# Phylogeography, systematics, and evolution of graylings (Salmonidae; Thymallinae; *Thymallus*) across their geographic range

## DISSERTATION

Zur Erlangung des akademischen Grades

**Dr. rer. nat**.

An der Naturwissenschaftlichen Fakultät der
Karl-Franzens Universität Graz
Institut für Biologie

Vorgelegt von

**Gernot K. Englmaier, MSc.**

Unter der Betreuung von

**Assoc. Univ. Prof. Dr. Steven Weiss**

Jänner 2023

## Preamble

The following dissertation includes four main chapters corresponding to the main objectives of the doctoral project. Chapter I was published in *Organisms Diversity & Evolution*, Chapter II was published in *Reviews in Fish Biology and Fisheries*, and Chapter III was published in *Molecular Phylogenetics and Evolution*, Chapter IV is currently being prepared to be submitted to a peer-reviewed scientific journal. I (Gernot K. Englmaier) am the first author in two of these research articles.

## Acknowledgements

This thesis would not have been possible without funding from the University of Graz in the form of a PhD grant within the Doctoral Academy Graz and the Ecology and Evolution in Changing Environments (EECE) consortium.

I would like to thank my PhD supervisor Steven Weiss for his continuous support, inspiration and advice, which has played an important role in my personal and scientific development. His support was especially important in challenging times when the COVID-19 pandemic restricted sampling expeditions and new research objectives had to be found. I am also grateful for interesting discussions and the freedom I had to develop independent research interests beyond the scope of my PhD thesis. Many thanks go to my mentor Boris Chagnaud and the speaker of the EECE consortium Kristina Sefc for their support and encouragement.

I would like to thank my colleagues and friends at the University of Graz in particular, among others, Anna Dünser, Lukas Zangl, Tamara Schenekar, Jaqueline Grimm, Maximilian Wagner, Sylvia Schäffer and Christoph Hahn who provided me with ideas and help whenever needed. Furthermore, I am thankful to Martina Strempfl and Karin Wagner for their continuous support with lab work at the Core Facility Molecular Biology (Medical University Graz). Many thanks go to Bernd Tatschl, Esentsengel Sukhee and Christian Beuger for their hospitality and organisation during fieldwork in Mongolia.

Lastly, my greatest thanks go to my family, especially to my parents. You have always been there to help me in every aspect and supported all my interests since childhood.

## Table of contents

## Summary

The family Salmonidae is one of the best-known families of freshwater or anadromous ray-finned fishes in the northern hemisphere due to its ecological and commercial importance. It is divided into three subfamilies and dependent on varying viewpoints includes > 250 species. The wide geographic distribution and high species diversity in temperate and arctic regions, together with its morphological plasticity, makes the group an interesting model in evolutionary biology and biogeography.

This dissertation focuses on the evolutionary history and phylogeography of graylings (Thymallinae, *Thymallus*). In contrast to other members of Salmonidae, graylings represent a rather poorly studied group in terms of both phylogenetic relationships and systematics. In the present study, mitochondrial and nuclear genomic DNA data form the basis of the main objectives of the doctoral thesis.

The first section of the dissertation addresses the phylogeography and systematic diversity of graylings based on whole mitochondrial genomes. For the first time, samples across the entire distribution range of the genus were analysed (78 samples) and the complex phylogenetic relationships were highlighted. Phylogenetic analyses and species delimitation methods supported the differentiation of 15 grayling species. Three geographical regions were particularly diverse: 1) the European Alps (*T. aeliani*, *T. ligericus*, *T. thymallus*), 2) the Altai-Sayan Mountains including the region of Lake Baikal (*T. baicalolenensis*, *T. baicalensis*, *T. nigrescens*, *T. brevirostris*, *T. brevicephalus*, *T. nikolskyi*, *T. svetovidovi*), and 3) the Amur River drainage (*T. baicalolenensis*, *T. burejensis*, *T. grubii*, *T. flavomaculatus*, *T. tugarinae*). In addition, divergence time estimates were obtained for the phylogeny based on three calibration points (two fossils and one paleohydrological event). The age of the genus *Thymallus* was dated at 7.3–10.9 Ma in the late Miocene. A biogeographical ancestral range estimation showed the Russian Far East and in particular the Amur River drainage as a potential origin of the genus. These findings represent a significant contribution to the understanding of the phylogenetic relationships within the genus *Thymallus* and subsequently provide the basis for comparative phylogenetic analyses, as well as the combination with nuclear genomic data to determine possible hybridisation events and mito-nuclear discordances.

The second part builds partly on the inferred mitochondrial phylogeny and provides insights into general sexually dimorphic traits in 11 of 15 known grayling species. For this purpose, an extensive morphological dataset with 1539 individuals (806 males 733 females) and 28 linear morphometric traits was analysed and presented in a comparative phylogenetic

perspective. All grayling species showed a very consistent pattern of sex-specific differences. particularly relating to size dimensions (length and height) of the dorsal-, anal-, pelvic- and pectoral fins. Interestingly, the display of the dorsal- and pelvic fins is an integral behavioural element in male-male competition and territorial behaviour during spawning season. In females, on the other hand, as in other gravel spawning salmonids, the anal fin seems to play an important role in oviposition or "probing" behaviour. Despite species-specific differences in the magnitude of sexual dimorphism, no significant phylogenetic signal was detected, which may indicate a differential strength of intrasexual selection across species. Subsequently, the sexually dimorphic traits in graylings were compared with other salmonids through a literature search. This allowed a comparison of anadromous and potamodromous species as well as semelparous and iteroparous species and thus provided insights into the driving forces of the evolution of sexually dimorphic traits in salmonids in more general.

