Molecular Ecology (2000) **9**, 1609–1624

Genetic lineages and postglacial colonization of grayling (*Thymallus thymallus*, Salmonidae) in Europe, as revealed by mitochondrial DNA analyses

M. T. KOSKINEN,* E. RANTA,* J. PIIRONEN,† A. VESELOV,‡ S. TITOV,§ T. O. HAUGEN,¶ J. NILSSON,** M. CARLSTEIN**†† and C. R. PRIMMER*

*Integrative Ecology Unit, Department of Ecology and Systematics, PO Box 17, 00014 University of Helsinki, Helsinki, Finland, †Finnish Game and Fisheries Research Institute, Enonkoski, Finland, ‡Institute of Biology, Karelian Research Centre, Russian Academy of Sciences, Petrozavodsk, Russia, §State Research Institute on Lake and River Fisheries, St. Petersburg, Russia, ¶Department of Biology, Division of Zoology, Oslo, Norway, **Department of Aquaculture, Swedish University of Agricultural Sciences, Umeå, Sweden*

Abstract

In stark contrast to other species within the Salmonidae family, phylogeographic information on European grayling, *Thymallus thymallus*, is virtually nonexistent. In this paper, we utilized mitochondrial DNA polymerase chain reaction–restriction fragment length polymorphism (mtDNA PCR–RFLP) and sequence variation to infer the postglacial dispersal routes of *T. thymallus* into and within northern Europe, and to locate geographically, potential evolutionarily distinct populations. Mitochondrial analyses revealed a total of 27 *T. thymallus* haplotypes which clustered into three distinct lineages. Average pairwise interlineage divergence was four and nine times higher than average intralineage divergence for RFLP and sequence data, respectively. Two European grayling individuals from the easternmost sample in Russia exhibited haplotypes more genetically diverged from any *T. thymallus* haplotype than *T. arcticus* haplotype, and suggested that hybridization/introgression zone of these two sister species may extend much further west than previously thought. Geographic division of the lineages was generally very clear with northern Europe comprising of two genetically differentiated areas: (i) Finland, Estonia and north-western Russia; and (ii) central Germany, Poland and western Fennoscandia. Average interpopulation divergence in North European *T. thymallus* was 10 times higher than that observed in a recent mtDNA study of North American *T. arcticus*. We conclude that (i) North European *T. thymallus* populations have survived dramatic Pleistocene temperature oscillations and originate from ancient eastern and central European refugia; (ii) genetic divergence of population groups within northern Europe is substantial and geographically distinct; and (iii) the remainder of Europe harbours additional differentiated assemblages that likely descend from a Danubian refugium. These findings should provide useful information for developing appropriate conservation strategies for European grayling and exemplify a case with a clear need for multinational co-operation for managing and conserving biodiversity.

Keywords: biodiversity, grayling, hybridization, mtDNA, PCR–RFLP, phylogeography, sequencing

Received 13 February 2000; revision received 29 May 2000; accepted 29 May 2000

Introduction

The use of molecular genetic markers for the study of phylogeography has had widespread effects on evolutionary and conservation genetics (Avise *et al*. 1987; Moritz 1994; Avise 2000). By enabling inference of the evolutionary relationships between populations, phylogeography studies have been increasingly used for elucidating appropriate conservation strategies of endangered or extensively harvested species (Avise 1992; Bernatchez & Wilson 1998; Avise 2000).

Correspondence: Mikko T. Koskinen. Fax: + 358 9191 28701; E-mail: mtkoskin@cc.helsinki.fi
††Present address: Älvdalen Educational Center, Älvdalen, Sweden.

© 2000 Blackwell Science Ltd

Phylogeography studies of fishes, over a wide range of geographical scales, have been numerous (reviewed in Bernatchez & Wilson 1998). Due to their economical importance and relevance for fisheries management, salmonid fishes have been given enhanced attention (e.g. Bernatchez *et al*. 1992, 1999; Wilson *et al*. 1996; Pigeon *et al*. 1997; Hansen & Mensberg 1998; Hansen *et al*. 1999; Taylor *et al*. 1999). These studies have revealed the dramatic effects of the Pleistocene glaciations on the present distribution and genetic composition of salmonid populations which are currently abundant in formerly glaciated areas. Accordingly in Europe, where at the height of the late Pleistocene period, around 18 000 BP, ice sheets covered major parts of the continent (Svendsen *et al*. 1999), postglacial colonization events undoubtedly have had significant impacts on the zoogeography and genetic structure of salmonid fishes (e.g. Bernatchez *et al*. 1992; Nielsen *et al*. 1996; García-Marín *et al*. 1999; Hansen *et al*. 1999; Verspoor *et al*. 1999).

The genus *Thymallus* (Salmonidae) can be divided into four species: the European grayling (*Thymallus thymallus*), which is distributed across a large part of the European continent including the majority of the northern region; the Arctic grayling (*T. arcticus*), which occurs in North America and Russia and has an overlapping distribution with *T. thymallus* around the Ural Mountains; and *T. brovirostris* and *T. nigrescens*, which are abundant only in Mongolia (e.g. Scott & Crossman 1973). The European grayling is a marinewater intolerant species with diverse life cycles, including the occurrence of river and lake spawning forms, as well as anadromous populations which may spend a number of years in a brackish environment before returning to freshwater to spawn (reviewed in Northcote 1995). Although not as important for commercial fisheries as, e.g. *Salmo salar* or *S. trutta*, grayling is nevertheless a culturally important species, with over 400 000 kilograms harvested by anglers annually in Finland alone (Nylander 1998). Despite its relative abundance compared to other salmonids, *T. thymallus* is becoming more endangered and has been listed among the 'Threatened Fishes of Europe' (Lelek 1984). Particularly due to locally declined population sizes (Maitland & Campbell 1992), stocking practices of *T. thymallus* are commencing, which emphasizes the need for developing appropriate strategies for its conservation. However, in stark contrast to other species within the Salmonidae family, studies of genetic population structure in *T. thymallus* have been limited to very small spatial scales (Bouvet *et al*. 1990). As a consequence, conservation of the species is currently proceeding with little or no information regarding the distribution of abundant genetic variation.

As northern Europe was entirely covered by ice during the late Pleistocene period, current species of the region must originate from ancestors formerly resident in nonglaciated areas. Despite this interesting aspect, only a limited number of phylogeography studies have included a dense sample of North European populations. In recent comparative phylogeography surveys across diverse taxa, Taberlet *et al*. (1998) and Hewitt (1999) demonstrated that regions in northern Europe were generally colonized from Iberic and Balkanic refugia. Furthermore, an eastern source likely played an important role for dispersal of some species into Fennoscandia and southern parts of the Baltic (Taberlet *et al*. 1998; Hewitt 1999). Complementing these interspecific results, recent data on brown trout and perch have implied three refugia as the main sources for the postglacial colonization of northern Europe: one likely residing in the Atlantic drainages of the Iberian Peninsula and southern France; a second in the Caspian basin (García-Marín *et al*. 1999; Nesbø *et al*. 1999); and a third very probably east of the late Pleistocene ice sheets (Osinov & Bernatchez 1996; Nesbø *et al*. 1999). In general, significant gene flow between the differentiated lineages, particularly in anadromous fish species, has complicated inferences on the dispersal history of populations (Verspoor *et al*. 1999; Nilsson *et al*. submitted).

In this paper, we utilized mitochondrial DNA polymerase chain reaction–restriction fragment length polymorphism (mtDNA PCR–RFLP) and sequence variation to examine the large scale phylogeographic patterns of European grayling populations, concentrating on the North European region. Unique characteristics of mtDNA, such as uniparental inheritance, lack of recombination and elevated mutation frequencies compared to most nuclear genes, make it optimal for assessment of population structure over large geographical scales and in areas which have been reproductively isolated for long time periods (Avise *et al*. 1987; Moritz *et al*. 1987). The principle objectives of the present work were to: (i) infer the postglacial dispersal routes of grayling into and within northern Europe; and (ii) identify geographically the potential evolutionarily distinct *T. thymallus* lineages in order to provide guidelines for conservation strategies of North European populations.

Materials and methods

Sampled populations

Specimens of *Thymallus thymallus* were caught from a range of populations across northern Europe by electrofishing, nets or fly fishing over a period from 1994 to 1999. Samples from additional European drainages were kindly provided by researchers listed in the acknowledgements (Table 1; Fig. 1). In addition, one population of *T. arcticus*, originating from British Columbia, Canada, was obtained. Populations from Holmön (Hol), Ulkokrunnit (Kru), and Oura (Our) represented anadromous grayling forms

Table 1 Investigated European grayling populations with respective locations, codes and sample sizes

