Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 34 (2005) 106–117
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

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

E. Froufe[a,b], I. Knizhin[c], S. Weiss[d,*]

[a] *CIBIO/UP, Campus Agrário de Vairão, 4480-661, Vairão, Portugal*
[b] *Faculdade de Ciências, Universidade do Porto, Praça Gomes Teixeira, 4000-002 Porto, Portugal*
[c] *Irkutsk State University, Biological Faculty, ul K. Marksa 1, Irkutsk, 664003, Russia*
[d] *Institute of Zoology, Karl-Franzens University Graz, Universitätsplatz 2, A-8010 Graz, Austria*

Received 21 March 2004; revised 20 September 2004
Available online 5 November 2004

Abstract

We present first insights into the molecular phylogeny of the grayling genus *Thymallus* (Salmonidae) using sequences from the mitochondrial control region and ATPase6 genes. A suite of analytical approaches were applied for each gene separately and for the combined data. The ATPase6 gene is shown to have a mean divergence rate across the genus of 2.46 times faster than the complete control region. Based on the combined data, four major (internal) clades, presumably originating in the Pliocene, were resolved with high support in all analyses and represented two distinct lineages in the Amur basin, one lineage in all remaining Siberian and Mongolian drainages, and one lineage corresponding to European grayling *Thymallus thymallus*. The resolution of multiple lineages, from both additional internal and terminal clades, within each major drainage basin underscores the complexity and effects that Pleistocene hydrological dynamics have had on the distribution of biodiversity in Siberia.

Keywords: mtDNA; Control region; ATPase6; *Thymallus*; Phylogeny; Salmonidae

Introduction

Fishes of the family Salmonidae hold the highest economic, ecological, and evolutionary interest among all temperate freshwater fishes. Among them, the genus *Thymallus* (some authors list as a distinct family Thymallidae: e.g., Osinov and Lebedev, 2002; Skurikhina et al., 1985) has received increasing attention, especially in Eurasia, both through advances in basic research on natural selection (Koskinen et al., 2002a) and due to concerns of seriously declining populations (Persat, 1996; Weiss et al., 2001). Nonetheless, even the most basic question of how many species the genus contains, or the breadth of diversity within each species, remains unanswered. Depending on the source, four or five species of grayling are generally recognized to be living in the cold freshwaters of Europe, Asia and North America. European grayling, *Thymallus thymallus* are found from southern France (Loire basin) (Persast et al., 1978) east to the Balkans as far south as the Luča River in Montenegro (Jankovic, 1962). In the north, European grayling are found from Great Britain across most of Scandinavia east to the Urals (see Fig. 1 in Weiss et al., 2002). The polytypic Arctic grayling *Thymallus arcticus* are currently found from the Urals in the west, across all of northern Eurasia and into N. America as far east as the Hudson bay, with an extant disjunct population in Montana, USA, and extinct populations as far east as

* Corresponding author. Fax: +43 316 380 9875.
E-mail address: steven.weiss@uni-graz.at (S. Weiss).

1055-7903/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2004.09.009

Michigan (see further range descriptions in Scott and Crossman (1998) and Redenbach and Taylor (1999)). The Amur grayling, *Thymallus grubii*, are found throughout the Amur drainage and some adjacent rivers flowing into Okhotsk Sea and Sea of Japan (Berg, 1949; Reshetnikov et al., 2002; Tugarina and Khramtzova, 1980). Mongolian grayling *Thymallus brevirostris* are limited to lakes of the closed central Asian basin in Western Mongolia and border regions of Kazakhstan and the Tuva Republic, and a fifth taxon, *Thymallus nigrescens*, listed as a species in Scott and Crossman (1998) but as a subspecies *T. arcticus nigrescens* in Reshetnikov et al. (2002) is present only in Lake Chovsgul, Mongolia (Shatunovskiy, 1983; Svetovidov, 1936; Tugarina and Anudarin, 1972).

The systematic status of these species, or additional putative taxa has repeatedly come into question, with recent genetic data playing a central role. For European grayling, only one taxon has been recognized (Kottelat, 1997), however Sušnik et al. (2001) suggest that Adriatic populations represent distinct lineages worthy of species status and Weiss et al. (2002) reveal an equally divergent lineage in the Loire basin, France. For Arctic grayling, some N. American lineages were originally given a species status, such as *Thymallus signifier* or *Thymallus montanus* but current views have supported a single species (see Redenbach and Taylor, 1999; and references therein). In Asia, there is a long history of subspecific and even infrasubspecific designations for Arctic grayling, including distinct taxa in Lake Baikal, the Lena River, and northeastern Russia. Within the Amur basin, genetic and morphological data presented in Froufe et al. (2003a) and Knizhin et al. (2004) clearly support the presence of multiple taxa, including the putative "large-scale" taxon (Antonov, 2000). Mongolian grayling, was once thought to represent the primitive member of the genus (Shatunovskiy, 1983; Svetovidov, 1936; Tugarina and Anudarin, 1972) but Koskinen et al. (2002b) show that its position reflects a relatively recent evolutionary event within the *T. arcticus* complex. *T. nigrescens*, which is closely related to populations in Lake Baikal, also underscores the paraphyletic status of the *T. arcticus* complex. There has been no systematic or comprehensive attempt at describing the phylogenetic relationships of major lineages within the grayling genus. Thus, this paper presents the first insights into the molecular phylogeny of *Thymallus*, using the mitochondrial gene sequences ATPase6 (ATP6) and the complete control region (CR).

Materials and methods

Sampling

Grayling ($N = 35$) were collected by angling or nets in 1998–2003 (a small fin clip was preserved in 96% ethanol) from 16 primarily remote Asian populations, three locations in Europe and two in N. America, covering the distribution of the genus and representing the five noted systematic units, several subspecies, and distinct major basins (see Table 1). At least two individuals were sequenced for each putative taxon or major basin (except that there is only one individual from the Ob basin). Outgroup taxa included two sheefish (*Stenodus leucichthys*), one whitefish (*Coregonus lavaretus*) and two mountain whitefish (*Prosopium williamsoni*). Sample locations are also provided in Table 1.

Amplification and sequencing

Whole genomic DNA was extracted using a standard high-salt protocol (Sambrook et al., 1989). Two mtDNA genes, ATP6 and the CR, were amplified using the polymerase chain reaction (PCR). The ATP6 gene was amplified in all individuals using the L8558 and H9208 primers described in Giuffra et al. (1994). The CR (including a partial segment of both flanking tRNA genes) was amplified in 13 *Thymallus* individuals, using the LRBT-25 and LRBT-1195 primers described in Uiblein et al. (2001). The remaining CR sequences were taken from previously published research (Froufe et al., 2003a; Koskinen et al., 2002b; Weiss et al., 2002; GenBank Accession Nos. in Table 1).

PCR conditions (25 μl reactions) were as follows: each reaction contained 19 μl H_2O, 2.5 μl of 10× Promega Buffer B, 0.5 μl of 10 mM of each primer, 1.5 μl of 25 mM $MgCl_2$, 0.5 μl of 10 mM dNTP's, 0.1 μl Promega *Taq* DNA polymerase, and 0.5 μl of 100 ng/μl DNA template. The cycle parameters were as follows: initial denaturation at 94 °C for 3 min, denaturation at 94 °C (40 s), annealing at 53 °C for ATP6 and 55 °C for CR (40 s), and extension at 72 °C (40 s) repeated for 30 cycles. Amplified DNA templates were purified using the NucleoSpin Extract Kit (Machery-Nagel) and approximately 100 ng of purified PCR product was used in cycle sequencing reactions following ABI (Applied Biosystems) PRISM BigDye Terminator protocols. Sequences were visualized on either an ABI-3100 or ABI-310 genotyping apparatus.

