0	0	0	0	0	0	
58.2	-101.11	1_33	Wollaston Lake	0	0	0	0	0	0	0	0	0	0	0	0	
45.15	-111.28	1_34	Madison River	0	0	0	0	0	0	0	0	0	0	0	0	
45.25	-112.15	1_35	Big Hole River	0	0	0	0	0	0	0	0	0	0	0	0	
60.00	150.00	-	Yana River	0	0	0	0	0	0	0	0	0	0	0	0	
Total number of individuals				3	1	3	1	1	1	1	3	1	1	1		

																				6	2	1	3	9	8	10	7								
Ana1	Ole2	Ole1	Yan2	Sel3	Lab7	Sja5	Mes7	Fr2	Tes2	Br4	Br9	Br5	Us2	Mr2	Kol7	Kol3	Yan3	Kol1	Total																
12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30																	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2																
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4																
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1																
0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4																
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2																
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3																
0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3																
0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4																
0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2																
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2																
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	3																
0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1																
0	0	0	0	0	0	0	0	0	0	11	3	3	0	0	0	0	0	0	17																
0	0	0	0	0	0	0	0	10	7	0	2	0	0	0	0	0	0	0	19																
0	0	0	0	0	0	0	0	10	5	0	0	0	0	0	0	0	0	0	15																
0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	5																
0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	10																
0	0	0	0	0	0	0	0	14	5	0	0	0	0	0	0	0	0	0	19																
0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	5																
0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	4																
0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	3																
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1																
0	0	0	0	0	1	0	0	2	6	0	0	0	0	0	0	0	0	0	9																
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1																
0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	20																
0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	20																
0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	20																
0	0	0	0	0	0	0	1	19	0	0	0	0	0	0	0	0	0	0	20																
0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	20																
0	0	0	0	0	0	0	0	22	0	0	0	0	0	2	0	0	0	0	24																
0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	14																
0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	10																
0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	20																
0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2																
0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	0	0	0	0	5																
0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2																
1	1	2	1	6	1	1	1	177	46	11	5	3	2	37	1	1	1	1	316																