Location				
Region	Water system	Geographical position	Code	Sample size*
Kola Peninsula	Varzuga River	67°05′ N 36°40′ E	Var	26 (1)
northern Finland	Tenojoki	69°50′ N 26°45′ E	Ten	28 (2)
north-western Finland	Lätäseno	68°50′ N 22°10′ E	Lät	26 (1)
north-eastern Finland	Oulankajoki	66°20′ N 29°25′ E	Oul	29 (1)
north-western Russia	Pistojoki	65°35′ N 30°30′ E	Pis	5 (2)
eastern Bothnian Bay	Ulkokrunnit	65°25′ N 24°50′ E	Kru	17 (–)
eastern Bothnian Bay	Oura	61°50′ N 21°20′ E	Our	10 (1)
eastern Finland	Lieksanjoki	63°10′ N 30°25′ E	Lie	29 (1)
western Finland	Isojoki	62°10′ N 21°50′ E	Iso	24 (–)
eastern Finland	Etelä-Saimaa	61°20′ N 28°20′ E	Esa	28 (1)
north-western Russia	Lake Onega	62°40′ N 30°55′ E	One	1 (–)
western Russia	Vruda River	59°20′ N 29°10′ E	Vru	3 (–)
north-eastern Estonia	Selja River	59°30′ N 25°05′ E	Sel	26 (1)
Archangelsk	Sjamzhenga River	63°70′ N 46°26′ E	Sja	13 (3)
Bothnian Bay	Holmön	63°47′ N 20°53′ E	Hol	23 (1)
north-eastern Sweden	Vindelälven	64°12′ N 19°44′ E	Vin	26 (2)
central Sweden	Tandsjön	61°44′ N 14°37′ E	Tan	25 (–)
southern Sweden	Vättern	58°18′ N 14°17′ E	Vät	29 (1)
southern Norway	Gudbrandsdalslågen	61°17′ N 10°40′ E	Gud	21 (2)
southern Norway	Lesjaskogsvatn	62°12′ N 08°29′ E	Les	24 (–)
central Norway	Vefsna	65°50′ N 13°10′ E	Vef	14 (1)
western Denmark	Skjern River	55°55′ N 08°50′ E	Skj	26 (–)
northern Poland	Radunia River	54°20′ N 18°10′ E	Rad	10 (1)
central Germany	Eger River	50°10′ N 12°55′ E	Ege	29 (1)
southern Germany	Ramsach River	47°10′ N 11°30′ E	Ram	25 (2)
Slovenia	Obrh River	45°42′ N 14°31′ E	Obr	17 (–)
United Kingdom	Avon/Derwent/Earn†		UK	7 (1)
British Columbia‡	Upper Peace River	55°30′ N 122°35′ W	Upp	20 (1)

*Indicates the numbers of individuals assayed by PCR–RFLP (sequence) analyses. †Individuals from three rivers pooled; coordinates therefore not included. ‡*T. arcticus*.

(Fig. 1). The majority of samples included in the study were wild-caught specimens, the exceptions being the Finnish Isojoki (Iso) population (1st generation hatchery stock produced from locally caught founders) and the Polish Radunia (Rad) hatchery stock whose history is unknown. Stocking of hatchery reared fish may have occurred in some areas from which fish were sampled, however, to our knowledge, all the performed introductions into these drainages have been with locally caught and reared fish.

Extraction and PCR amplification of the mtDNA

DNA extractions were performed from ethanol stored fin clips, muscle or liver tissues using chelex-proteinase K based treatments as in Estoup *et al*. (1996) or standard phenol/chloroform procedures. Six mtDNA regions, covering approximately 8.8 kilobases (kb) of the mtDNA genome, were PCR amplified in four separate reactions with primers which have been previously reported as follows: ND 1, ND 3/4 and ND 5/6 (Nielsen *et al*. 1998); cyt *b*/D-loop with primers C-Glu (Cronin *et al*. 1993; Bernatchez & Osinov 1995) and HN20 (Bernatchez & Danzmann 1993). PCRs consisted of 24 pmol of each of the forward and reverse primers, 0.2 mM dNTP, 1.5 mM (ND 1, ND 3/4, ND 5/6) or 4.0 mM (cyt *b*/D-loop) $MgCl_2$ and 2 U of Amplitaq (ND1, 3/4; PE Biosystems) or 0.5 U of Biotaq (ND 5/6, cyt *b*/D-loop; Bioline) DNA polymerase. All amplifications were performed in 40 µL volumes using PTC 100 or PTC 200 thermal cyclers (MJ Research) with a 3 min predenaturation at 94 °C, 30 cycles of 94 °C for 1 min, 55 °C (ND 1, 3/4, 5/6) or 60 °C (cyt *b*/D-loop) for 1.5 min and 72 °C for 1.5 min, followed by a final extension at 72 °C for 5 min.

RFLP analysis

PCR amplified mitochondrial fragments were screened for polymorphism in a panel of nine individuals (one

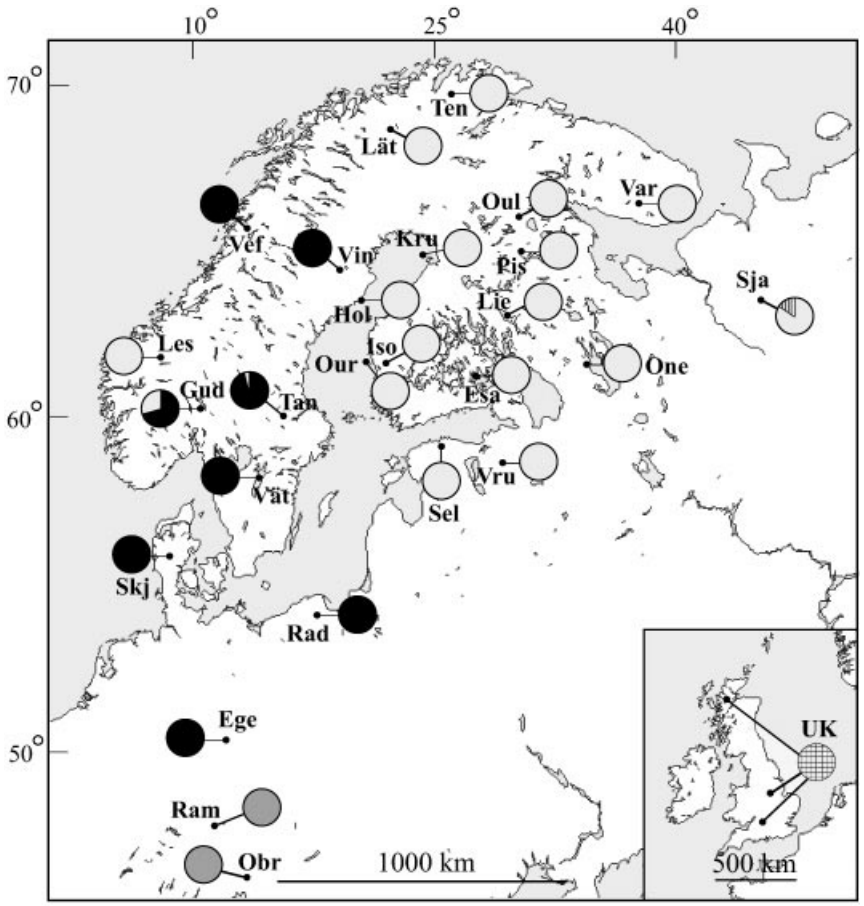

Fig. 1 Locations of the investigated *Thymallus thymallus* populations and their relative frequencies for the defined mtDNA haplotype lineages resolved by phylogenetic analysis of the restriction site and sequence data: lineage I (haplotypes 1–12), light grey; lineage II (haplotypes 17–26), black; lineage III (haplotypes 13, 15, 16), dark grey. Population codes according to Table 1. The haplotype observed in the UK samples could not be reliably grouped with any of the three lineages (haplotype cross-hatched). Two individuals from the Sjamzhenga (Sja) population displayed a haplotype (hatched) which was very distantly related to all other *T. thymallus* haplotypes (see Table 4 and Figs 2, 3 for details).

each from the Tenojoki, Lieksanjoki, Etelä-Saimaa, Varzuga River, Lake Onega, Skjern River, Gudbrandsdalslågen, Radunia River and Obrh River populations) using the restriction endonucleases *Hae*III, *Hin*cII, *Eco*RI, *Hin*fI, *Sty*I, *Rsa*I, *BamH*I, *Apa*I, *Sma*I, *Alu*I, *Xba*I, *Mbo*I, *Dra*I and *Taq*I (Promega). Restriction digestions were carried out in 30 μL volumes using 5 μL of the PCR amplified DNA and 5 U of the enzymes and buffers according to the manufacturer's recommendations (Promega). Resulting fragments were electrophoresed in 1.3% agarose gels (Seakem; FMC Bio Products) using ethidium bromide staining, alongside 100 base pair (bp) DNA ladder (Amersham Pharmacia) and visualized and photographed under an ultraviolet transilluminator. Thirteen polymorphic fragment-enzyme combinations were observed of which nine were chosen for further analysis (ND 5/6: *Hae*III, *Hin*fI, *Sty*I, *Rsa*I, *Apa*I and *Alu*I; cyt *b*/D-loop: *Hin*fI, *Rsa*I and *Taq*I). Additional polymorphic fragment-enzyme combinations, not further investigated here were ND 1 restricted with enzyme *Hae*III and ND 3/4 with enzymes *Hin*fI, *Rsa*I and *Taq*I. The selected restriction endonucleases recognized cut sites of four (*Hae*III, *Rsa*I, *Alu*I, *Taq*I), five (*Hin*fI) or six (*Sty*I, *Apa*I) base pairs.

Fragments over 200 bp in length could be generally interpreted unambiguously after agarose gel electrophoresis. To enable accurate scoring of shorter restriction fragments and to confirm the presence/absence of multiple fragments of similar length, at least one individual for each enzyme-fragment pattern was fluorescently analysed on an ABI 377 DNA sequencer. This was accomplished by including 1 mM of R6G dCTP nucleotides (PE Biosystems) into the amplification reactions prior to digestion. PCR–RFLP generated fragment profiles were identified by letters which were then combined to define composite mtDNA haplotype patterns (Table 2).