Sequence alignment and analysis

The ATP6 sequences were aligned by eye based on the amino acid sequences using the standard mtDNA code for vertebrates in BioEdit program (Hall, 1999). This gene was selected to have a coding region that could be easily aligned to outgroup taxa, and following the presumption that it would have a lower substitution rate than the CR, based on general knowledge of fish mtDNA (Meyer, 1993) and comparative data for

Table 1
Sample locations including major river basins, population code with the number of individuals sampled shown in brackets, geographical coordinates, and GenBank accession numbers for *Thymallus* and the three outgroups used in this study: *Coregonus lavaretus*, *Stenodus leucichthys*, and *Prosopium williamsoni*

Population	Basin	Pop. code (*N*)	Lat. (N)	Long. (E)	Genbank accession numbers ATPase 6	GenBank accession numbers control region
Onon River	Shilka R. → Amur R. → Tatar Strait, Pacific Ocean	Ogb (1)	48° 75′	110° 25′	AY778972	AY168390
Sypchergurka River	Olengui R. → Ingoda R.→ Shilka R. → Amur R. → Tatar Strait, Pacific Ocean	Syp (1)	51° 20′	113° 26′	AY778973	AY168397
Bureya River	Amur R. → Tatar Strait, Pacific Ocean	Bur (4)	51° 55′	134 ° 53′	AY778974–AY778977	AY779007–AY779010
Botchi River	Tatar Strait, Pacific Ocean	Bot (2)	48° 05′	139° 21′	AY778978–AY778979	AY779011–AY779012
Buta River	Khutu R.→ Tumnin R. → Tatar Strait, Pacific Ocean	But (3)	49° 09′	138° 56′	AY778980–AY778982	AY246417–AY246419
Merek River	Amgun River → Amur → Tatar Strait, Pacific Ocean	Mer (2)	51° 17′	134° 47′	AY778983–AY778984	AY246410–AY246411
Taltzinka River	Irkutsk Reservoir → Angara→ Enisey	Irt (2)	51° 59′	104° 36′	AY778985–AY778986	AY168361; AY168363
Ushkaniy Islands	Lake Baikal (Central)	Uib (1)	53° 40′	108° 37′	AY778987	AY168370
South Fork Terkhyn R.	Selenga River → Lake Baikal	Sft (1)	47° 45′	99° 20′	AY778988	AY168377
Shishkhid River	Enisey River (upper)	Jb (1)	51° 30′	98° 30′	AY778989	AY168403
Lake Khökh Nuur	Central Asia	Kn (2)	47° 31′	98° 27′	AY778990–AY778991	AY168387; AY168389
Biya River	Ob' River	Byi (1)	51′47	87′14	AY778992	AY779013
Lena–Delta	Lena River–Laptev Sea	Del (2)	71′08	127′08	AY778993–AY778994	AY779015–AY779016
Sjamzhenga River	Dvina River → White Sea	Sja (1)	63° 70′	46° 26′	AY778995	AY779014
Pierce Lake	Yukon Territory, Pacific, North Am.	Pac (1)	—	—	AY778996	AY168401
Burnt River	British Columbia, Pacific, North Am.	Pac (1)	122° 20′	55° 25′	AY778997	AY168402
Anui River	Anui River → Amur–Tatar Strait, Pacific Ocean	Anu (2)	49° 17′	137° 55′	AY778998–AY778999	AY246404–AY246389
Lake Chovsgul	Selenga River → Lake Baikal	Chv (2)	51° 27′	100° 39′	AY779000–AY779001	AY168348; AY168349
Chitkanda Lake	Olekma River → Lena River	Chi (2)	57° 00′	119° 33′	AY779002–AY779003	AY779017
Soca River	Soca, Isonzo → Adriatic	Soc (1)	47° 50′	13° 20′	AY779004	AF522419
Loire River	Loire → Atlantic, France	Lob (1)	44° 51′	03° 55′	AY779006	AF522425
Drina River	Drina → Danube	Dri (1)	43° 24′	18° 49′	AY779005	AF522425
Coregonus		*Coregonus* (1)	—	—	AB034824	AB034824
Stenodus (Yukon River)		*Stenodus* (2)	—	—	AY778968–AY778969	—
Prosopium (Idaho, USA)	Clearwater Drainage	*Prosopium* (2)	—	—	AY778970–AY778971	—

several other salmonid genera (Churikov et al., 2001; Giuffra et al., 1994). The CR was chosen based on our previous work and the availability of GenBank sequences. The usefulness of the CR for the outgroup was examined using one complete *Coregonus* CR (GenBank Accession No. AB034824), whereby both the entire CR as well as several highly conserved regions alone were aligned with *Thymallus*. Sequences were aligned with Clustal W (Thompson et al., 1994) applying several different gap costs, and the alignment was chosen that produced the shortest maximum parsimony tree. Another approach to using the CR was followed by first identifying several of the most conserved sequence blocks (50–200 bp) in our data set, and adding these blocks to the ATP6 sequences in an attempt to use additional phylogenetic information.

Phylogenetic analysis

Sequences were imported into PAUP*4.0b10 (Swofford, 2002) for phylogenetic analyses as well as obtaining the observed pairwise sequence divergence (uncorrected *p* distances) and the number of transitions and transversions. Between-group variation (corrected for within-group variation) was calculated using the net nucleotide divergence (*Da*) in MEGA version 2.1 (Kumar et al., 2001). To assess the degree of saturation in each codon position of the ATP6 sequences, the number of transitions and transversions were plotted against the uncorrected pairwise distances, for first, second, and third positions. Base composition homogeneity was tested using a chi-square (χ^2) test for equal base frequencies across taxa, examining each coding position separately for the ATP6 data.

Maximum parsimony (MP), maximum-likelihood (ML), and Bayesian analysis were used for phylogenetic reconstruction with the ATP6 and CR data sets alone, and for both genes together. Modeltest 3.0 (Posada and Crandall, 1998) was used to choose the most likely of 56 models of nucleotide evolution for each data set alone and for the combined data set. The best-fit models estimated by Modeltest 3.0 were used to estimate a tree using ML and Bayesian inference. A heuristic search (10 replicates) was used to estimate the most likely topology for ML and MP methodologies. Heuristic searches started with stepwise addition trees and were replicated 10 times, with each replicate beginning with a random order of sequences. Branch swapping was performed by the tree-bisection-reconnection (TBR) method using default parameters. For MP analyses of the CR data, results were compared when gaps were treated as "missing data" and when coded as a fifth base. Bootstrap analysis (Felsenstein, 1981) was used to estimate support for the resulting topologies, with 1000 replicates, each with 100 random additions of sequences. Full heuristic search algorithms were applied for the MP and the "fast" stepwise addition method for the ML analysis.

The Bayesian analysis was implemented using MrBayes v3.0 (Huelsenback and Ronquist, 2001) which calculates Bayesian posterior probabilities using a Metropolis-coupled, Markov chain Monte Carlo (MC^3) sampling approach. Analyses were carried out assuming the optimal model determined by Modeltest. Chains were run for 1×10^6 generations, and sampled every 100. In both searches, stationarity of the Markov chain was determined as the point when sampled log likelihood values plotted against generation time reached a stable mean equilibrium value; "burn-in" data sampled from generations preceding this point were discarded. All data collected at stationarity were used to estimate posterior nodal probabilities and a summary phylogeny. Two independent replicates were conducted and inspected for consistency to check for local optima (Huelsenback and Bollback, 2001). We used the Shimodaira–Hasegawa test (SH-test) (Shimodaira and Hasegawa, 1999) as implemented in PAUP*4.0b10 to compare the topologies of trees within each data set.