Finally, the third and fourth sections provide new phylogenetic and phylogeographic insights based on nuclear genomic data (ddRADseq) in combination with mitochondrial sequences. New insights into the evolutionary history of graylings are provided by both a global analysis (entire genus, 128 samples) and a specific analysis of the European grayling species (178 samples). A major focus of this work was the inference of hybridisation events among distinct grayling species and their influence on phylogenetic relationships (mitochondrial vs. nuclear genomic data). These analyses showed that species or phylogenetic lineages may have frequently come into secondary contact, which was supported by signals of both recent and historical introgression, as well as examples of mitochondrial capture, leading to mito-nuclear discordance. Hybridisation and introgression upon secondary contact may have thus played an important role in the evolutionary history of this group of salmonids, opening new perspectives for biogeographical considerations in freshwater fishes of temperate and arctic regions, which until now have been based primarily on mitochondrial data. In addition, divergence times were estimated for the European dataset which was based on two biogeographical calibration points. This analysis showed that the splits between some of the modern European lineages might be significantly older than the late phase of the Pleistocene glaciations, challenging earlier hypotheses on the role of late Pleistocene glacial cycles as a general driving force for lineage differentiation.

## Zusammenfassung

Die Familie der Salmonidae gehört aufgrund ihrer ökologischen und kommerziellen Bedeutung zu den bekanntesten Fischfamilien der nördlichen Hemisphäre. Sie ist in drei Unterfamilien unterteilt und umfasst je nach Ansicht > 250 Arten. Ihre weite geographische Verbreitung und hohe Artendiversität in temperaten und arktischen Gebieten, zusammen mit einer erstaunlichen morphologischen Plastizität, macht die Gruppe zu einem interessanten Modellsystem der Evolutionsbiologie und Biogeographie.

Die vorgelegte Dissertation befasst sich mit der Evolutionsgeschichte und Phylogeographie der Äschen (Thymallinae, *Thymallus*). Im Gegensatz zu anderen Vertretern der Salmonidae stellen die Äschen eine bis dato eher wenig erforschte Gruppe dar was sowohl die phylogenetischen Verwandtschaftsverhältnisse also auch die Systematik betrifft. In der vorliegenden Arbeit bilden mitochondrielle und kerngenomische DNS Daten die Grundlage der behandelten Fragestellungen.

Der erste Teil der Dissertation beschäftigt sich mit der Phylogeographie und der systematischen Diversität der Äschen basierend auf ganzen mitochondriellen Genomen. Dabei wurden erstmals Proben über das gesamte Verbreitungsgebiet der Gattung analysiert (78 Individuen) und die komplexen phylogenetischen Zusammenhänge dargestellt. Phylogenetische Analysen und Methoden zur Artabgrenzung unterstützen die Unterscheidung von 15 Äschenarten. Auffallend dabei sind drei geographische Regionen mit besonders hoher Artendiversität: 1) die Europäischen Alpen (*T. aeliani, T. ligericus, T. thymallus*), 2) das Altai-Sayan Gebirge inklusive der Region des Baikal Sees (*T. baicalolenensis, T. baicalensis, T. nigrescens, T. brevirostris, T. brevicephalus, T. nikolskyi, T. svetovidovi*), und 3) das Einzugsgebiet des Amur (*T. baicalolenensis, T. burejensis, T. grubii, T. flavomaculatus, T. tugarinae*). Darüber hinaus wurde eine Altersabschätzung für die Phylogenie der Äschen berechnet, die auf drei Kalibrierungspunkten (zwei Fossilien und ein paleohydrologisches Ereignis) basiert. Dabei wurde die Gattung *Thymallus* auf ein Alter von 7.3–10.9 Ma im späten Miozän datiert. Eine biogeographische „ancestral range estimation“ zeigt den fernen Osten Russlands und im Speziellen das Einzugsgebiet des Amur als potenziellen Ursprung der Gattung. Diese Erkenntnisse stellen einen wesentlichen Beitrag zum Verständnis der Verwandtschaftsbeziehungen innerhalb der Äschen dar und ermöglichen in weiterer Folge Einblicke in makroevolutionäre Trends durch vergleichende phylogenetische Analysen. In Kombination mit kerngenomischen Daten ermöglichen sie außerdem die Bestimmung von

Hybridisierungs-Ereignissen und erlauben damit neue Einblicke in die Evolutionsgeschichte der Äschen.

Der zweite Teil der Arbeit baut teilweise auf der zuvor erarbeiteten mitochondriellen Phylogenie auf und gibt Einblicke in allgemeine geschlechtsdimorphe Merkmale bei 11 von 15 bekannten Äschenarten. Dafür wurde ein umfangreicher morphologischer Datensatz mit 1539 Individuen (806 Männchen 733 Weibchen) und 28 linearen morphometrischen Merkmalen analysiert und in einem vergleichenden phylogenetischen Ansatz dargestellt. Alle Äschenarten zeigen dabei ein sehr einheitliches Muster von geschlechtsspezifischen Unterschieden. Diese betreffen vor allem Längen- und Höhenverhältnisse der Rücken-, Anal-, Bauch-, und Brustflossen. Trotz artspezifischer Unterschiede in der Stärke des Geschlechtsdimorphismus, wurde kein signifikantes phylogenetisches Signal innerhalb der Gattung nachgewiesen, was auf eine unterschiedlich starke intrasexuelle Selektion zwischen den Arten hinweisen könnte. Ganz allgemein stellt die Zurschaustellung der Rücken- und Bauchflossen für Männchen ein wesentliches Verhaltenselement während der Laichzeit dar (Konkurrenz zwischen den Männchen um Zugang zu fortpflanzungsfähigen Weibchen). Bei Weibchen wiederum scheint die Analflosse, so wie bei anderen interstitiallaichenden Salmoniden, eine wichtige Funktion bei der Eiablage oder dem sogenannten „probing“ Verhalten (Sondierung des Schotteruntergrundes) zu spielen. In weiterer Folge wurden die geschlechtsdimorphen Merkmale bei Äschen durch eine Literaturrecherche mit Merkmalen anderer Salmoniden verglichen. Dies ermöglicht sowohl einen Vergleich von anadromen und potamodromen Arten also auch semelparen und iteroparen Arten und liefert damit Einblicke in die treibenden Kräfte der Evolution geschlechtsdimorpher Merkmale bei Salmoniden.