Sequencing

Individuals were selected for sequencing in order to maximize the number of different RFLP haplotypes sequenced. In total, 529 bp from the 5′ end of the ND 5 gene were sequenced for all of the composite haplotypes except haplotypes 13, 20 and 21. ND 5/6 PCR products (see above) were purified with exonuclease-shrimp alcaline phosphate treatment and 2 μL of the purified PCR products were included in 10 μL dideoxy termination sequencing

Table 2 Characterization and distribution of the PCR–RFLP composite haplotypes in 27 *Thymallus thymallus* populations and one population of *T. arcticus*. Population codes are indicated in Table 1. The composite haplotypes comprising three distinct lineages are grouped together: lineage I, haplotypes 1–12; lineage II, haplotypes 17–26; lineage III, haplotypes 13, 15, 16 (see Fig. 2 for details)

		Population by geographical area*																												
		A														B											C		
Hapl. number	Composite haplotype†	Var	Ten	Lät	Oul	Pis	Ulk	Our	Lie	Iso	Esa	One	Vru	Sel	Sja	Hol	Vin	Tan	Vät	Gud	Les	Vef	Skj	Rad	Ege	UK	Ram	Obr	Upp‡
1	aaaaaaaaa	25	27	24	27	1	17	7	5	4	28	1	3			19		1		5	24								
2	aaaaaaaac					1																							
3	acaaaaaaa					3																							
4	aaaaaabaa				2																								
5	aaabaaaaa		1																										
6	aaaaaaaae			2																									
7	aaaaaaaga													3															
8	baaabaaaa							3																					
9	bacabaaaa													23	11														
10	baaabaaab									20						4													
11	baaabbaab								24																				
12	baaabbaaa	1																											
13	daccbaaaa																											17	
15	dadcbecca																										12		
16	daddbecca																										13		
17	cbbbadaba																						26		29				
18	cbbbadbba																1												
19	cbbbadbbc																25	11	8										
20	ebbfadbbc																	13											
21	cbbhadbbc																		4										
22	cbbbadbbd																		17										
23	cbbbaabbc																			9		9							
24	cbbbafbbc																			7									
25	gbbbaabbc																					5							
26	cbbbaabdc																							10					
14§	dccbbacba																									7			
27¶	ecdiageac														2														
28	edeeacefa																												20

*Geographical areas A, B and C roughly correspond to the areas summarized to harbour distinct populations dispersed via the main colonization routes and divided by the subsequent suture zones (Taberlet *et al*. 1998; Hewitt 1999). †Produced by the restriction endonucleases *Hae*III, *Hin*fI, *Sty*I, *Rsa*I, *Apa*I, *Alu*I (ND 5/6) and *Hin*fI, *Rsa*I and *Taq*I (cyt *b*/D-loop), respectively. ‡*T. arcticus* population. §Haplotype 14 could not be confidently placed in any of the three lineages (see Fig. 2). ¶The very diverged haplotype 27 was most closely related to the Canadian haplotype 28 (see Figs 2, 3).

© 2000 Blackwell Science Ltd, *Molecular Ecology*, 9, 1609–1624

reactions, using 10 pmol of the ND 5 forward primer (Nielsen *et al*. 1998) and 4 μL of the BigDye Terminator Cycle Sequencing Ready Reaction Kit premix (PE Biosystems). Sequencing reactions consisted of a 1-min predenaturation at 96 °C and 29 cycles of 96 °C for 30 s, 50 °C for 15 s and 60 °C for 4 min. Sequenced products were purified with ethanol/sodium acetate precipitation following recommended guidelines (PE Biosystems) and electrophoresed with an ABI 377 automated sequencer (PE Biosystems). Sequences were analysed using ABI PRISM sequencing analysis software version 3.3 (PE Biosystems) and aligned and manually checked with SEQUENCHER version 3.0. Sequences included in this paper have been deposited in GenBank under accession nos AF270846–AF270858.

Statistical analyses

Fragment patterns produced by the restriction endonucleases were used to infer the presence or absence of restriction sites in the haplotypes, i.e. restriction site data. This transformation was performed because restriction fragment data violate the assumption of independence among characters and may consequently produce erroneous results (Swofford *et al*. 1996). When compared to their most closely related haplotype, all restriction fragment patterns observed in *T. thymallus*, as well as when including data from *T. arcticus*, could be accounted for by single or double restriction site changes. The inferred restriction site data were converted into binary (1, 0) format and the resulting matrix was imported into the REAP computer analysis package (McElroy *et al*. 1992) to calculate the number of nucleotide substitutions per site, i.e. d (program D; Nei & Tajima 1981; Nei & Miller 1990), weighted by the employed enzyme class (Nei & Tajima 1983). REAP was further used to estimate haplotype diversity (Nei 1987) and nucleotide diversity (Nei & Tajima 1981) within populations and nucleotide divergence (da) between all pairs of populations (program DA; Nei & Tajima 1981).

Affinities among grayling mitochondrial haplotypes were summarized using several different analyses. First, the d-values produced by REAP were analysed using Fitch-Margoliash method (Fitch & Margoliash 1967), implemented by the program FITCH of PHYLIP version 3.57c (Felsenstein 1995). Second, 529 bp of the ND 5 gene information from the sequenced haplotypes were replicated 1000 times with the program SEQBOOT (PHYLIP) and, as transitions greatly outnumbered transversions, divergence between the haplotypes was estimated using Kimura 2-parameter distances (Kimura 1980) in the program DNADIST (PHYLIP). Distances were then imported into NEIGHBOUR (PHYLIP) to create neighbour-joining (NJ) dendrograms, which were combined employing consensus tree analysis carried out by the program CONSENSE (PHYLIP). Finally, the sequence information was transformed into binary format and combined with the PCR–RFLP site data. To prevent analysis of the same mutational information twice, variable nucleotide positions already covered by the recognition sites of the employed restriction enzymes were excluded (see Table 3 for excluded positions). The resulting matrix was replicated 1000 times with SEQBOOT and analysed by Wagner parsimony method in the program MIX (PHYLIP). Again, the trees were combined using the program CONSENSE.

Hierarchical analysis of molecular variance (AMOVA) was performed for the PCR–RFLP inferred site data following the approach of Excoffier *et al*. (1992) implemented by the program ARLEQUIN version 1.1 (Schneider *et al*. 1997). AMOVA analysis was conducted using the distance matrix among haplotypes. Variance components were extracted for three hierarchical levels: among groups; between sampling sites within groups; and within sampling sites. Populations were partitioned into groups based on two alternative hypotheses: (i) based on the main postglacial colonization routes and subsequent suture zones depicted across a diverse set of European taxa (Taberlet *et al*. 1998; Hewitt 1999). These summaries suggested three approximate genetic divisions corresponding to eastern/north-eastern, central/north-western and southern lineages in the study region. Pooling the samples accordingly corresponded to divisions based on the defined geographical areas A, B and C (Table 2). And (ii) based on earlier phylogeography studies of European freshwater fish populations (Bernatchez *et al*. 1992; Bernatchez & Dodson 1994; Bernatchez & Osinov 1995; Durand *et al*. 1999; García-Marín *et al*. 1999; Nesbø *et al*. 1999). Perhaps due to limited sampling coverage of North European populations in many studies, the only clear and consistent genetic population division across European freshwater fishes could be defined between the southern (south of the Alps) and northern regions. This grouping was made by merging populations within the geographical areas A and B and retaining the group of populations within area C as above (Table 2). The statistical significance values for the attained variance components were estimated by permuting the haplotypes according to each hierarchical level. Associated Φ statistics were estimated for these levels (and the two alternative groupings) and are defined as follows: Φ_{CT} is the correlation of random haplotypes within a population group, relative to that of random pairs of haplotypes from the whole data set; Φ_{SC} is the correlation of the molecular diversity of random haplotypes within populations, relative to that of random pairs of haplotypes from within the region; Φ_{ST} is the correlation of random haplotypes within populations, relative to that of random pairs of haplotypes from the whole data set (Excoffier *et al*. 1992).

Table 3 Variable nucleotide sites across 529 base pairs of the ND 5 gene 5′ end sequence among the resolved PCR–RFLP haplotypes. Dots indicate homology with the reference sequence (PCR–RFLP haplotypes 1, 3–6). The complete sequences have been entered into GenBank under accession nos AF270846–AF270858

	Sequence position and change†																														
PCR–RFLP haplotype*	4 9	6 1	6 2	6 3	6 9	7 6	8 5	8 8	8 9	9 0	9 7	1 0 0	1 0 2	1 1 9	1 2 1	1 3 0	1 3 6	1 3 9	1 4 0	1 4 5	1 5 1	1 7 3	1 8 4	1 8 7	2 0 3	2 1 5	2 2 0	2 2 3	2 3 2	2 5 0
1, 3–6	C	C	A	T	T	C	C	A	G	A	C	A	G	C	T	C	C	A	A	T	A	A	T	G	A	C	A	A	A	T
2, 7	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
8, 10, 12	•	T	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
9	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	C
11	•	T	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	A	•	•	•	•
15	•	•	T	C	•	•	•	•	•	•	T	•	•	•	•	•	A	G	•	•	G	G	C	•	G	•	•	•	G	•
16	•	•	T	C	•	•	•	•	•	•	T	•	•	•	•	•	A	G	•	•	•	•	C	•	G	•	•	•	G	•
17	•	•	•	•	•	•	T	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	G	•
18, 19, 22–26	•	•	•	•	•	•	T	•	A	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	G	•
14	A	•	•	•	A	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	G	•
27a	•	T	•	•	•	A	T	•	A	C	T	G	A	A	C	T	•	•	C	•	•	•	•	A	G	•	G	G	•	A
27b	•	T	•	•	•	A	T	•	A	C	T	G	A	A	C	T	•	•	C	C	•	•	•	A	G	•	G	G	•	A
28‡	•	T	•	•	•	A	T	G	A	•	T	G	A	A	C	T	•	•	C	•	•	•	•	A	G	•	G	•	•	A