Results

ATP6

The final alignment of the ATP6 gene yielded 661 bp (12 codons short of the complete gene) for 40 individuals. There was no significant difference in base frequencies across taxa and plots of the number of substitutions against uncorrected *p* distances revealed no saturation for transversions or transitions for both total positions and third codon positions alone (Figs. 1A and B). Pairwise sequence divergence within the ingroup ranged from 0 to 9.2%. A total of 200 variable sites were found of which 191 were parsimony informative and there were no stop codons. Within the ingroup the transition/transversion ratio was 3.8, and there were 3 amino acid changes (base pair position five: Ile-Val; 13: Leu-Ser-Thr; 14: Leu-Val). There were also 3 amino acid changes between the ingroup and outgroup (position 3: Leu-Met;

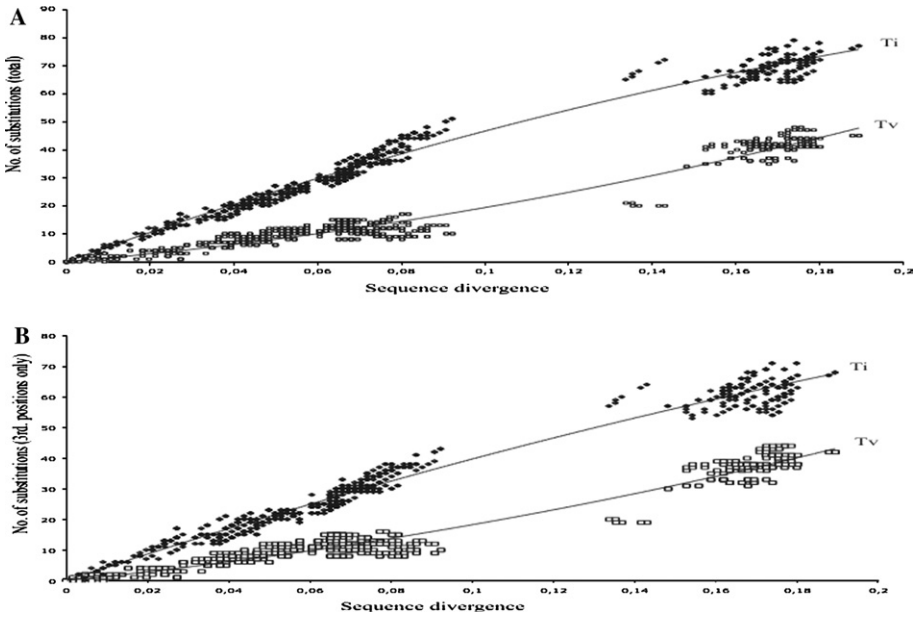

Fig. 1. The number of transitions and transversions plotted against uncorrected *p* distances. (A) Total number of substitutions; (B) third codon positions alone.

60: Leu-Met; 183: Ile-Leu) and one within the outgroup (position 173: Leu-Met). Considering these results, there is no reason to suspect nuclear copies or mtDNA pseudogenes.

The MP analysis resulted in 33 equally parsimonious trees of 400 steps (CI = 0.655; RI = 0.8752). Based on Modeltest, the HKY model with an estimate of invariable sites (0.568) and a discrete approximation of the gamma distribution (1.147) was chosen. Using this model, we obtained one most likely tree with ML analysis. Bayesian analysis produced a similar topology, which differed primarily in varying levels of node support. The strict consensus tree derived from MP gave nearly identical results for the well-supported clades obtained in ML, and likewise no further resolution in branch order (Fig. 2). The SH-test revealed no significant differences between tree topologies based on different phylogenetic methodologies. In all analyses the most internal nodes were short with little or no support. However, four major clades (most internal well-supported nodes) as well as a number of more terminal clades were identified with high-node support values (Fig. 2). These clades will be characterized below, considering as well the results from the CR.

The net divergence between the monophyletic ingroup and outgroup taxa ranged from 14.3% for *Prosopium* to 13.7% for one *Stenodus* sequence (Stenodus_2) and 14.0% for *Coregonus*. The other *Stenodus* sequence grouped with *Coregonus* with high-bootstrap support, differing by only 0.9%.

CR

Highly divergent sequence for the CR was noted between *Coregonus* and *Thymallus*, and the shortest MP tree obtained resulted in no additional internal nodes (with or without support) and thus the complete CR was determined not to be useful for the outgroup taxa with our aim of resolving relationships within the ingroup. The alignment for the CR across *Thymallus* included 1107 bp (including 87 bp of tRNA proline and 11 bp of tRNA phenylalanine genes). There were a total of 16 indels of 1–2 bp in length and the transition/transversion ratio was 2.0. Including indels there were a total of 32 unique sequences among the 34 individuals analyzed. Pairwise sequence divergence ranged from 0 to 5.7%. The MP analysis resulted in 18 equally parsimonious trees of 308 steps (CI = 0.6461; RI = 0.8584) when gaps

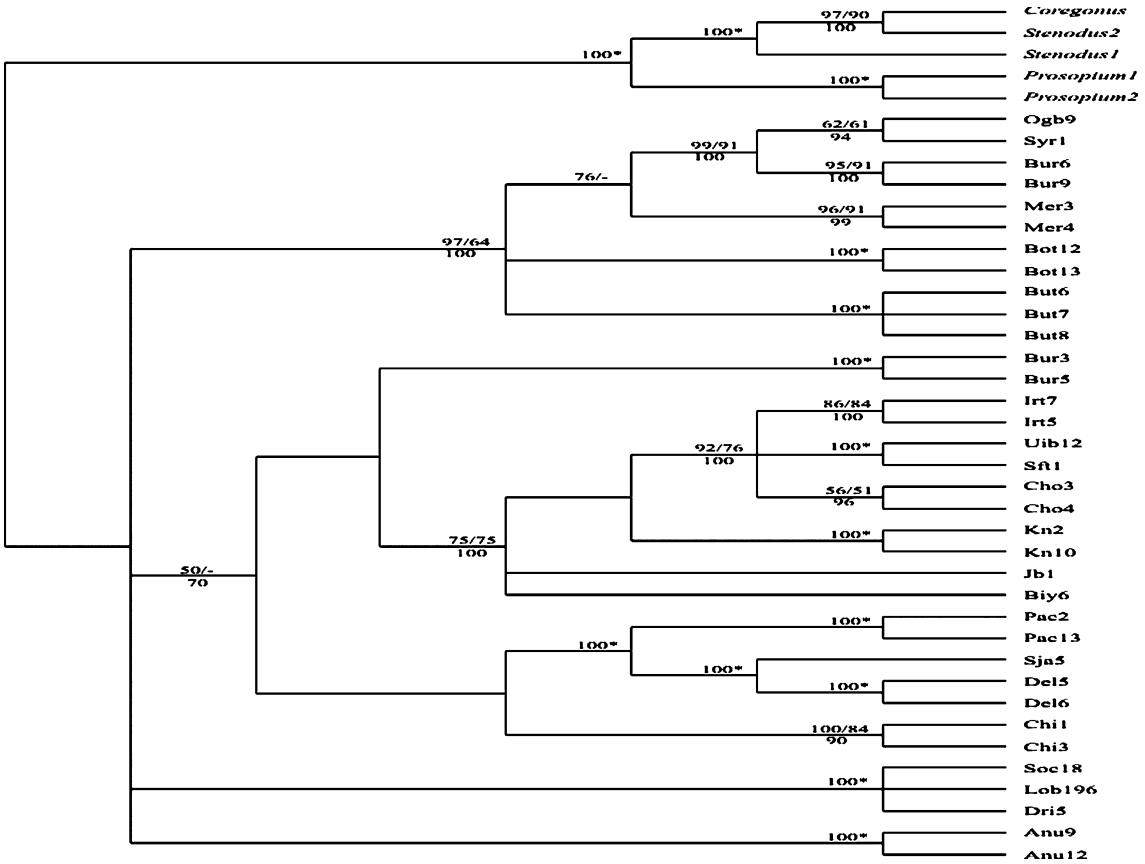

Fig. 2. Strict consensus tree of the 33 most-parsimonious trees for the ATPase6 sequences (see text for details). Bootstrap values (over 50%) for MP (above, left), for ML (HKY + G + I model) (above, right) and Bayesian probabilities (below the nodes). 100* means that all bootstrap values are higher than 95. The tree was rooted using *Coregonus*, *Stenodus*, and *Prosopium*.

were treated as "missing data" and in 28 equally parsimonious trees of 334 steps (CI = 0.6617; RI = 0.8653) when gaps were treated as fifth base. Based on Modeltest, the HKY model with an estimate of invariable sites (0.5938) and a discrete approximation of the gamma distribution (0.573) was chosen. Similar topologies were revealed for all methodologies and the SH-test revealed no significant differences between them. Thus only the unrooted ML tree is presented (Fig. 3C) together with major node support values for MP and Bayesian analysis. Description of the clades will be provided below when considering both genes.