Der dritte und vierte Teil dieser Arbeit liefert schlussendlich neue phylogenetische und phylogeographische Erkenntnisse basierend auf kerngenomischen Daten (ddRADseq) in Kombination mit mitochondriellen Sequenzen. Dabei werden neue Einblicke in die Evolutionsgeschichte der Äschen sowohl durch eine globale Analyse (gesamte Gattung, 128 Individuen) als auch durch eine spezifische Analyse der Europäischen Äschenarten (178 Individuen) gegeben. Ein wesentlicher Fokus dieser Arbeiten liegt auf der Bestimmung von Hybridisations-Ereignissen und deren Einfluss auf die bestehenden Verwandtschaftsverhältnisse (mitochondrielle vs. kerngenomische Daten). Dabei zeigt sich, dass Arten bzw. phylogenetische Linien im Laufe ihrer Evolutionsgeschichte häufig in Kontakt gekommen sein dürften, was sowohl durch Ergebnisse rezenter- als auch historischer Introgression und „mitochondrial capture“ unterstützt wird, und zu mito-nuklearen Diskordanzen führt. Im Kontext der Verbreitungsgeschichte lassen diese Resultate auf eine

komplexe Evolutionsgeschichte der Äschen schließen, die stark mit einer dynamischen Paleohydrologie in Eurasien und Nordamerika verbunden scheint. Diese Ergebnisse zeigen außerdem, dass Hybridisierung eine wichtige Rolle in der Evolution der Äschen gespielt hat, und eröffnen damit neue Perspektiven für biogeographische Überlegungen von Süßwasserfischen in temperaten und arktischen Gebieten, die bis dato vor allem auf mitochondriellen Daten beruhen. Für den Europäischen Datensatz wurde zusätzlich eine Altersabschätzung für die Phylogenie berechnet. Diese beruhte auf zwei sorgfältig ausgewählten biogeographischen Kalibrierungspunkten. Dabei zeigte sich, dass die Aufspaltung zwischen den rezenten phylogeographischen Linien zum Teil deutlich älter als die Spätphase der pleistozänen Vergletscherungen zu sein scheint. Basierend auf mitochondriellen Daten wurde schon im ersten Teil der Dissertation das Alter der Europäischen Äschen in das späte Miozän bis frühes Pliozän datiert.

# Introduction

The family Salmonidae represents a widespread group of freshwater or anadromous fishes in temperate and artic regions of the northern hemisphere. Within the ray-finned fishes (Actinopterygii), salmonids are grouped with Esociformes, Galaxiiformes, and Argentiniformes in the superorder Protacanthopterygii (Betancur-R et al. 2017). Currently, three salmonid subfamilies are recognized (Nelson, 2006): 1) the Salmoninae (genera *Oncorhynchus* Suckley, 1861, *Salmo* Linnaeus, 1758, *Salvelinus* Richardson, 1836 *Brachymystax* Günther, 1866, *Hucho* Günther, 1866, *Parahucho* Vladykov, 1963; with 142 valid species according to Fricke et al. 2022), the Coregoninae (genera *Coregonus* Linnaeus, 1758, *Prosopium* Jordan, 1878, *Stenodus* Richardson, 1836; with 91 valid species according to Fricke et al. 2022), and 3) the Thymallinae (genus *Thymallus* Linck, 1790; with 19 valid species according to Fricke et al. 2022, but see Knizhin, 2009, and Dyldin et al. 2017). Salmonids are famous for their diverse life-history strategies and the ecological and morphological diversity (Fleming and Reynolds, 2004), which makes them an interesting model in evolutionary biology. Some species such as the Atlantic Salmon (*Salmo salar* Linnaeus, 1758), the Rainbow trout (*O. mykiss* (Walbaum, 1792)) or the various Pacific salmon (*Oncorhynchus* spp.) are also of great importance for commercial and sport fisheries as well as aquaculture (Nelson, 2006) and have thus received considerable research interest.

Compared to these species, the genus *Thymallus* (graylings) is less well known, although morphological and genetic work in the last 20 years has greatly expanded the understanding of species diversity and distribution ranges (Weiss et al. 2002; Antonov, 2004; Stamford and Taylor, 2004; Knizhin et al. 2006a, 2007; Knizhin and Weiss, 2009; Persat et al. 2019). With few exceptions (e.g. populations of European grayling *T. thymallus* (Linnaeus, 1758) in the northern Bothnian Bay), graylings are exclusively freshwater fishes and distributed throughout much of the temperate and arctic regions in the northern hemisphere. Members of the genus mainly inhabit cool and oxygen-rich rivers and lakes, although some species are specifically adapted to lacustrine habitats (Knizhin et al. 2008a; Olson et al. 2019). During the spawning season in spring (usually between March and May), graylings migrate over short distances to suitable spawning habitats. Like most Salmoninae, graylings are gravel spawners, but they do not construct spawning redds or actively cover the eggs with substrate after fertilization (Fabricius and Gustafson, 1955). As r-selected species, female graylings usually produce several thousand eggs (Bishop, 1967) which develop in only a few weeks; about 130–140

degree-days in *T. thymallus* (see Haugen and Vøllestad, 2001), or 216.5–256.7 degree-days in Arctic grayling *T. arcticus* (Pallas, 1776) (see Bishop, 1971; Ward, 1951).

Due to their specific habitat requirements in all life stages, European grayling species are important biological indicators for assessing the status of fluvial ecosystems. They are particularly characteristic for the middle reaches of rivers (the hyporhithral), collectively referred to as the "grayling zone" (Illies, 1961). However, the ecological niche of European grayling species (*T. thymallus*, Adriatic grayling *T. aeliani* Valenciennes, 1848, and Loire grayling *T. ligericus* Persat, Weiss, Froufe, Secci-Petretto & Denys, 2019) is not identical with other representatives of the genus (Knizhin, 2009); some grayling species are characteristically found in small headwater streams (e.g. Baikal-Lena grayling *T. baicalolenensis* Matveev, Samusenok, Pronin & Tel'pukhovsky, 2005) or are specifically adapted to lacustrine environments (e.g. Chovsgul grayling *T. nigrescens* Dorogostaisky, 1923).