	Sequence position and change†																													
PCR–RFLP haplotype*	2 6 5	2 6 8	2 7 6	2 7 7	2 9 2	3 1 3	3 2 4	3 4 9	3 5 2	3 7 6	3 8 5	3 8 7	3 9 1	3 9 4	3 9 7	4 0 0	4 0 9	4 3 3	4 3 6	4 3 9	4 4 2	4 4 8	4 6 6	4 7 8	4 8 4	4 8 7	5 0 5	5 1 1	5 1 7
1, 3–6	T	C	C	C	T	A	T	T	C	T	T	T	C	A	C	G	G	G	G	C	C	A	A	C	G	A	T	T	T
2, 7	•	•	T	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
8, 10, 12	•	•	T	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
9	•	T	T	•	•	•	A	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
11	•	•	T	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
15	•	•	T	•	C	•	•	C	T	•	•	•	•	•	T	•	•	•	•	•	T	G	•	T	•	•	•	•	•	•
16	•	•	T	•	C	•	•	C	T	•	•	•	•	•	T	•	•	•	•	•	T	G	•	T	•	•	•	•	•	•
17	•	•	T	•	•	•	•	C	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
18, 19, 22–26	•	•	T	•	•	•	•	•	•	•	•	•	•	C	•	•	•	•	•	•	•	•	•	•	•	•	•	C	•	•
14	•	•	T	•	•	•	•	C	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	C	•	•
27a	•	•	T	T	•	•	•	C	•	C	C	C	A	•	•	•	A	•	A	T	•	G	G	T	A	G	C	•	C	•
27b	•	•	T	T	•	•	•	C	•	C	C	C	A	•	•	•	A	•	A	T	•	G	G	T	A	G	C	•	C	•
28‡	C	•	T	T	•	C	•	C	•	C	C	C	A	•	•	A	A	A	A	T	T	G	G	T	A	•	C	•	C	•

*PCR–RFLP haplotype nomenclature as in Table 2. The two individuals harbouring the highly diverged haplotype 27 exhibited different sequences (27a and b) †The variable position 250 was excluded from the pooled phylogenetic analysis of the PCR–RFLP and sequence data as three character states were present. The variable positions 97, 100, 102, 145 and 517 were excluded from the pooled phylogenetic analysis of the PCR–RFLP and sequence data to prevent assessment of duplicated mutational information (recognition sites of the employed restriction enzymes covered these positions). ‡*T. arcticus* haplotype.

Results

PCR–RFLP mtDNA haplotype polymorphism in Thymallus thymallus

PCR–RFLP analysis of the ND 5/6 and cyt *b*/D-loop regions of 541 *Thymallus thymallus* individuals from 27 distinct sampling locations and 20 *T. arcticus* individuals from a single Canadian population revealed 27 and one composite haplotype patterns, respectively (Table 2), with 121 distinct restriction fragments and a total of 47 variable cut sites. The fragment patterns and approximate sizes of the fragments are available from the corresponding author on request. In *T. thymallus*, over the ND 5/6 and cyt *b*/D-loop segments, the nine restriction enzymes recognized an average of 44 cut sites per composite haplotype, corresponding to an average of 204 bases of the European grayling mitochondrial genome. Restriction enzyme haplotypes were

distinguished by one to six cut sites between patterns of *T. thymallus* and by two to four cut sites between haplotypes of *T. arcticus* and *T. thymallus*.

ND 5 sequence variation

In total, 529 bp were sequenced from the 5′ end of the ND 5 gene for 26 individuals (originating from 19 populations) representing 24 different *T. thymallus* PCR–RFLP haplotypes (two individuals possessing the RFLP haplotypes 1 and 27 were sequenced) and one individual representing the *T. arcticus* haplotype. Of these 27 sequences, 13 different sequence haplotypes were observed and a total of 59 positions were polymorphic of which 54 were found among the *T. thymallus* haplotypes (Table 3). Within the whole data set, the transition/transversion ratio of the observed substitutions was 3.5. After excluding the very differentiated sequences of *T. thymallus* RFLP haplotype 27 (see below), there were 29 variable sites within the *T. thymallus* data set. Between the *T. arcticus* sequence haplotype and the haplotype 27, only 10 variable sites were detected (Table 3).

Genetic affinities and differences among haplotypes

Clustering analyses of European grayling mitochondrial haplotypes produced three distinct lineages (lineage I, haplotypes 1–12; lineage II, haplotypes 17–26; and lineage III, haplotypes 13, 15, 16; Figs 2 and 3). Considering the pooled RFLP and sequence data, grouping of haplotypes defining lineage I occurred in 70% of the bootstrap replicates, while bootstrap support of 92% and 78% was observed for clustering lineage II and III haplotypes as monophyletic assemblages, respectively. Based on the ND 5 sequence data alone, statistical support for definition of these three lineages was similarly high (with sequence data alone, however, haplotype 17 could not be confidently assigned within lineage II) (Fig. 2). The placement of haplotype 14, the only haplotype observed in the UK sample, was contradictory: in the parsimony analysis haplotype 14 clustered within lineage II with relatively low bootstrap support (Fig. 2) and in the Fitch-Margoliash analysis of the RFLP site data, it related most closely to lineage III haplotypes (Fig. 3).

Estimates of European grayling interlineage haplotype divergence were substantially larger than those observed within lineages. The average pairwise RFLP evolutionary divergence (d of REAP) between the three lineages ranged from 3.1% to 4.6%, whereas within lineages an average estimate of 0.8% was observed (Table 4a; Fig. 3). Similarly, sequence data indicated markedly higher interlineage divergence with the pairwise Kimura 2-parameter distances averaging from 1.1% to 3.6% between lineages, while within lineages, an average estimate of only 0.3% was detected (Table 4b). The *T. arcticus* haplotype was separated from the *T. thymallus* RFLP lineages I, II and III by average

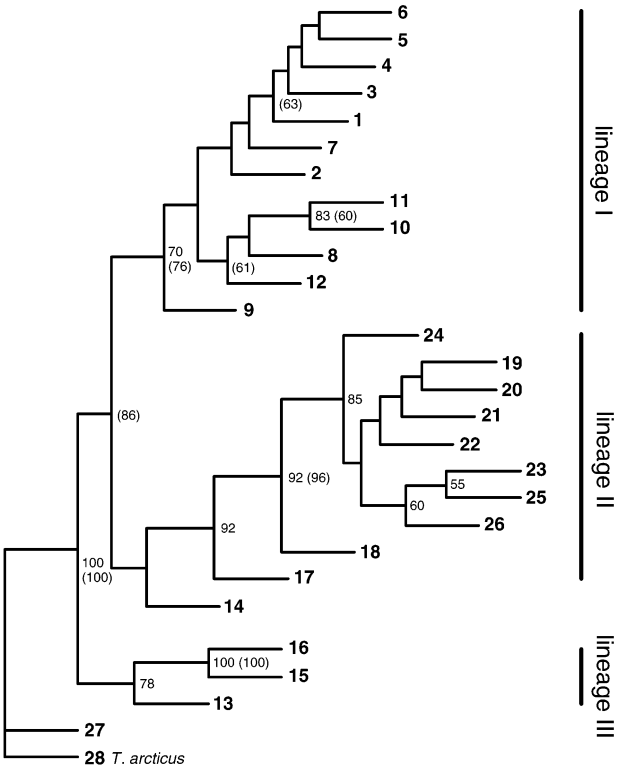

Fig. 2 Grayling mtDNA haplotype consensus tree based on Wagner parsimony analysis of pooled PCR–RFLP site data covering the ND 5/6 and cyt *b*/D-loop regions and 529 bp of ND 5 sequence data transformed into binary format. Numbering of the haplotypes (in bold) according to Table 2. Bootstrap support estimates are indicated for statistically supported groupings (≥50%). In cases where sole analysis of the ND 5 sequence data using Kimura 2-parameter distances and Neighbour-Joining method produced congruent and statistically supported topology bootstrap estimates are shown in parentheses. Thick vertical bars demonstrate the three *Thymallus thymallus* haplotype lineages.

pairwise differences of 4.9%, 6.7% and 6.0%, respectively (Table 4a; Fig. 3). Based on the ND 5 sequence data, the *T. arcticus* haplotype was separated from the *T. thymallus* lineages I, II and III with average divergence of 7.2%, 7.3% and 8.0%, respectively (Table 4b).

Lineage I included an internal cluster, formed by the haplotypes 10 and 11 and found in 83% of the respective bootstrap replicates of the pooled RFLP and sequence data (Fig. 2). Sequence data alone also revealed the haplotypes 8 and 12 to likely group within this substructure (Fig. 2). Average evolutionary divergence between this internal cluster and the remaining haplotypes in the lineage was 1.4% and 0.34% for the RFLP and sequence data, respectively. Also lineages II and III displayed substructuring with several internal groups exhibiting relatively high or high (up to 100%) bootstrap support (Fig. 2).

Interestingly, one very distinct haplotype (haplotype 27) was observed in the easternmost *T. thymallus* population

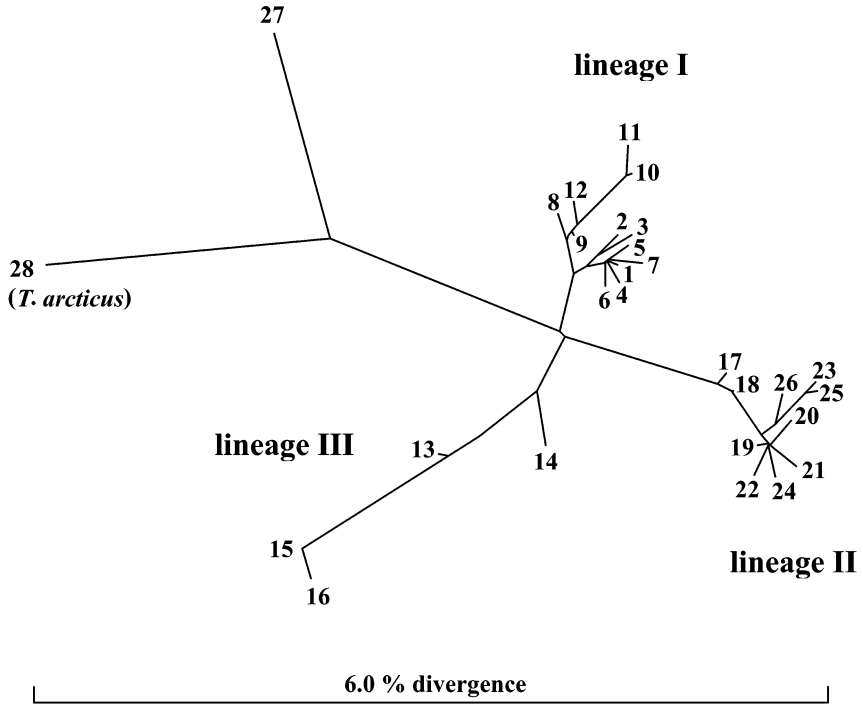

Fig. 3 Fitch-Margoliash dendrogram relating the d-values of REAP, calculated from the mtDNA PCR–RFLP inferred site matrix of the resolved 27 *Thymallus thymallus* composite haplotypes and one *T. arcticus* composite haplotype. Haplotype numbers according to Table 2.