CR and ATP6 comparison and combined analysis

Pairwise sequence divergence estimates for the ATP6 gene were clearly higher than for the CR (Table 2). To compare divergence rates we calculated the mean divergence ratio for the two genes. This comparison was done excluding three OTU's for which the same individual was not sequenced for both genes. Based on 32 OTU's sequenced at the individual level for both genes, mean pairwise divergence of ATP6 sequences was 2.46 greater (SD = 0.99) than for the CR. Plotting divergence estimates for the two genes reveals a nonlinear relation with

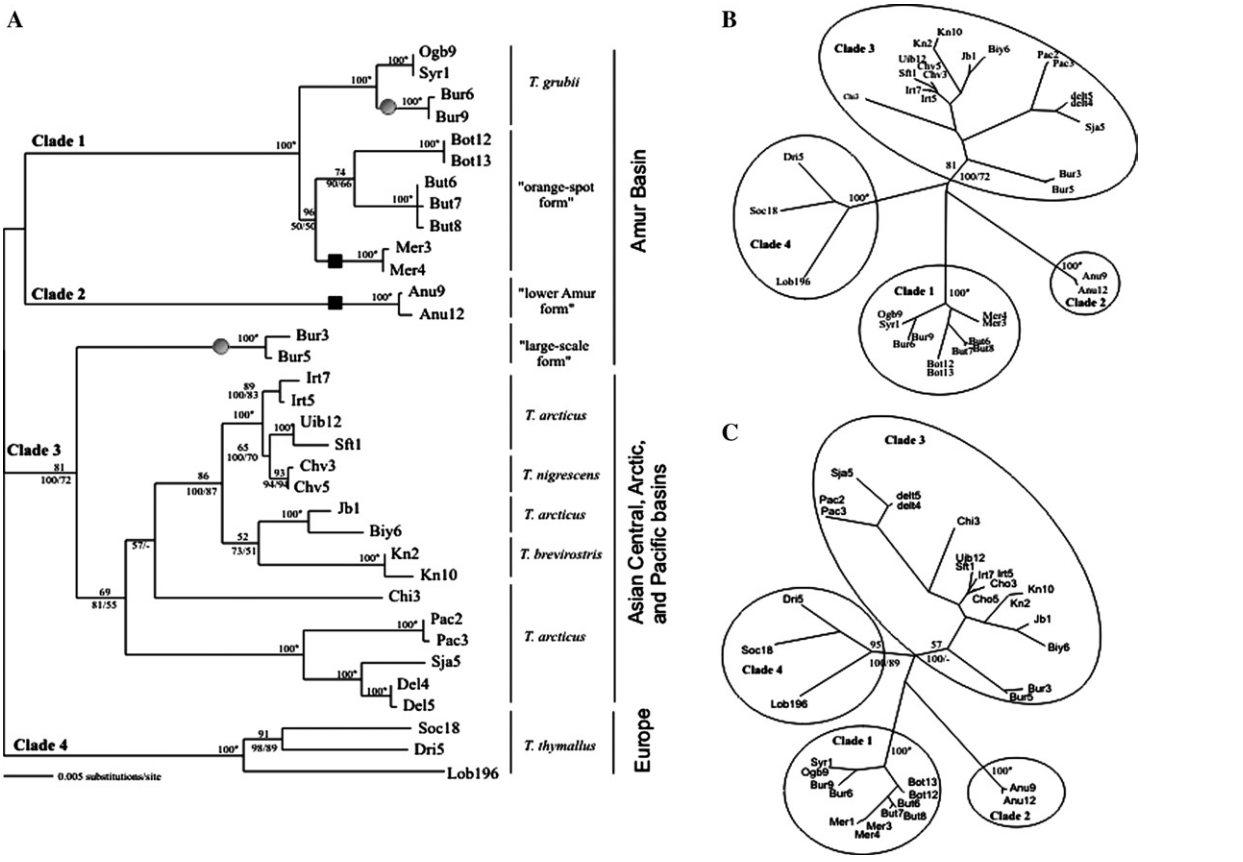

Fig. 3. (A) A mid-point rooted tree derived from a ML search using the HKY + G + I model for the combined CR and ATP6 sequences (see text for details). This rooted tree is shown for pictorial purposes only and represents the same topology shown in the unrooted tree (B). All analyses (ML, MP, and Bayesian) gave similar estimates of relationships. For the major clades, bootstrap values (over 50%) are shown for ML (above); MP (without gaps) (below, left) and Bayesian probabilities (below, right). 100* means that all bootstrap values are higher than 95. The symbol ● and ■ refer to two pairs of taxa that are found in sympatry in the Amur basin. (C) An unrooted tree derived from a ML search using the HKY + G + I model for CR sequences only, (see text for details). All analyses (ML, MP, and Bayesian) gave similar estimates of relationships. For the major clades, bootstrap values (over 50%) are shown for ML (above); MP (gaps treated as "missing" data) (below, left) and Bayesian probabilities (below, right). 100* means that all values are higher than 95.

Table 2
Net nucleotide divergence (*Da*) between groups (uncorrected *p* distance) (corrected for within-group variation) for the four major clades (shown in Fig. 3) derived from each gene data set separately as well as together

	Clade 1	Clade 2	Clade 3	Clade 4	Clade 1	Clade 2	Clade 3
Clade 1	—	0.053	0.044	0.066	—		
Clade 2	0.048	—	0.037	0.058	0.050	—	
Clade 3	0.027	0.035	—	0.048	0.033	0.036	—
Clade 4	0.027	0.038	0.016	—	0.041	0.046	0.028

Left table, ATP6 values in upper diagonal; CR values in lower diagonal. Right table, values for both genes together.

CR divergence estimates, which level off in relation to those for ATP6 (Fig. 4).

Preliminary analysis with the ATP6 gene combined with several 50–200 bp blocks of the most conserved regions of the CR, revealed no further resolution in branch order within the ingroup, and thus outgroup analysis was abandoned with CR sequence, and analysis of the combined data set was carried out without rooting. The final alignment for both genes yielded a 1768 bp fragment. Pairwise sequence divergence among all taxa ranged from 0 to 6.5%. The MP analysis resulted in 1 tree of 541 steps treating gaps as "missing data" (250 parsimony informative characters; CI = 0.6525, RI = 0.8652) and three trees of 567 steps when gaps were treated as fifth base (262 parsimony informative characters; CI = 0.6220, RI = 0.8686). For the combined data set, the HKY model was again chosen, with an estimate of invariable sites (0.593) and a discrete approximation of the gamma distribution (0.573). One unrooted ML tree was found (Fig. 3B). For pictorial purposes only, this tree is also shown with mid-point rooting (no change in branch order occurs) together with the node support values for ML, MP (without gaps), and Bayesian analysis (Fig. 3A). The SH-test revealed no significant differences between tree topologies based on different phylogenetic methodologies.

Beginning with the most internal nodes, four major lineages are resolved with consistently high support across all methodologies (Fig. 3A). Two lineages represent taxa from within the Amur basin, including the population sampled from the Okhotsk basin, adjacent to the Amur. One lineage includes all N. American and Siberian samples outside the Amur basin in addition to two highly divergent haplotypes from the Bureya River (Amur basin). The fourth lineage consists of European grayling *T. thymallus*. These lineages represent the earliest divisions within the genus for which there is reliable phylogenetic signal.

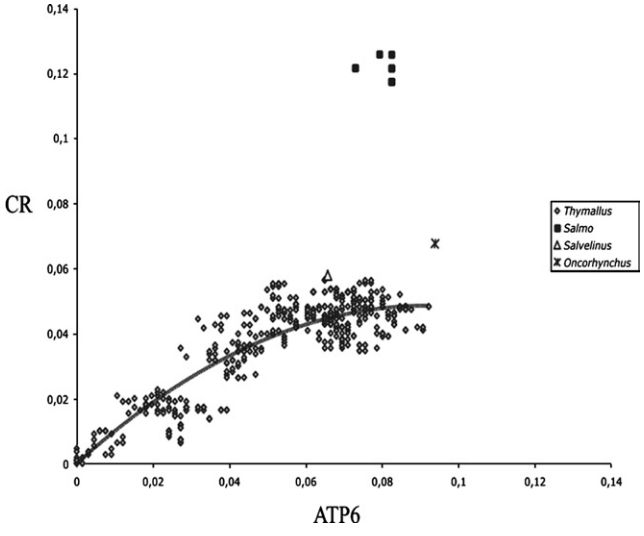

Fig. 4. Plot of sequence divergences for ATP6 against the CR for *Thymallus*, *Salmo*, *Salvelinus*, and *Oncorhychus* (see text for details).