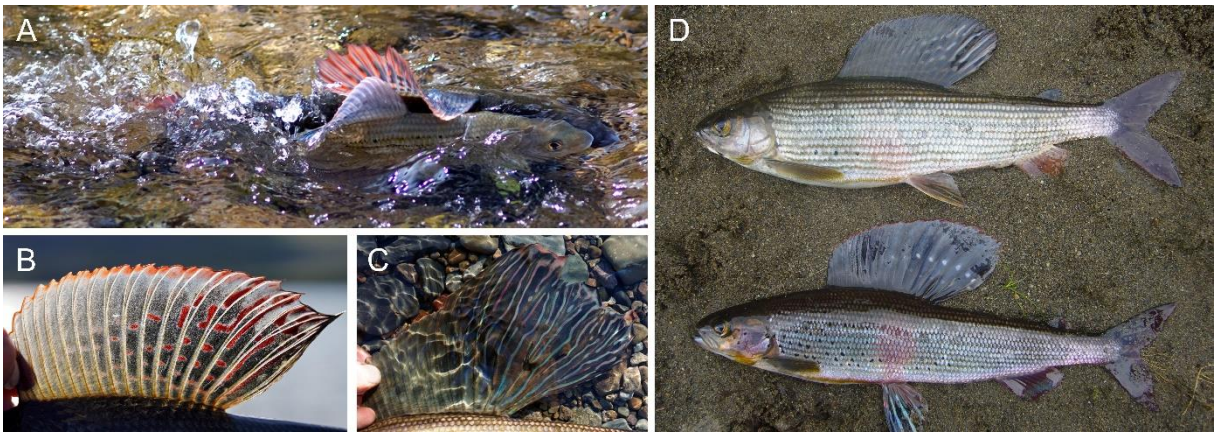



**Figure 1. A** spawning behaviour of *Thymallus*; male (back) and female (front) of *T. thymallus*. During spawning, male grayling bend sideward covering the dorsal region of the female with their enlarged dorsal fin and place their caudal peduncle on top the females'. Vigorous quivering by both sexes helps forcing the females' caudal region into the substrate where eggs are released and fertilized (Fabricius and Gustafson, 1955). **B** and **C** showing examples of the dorsal fin colouration pattern in two grayling species: *T. thymallus* (**B**) and *T. arcticus* (**C**). **D** two sympatric species of grayling – *T. burejensis* (above) and *T. grubii* (below) from the Bureya River, a tributary of the lower Amur River. Photos by Clemens Ratschan (**A**, **B**, **C**) and Alexander Antonov (**D**).

A characteristic feature of all grayling species is the large and colourful dorsal fin (Figure 1), which in some species and/or populations can extend beyond the origin of the adipose fin (Knizhin, 2009). In many species, the dorsal fin of males appears to be significantly larger and more colourful than in conspecific females (Persat, 1977; Semenchenko, 2005; Mikheev, 2009; Romanov, 2016), which points to its specific functions in male-male contests as well as courtship and spawning (Figure 1) where the display of the colourful dorsal fin, often together

with the pelvic fins, is an important behavioural element (Fabricius and Gustafson, 1955; Kratt and Smith, 1980). Outside the spawning season, the display of the dorsal fin is part of the intraspecific competition for feeding positions (Hughes and Dill, 1990; Hughes, 1992).

No conclusion has yet been reached on the phylogenetic relationships between distinct grayling species (Froufe et al. 2005, Weiss et al. 2020a) as well as the number of species within the genus (Knizhin, 2009, Dyldin et al. 2017) and currently 19 valid grayling species are recognized (Fricke et al. 2022). The type species of the genus, *T. thymallus*, originally described as *Salmo thymallus*, was commonly recognised as the only European species (Weiss et al. 2002). Subsequent morphological and genetic studies (Persat et al. 2019, Bravničar et al. 2020) supported the distinction of two additional species in Europe, *T. ligericus* and *T. aeliani*, rendering *T. thymallus* as paraphyletic (Marić et al. 2014). In the extreme North-East of Europe, *T. thymallus* forms a poorly defined contact zone with *T. arcticus* in river systems of the Pechora (Shubin and Zakharov, 1984) and Dvina (Koskinen et al. 2000). *T. arcticus* has probably the largest distribution range of all grayling species, extending from the Kola Peninsula in the North-West of Russia to North America (Knizhin, 2009). However, the taxonomic status of several nominal species possibly closely related to *T. arcticus* (e.g. *T. signifer* (Richardson, 1823), *T. pallasii* Valenciennes, 1848, *T. mertensii* Valenciennes, 1848, *T. ontariensis* Valenciennes, 1848, *T. tricolor* Cope, 1865, *T. montanus* Milner, 1874, *T. lewisi* Henshall, 1898) is uncertain (Knizhin, 2009).

The greatest diversity of grayling species is found in Siberia and the Russian Far East, especially in the Amur River drainage with five distinct species (*T. baicalolenensis*, Bureya grayling *T. burejensis* Antonov, 2004, Amur grayling *T. grubii* Dybowski, 1869, Yellow-spotted grayling *T. flavomaculatus* Knizhin, Antonov & Weiss, 2006, and Lower Amur grayling *T. tugarinae* Knizhin, Antonov, Safronov & Weiss, 2007) (Figure 1). Several studies have shown sympatric occurrence of multiple grayling species in the Bureya River (e.g. Weiss et al. 2020b), a tributary of the lower Amur River. Interestingly, hybridisation between these species seems to be rather rare (Weiss et al. 2020b). Otherwise, the contact zones between two or more grayling species are limited to small geographically defined areas and can be found, for example, in the lower reaches of the Yenisei River drainage (Weiss et al. 2007, Romanov, 2020) or in some tributaries of Lake Baikal (Knizhin et al. 2006b).