Table 4 (**a**) Pairwise nucleotide substitution matrix (d-values) for the three defined mtDNA PCR–RFLP composite haplotype lineages, haplotype 27 and the *Thymallus arcticus* haplotype. Values were calculated according to Nei & Tajima (1981) as implemented by the program D in the package REAP; (**b**) Pairwise nucleotide substitution matrix for 529 bp of ND 5 sequences from representative individuals of 25 different PCR–RFLP haplotypes. Values were calculated using Kimura 2-parameter distances with the program DNADIST of the package PHYLIP. Indicated are the average (range) of all pairwise combinations in percentages. Intra-group divergence for the three defined lineages are indicated along the diagonal in italics

(**a**)

	lineage I	lineage II	lineage III	haplotype 27
lineage I	*0.64 (0.25–1.57)*			
lineage II	3.11 (1.86–4.11)	*0.70 (0.24–1.57)*		
lineage III	3.30 (1.20–4.67)	4.55 (2.49–5.79)	*1.10 (0.23–1.65)*	
haplotype 27	4.75 (4.31–5.24)	5.25 (4.80–5.59)	5.52 (4.58–6.12)	
T. arcticus	4.91 (4.31–5.59)	6.74 (5.82–7.48)	5.99 (4.98–6.61)	3.82*

*Range not indicated as only one comparison was made.

(**b**)

	lineage I	lineage II	lineage III	haplotype 27
lineage I	*0.36 (0.00–0.95)*			
lineage II	1.11 (0.76–1.53)	*0.14 (0.00–0.57)*		
lineage III	3.30 (2.89–3.87)	3.59 (2.89–3.89)	*0.38**	
haplotype 27	6.58 (6.30–6.93)	6.72 (6.51–6.87)	7.76 (7.55–7.97)	
T. arcticus	7.19 (6.91–7.54)	7.34 (7.12–7.50)	7.97 (7.76–8.17)	1.82 (1.72–1.92)

*Range not indicated as only one comparison was made.

(Fig. 1) sampled (100% bootstrap support for grouping of all other *T. thymallus* haplotypes separate from haplotype 27; Fig. 2). Haplotype 27 exhibited an average evolutionary divergence of 6.5% and 5.0% from the other *T. thymallus* haplotypes for sequence and RFLP data, respectively, and was in fact more closely related to the *T. arcticus* haplotype than to any other *T. thymallus* haplotype: haplotype 27 and the *T. arcticus* haplotype differed by only 1.8% in their ND 5 sequence and by 3.8% in their RFLP data (Table 4).

Distinct geographical distribution of mtDNA haplotypes in T. thymallus *populations*

The mtDNA composite haplotypes observed in European grayling populations exhibited very clear geographical divisions, with one of the three lineages being dominant in each population (Table 2; Fig. 1). Haplotypes 1–12 (lineage I; Table 2) were observed primarily across Finland, Estonia and north-western Russia, including the easternmost

Table 5 Percentage of the total molecular variance explained (%), statistical significance estimates (P) and Φ statistics for two alternative groupings and three hierarchical levels of European grayling populations from the AMOVA analysis. Populations within groups A, B and C (geographical areas) are described in Table 2 and definitions of the Φ statistics in the text. The results were obtained with the ARLEQUIN computer package (Schneider *et al*. 1997) following the approach of Excoffier *et al*. (1992)

Partitioning of populations	Among groups			Within groups			Within populations		
	%	P	Φ_{CT}	%	P	Φ_{SC}	%	P	Φ_{ST}
A/B/C	59	<0.001	0.59	36	<0.001	0.88	5	<0.001	0.95
A + B/C	51	0.003	0.51	45	<0.001	0.92	4	<0.001	0.96

population sampled from the Sjamzhenga River, a tributary of the river Dvina (Fig. 1). Lineage II haplotypes (haplotypes 17–26; Table 2) occurred in central Germany, Poland and throughout western Fennoscandia (Fig. 1). Finally, the lineage III haplotypes (Table 2) were found in the Slovenian population sampled from a Danubian tributary and the Ramsach River in southern Germany. Haplotype 14 was displayed only by individuals sampled from the UK drainages. The genetically very diverged *T. thymallus* haplotype 27 occurred in the easternmost sample, the Russian Sjamzhenga population (Fig. 1).

Only three sampling locations were observed to harbour haplotypes from more than one major lineage revealed by the phylogenetic analyses: in addition to the lineage II haplotypes, populations from Tandsjön and Gudbrandsdalslågen in southern Sweden and southern Norway displayed the lineage I haplotype 1 with frequencies of 0.04 and 0.24, respectively. Furthermore, all 24 fish from the lake Lesjaskogsvatn (southern Norway) were found to have the lineage I haplotype 1. Based on the occurence of mtDNA haplotypes, the three geographical areas harbouring the distinct European grayling lineages are referred to as the geographical areas A, B and C. Table 2 indicates occurrence of the populations within these areas.

mtDNA RFLP variation within and among populations

Excluding populations where less than 10 individuals were analysed, only one population of 23 displayed more than two composite haplotypes (Table 2). Accordingly, calculations of haplotype and nucleotide diversities revealed low European grayling intrapopulation variation. Populations which had more than one RFLP haplotype (a total of 17 in the *T. thymallus* data set) exhibited haplotype diversities between 0.07 and 0.66 (mean 0.22 ± 0.002) and nucleotide diversities between 0.20×10^{-3} and 12.71×10^{-3} (mean $2.20 \times 10^{-3} \pm 0.0004 \times 10^{-3}$). Nucleotide divergence values (da; program DA of REAP) observed between European grayling populations were from < 0.0001–0.0529 with a mean estimate among all population pairs of 0.0173.

Hierarchical AMOVA analysis of the mtDNA PCR–RFLP inferred site data also revealed that the great majority of the observed variation resided among populations. When the investigated populations were pooled into eastern/north-eastern, central/north-western and southern groups (thereby corresponding to geographical areas A, B and C; see Table 2 for populations within these areas), as suggested by the general colonization routes and suture zones across diverse taxa (Taberlet *et al*. 1998; Hewitt 1999), 59% of the total variation was explained by the among-group component, while 36% and 5% resulted from the within-group and within-population components, respectively (Table 5). Pooling sampling sites into northern and southern groups, as implied by earlier refugia findings across European freshwater fishes (Bernatchez *et al*. 1992; Bernatchez & Dodson 1994; Bernatchez & Osinov 1995; Durand *et al*. 1999; García-Marín *et al*. 1999; Nesbø *et al*. 1999), resulted in a slightly smaller proportion of the total variation (51%) to be due to the among-group component. Nevertheless, the among-group component still explained a substantially larger proportion of the total variation than did the within-group or within-population components (Table 5). With all of the above arrangements, the variance components were statistically highly significant ($P \leq 0.003$).

Discussion

Divergence of grayling mtDNA

PCR–RFLP analysis resolved 27 *Thymallus thymallus* composite haplotypes defined by an average pairwise evolutionary difference (d of REAP) of 2.6%. Roughly in line with this estimate, sequencing of 529 bp of the ND 5 gene from individuals representative for the RFLP haplotypes indicated an average divergence of 2.4% between the *T. thymallus* haplotypes. The European grayling mtDNA haplotypes clustered into three lineages supported by relatively high to high bootstrap percentages (Fig. 2). Based on the PCR–RFLP data, the three lineages were separated by pairwise differences from 1.2% to 5.8%

(Table 4a; Fig. 3). Sequence data of the ND 5 gene suggested minimum divergence of 0.8% between the three lineages (Table 4b). In fish studies, an mtDNA divergence rate of 2% per million years (Myr) is commonly employed to date divergence times (e.g. Becker *et al*. 1988; Bernatchez & Dodson 1991; Bernatchez *et al*. 1992). If we direct this rate to the minimum divergence between *T. thymallus* haplotypes of our study, the three European lineages could have separated approximately 400 000 years ago. This indeed must be considered a conservative time estimate, because a low calibration of 1% sequence divergence per Myr has been suggested (Smith 1992). In any case, the divergence event clearly predates the late Pleistocene period (130 000–10 000 years) implying that *T. thymallus* has survived dramatic temperature oscillations in distinct refugia. Importantly, this time estimate refutes the view, based on the present distribution of *T. thymallus*, of solely post-Pleistocene divergence of lineages (Seppovaara 1982). It also roughly parallels with the separation times reported for differentiated *T. arcticus* mtDNA haplotypes (Redenbach & Taylor 1999) and divergence estimates observed among the phylogenetic assemblages within the European brown trout species complex (Bernatchez *et al*. 1992). Furthermore, the result agrees with fossil evidence, of *T. thymallus* inhabiting Europe long before the late Pleistocene period (Cumbaa *et al*. 1981; Banarescu 1990).