Within clade 1 (Amur) there is high-node support for samples representing Amur grayling *T. grubii*. A sister clade to Amur grayling consists of samples from three different populations of a recently described but unnamed taxon ("orange-spot") (Antonov, 2000; Froufe et al., 2003a; Knizhin et al., 2004). Clade 2 consists of two samples from the lower Amur, representing a distinct taxon ("lower Amur form") sympatric to the "orange-spot" grayling characterized in Froufe et al. (2003a) and Knizhin et al. (2004). Within clade 3 there is support (from all four methodologies) for a first split between the haplotypes from the Bureya River representing "large-scale" grayling (Antonov, 2000; Knizhin et al., 2004), and all other OTU's. These OTU's further form 3 well-supported lineages, one of which includes N. American, Ural, and lower Lena populations of *T. arcticus*; one single haplotype from the upper Lena basin, and a diverse group of *T. arcticus* populations residing in the Baikal basin, as well as the upper Ob and Enisey rivers, but also *T. nigrescens* and *T. brevirostris*. Thus, as shown in Koskinen et al. (2002b), *T. arcticus* is paraphyletic, relative to *T. brevirostris* and *T. nigrescens*. The forth clade consists of European grayling *T. thymallus*.

Discussion

Phylogenetic relationships among grayling lineages

For comprehensive characterization of the phylogenetic relation of all grayling lineages, we use the combined data set, noting that there are no significant conflicts in well-supported nodes between any of the trees for either gene. However, discussion of divergences and divergence times is based on CR sequences only, relying on the 1%/MY calibration, which provides consistency with previously reported divergences (Froufe et al., 2003a; Koskinen et al., 2002b; Weiss et al., 2002). Although we demonstrate the lack of linearity in CR divergence (rejection of the molecular clock), the calibration is based on lineages diverging in the past 110–450,000 years, and was compatible with secondary paleo-hydrological assumptions involving lineages diverging in the past 2.5 MY (Koskinen et al., 2002b). Thus, the molecular clock heterogeneity that we observe (due to CR constraints) results in underestimates for the most divergent lineages, but perhaps have little effect on divergence estimates involving lineages evolving within at least the last 2.5 MY.

The most fundamental and consistent result of the phylogenetic reconstruction was the resolution of four major lineages, which, relative to the resolution of our sequence data, have split contemporaneously. This can be seen in the outgroup-rooted trees (Fig. 2) as a poly-

chotomy (as nodes more internal than those supporting these four lineages are poorly supported). These lineages have presumably evolved in isolation between 1.6 and 4.8 Mya, predominately in the Pliocene epoch (Table 2). The maximum divergence (4.8%) is seen between two lineages within the Amur basin, and the minimum between European grayling and the diverse clade 3 (1.6%). This underscores the historical depth of major phylogeographic events for *Thymallus* within far-eastern Siberia, compared to the probable Pleistocene/Pliocene boundary colonization of Europe. This also contrasts with much shallower divergences and simpler phylogeographic patterns for several other salmonid taxa in far-eastern Siberia (Froufe et al., 2003b).

A more detailed examination of the OTU's within these four mitochondrial lineages reveals considerable complexity in terms of the distribution of haplotypes within and among drainages as well as their concordance with existing systematics and taxonomy. Clade 1 (Fig. 3A) consists of two individuals of Amur grayling *T. grubii* sampled from tributaries of the upper Amur drainage, corresponding roughly to the original type specimen locality for the species (Dybowski, 1869). A highly supported sister group to these upper Amur samples, are two individuals sampled from the Bureya River (an Amur tributary), which also represent *T. grubii*, though from a divergent population as they are ca. 2000 km apart. For "orange-spot" grayling, the three populations originate from presently disjunct but adjoining basins (Merek–lower Amur; Buta–Okhotsk Sea; Botchi–Sea of Japan). We hypothesize that this taxon has a monophyletic origin with its present distribution reflecting the dynamics of the postglacial hydrological landscape. Clade 2 consists of two specimens from the lower Amur ("lower Amur form") found in sympatry with "orange-spot" grayling as described in Froufe et al. (2003a) and Knizhin et al. (2004). While in sympatry, we emphasize that no hybridization was documented between "orange-spot" and the "lower Amur form" and sequence divergence (CR) is 4.6% (Froufe et al., 2003a). Grayling sampled from the lower Amur were previously described as *T. arcticus grubei* (Tugarina and Khramtzova, 1980).

Clade 3 also contains samples from the Amur basin (Bureya) representing "large-scale" grayling. These haplotypes are highly divergent from all other OTU's in Clade 3 (2.4–3.7%), and is 3.2–5.0 % divergent from all other Amur basin OTU's (including samples from the adjoining Okhotsk sea and Sea of Japan). The hypothesis of Knizhin et al. (2004) suggesting that these "large-scale" grayling are more closely related to Siberian lineages (i.e., *T. arcticus* of Kolyma and lower Lena drainages referred to as the *pallasi* subspecies) north of the Amur basin is supported by all four phylogenetic approaches when both genes are considered (Fig. 3A). The existence of this lineage in the Amur can be explained by river capture of north flowing rivers (Zeya and Amgun/Bureya) by the Amur basin following the melting of the Stanovoy glacier complex and Okhotsk ice sheet as depicted in Grosswald (1998). Thus, there appears to be two sympatric taxa in the Bureya River with 3.4% CR divergence. Antonov (2000) suggests that there are as many as 3 taxa in the Bureya River, though our limited number of samples cannot yet support this hypothesis. Clearly, as suggested in Froufe et al. (2003a) and further characterized by detailed morphological description (Knizhin et al., 2004) the Amur basin is a complex system of multiple sympatric and allopatric lineages of grayling, even based on our relatively sparse geographic sampling.

Among the remaining lineages of clade 3 is a subclade representing all samples within the Baikal/Enisey basin. These individuals (Irt, Uib, Sft, and Chv) represent a shallow diverse clade described in Koskinen et al. (2002b), presumed to have colonized Lake Baikal and its tributaries in the mid-Pleistocene (110–450,000 years ago) following the "break-out" of the rising Lake Baikal into the Angara River. Populations from the Angara River (Irt), Lake Baikal (Uib), and the upper Selenga basin (Sft) all represent *T. arcticus* whereby the population from ancient Lake Chovsgul in Mongolia (Chv), part of the upper Selenga basin, represents *T. nigrescens*. These populations await morphological analysis aimed at supporting or refuting various subspecific or intrasubspecfic designations. A sister clade from Baikal/Enisey basin lineages contains two sub-clades one of which contains one sample from the uppermost Enisey basin in Mongolia (Jb1) and one sample from the uppermost Ob basin (Biy) in the Altai region, both considered to represent *T. arcticus*. As shown in Koskinen et al. (2002b) the uppermost Enisey contains lineages distantly related from populations of the mid or lower Enisey basin, which are now directly connected to Baikal via its only outflow, the Angara River. The second sub-clade contains two individuals of Mongolian grayling *T. brevirostris* (Kn), from Khökh Nuur, a small lake at the easternmost range of the species. Divergence between the Baikal/Enisey clade and its sister clade is only 0.9% underscoring relatively recent phenotypic diversification marked by the highly morphologically distinct Mongolian grayling.

One individual represents a highly divergent lineage from the upper Lena basin (Chi3) and is poorly defined within clade 3, as its position is never well supported and changes depending on the methodology applied. This single sample was taken as a representative (data not shown) of the Lena basin clade depicted in Koskinen et al. (2002a). Individuals from the lower Lena delta region (Del) cluster together with one individual sampled from the Dvina River drainage west of the Ural mountains (Sja5), and these three individuals form a sister group to two individuals from Pacific drainage popu-

lations in North America (Pac). The Sja5 sample is the identical specimen sequenced for the NADH 5 gene in Koskinen et al. (2000), and apparently comes from within the hybrid zone of *T. thymallus* and *T. arcticus* (Shubin and Zakharov, 1984; Zinoviev, 1980), though the population is more than 500 km west of the Urals. The divergence between the upper Lena individual (Chi3) and the latter described clade is 3.1%. While only a few individuals are used in this analysis, extensive sampling of the Lena basin demonstrates that these clades represent two highly distinct monophyletic lineages (Weiss et al. unpublished data). While all of these populations are considered to represent *T. arcticus*, opinions differ concerning subspecific status of phenotypically diverse grayling of Arctic draining rivers.