This dissertation aims to clarify questions on the evolutionary history and phylogenetic relationships in graylings. Both genomic (mitochondrial and nuclear genomic DNA) and morphological data are used to address three major objectives: 1) the phylogeography and systematic diversity of graylings based on whole mitochondrial genomes, which includes a time

calibrated phylogeny of the genus as well as an assessment of both the conservation and taxonomic status of each species; 2) the evaluation of general sexually dimorphic traits in graylings in a comparative phylogenetic approach, and in comparison to other representatives of the family Slamonidae; and 3) the phylogeny and phylogeography of graylings based on nuclear genomic data (ddRADseq), which includes the analysis of both a global dataset of all grayling species and a specific dataset of European grayling species.

# Chapter 1

## Global systematic diversity, range distributions, conservation and taxonomic assessments of graylings (Teleostei: Salmonidae; *Thymallus* spp.)

Steven J. Weiss, Duarte V. Gonçalves, Giulia Secci-Petretto, Gernot K. Englmaier, André Gomes-Dos-Santos, Gael P. J. Denys, Henri Persat, Alexander Antonov, Christoph Hahn, Eric B. Taylor, Elsa Froufe

**Citation:**

Weiss SJ, Gonçalves DV, Secci-Petretto G, Englmaier GK, Gomes-Dos-Santos A, Denys GPJ, Persat H, Antonov A, Hahn C, Taylor EB, Elsa F (2021). Global systematic diversity, range distributions, conservation and taxonomic assessments of graylings (Teleostei: Salmonidae; *Thymallus* spp.). Organisms Diversity & Evolution, 21, 25–42. https://doi.org/10.1007/s13127-020-00468-7

### Abstract

Graylings (*Thymallus*) are among the less well-studied groups of salmonid fishes, especially across their Asian distribution range. Here we perform a comprehensive global review of their phylogeography, systematic diversity and range distributions, including biogeographic reconstruction and assessment of both conservation and taxonomic status of each species. Based on a mitogenomic phylogenetic analysis, three approaches to the delineation of molecular operational units, and evaluation of 15 a-priori defined species, we provide biological support for the recognition of 13 grayling species, plus two additional species tentatively. Several instances of paraphyly and its potential effect on systematic inferences are discussed. Overall, the genus displays increasing species diversity and decreasing range size from higher to lower latitudes and ancestral trait reconstruction supports an East Asian origin for extant diversity, most likely centred in the Amur River drainage. Europe's colonization by *Thymallus* took place as early as the late Miocene, at least two colonisations of North America are supported, and multiple dispersal events likely took place into Western Siberia. The conservation status for the 15 taxa was estimated to be: 6 least concern, 1 near-threatened, 2 vulnerable, 3 endangered and 3 data deficient.

**Keywords:** Mitogenome phylogeny, Species delineation, Conservation assessment, Time-calibrated phylogeny, Siberian biogeography, Grayling

## Introduction

Salmonidae contains many of the most economically and ecologically important temperate freshwater fishes, which, due to their common tetraploid ancestry (Allendorf and Thorgaard 1984), attract considerable attention in evolutionary research. The subfamily Thymallinae (graylings), considered a distinct family by Skurikhina et al. (1985) and Osinov and Lebedev (2000), is less well known, although the recognized species diversity has increased considerably over the past 15 years (Antonov 2004; Knizhin et al. 2006a, 2007; Knizhin and Weiss 2009). Additionally, increased conservation concerns over European grayling *Thymallus thymallus* (Dawnay et al. 2011; Weiss et al. 2013; Mueller et al. 2018), as well as its recent whole-genome sequencing (Varadharajan et al. 2018; Sävilammi et al. 2019), have opened the doors to more focused evolutionary and conservation management research (Papakostas et al. 2014; Mäkinen et al. 2016; Huml et al. 2018). However, a pre-requisite to detailed evolutionary research and conservation is a clear understanding of species diversity, systematic relationships and distributions. The last comprehensive phylogenetic analysis of *Thymallus* (Froufe et al. 2005) became obsolete due to the description of a number of new species (e.g. *T. ligericus*) (Table 1). A recent annotated checklist of grayling species diversity noted that several species are of questionable status, or require additional research (Dyldin et al. 2017).

The majority of grayling species occur in Asia, occupying some of the world's largest river drainages. Siberia's complex paleo-hydrology has shaped the diversity and distribution of grayling throughout Eurasia (Koskinen et al. 2002; Froufe et al. 2003a, 2005). This is illustrated by the phylogeographic patterns in the Altai-Sayan mountain region (Knizhin and Weiss 2009; Weiss et al. 2020a) or the three sympatric species with allopatric origins found in the Bureya River, Amur River drainage (Weiss et al. 2020b). High phylogeographic complexity is also found in Europe within the Danube River drainage as well as the larger Ponto-Caspian basin (Weiss et al. 2002, 2013; Gum et al. 2009; Marić et al. 2012, 2014).

Regional phylogeographic studies have further contributed to understanding within-genus diversity. For instance, Knizhin et al. (2008a) demonstrated that different grayling phenotypes in the endorheic basin of Western Mongolia can be phylogenetically assigned to *T. brevirostris*. *Thymallus brevicephalus* found in the upper Irtysh River drainage in Kazakhstan is the sister taxon to *T. brevirostris* (Weiss et al. 2020a). While such studies combined single-gene mtDNA

analysis with population genetic and morphological analysis, they may lack phylogenetic resolution due to the limited length of the mtDNA sequence analysed. Population genetic investigations of grayling across their range support that most, if not all species thus far described are either allopatric or when found in sympatry show little to no introgression (Koskinen et al. 2002; Froufe et al. 2003b; Weiss et al. 2006, 2007, 2020b). Thus, it is likely that mtDNA-based phylogenetic analyses of grayling reflect evolutionary relationships of the entire genome. Mitochondrial genomes (henceforth 'mitogenomes') present a significant improvement in resolution over analyses based on one or a few mtDNA genes (Miya and Nishida 2015; He et al. 2018; Boo and Hughey 2019). Low coverage genome-wide short-read sequencing can be used to recover mtDNA even from relatively degraded samples, due to the high mtDNA copy number per cell (Miller et al. 2012; Liedigk et al. 2015). Despite such advantages, to date, only a single application of mitogenomes to grayling phylogeny was carried out (Ma et al. 2016), whereby only eight mitogenomes were sequenced across seven species.