Potential interspecific hybridization/introgression

It was notable that haplotype 27, discovered in two specimens (out of 13) from the easternmost population in Sjamzhenga River, was much more diverged from any other *T. thymallus* haplotype than the *T. arcticus* haplotype (haplotype 27 displayed an average sequence divergence of over 6.5% from any other *T. thymallus* haplotype but only 1.8% difference from the *T. arcticus* haplotype; Table 4b). It is possible that the individuals exhibiting haplotype 27 represent historical *T. thymallus*/*T. arcticus* hybrids which have subsequently back-crossed with *T. thymallus* males. Alternatively, haplotype 27 could represent ancestral polymorphism shared between the two species. Although our results may suggest that the hybridization/introgression zone of *T. arcticus* and *T. thymallus* potentially extends much further west than earlier considered (the two species are reported to hybridize only in the Ural Mountains; Shubin & Zakharov 1984), molecular investigation, including assessment of nuclear loci, for a dense sample from the Archangelsk region and east of the Urals needs to be conducted to conclusively resolve this issue.

Division of populations in northern Europe

Two very distinct mtDNA haplotype lineages were revealed in North European *T. thymallus* populations which is in contrast to data from some other salmonids, e.g. the extremely diverse brown trout species complex where a single colonization source into northern Europe has been proposed (Bernatchez *et al*. 1992; Bernatchez & Osinov 1995). mtDNA evolutionary divergence between the two grayling lineages (RFLP average 3.1%, ND 5 sequence average 1.1%; Table 4) is equivalent for instance to those observed between trans-Atlantic haplotypes in Atlantic salmon (Verspoor *et al*. 1999; Nilsson *et al*. submitted). Such deep divergence of North European grayling populations may stem from: (i) early division and low level of introgression between the lineages; and/or (ii) low effective population size and high related genetic drift in *T. thymallus* populations contributing to the postglacial colonization (and present-day genetic diversity) of this species.

Comparison of population based mtDNA PCR–RFLP divergence estimates (da of REAP) of the present data and a recent study of Nearctic Arctic grayling (Redenbach & Taylor 1999) suggested a marked difference in the level of intraspecific genetic differentiation of *T. thymallus* and *T. arcticus*; North European grayling populations (above 50° latitude) exhibited, on average, over 10-fold higher nucleotide divergence compared to that between drainage basins of North American Arctic grayling (Redenbach & Taylor 1999). This was somewhat surprising given that the same mtDNA segments were assayed in both studies and that the two species are thought to display very similar life history strategies and dispersal abilities (reviewed in Northcote 1995), with even a capability to hybridize (Shubin & Zakharov 1984). One potential source of bias to this comparison is the fact that only six of the eight restriction endonucleases employed by Redenbach & Taylor (1999) were assessed in this work. Also, in the present paper, only polymorphic enzymes were selected for each fragment, whereas Redenbach & Taylor (1999) restricted each mitochondrial area with all selected enzymes, a number of which were monomorphic. However, re-analysis of the *T. thymallus* data with only the amplicon and restriction enzymes shared with the study of Redenbach & Taylor (1999) yielded a similar result (data not shown). Furthermore, by conservatively assuming that all additional fragment-enzyme combinations studied by Redenbach & Taylor (1999) were monomorphic across all *T. thymallus* populations (541 individuals), the average divergence between European grayling populations was still over eight times higher than that among Arctic grayling. The inherent difference in interpopulation differentiation between these two closely related species may reflect lower dispersal/migration between grayling populations in northern Europe, perhaps connected with the distinctiveness of *T. thymallus* refugia during Pleistocene interglacials. The lower divergence of Arctic grayling in North America may also be due to the possible survival and dispersal of most of the abundant populations from only a single

major Wisconsinan glacial refugium, as suggested by Redenbach & Taylor (1999). The above comparison, however, is valid only for the North American populations of *T. arcticus*. Indeed, the relatively deep divergence among Siberian *T. arcticus* and North American *T. arcticus* haplotypes, resolved by Redenbach & Taylor (1999), may indicate substantial diversity across the entire geographical range of Arctic grayling as well.

Colonization of T. thymallus *into northern Europe*

The profound genetic divergence of *T. thymallus* haplotype lineages and their strikingly distinct geographical occurrence, together with the dense sampling strategy across northern Europe may provide a robust insight into the historical colonization routes of the populations. It has been suggested that Mongolia is the original source of all grayling diversity (e.g. Schöffmann 2000). Assuming this origin, and taking into account the present distribution of European grayling, it is possible that separation of the observed mtDNA lineages occurred close to the Ural Mountains, perhaps near the Caspian basin. Consequently, one refugium probably persisted north of that, contributing to the colonization of the Archangelsk area and subsequently the remainder of north-western Russia, Finland and Estonia, as shown by the predominance of lineage I haplotypes in these regions (Table 2; Fig. 1). Westward movement from Archangelsk would have been feasible starting from the late Weichselian, ≈ 13 000 BP, by the edge of the retreating ice sheet, at the time extending to the southern parts of the White Sea (Kvasov 1979; Rainio 1991; Svendsen *et al*. 1999), or during the Baltic Ice-lake stage, ≈ 10 500 BP, when most of the present White Sea was freshwater (Donner 1995; Svendsen *et al*. 1999). These results provide further support to a highly probable multispecies refugium area located east of the ice sheets (Osinov & Bernatchez 1996; Nesbø *et al*. 1999) at the height of the last Pleistocene event, about 18 000 BP.

A second grayling refugium possibly resided in the central area of Europe and acted as a source for the colonization of central Germany, Denmark, Poland, Sweden and southern and central parts of Norway. Individuals bearing the central German/Danish haplotype (haplotype 17) were separated recently from the other representatives of lineage II, as evidenced by the shallow divergence between haplotype 17 and haplotypes 18–26 (Fig. 3), and probably utilized the Elbe River system for their dispersal (see Hansen *et al*. 1999). This was followed by a northward expansion of grayling, possibly beginning from the onset of the Baltic Ice-lake stage, approximately 12 000 BP (Donner 1995). In southern Norway, grayling are naturally distributed only in the Glomma River system. Shortly after the major ice sheets started retreating (≈ 10 200 BP, Donner 1995), lake Vänern (southern Sweden) was connected to Glomma (Jonsson 1995) and we hypothesize dispersal of grayling via this connection. Alternative routes, e.g. utilizing the Baltic Sea or the Atlantic Ocean are less likely due to grayling's intolerance to the related salt concentrations (*T. thymallus* tolerates only under 4‰ salinity levels; Nykänen & Huusko 1999). Indeed, colonization of freshwater fishes into Norway via the Swedish watercourses has been suggested based on the distribution of species (Huitfeldt-Kaas 1918; Jonsson 1995) and molecular data (Refseth *et al*. 1998; Nesbø *et al*. 1999). Further dispersal towards northern Sweden (populations Tandsjön, Vindelälven) would have been possible particularly during the Ancylus lake stage, ≈ 9000 BP, as isostatic land upheaval had obstructed the connection between the Baltic basin and the Atlantic Ocean, and the melting ice had created a nonsaline waterbody covering the entire Baltic Sea area (e.g. Donner 1995). Also the Norwegian Vefsna River was perhaps colonized during this period. Alternatively, grayling in this river may have been artificially introduced, e.g. by the Lapps.

Colonization of the 4400 km^2 greater Lake Saimaa (south-eastern Finland) has puzzled zoogeographists for a long time, and it is currently argued that dispersal from the west has contributed significantly to the genetic diversity of salmonid species found in this lake system, such as Atlantic salmon (Koljonen *et al*. 1999). However, identical *T. thymallus* mtDNA composite haplotypes are found in two sites at Lake Saimaa and also at a high frequency (0.96; Table 2) in the Varzuga River (population Var; Fig. 1) in easternmost Fennoscandia. Furthermore, composite haplotypes 12 (found only in Varzuga River) and 11 (found only in Lake Saimaa) are apparently closely related (Table 2; Figs 2, 3). As our data provide no evidence for the eastward dispersal of *T. thymallus* within the Baltic Sea area, we suggest that dispersal from the east has contributed significantly to the colonization of Lake Saimaa grayling, perhaps via the possible connection between Lake Saimaa and the present White Sea during the Baltic Ice-lake stage, ≈ 10 200 BP (reviewed in Rainio 1991). Also in line with this hypothesis, are recent DNA-based results on Atlantic salmon (Nilsson *et al*. submitted).

The geographically very distinct haplotype pattern of *T. thymallus* clearly demonstrates a two-way colonization event into the Baltic Sea area. Similarly for perch, Refseth *et al*. (1998) and Nesbø *et al*. (1999) suggested two distinct colonizations, one from the east and one from the south. The suggestion of three refugia contributing for the dispersal of perch into the entire North European region (two refugia for the north-eastern area; Nesbø *et al*. 1999) is, however, not supported for *T. thymallus*. Instead, our results imply only two significant dispersal routes of grayling into the North European region, including the eastern Fennoscandia and Archangelsk.

The Danubian water system has been suggested to be an important contributor for postglacial colonization of

many fish species (Bernatchez *et al*. 1992; Bernatchez & Osinov 1995; Durand *et al*. 1999; Nesbø *et al*. 1999) and it has also been considered a potential source for dispersal of *T. thymallus* populations (e.g. Gardiner 2000). Indeed, in our analyses the populations in southern Germany were fixed for haplotypes related with the Danubian Obrh population haplotype (Table 2; Fig. 1) and therefore likely to be descended from a Danubian refugium. However, all North European haplotypes were highly differentiated from the Obrh haplotype (Table 2; Figs 2, 3), indicating that the Danubian area did not contribute markedly to the recolonization of grayling into northern glaciated regions. This conclusion is congruent with the data on brown trout (Bernatchez *et al*. 1992) and perch (Nesbø *et al*. 1999). Phylogenetic placement of the haplotype 14, observed only in the UK sample, was ambiguous (Figs 2, 3) and the origins of the UK grayling cannot be reliably speculated based on the present data. Our analyses indicate however, that northern and southern parts of Europe harbour differentiated and geographically distinct grayling haplotypes, and consequently dense sampling of *T. thymallus* from its westernmost range should aid in resolving these issues.