In summary, clade 3 (exclusive of "large-scale" grayling) presents a confounding picture of a paraphyletic *T. arcticus* complex, containing several distinct species, as well as a number of controversial subspecies. European grayling, *T. thymallus* are represented by clade 4. Though simplistically depicted here, the species is indeed phylogeographically complex (Weiss et al., 2002), exhibits 4% CR sequence divergence across its range, and may contain at least one additional taxon, from Adriatic draining rivers (Sušnik et al., 1999, 2001), represented here by the Soc18 sample.

This phylogenetic framework provides the first broad-scale overview of the genus *Thymallus*, reporting on several locations and or taxa that have not been previously analyzed, expanding on the most recent findings in the Amur basin, and addressing the substantial remaining systematic uncertainties particularly concerning the *T. arcticus* complex. Based on our knowledge of the within drainage genetic and phenotypic diversity in the Amur, Lena, and even Danube River in Europe, it is clear that few conclusions can be drawn from a basin with few samples. For instance, while some authors assume that Lena river *T. arcticus* represent a subspecies (*pallasi*), we have found two highly divergent lineages in this drainage. For the Ob, another enormous hydrological system, we cannot assume that only one lineage is present based on our single mtDNA sequence. In fact, all other major Siberian drainages (Lena, Enisey and Amur) that have been sampled more extensively, exhibit multiple lineages with presumably deep and most likely allopatric origins. Thus, while significant advances have been recently made in the description of systematic diversity in *Thymallus* populations throughout their range, there are some important gaps in our knowledge. Together with the poorly sampled Ob, areas of both the Enisey and Amur remain unexplored; some taxa (*T. brevirostris*, *T. nigrescens*) are only represented in our data by single populations; and some of the peripheral range of the *T. arcticus* complex have not been sampled. Additionally, we stress the need for both population level sampling and parallel collection of reliable phenotypic and morphological characters together with genetic data to provide the basis for comprehensive systematic revision.

The authors are engaged in such a work which involves the collection of 12 meristic and 46 morphological characters, sequencing of the complete CR, and where necessary screening of multiple microsatellite loci (to determine reproductive isolation of putative sympatric taxa), throughout the Eurasian distribution of the genus. It is our opinion that, especially for a popular well-studied group of fishes like salmonids, such comprehensive and exhaustive work is necessary to provide both a clear understanding of evolutionary diversity and to develop a broadly acceptable and practical systematic and taxonomic system.

Outgroup significance

The use of a coding gene allowed us to incorporate multiple outgroups, and validate the monophyletic status of *Thymallus*. However, the use of the closest available outgroups, representing the three ancestral genera within Salmonoidei (Ramsden et al., 2003; Sanford, 1990; and references therein), provided no resolution of the first splits within the ingroup, the outgroup taxon did not support any branch ordering among the four major lineages described above. Preliminary analysis (not shown) adding more derived genera (*Salvelinus*, *Salmo*, and *Oncorhynchus*) as outgroups, demonstrated that they were even more distant from *Thymallus*, and again, provided no phylogenetic information for resolving the branch order of the most internal nodes of the ingroup. Within the outgroup, we further note the relative proximity (0.9%) of one *Stenodus* sequence with *Coregonus* whereas the other is 3.3% divergent. As both *Stenodus* individuals are from the sample population in the Yukon drainage, and hybridization between these two genera is frequently reported (Alt, 1971; Reist et al., 1992) we presume this to be the signal of past introgression.

Comparison of divergence rates between ATP6 and CR

The ATP6 gene evolves faster than the complete CR within *Thymallus*. This is surprising consider the results within a study of brown trout *Salmo trutta*, where 11 haplotypes were found screening ATP6 in 24 individuals whereby 17 haplotypes were found using the first 310 bp of the CR (Giuffra et al., 1994). In a comparative study of four taxa of Pacific salmon (*Oncorhynchus*), Churikov et al., 2001 showed that the combined fragment of ATPase 8 and 6, COIII, and ND3 evolved slower in 3 of 4 taxa compared to the combined fragment of cytochrome *b* (cyt *b*) and the CR. In a well-calibrated framework, Bermingham et al. (1997) derive a 1.3%/MY rate for the ATP6 gene across a diverse range of teleost taxa,

and Sivasundar et al. (2001) report conspecific maximum divergence rates in the CR nearly 3 times higher than for ATP6 in the characiform genus, *Prochilodus*.

We evaluated the relative divergence rate of ATP6 and the CR in available salmonid GenBank sequences. Choosing individuals for which both genes were sequenced, the relative divergence rate across two species of charrs (*Salvelinus alpinus*, *S. fontinalis*; GenBank Accession Nos. AF154851, AF154850), two species of *Oncorhynchus* (*O. mykiss* and *O. tshawytscha*, GenBank Accession Nos. L29771, AF392054), and two species of *Salmo* (*S. salar*, GenBank Accession No. NC_001960, and *S. trutta*, GenBank Accession Nos. X74240–X74245; M97962, M97966, M97972–M97984) was assessed. Calculating the ratio of mean divergences, as above, the ATP6 evolves 1.57 times faster than the CR in *Salvelinus* and 1.93 times faster in *Oncorhynchus* (Fig. 4). For the *Salmo* comparison, only partial sequences were available (315 bp for ATP6 and 241 alignable bp for CR) and for this more limited comparison the CR evolved 1.98 times faster than ATP6. Thus, in 3 of the 4 salmonid genera evaluated, ATP6 is shown to evolve considerably faster than the CR. The ATP6 is thus another coding mtDNA segment, together with NADH subunits 1, 2, 5, and 6 (see Hansen and Loeschcke, 1996 and Apostolidis et al., 1997) that evolves faster than the CR in salmonid fishes.

The CR in salmonid fishes thus exhibits substantial evolutionary constraints. It is noteworthy that some authors, influenced by the obviously higher substitution rates found in other fishes or endothermic vertebrates, have entertained divergence rates as high as 10%/MY in a salmonid (Brunner et al., 2001) We have not found any empirical evidence for such CR divergence rates in salmonid fishes, and reiterate that the approximate 1% rate given for *Thymallus* is inferred from multiple independent sources (see discussions in Koskinen et al., 2002b and Weiss et al., 2002). This rate is also equal to that estimated for the entire mtDNA molecule for the salmonid family (Smith, 1992).

While we calculated a mean relative rate for the ATP6 and CR (2.46:1) the substitution rate of one or both genes is clearly not constant across all lineages, and thus calculations of divergence estimates (*Da*) among groups for both genes are not consistent. For example, both CR (2.7%) and ATP6 (6.6%) divergence estimates correspond to ca. 2.7 MY between European grayling and upper Amur grayling (clade 1), and results are similar between European grayling and the diverse clade 3 (1.6% for CR = 1.6 MY vs. 4.8% for ATP6 = 1.9 MY) (Table 2). However, for all other comparisons, the apparently faster mutating ATP6 gene results in lower divergence estimates (50% or more) between the major lineages. Thus, our mean relative estimate is very approximate as the relation between the two genes is not linear, with the CR exhibiting constraints after about 5.8% divergence (i.e., saturation) whereas pairwise divergences for ATP6 reach over 9%. This suggests species-specific constraints on the non-coding CR within the genus *Thymallus*. While results for *Salvelinus* and *Oncorhynchus* are congruent with our comparative rates within *Thymallus*, it is clear that there is no overall constraint on the salmonid CR as sequences between genera are indeed difficult to align. Thus, it appears that the CR's usefulness in salmonid phylogenetics is indeed limited to within genus comparisons. In another study, the CR was one of the only gene segments that did not support the monophyly of the large genus *Oncorhynchus*, so even this generalization must be taken with caution (Crespi and Fulton, 2004).