DNA-based species delineation methods have become standard tools for delineating molecular operational taxonomic units (MOTUs) (e.g. Lopes-Lima et al. 2019) or uncovering potentially cryptic species (e.g. Pan et al. 2019) in many organisms, including freshwater fishes (Patil et al. 2018; Berbel-Filho et al. 2018; Corral-Lou et al. 2019). The results of such methods are dependent on the choice of samples and loci (Ritchie et al. 2016; Sukumaran and Knowles 2017) and must be carefully interpreted. However, for a widespread genus like *Thymallus*, such analyses should allow a standardized evaluation of historically described taxa, which have received limited attention from modern molecular analyses.

In this context, a phylogenetic analysis of the genus *Thymallus* was carried out, using newly obtained mitogenomes via low-coverage whole-genome sequencing, together with already available mitogenome sequences. Using comprehensive taxon sampling and literature review, we aimed to: (a) produce an updated robust phylogeny of graylings across their global range; (b) evaluate current systematics using a set of species delineation methods; (c) provide an overview of the distribution and conservation status of each species; (d) reconstruct ancestral biogeographic patterns and (e) provide an up-to-date taxonomic overview. More broadly, we aim to provide a baseline for future research on evolutionary patterns, taxonomic diversity, conservation and biogeography of grayling throughout the world.

**Table 1** Overview of *Thymallus* species following the latest review of the genus by Dyldin et al. (2017), including synonyms listed in Fricke et al. (2020) (see also Dyldin et al. 2017 and references therein), type localities and location of type specimens (where known). Drainage systems of type localities are given in parentheses.

| Species | Synonyms | Type locality | Types |
|---|---|---|---|
| *Thymallus aeliani* Valenciennes, 1848 | | Lago Maggiore [Lac Majeur] (Po), Italy | unknown |
| *Thymallus arcticus* (Pallas, 1776) | | Tributaries of the Sob River (Ob), Russia | unknown |
| | *Salmo digitalis* Bloch & Schneider, 1801 | unknown | unknown |
| *Thymallus baicalensis* Dybowski, 1874 | *Thymallus grubii* var. *baicalensis* Dybowski, 1874 | Lake Baikal [Baical], Selenga [Sielenga] and Angara rivers (Enisei), Russia | BMNH 1871.7.19.3, BMNH 1897.7.5.20 |
| *Thymallus baicalolenensis* Matveev, Samusenok, Pronin & Tel'pukhovsky, 2005 | *Thymallus arcticus baicalolenensis* Matveev, Samusenok, Pronin & Tel'pukhovsky, 2005 | Streams and lakes in North Baikal (Enisei) and the upper Lena River (Lena), Russia | ZMISU V-19 |
| *Thymallus brevicephalus* Mitrofanov, 1971 | *Thymallus arcticus brevicephalus* Mitrofanov, 1971 | Lake Markakol (Ob), Kazakhstan | unknown |
| *Thymallus brevipinnis* Svetovidov, 1931 | *Thymallus arcticus brevipinnis* Svetovidov, 1931 | Lake Baikal (Enisei), Russia | unknown |
| *Thymallus brevirostris* Kessler, 1879 | | Tributaries of Daingol [Dayan Nuur] (Khovd), Dsabchyn [Zavkhan] River (Zavkhan), Mongolia | ZIN 4212–13 |
| | *Phylogephyra altaica* Boulenger, 1898 | South side of Altai Mountains (unknown drainage) China | BMNH 1898.2.17.1 |
| | *Thymallus brevirostris altaicus* Dashdorj, Dulmaa & Tsendayush, 1968 | Lakes Khoton and Khorgon (Khovd), Mongolia | unknown |
| | *Thymallus brevirostris kozovi* Dashdorj, Dulmaa & Tsendayush, 1968 | Lakes Khoton and Khorgon (Khovd), Mongolia | unknown |
| *Thymallus burejensis* Antonov, 2004 | | Levaya Bureya (Amur), Russia | MGU r-20 928, MGU r-20 929 |
| *Thymallus flavomaculatus* Knizhin, Antonov & Weiss, 2006 | *Thymallus grubii flavomaculatus* Knizhin, Antonov & Weiss, 2006 | Anyui River (Amur), Russia | ZMISU R-3–R-4 |