The putative refugia and colonization routes presented for European grayling parallel with many of those reported in recent comparative phylogeography surveys by Taberlet *et al*. (1998) and Hewitt (1999), neither of which included fish data. The current results further indicate the hypothesized contact zones of the differentiated *T. thymallus* lineages I/II and II/III to be situated similarly with the main northern and southern suture zones identified across diverse taxa (Taberlet *et al*. 1998; Hewitt 1999). A clear east–west genetic split occurs between populations of European grayling around 20° longitude (Fig. 1). A similar division seems evident also for chub (Durand *et al*. 1999), but has previously been demonstrated only for the central and southern populations of European fishes (Durand *et al*. 1999). Our data strongly indicate that in *T. thymallus* (and potentially in many other freshwater fishes) this split extends as north as 66° latitude and, within Fennoscandia, is surprisingly well in line with data from some mammals, e.g. the brown bear, *Ursus arctos* (Taberlet *et al*. 1995) and the field vole, *Migrotus agrestis* (reviewed in Jaarola *et al*. 1999).

Maternal geneflow between geographical areas in northern Europe

In contrast to the results of many earlier studies of fishes, particularly those of anadromous salmonids (Verspoor *et al*. 1999; Nilsson *et al*. submitted), maternal geneflow between the *T. thymallus* lineages within the North European region was very limited with only four of the 24 sampled locations providing evidence of dispersal across geographical areas A and B (Table 2; Fig. 1). Geographic distinctiveness of haplotype distribution was disrupted in central Sweden and southern Norway (Table 2; Fig. 1). Again, the Ancylus lake stage is envisaged as the potential period for dispersal of individuals displaying the haplotype 1 into coastal streams along the Swedish coast. Moreover, as haplotype 1 occurs in southern Norway, these individuals (in addition to those colonizing from the south) could have also dispersed via the Vänern-Glomma connection during the short time period after the Baltic Ice-lake stage. Eventually, Glomma lost its connection with Vänern and joined with southern parts of Gudbrandsdalslågen, thereby enabling colonization of grayling. The fixation of haplotype 1 in the Lesjaskogsvatn population could be explained by a random founder event, perhaps resulting from historical local stocking; Lesjaskogsvatn is located in the upper regions of Gudbrandsdalslågen, separated from the lower regions by several waterfalls hampering upward migration of grayling. Although not documented, it is possible that dispersal has been induced by humans.

Interestingly, the anadromous Holmön population on the north-eastern coast of Sweden exhibited solely haplotypes found in area A (Table 2; Fig. 1). The sea area, from where anadromous populations included in this study arise, contains much lower salt concentrations (≤4‰) than southernmost parts of the Baltic (Auer *et al*. 1960) and it is certainly feasible that geneflow occurs between the Oura (Our)/Ulkokrunnit (Kru)/Holmön (Hol) grayling (Fig. 1), across the brackish Bothnian Bay.

Structuring of population genetic diversity: implications for conservation of T. thymallus

In general, estimates of within-population haplotype and nucleotide diversity were low. This is not surprising given that the focus of this study was on populations from the North European region: low levels of genetic diversities within populations resident in formerly glaciated areas are often observed due to their displacement by glacial advances and subsequently reducing effective numbers of individuals (e.g. Hewitt 1996, 1999; Bernatchez & Wilson 1998). In contrast, the level of mtDNA variation among *T. thymallus* populations and, particularly, between the defined geographical groupings was high, regardless of the bases for pooling the populations (Table 5).

The accumulation of molecular variance among grayling populations likely stems from historical colonization events (see above). However, in many areas covered by the present work, postcolonization mixing of the genetically differentiated populations should have been possible. For instance in the Lake Saimaa water system relatively differentiated mtDNA haplotypes dominated in the two respective populations analysed (Table 2; Fig. 3). It is therefore possible that the maintenance of high

interpopulation level divergence in *T. thymallus* is due to biological characteristics of the species, e.g. poor dispersal ability or very strong behaviour for homing. The latter alternative seems more likely, as number of tagging studies have revealed moderately distant trophic (up to 120 km in 21 days; Andersen 1968) and reproductive migrations (e.g. Woolland 1986; reviewed in Nykänen & Huusko 1999) of *T. thymallus*. Such highly site fidelic and reproductively isolated populations may be strongly adapted to their local habitats in a number of quantitative traits that have a marked effect to the fitness of individuals (Allendorf & Waples 1996). Therefore, in the development of conservation strategies for European grayling it is potentially very important to recognize the genetically different populations as distinct management units.

It is commonly stated that populations with independent evolutionary histories, therefore contributing most to the overall intraspecific genetic diversity, should be prioritized for conservation (e.g. Moritz 1994; Bernatchez & Wilson 1998). While criteria for defining such populations are still debated (e.g. Ryder 1986; Moritz 1994; Paetkau 1999), assessment of the genetic distinctiveness of populations has often been more straightforward, particularly in many fish species (e.g. Bernatchez & Dodson 1991). Similarly, mtDNA analysis of European grayling populations provided two evolutionarily distinct North European lineages (Figs 2, 3), which deserve recognition as independent groups. Moreover, our analyses suggest that central and southern Europe harbour additional genetically differentiated lineages (Figs 1, 2, 3). In addition to the clear genetic division of *T. thymallus* over a broad geographical scale, it is also important to note the occurrence of variation within lineages. For instance, within lineage I haplotypes 8, 10, 11 and 12 (found in the Isojoki and Lieksanjoki populations) formed their own statistically reliable subgroup (Fig. 2) which was moderately differentiated from the remaining haplotypes in this lineage (Fig. 3). Similarly, within lineages II and III several haplotypes, fixed for some populations, were placed into separate groups with substantial bootstrap support (Fig. 2).

Observed discrepancies between mitochondrial and nuclear DNA diversity within species and populations (e.g. Palumbi & Baker 1994) indicate that mtDNA-based conclusions on within-group or within-population variability should be complemented by nuclear marker data. Also recent results on salmonids support this view (Brunner *et al*. 1998). Therefore, analysis of genetic data from nuclear DNA markers, such as microsatellites, may prove useful for developing conservation guidelines for *T. thymallus* on a finer scale.

Conservation strategies are normally defined from a national perspective. This study poses a challenge to this practise as, for example, the Holmön grayling population on the north-eastern coast of Sweden and the Lesjaskogsvatn drainage in southern Norway were observed to have only individuals with mitochondrial haplotypes dominantly found in Finland and Russia. Similarly in Germany, mtDNA haplotype fixed in the Eger River population from the central region was very closely related to the haplotypes found in Poland, Denmark, Sweden and Norway (Figs 1, 3), but extremely diverged from both of the haplotypes observed in the southern part of the country (Ramsach population). Indeed, countries with native populations of European grayling clearly are required to work together to preserve the genetic diversity and to ensure the long-term sustainability of this culturally very important salmonid species. This paper provides a good example for comparing political vs. evolutionary significant units, and drives home the need for multinational co-operation for managing and conserving biodiversity.

Acknowledgements

This study was made possible due to generosity of numerous persons providing *Thymallus thymallus* samples. We especially acknowledge Teija Aho and Jaakko Lumme (Finland), Michael Hansen (Denmark), Kjetil Hindar (Norway), Ross Gardiner (UK), Simona Sušnik and Aleš Snoj (Slovenia), Riho Gross (Estonia), Bernhard Gum and Ralph Kühn (Germany) and Pawel Brzuzan (Poland). We also thank Eric Taylor and Mike Stamford for the North American *T. arcticus* DNA and Zoë Redenbach for the detailed information on restriction fragment patterns in *T. arcticus*. Jaakko Lumme, Louis Bernatchez and two anonymous referees are acknowledged for their helpful comments on an earlier version of the manuscript. One anonymous referee coined the phrase 'political vs. evolutionary significant units'. The work was supported in part by grants from the University of Helsinki, the Finnish Game and Fisheries Research Institute and the Ministry of Agriculture and Forestry.

References

Allendorf FW, Waples RS (1996) Conservation and genetics of salmonid fishes. In: *Conservation Genetics: Case Histories from Nature* (eds Avise JC, Hamrick JL), pp. 238–280. Chapman & Hall, New York.

Andersen C (1968) *Vandring hos harr i Trysilvassdraget belyst ved merkingsforsøk*. MSc Thesis. University of Oslo, Norway (in Norwegian).

Auer V, Ajo R, Jutikkala E *et al*. (1960) *Suomen Kartasto* (ed. Aario L). Otava, Helsinki (in Finnish).

Avise JC, Arnold J, Ball RM *et al*. (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

Avise JC (1992) Molecular population structure and biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos*, **63**, 62–76.

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

Banarescu P (1990) General distribution and dispersal of freshwater animals. In: *Zoogeography of Freshwaters 1* (ed. Banarescu P). AULA-Verlag, Wiesbaden.

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

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

Bernatchez L, Guyomard R, Bonhomme F (1992) DNA sequence variation of the mitochondrial control region among geographically and morphologically remote European brown trout *Salmo trutta* populations. *Molecular Ecology*, **1**, 161–173.