While the vertebrate CR contains both hyper-variable and conserved sequence blocks, the blocks are limited in length and are not necessarily conserved across all teleost lineages (Lee et al., 1995). Additionally, the conserved blocks cannot explain the relaxation of our hypothesized species-specific constraints among salmonid genera. A related finding demonstrates both among and within genus-specific differences in CR versus cyt *b* divergences in a survey of 68 avian species (Roukonen and Kvist, 2002). Whereby many avian lineages were shown to have more rapidly evolving CR's (in one genus 5.14–21.65 times faster, dependant on species), at least two genera displayed CR/cytb ratios of less than 1 (0.46:0.94 and 0.36:0.81). Such lineage-specific constraints suggest yet-to-be identified selectively advantageous functions for the CR.

Pliocene/Pleistocene phylogeographic perspectives

More frequent and improved calibrations of divergence estimates, have suggested that the origins of many extant lineages reach into the Pliocene, contradicting earlier notions of the importance of the Pleistocene epoch in generating present biodiversity. This is true for the N. American avian fauna (Klicka and Zink, 1997), diverse organisms on the Iberian Peninsula (Gómez and Lunt, in press), as well as for N. American freshwater fishes (Avise et al., 1998; Avise and Walker, 1998). For European freshwater fishes, Weiss et al. (2002) suggested a pre-Pleistocene spread of *T. thymallus*, Kotlik and Berrebi (2001) showed a similar result for barbel *Barbus barbus* and Durand et al. (1999) for chub *Leuciscus cephalus*, and maximum divergences among major lineages of *Salmo trutta* point to at least a Pliocene/Pleistocene origin (Bernatchez, 2001). In a broader phylogenetic work on European cyprinids, Zardoya and Doadrio (1999) report that most speciation events within the genera *Barbus* and *Luciobarbus* occurred during the Pliocene when the major existing European river drainages first formed.

In northern Asia, the overall understanding of glacial dynamics is far less clear than for N. America and

Europe, and, due to the paucity of studies, the phylogeographic patterns of most Asian freshwater organisms is unknown. The historical perspective of Siberian paleohistory was that the region was too dry to accumulate the amounts of precipitation needed to build glaciers as extensive as in N. America and Europe. Most recent advances, however, repeatedly stress that Siberian glaciations were at least as extensive, and corresponding Pleistocene dynamics at least volatile as elsewhere (see Grosswald, 1998 and references therein). While this new paradigm is being more generally accepted, its effects on Siberian fauna and particularly freshwater fishes is little appreciated. At least for the genus *Thymallus*, the overall pattern of Pliocene origins of major lineages and extensive Pleistocene paleohydrological dynamics resulting in a mosaic of highly divergent lineages in each of the major river basins investigated has been clearly presented (Froufe et al., 2003a,b; Koskinen et al., 2002b; this study).

Acknowledgments

We thank D.J. Harris, C. Sturmbauer, K. Sefc, S. Koblmüller, and N. Duftner for comments, suggestions or help with some analysis; P. Spruel and T. Otis for providing the *Prosopium* and *Stenodus* samples, respectively; C. Primmer for use of the Sja5 (Ural) sample; E. Taylor for the N. American samples; and A. Snoj for the Drina River sample, and for various Russian samples, A.F. Kirillov. We also give thanks for field assistance to B.E. Bogdanov, A.L. Antonov, G.P. Safronov, A.N. Matveev, and V.P. Samusenok. 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 Ph.D. grant to E. Froufe (SFRH/BD/11377/2002); Russian Ministry of Education (Grant A03-2.12-265), Russian program "Universities of Russia" Grant No. UR.07.01.009.

References

Avise, J.C., Walker, D., Johns, G.C., 1998. Speciation durations and Pleistocene effects on vertebrates phylogeography. Proc. R. Soc. Lond. B 265, 1707–1712.

Avise, J.C., Walker, D., 1998. Pleistocene phylogeographic effects on avian populations and the speciation process. Proc. R. Soc. Lond. B 265, 457–463.

Alt, K.T., 1971. Occurrence of hybrids between Inconnu, *Stenodus leucichthys nelma* (Pallas), and Humpback whitefish, *Coregonus pidschian* (Linnaeus) in Chatanika River, Alaska. Trans. Am. Fish Soc. 100, 362–365.

Antonov, A.L., 2000. New salmonoidei fishes from the Amur basin. In: Proceedings Biodiversity and Dynamics of Ecosystem in North Eurasia (BDENE) Novosibirsk, vol. 3. Part 1, August 21–26, RAS SB, IC and G Novosibirsk, pp. 120–122.

Apostolidis, A.P., Triantaphyllidis, C., Kouvatsi, A., Economidis, P.S., 1997. Mitochondrial DNA sequence variation and phylogeography among *Salmo trutta* L. (Greek brown trout) populations. Mol. Ecol. 6, 531–542.

Berg, L.S., 1949. Freshwater Fishes of the USSR and adjacent countries, fourth ed. Part 3. USSR Academy of Sciences, Moscow-Leningrad, pp. 929–1382 (in Russian, Translation available, 1965, Smithson. Inst. by Israel Prog. Sci. Transl., Jerusalem).

Bermingham, E., McCafferty, S., Martin, A.P., 1997. Fish biogeography and molecular clocks: perspectives from the Panamanian Isthmus. In: Kocher, T.D., Stepien, C.A. (Eds.), Molecular Systematics of Fishes. Academic Press, San Diego, pp. 113–128.

Bernatchez, L., 2001. The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analysis of mitochondrial DNA variation. Evolution 55 (2), 351–379.

Brunner, P.C., Douglas, M.R., Osinov, A., Wilson, C.C., Bernatchez, L., 2001. Holarctic phylogeography of arctic charr (*Salvelinus alpinus* L.) inferred from mitochondrial DNA sequences. Evolution 55 (3), 573–586.

Churikov, D., Matsuoka, M., Luan, X., Gray, A.K., Brykov, V.L.A., Gharrett, A.J., 2001. Assessment of concordance among genealogical reconstructions from various mtDNA segments in three species of Pacific salmon (genus *Oncorhynchus*). Mol. Ecol. 10, 2329–2339.

Crespi, B.J., Fulton, M.J., 2004. Molecular systematics of Salmonidae: combined nuclear data yields a robust phylogeny. Mol. Phyl. Evol. 31 (2), 658–679.

Durand, J.D., Persat, H., Bouvet, Y., 1999. Phylogeography and postglacial dispersion of the chub (*Leuciscus cephalus*) in Europe. Mol. Ecol. 8, 987–989.

Dybowski, B., 1869. Vorläufige Mitteilungen über die Fischfauna des Onon flusses und des Ingoda in Transbaikalien. Wiener Mitteilung, 946–951 (in German).

Felsenstein, J., 1981. Evolutionary trees from DANN sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368–376.

Froufe, E., Knizhin, I., Koskinen, M.T., Primmer, C.R., Weiss, S., 2003a. Identification of reproductively isolated lineages of Amur grayling (*Thymallus grubii* Dybowski 1869): concordance between phenotypic and genetic variation. Mol. Ecol. 12 (9), 2345–2355.

Froufe, E., Alekseyev, S., Knizhin, I., Alexandrino, P., Weiss, S., 2003b. Comparative phylogeography of salmonid fishes (Salmonidae) reveals late to post-Pleistocene exchange between three now-disjunct rivers basins in Siberia. Diver. Distrib. 9, 269–282.

Giuffra, E., Bernatchez, L., Guyomard, R., 1994. Mitochondrial CR and protein coding genes sequence variation among phenotypic forms of brown trout *Salmo trutta* from northern Italy. Mol. Ecol. 3, 161–1171.

Gómez, A., Lunt, D.H., in press. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss, S., Ferrand, N. (Eds.), Phylogeography of Southern European Refugia, Kluwer Academic Press, Dordrecht, Netherlands.

Grosswald, M.G., 1998. New approach to the ice age paleohydrology of Northern Eurasia. In: Benito, G., Baker, V.R., Gregory, K.J. (Eds.), Paleohydrology and Environmental Change. Wiley, Chichester, England, pp. 199–214.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Res. Symp. Ser. 41, 95–98.

Hansen, M.M., Loeschcke, V., 1996. Genetic differentiation among Danish brown trout populations, as detected by RFLP analysis of PCR amplified mitochondrial DNA segments. J. Fish Biol. 48, 422–436.

Huelsenback, J.P., Bollback, J.P., 2001. Empirical and hierarchical Bayesian estimation of ancestral states. Syst. Biol. 50, 351–366.