| | | | |
|---|---|---|---|
| *Thymallus grubii* Dybowski, 1869 | | Onon and Ingoda rivers (Amur), Russia | BMNH 1892.11.24.6 |
| | *Thymallus arcticus yaluensis* Mori, 1928 | Upper Yalu River at Kozan (Yalu), North Korea | unknown |
| *Thymallus ligericus* Persat, Weiss, Froufe, Secci-Petretto & Denys, 2019 | | Alagnon River at La Chapelle d'Alagnon (Loire), France | MNHN 2018-0722–2018-0728, MNHN 2019-0266–2019-0267 |
| *Thymallus mertensii* Valenciennes, 1848 | | Kamtchatka (unknown drainage), Russia | unknown |
| *Thymallus nigrescens* Dorogostaisky, 1923 | *Thymallus arcticus nigrescens* Dorogostaisky, 1923 | Lake Khuvsgul [Hövsgöl, Hovsgol] (Enisei), Mongolia | unknown |
| *Thymallus nikolskyi* Kaschenko, 1899 | | Rybnusska stream at Ryblushka, Urusul River at Ongudai, Tcharysh River at Ust-Kan, Katun River at Nizhnii Uimon, Lake Talmenie, Tom River above Kusnetsk (Ob), Russia | unknown |
| | *Thymallus nikolskyi* var. *ongudajensis* Kaschenko, 1899 | Urusul River at Ongudai (Ob), Russia | unknown |
| | *Thymallus sellatus* Kaschenko, 1899 | Lake Tenga [Kenga], Ursul River drainage (Ob), Russia | unknown |
| *Thymallus pallasii* Valenciennes, 1848 | | Russia (unknown drainage) | MNHN 0000-3664; ZMB 23549–23550 |
| *Thymallus signifer* Richardson, 1823 | | Rivers north of Great Slave Lake (Mackenzie), Canada | unknown |
| | *Coregonus thymalloides* Richardson, 1823 | Winter River, north of Great Slave Lake (Mackenzie), Canada | unknown |
| | *Thymallus ontariensis* Valenciennes, 1848 | Lake Ontario (Saint Lawrence), U.S.A. | MNHN 0000-3662–0000-3663 |
| | *Thymallus tricolor* Cope, 1865 | Michigan (unknown drainage), U.S.A. | ANSP 7796; UMMZ 157347, 213814; USNM 23217, 37862 |
| | *Thymallus montanus* Milner, 1874 | Camp Baker, tributary of the Missouri River, Montana Territory (Mississippi), U.S.A. | USNM 13090–13091 |

| | | | |
|---|---|---|---|
| | *Thymallus lewisi* Henshall, 1898 | Jefferson, Madison, and Gallatin rivers (Mississippi), U.S.A. | unknown |
| *Thymallus svetovidovi* Knizhin & Weiss, 2009 | | Sharga Gol River (Enisei), Mongolia | ZMMU R-21992–R-21993; ZMISU R-8 |
| *Thymallus thymallus* (Linnaeus, 1758) | | Coastal rivers of Europe (unknown drainage) | BMNH 1853.11.12.159 |
| | *Salmo thymus* Bonnaterre (ex Salviani), 1788 | Rivers of England and Germany (unknown drainage) | unknown |
| | *Salmo striatus* Reisinger, 1830; preoccupied by *Salmo striatus* Bloch & Schneider, 1801 | Liptoviae [Liptov] (Danube), Hungary (now Slovakia) | unknown |
| | *Thymallus vexillifer* Fitzinger (ex Agassiz), 1832 | Schwarza, Fischer, and Traun rivers (Danube), Austria | unknown |
| | *Thymallus vulgaris* Nilsson, 1832 | Dalelfven [Dalälven], Klarelfven [Klarälven], Glommen [Glomma], and Laugenelf rivers (various drainages), Sweden and Norway | unknown |
| | *Thymallus decorus* Koch, 1840 | Laaber [Laber] River (Danube), Germany | unknown |
| | *Thymalus gymnogaster* Valenciennes, 1848 | Newa [Neva] River at St. Petersburg (Neva), Russia | MNHN 0000-3665 |
| | *Thymalus gymnothorax* Valenciennes, 1848 | Berlin market (unknown origin), Germany? | unknown |
| | *Thymallus umbrosa* Gistel, 1848 | Brooks and rivers (unknown drainage), Germany? | unknown |
| | *Salmo punctatus* Gronow, 1854 | Germany (unknown drainage) | unknown |
| | *Thymallus thymallus kamensis* Lukasch, 1929 | Upper reaches of the Kama River (Volga), Russia | unknown |
| *Thymallus tugarinae* Knizhin, Antonov, Safronov & Weiss, 2007 | | Anyui River (Amur), Russia | ZMISU P-1; ZMISK P-2, P-5–P-6 |

*ANSP* Academy of Natural Sciences of Philadelphia, *BMNH* British Museum of Natural History, *MGU/ZMMU* Zoological Museum of Moscow State University, *MNHN* Muséum national d'histoire naturelle, *UMMZ* Museum of Zoology University of Michigan, *USNM* the United States National Museum, *ZMISK/ZMISU* Zoological Museum Irkutsk State University, *ZIN* Zoological Institute Russian Academy of Sciences

## Methods

### DNA extraction and sequencing

Whole genomic DNA of 47 grayling representing species across their global range was extracted from fin-clips stored in 96% ethanol using a high-salt protocol (Sambrook et al. 1989) (Table S1, Electronic Supplementary Material). Sample choice was influenced by the availability of published mitogenomes, aiming for a data set covering the global diversity of the genus. Included are most species listed in Dyldin et al. (2017) and the newly described *T. ligericus* (Persat et al. 2019) (Table 1). For each sample, Illumina paired-end libraries (TrueSeq DNA PCR-Free kit) were prepared and ~2 Gb of raw paired-end 150 bp (PE150) reads were sequenced on an Illumina Platform at Macrogen (Seoul, South Korea) and Novogene (UK, Europe).

### Mitogenome assembly, annotation and characterization

Whole mitogenome assemblies were performed using a seed and an extend algorithm in NOVOplasty v2.7.2 (Dierckxsens et al. 2017), and the baiting and iterative approach in MITObim v1.9.1 (Hahn et al. 2013) following the authors' guidelines and using default parameters. Additionally, six mitogenomes of *T. thymallus* and eight from *T. ligericus* were produced with the long-range PCR protocol described in Denys et al. (2020a). The 61 new mitogenomes were annotated with MitoAnnotator (Iwasaki et al. 2013) via the webinterface (http://mitofish.aori.u-tokyo.ac.jp/, last accessed 28.06.2019). Annotations were converted to standard feature table format using a custom script (https://github.com/chrishah/MitoFish2tbl) and deposited to GenBank (MT062993-MT063053) and validated by comparison to published data (Table S1).