Bernatchez L, Danzmann RG (1993) Congruence in control-region sequence and restriction-site variation in mitochondrial DNA of Brook charr (*Salvelinus fontinalis* Mitchill). *Molecular Biology and Evolution*, **10**, 1002–1014.

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

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

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

Bouvet Y, Soewardi K, Pattee E (1990) Genetic divergence within natural populations of grayling (*Thymallus thymallus*) from two French river systems. *Archiv Fur Hydrobiologie*, **119**, 89–101.

Brunner PC, Douglas MR, Bernatchez L (1998) Microsatellite and mitochondrial DNA assessment of population structure and stocking effects in Arctic charr *Salvelinus alpinus* (Teleostei: Salmonidae) from central alpine lakes. *Molecular Ecology*, **7**, 209–223.

Cronin MA, Spearman WJ, Wilmot RL (1993) Mitochondrial variation in Chinook (*Oncorhynchus tsawytscha*) and Chum salmon (*O. keta*) detected by restriction enzyme analysis of Polymerase Chain Reaction (PCR) products. *Canadian Journal of Fisheries and Aquatic Sciences*, **50**, 708–715.

Cumbaa SL, McAllister DE, Morlan RE (1981) Late Pleistocene fish fossils of *Coregonus*, *Stenodus*, *Thymallus*, *Catostomus*, *Lota* and *Cottus* from the Old Crow Basin, northern Yukon, Canada. *Canadian Journal of Earth Sciences*, **18**, 1740–1754.

Donner J (1995) *The Quaternary History of Scandinavia*. Cambridge University Press, New York.

Durand JD, Persat H, Bouvet Y (1999) Phylogeography and postglacial dispersion of the chub (*Leuciscus cephalus*) in Europe. *Molecular Ecology*, **8**, 989–997.

Estoup A, Largiadèr CR, Perrot E, Chourrout D (1996) Rapid one-tube DNA extraction for reliable PCR detection of fish polymorphic markers and transgenes. *Molecular Marine Biology and Biotechnology*, **5**, 295–298.

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

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

Fitch WM, Margoliash E (1967) Construction of phylogenetic trees. *Science*, **155**, 279–284.

García-Marín J-L, Utter FM, Pla C (1999) Postglacial colonization of brown trout in Europe based on distribution of allozyme variants. *Heredity*, **82**, 46–56.

Gardiner R (2000) The origins and present distribution of grayling. In: *The Complete Book on Grayling* (ed. Broughton R). Robert Hale Publishers, London (in press).

Hansen MM, Mensberg K-LD (1998) Genetic differentiation and relationship between genetic and geographic distance in Danish sea trout (*Salmo trutta* L.) populations. *Heredity*, **81**, 493–504.

Hansen MM, Mensberg K-LD, Berg S (1999) Postglacial recolonization patterns and genetic relationships among whitefish (*Coregonus* sp.) populations in Denmark, inferred from mitochondrial DNA and microsatellite markers. *Molecular Ecology*, **8**, 239–252.

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 GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Huitfeldt-Kaas H (1918) *Ferskvandfiskenes utbredelse og indvandring i Norge, med et tillaeg om krebsen*. pp. 16–27 and pp. 52–57. Centraltrykkeriet-Kristiania, Oslo, Norway (in Norwegian).

Jaarola M, Tegelström H, Fredga K (1999) Colonization history in Fennoscandian rodents. *Biological Journal of the Linnean Society*, **68**, 113–127.

Jonsson L (1995) Vertebrate fauna during the Mosolithic on the Swedish west coast. In: *Man and the Sea in the Mesolithic. Coastal Settlement Above and Below Present Sea Level* (ed. Fischer A), pp. 147–160. The Short Run Press, Great Britain.

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

Koljonen M-L, Jansson H, Paaver T, Vasin O, Koskiniemi J (1999) Phylogeographic lineages and differentiation pattern of Atlantic salmon (*Salmo salar*) in the Baltic Sea with management implications. *Canadian Journal of Fisheries and Aquatic Sciences*, **56**, 1766–1780.

Kvasov DD (1979) The Late-Quaternary history of large lakes and inland seas of Eastern Europe. *Annales Academiae Scientiarum Fennicae* A III **127**.

Lelek A (1984) Threatened fishes of Europe. In: *The Freshwater Fishes of Europe 9* (eds European Committee for the Conservation of Nature, Natural Resources-Council of Europe), pp. 93–96. AULA-Verlag, Wiesbaden.

Maitland PS, Campbell RN (1992) *Freshwater Fishes of the British Isles* (eds Walters M, West R, Streeter D, Corbet SA). Butler & Tanner, Somerset.

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.

Moritz C, Dowling TE, Brown WM (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics*, **18**, 269–292.

Moritz C (1994) Defining 'evolutionary significant units' for conservation. *Trends in Ecology and Evolution*, **9**, 373–375.

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

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

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

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.

Nesbø CL, Fossheim T, Vøllestad LA, Jakobsen KS (1999) Genetic divergence and phylogeographic relationships among European perch (*Perca fluviatilis*) populations reflect glacial refugia and postglacial colonization. *Molecular Ecology*, **8**, 1387–1404.

Nielsen EE, Hansen MM, Loeschcke V (1996) Genetic structure of European populations of *Salmo salar* L. (Atlantic salmon) inferred from mitochondrial DNA. *Heredity*, **77**, 351–358.

Nielsen EE, Hansen MM, Mensberg K-LD (1998) Improved primer sequences for the mitochondrial ND1, ND3/4 and ND5/6 segments in salmonid fishes: application to RFLP analysis of Atlantic salmon. *Journal of Fish Biology*, **53**, 216–220.

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

Nykänen M, Huusko A (1999) Harjuksen elinympäristövaatimukset virtavesissä. *Riista-ja kalatalouden tutkimuslaitos, kalatutkimuksia*, **156**. Helsinki (in Finnish).

Nylander E (1998) *Finnish Fisheries. Facts and Statistics*. Oy Edita Ab, Helsinki.

Osinov AG, Bernatchez L (1996) 'Atlantic' and 'Danubian' phylogenetic groupings of brown trout *Salmo trutta* complex: Genetic divergence, evolution and conservation. *Journal of Ichthyology*, **36**, 723–746.

Paetkau D (1999) Using genetics to identify intraspecific conservation units: a critique of current methods. *Conservation Biology*, **13**, 1507–1509.

Palumbi SR, Baker CS (1994) EPIC amplifications of nuclear DNA introns: Opposing views of population structure using mitohondrial and nuclear sequences from humpback whales. *Molecular Biology and Evolution*, **11**, 426–435.

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

Rainio H (1991) The younger dryas ice-marginal formations of southern Finland. In: *Eastern Fennoscandian Younger Dryas End Moraines* (eds Rainio H, Saarnisto M), pp. 25–72. Geological Survey of Finland, Espoo.

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

Refseth UH, Nesbø CL, Stacy JE *et al.* (1998) Genetic evidence for different migration routes of freshwater fish into Norway revealed by analysis of current perch (*Perca fluviatilis*) populations in Scandinavia. *Molecular Ecology*, **7**, 1015–1027.

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution*, **1**, 9–10.

Schneider S, Kueffer J-M, Roessli D, Excoffier L (1997) *ARLEQUIN, Version 1.1*. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva, Switzerland.

Schöffmann J (2000) The grayling species (Thymallinae) of three different catchment areas of Mongolia. *Journal of the Grayling Society*, **Spring**, 41–44.

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

Seppovaara O (1982) Harjuksen (*Thymallus thymallus* L.) levinneisyys, biologia, kalastus ja hoitotoimet Suomessa. *Riista-ja kalatalouden tutkimuslaitos, kalantutkimusosasto, monistettuja julkaisuja*, **5**. Helsinki (in Finnish).

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.

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

Svendsen JI, Astakhov VI, Bolshiyanov DYU *et al.* (1999) Maximum extent of the Eurasian ice sheets in the Barents and Kara Sea region during the Weichselian. *Boreas*, **28**, 234–242.

Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996) Phylogenetic inference. In: *Molecular Systematics*, 2nd edn (eds Hillis DM, Moritz C, Mable B), pp. 407–514. Sinauer Associates, Inc., Sunderland.

Taberlet P, Swenson JE, Sandegren F, Bjärvall A (1995) Localization of a contact zone between two highly divergent mitochondrial DNA lineages of the brown bear (*Ursus arctos*) in Scandinavia. *Conservation Biology*, **9**, 1255–1261.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Taylor EB, Pollard S, Louie D (1999) Mitochondrial DNA variation in bull trout (*Salvelinus confluentus*) from northwestern North America: implications for zoogeography and conservation. *Molecular Ecology*, **8**, 1155–1170.

Verspoor E, McCarthy EM, Knox D *et al.* (1999) The phylogeography of European Atlantic salmon (*Salmo salar* L.) based on RFLP analysis of the ND1/16sRNA region of the mtDNA. *Biological Journal of the Linnean Society*, **68**, 129–146.

Wilson C, Hebert PDN, Reist JD, Dempson JB (1996) Phylogeography and postglacial dispersal of arctic charr *Salvelinus alpinus* in North America. *Molecular Ecology*, **5**, 187–197.

Woolland JV (1986) Grayling in the Welsh Dee. Part 2. Movements. *Journal of the Grayling Society*, **Autumn**, 35–37.

This investigation represents one in a series carried out at the IKP unit of the Department of Ecology and Systematics, University of Helsinki, with an aim of studying evolutionary and conservation genetics of salmonid fishes in the North European region. The work is part of MTK's PhD research being conducted under the supervision of CRP and stems from a joint effort between Finnish (MTK, ER, JP, CRP), Russian (AV, ST), Norwegian (TOH) and Swedish (JN, MC) collaborators. Many of these collaborators belong to the Baltic-White Sea-Barents Sea (BWB) salmonid consortium.