Huelsenback, J.P., Ronquist, F., 2001. MR-BAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Jankovic, D., 1962. Synopsis of biological data on European grayling, *Thymallus thymallus* (L. 1758): FAO Fisheries Biology Synopses, vol. 24. FAO, Rome.

Klicka, J., Zink, R.M., 1997. The importance of recent ice ages in speciation: a failed paradigm. Science 277 (12), 1666–1669.

Knizhin, I., Weiss, S., Antonov, A.L., Froufe, E., 2004. Morphological and genetic diversity of Amur graylings (*Thymallus*; Thymallidae). J. Ichthyol. 44, 59–76.

Kotlik, P., Berrebi, P., 2001. Phylogeography of the barbel (*Barbus barbus*) assessed by mitochondrial DNA variation. Mol. Ecol. 10, 2177–2185.

Koskinen, M.T., Ranta, E., Piironen, J., Veselov, A., Titov, S., Haugen, T.O., Nilsson, J., Carlstein, M., Primmer, C.R., 2000. Genetic lineages and postglacial colonization of the grayling (*Thymallus thymallus*, Salmonidae) in Europe, as revealed by mitochondrial DNA analyses. Mol. Ecol. 9, 1609–1624.

Koskinen, M., Haugen, T.O., Primmer, C.R., 2002a. Contemporary fisherian life-history evolution in small salmonid populations. Nature 419, 826–830.

Koskinen, M.T., Knizhin, I., Primmer, C.R., Schlötterer, C., Weiss, S., 2002b. 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. Mol. Ecol. 11, 2599–2611.

Kottelat, M., 1997. European fresh water fishes. An heuristic checklist of the fresh water fishes of Europe (exclusive of former USSR), with an introduction for non-systematists and comments on nomenclature and conservation. Biology 52 (Suppl. 5), 1–271.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17, 1244–1245.

Lee, W.J., Conroy, J., Howell, W.H., Kocher, T.D., 1995. Structure and evolution of teleost mitochondrial control regions. J. Mol. Evol. 41, 54–66.

Meyer, A., 1993. Evolution of mitochondrial DNA of fishes. In: Hochackka, P.W., Mommsen, P. (Eds.), Molecular Biology Frontiers, Biochemistry and Molecular Biology of Fishes. Elsevier Press, Amsterdam, pp. 1–38.

Osinov, A.G., Lebedev, V.S., 2002. Genetic divergence and phylogeny of the Salmoninae based on allozyme data. J. Fish Biol. 57 (2), 354–381.

Persast, H., Pattee, E., Roux, A.L., 1978. In: Origine et caractérstiques de la distrubtion de l'Ómbre commun, *Thymallus thymallus* L., (1758), en Europe et en France, vol. 20. International Association of Theoretical and Applied Limnology, Verhandlugnen, pp. 217–2121.

Persat, H., 1996. Threatened populations and conservation of the European grayling, *Thymallus thymallus* (L., 1758). In: Hefti, A., Kirchofer, D. (Eds.), Conservation of Endangered of Europe, pp. 233–247.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Ramsden, S.D., Brinkmann, H., Hawryshyn, C.W., Taylor, J.S., 2003. Mitogenomics and the sister of Salmonidae. TREE vol. 18, No. 12.

Redenbach, Z., Taylor, E.B., 1999. Zoogeographical implications of variation in mitochondrial DNA of Arctic grayling (*Thymallus arcticus*). Mol. Ecol. 8, 23–25.

Reist, J.D., Vuorinen, J., Bodaly, R.A., 1992. Genetic and morphological identification of Coregonid hybrids fishes from Arctic Canada. Pol. Arch. Hydro. 39, 551–561.

Reshetnikov, Yu.S., Popova, O.A., Sokolov, L.I., Tsepkin, E.A., Sideleva, V.G., Dorofeeva, E.A., Chereshnev, I.A., Moskal'kova, K.I., Dgebuadze, Yu.Yu., Ruban, G.I., Korolev, V.V., 2002. Atlas presnovodnykh ryb Rossii v dvukh tomakh (Atlas of Russian Freshwater Fishes in Two Volumes), vol. 1, Moscow, Nauka, 379 pp. (in Russian).

Roukonen, M., Kvist, L., 2002. Structure and evolution of the avian mitochondrial control region. Mol. Phylogenet. Evol. 23, 422–432.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, second ed. Cold Spring Harbor Press, Cold Spring Harbour, NY.

Sanford, C.P.J., 1990. Phylogenetic relationships of salmonid fishes. Bull. Br. Mus. Nat. Hist. (Zool.) 56 (2), 145–153.

Scott, W.B., Crossman, E.J., 1998. Freshwater Fishes of Canada, fifth ed. Galf House Publications, Oakville, Ont., Canada.

Shatunovskiy, M.I., 1983. Fishes of Mongolian People Republic Moscow. Nauka Press p. 273.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Shubin, P., Zakharov, A., 1984. Hybridization between European grayling, *Thymallus thymallus*, and Arctic grayling, *Thymallus arcticus*, in the contact zone of the species. J. Ichthyol. 4, 159–163.

Skurikhina, L.A., Mednikov, B.M., Tugarina, P.Ya., 1985. Geneticheskaya divergentziya khariusov (Thymallus Cuvier) Evrazii i seti vidov. Zool. Zh. 64 (2), 245 (in Russian).

Sivasundar, A., Bermingham, E., Orti, G., 2001. Population structure and biogeography of migratory freshwater fishes (*Prochilodus* : Characiformes) in major South American rivers. Mol. Ecol. 10, 407–417.

Smith, G.R., 1992. Introgression in fishes: significance for paleontology, cladistics, and evolutionary rates. Syst. Biol. 41, 41–57.

Sušnik, S., Snoj, A., Dovc, P., 1999. Microsatellites in grayling (*Thymallus thymallus*): Comparison of two geographically remote populations from the Danubian and Adriatic river basin in Slovenia. Mol. Ecol. 8, 1756–1758.

Sušnik, S., Snoj, A., Dovc, P., 2001. Evolutionary distinctiveness of grayling *(Thymallus thymallus*) inhabiting the Adriatic river system, as based on mtDNA variation. Biol. J. Lin. Soc. 74, 375–385.

Svetovidov, A.N., 1936. Graylings (genus *Thymallus*, Cuvier) of Europe and Asia. In: Proceedings of Institute of Zoology Academy of Sciences USSR, III, pp. 183–301 (in Russian).

Swofford, D.L., 2002. PAUP* ver 4.0.b10. Phylogenetic analysis using parsimony and other methods. Sinauer Associates, Sunderland, MA.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choise. Nucleic. Acids Res. 22, 4673–4680.

Tugarina, P.Ya., Khramtzova, V.S., 1980. Morphizioloogicheskaya kharakteristka amurskogo khariusa *Thymalllus grubei* Dyb. J. Ichthyol. 20 (N4 (123)), 590–605 (in Russian).

Tugarina, P.Ya., Anudarin, D., 1972. Mongolian grayling *Thymallus brevirostris* (Kessler) of Dzabakhan River Basin. J. Ichthyol. 12 (N5), 843–856 (in Russian).

Uiblein, F., Jagsch, A., Honsig-Erlenburg, W., Weiss, S., 2001. Status, habitat use, and vulnerability of the European grayling in Austrian waters. J. Fish Biol. 59 (A), 223–247.

Weiss, S., Uiblein, F., Jagsch, A., 2001. Conservation of salmonid fishes in Austria: current status and future strategies. In: Collares-Pereira, M.J., Coelho, M.M., Cowx, I.G. (Eds.), Freshwater Fish Conservation: Options for the Future. Fishing News Books, Blackwell Science, Oxford, pp. 72–78.

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 control region. Mol. Ecol. 11, 1393–1407.

Zardoya, R., Doadrio, I., 1999. Molecular evidence on the evolutionary and biogeographical patterns of European Cyprinids. J. Mol. Evol. 49, 227–237.

Zinoviev, E.A., 1980 Parallelizm izmenchivosti u evropeiskogo i sibirskogo khariusov//Lososevidniye rybi. ZIN AN SSSR, Leningrad, pp. 69–80 (in Russian).