### Phylogenetic analyses

The sequenced mitogenomes were aligned in MAFFT (Katoh and Standley 2013) together with mitogenomes retrieved from GenBank for 31 *Thymallus* spp. and three outgroup taxa (Table S1). For subsequent analyses, the mtDNA control region (CR) was removed due to several repeat motifs. The best partition scheme for each gene or group of genes was estimated using PartitionFinder2 v2.1.1 (Lanfear et al. 2012). Maximum likelihood analyses (ML) were

performed with the program RAxML-HPC2 Workflow v8.2.10 (Stamatakis 2016), using the general time reversible substitution model, with a gamma parameter and proportion of invariable sites (GRT + I + G) and 1000 bootstrap replicates. Bayesian phylogenetic inference (BI) was carried out in MrBayes v3.2.6 (Ronquist and Huelsenbeck 2003), using the best substitution models previously obtained in PartitionFinder2. Two independent runs were carried out (107 generations, one tree sampled every 100 generations), with default chains. PartitionFinder2, MrBayes and RAxML were run in CIPRES Science Gateway (Miller et al. 2010). Tracer v1.7.1 (Rambaut et al. 2018) was used to assess the effective sample size values (ESS) and determine the appropriate burn-in.

Species delineation methods

Two distance-based and one tree-based method were applied to determine molecular operational taxonomic units (MOTUs): the cluster tool implemented in BOLD (Ratnasingham and Hebert 2007), the Automatic Barcode Gap Discovery (ABGD) (Puillandre et al. 2012) using both an initial ABGD(i) and a recursive ABGD(r) partition, and an improved multi-rate version (mPTP; Kapli et al. 2017) of the Poisson Tree Processes (Zhang et al. 2013). The BOLD method accepts COI sequences only, but ABGD and mPTP were applied to the mitogenome dataset. For BOLD, COI sequences were retrieved from aligned mitogenomes and uploaded on the BOLD platform. ABGD was run online (https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html; last accessed 20.02.2020), on the same dataset used for the phylogeny. The relative gap width was set to 1.0 and uncorrected p distances were used as a substitution model; remaining parameters were set as default. Finally, the mPTP method was applied to the Bayesian tree (without outgroups) using an online platform (https://mptp.h-its.org/#/tree; last accessed 23.01.2020). Samples were also assigned to species based on a combination of phenotypic identification, type locality, a-priori range knowledge and GenBank annotations. Uncorrected p distances were calculated among species or MOTUs for each delineation method (putative species, BOLD, ABGD (i), ABGD(r), mPTP using MEGA v10.0.5 (Kumar et al. 2018)) and frequency histograms of each were made using SPSS v26. For these calculations, four taxa listed in Table 1 were excluded due to the lack of evidence supporting their biological meaningfulness and/or our inability to assign a sample to them due to ambiguity concerning their type localities. Three of these taxa (*T. mertensii*, *T. pallasii*, *T. signifer*) are herein grouped within *T. arcticus* sensu lato (s.l.) and all are discussed in the results and discussion as well as Appendix 1 (Electronic Supplementary Materials). MOTU delineation

was used as one line of potential evidence (sensu de Queiroz 2007) to support the existence of species. MOTU support by multiple algorithms together with population genetic and phylogeographic data, phenotypic distinction and ecological niche occupation were used to infer a taxon's biological recognition. Lack of MOTU support was not necessarily used to argue against species recognition if multiple lines of other biological evidence supporting a species status exist.

Time-calibrated phylogeny

A time-calibrated phylogeny was produced using BEAST v1.10.4 (Drummond and Rambaut 2007; Suchard and Rambaut 2009) in CIPRES science gateway (Miller et al. 2010). Mitogenomes of eight individuals from Salmoninae and Coregoninae were added to the dataset as outgroups needed for the calibration scheme. The best partition scheme was selected as mentioned above. BEAST analysis was done with unlinked substitution and clock models, and a linked tree model, using uncorrelated relaxed molecular clock priors with log-normal distributions (Drummond et al. 2006). The birth-death speciation process (Gernhard 2008) was chosen as a tree prior, given the dataset comprised intra- and inter-specific relationships. One calibration point was set at 50 MY for the most recent common ancestor (MRCA) of Salmonidae, using the fossil †*Eosalmo driftwoodensis* Wilson, 1977 as a minimum time constraint for the family (Crête-Lafrenière et al. 2012; Lecaudey et al. 2018). Prior parameters were implemented following Lecaudey et al. (2018) (lognormal distribution, offset: 50, mean: 10, SD: 1). Another calibration point was set using the expansion of *T. baicalensis* into Lake Baikal (Koskinen et al. 2002), adjusted to 0.13 MY (Normal distribution, mean: 0.12, SD: 0.1), following the most recent dating of the 'Angara breakout' (Arzhannikov et al. 2018).

A third calibration point was set at 7.6 MY (stem dating; lognormal distribution; mean: 7.6; SD: 1.5) based on fossil remains of †*Thymallus latisulcatus* Rückert-Ülkümen & Kaya, 1993 found in Yalova (otolith used for species description) and Yalakdere, Turkey (Rückert-Ülkümen and Kaya 1993). The fossils were located in the strata corresponding to the Khersonian-Maeotian transition, whose date was recently revised to 7.6 MY (Lazarev et al. 2020; Palcu et al. 2019). Since this fossil may or may not correspond to a direct ancestor of the extant European lineages, the calibration was placed on the stem. The substitution rate prior for each partition was set with a normal distribution and 1%/MY (average 0.01 and 0.004 SD) following the mtDNA molecule calibration for salmonids (Smith 1992) and the reported lower bound of the *T. baicalensis* expansion (Koskinen et al. 2002). Additional Quaternary fossils of

*T. arcticus* have been reported from North America (Cumbaa et al. 1981; McAllister and Harington 1969), but these were not included, since it was not possible to determine with certainty where these fossils would fall in the phylogeny. They could represent an ancestor of all North America populations, or only one of the lineages, or be placed within one of the lineages. Since the North American lineages are not monophyletic, using these fossils to delimit a minimum or maximum age for the split of both NA lineages would also require assuming where the split took place, and such an assumption might be less than informative. Analysis was performed with three 30 million MCMC iteration runs, sampling every 3000 runs. Burn-in and run convergence (ESS > 200) were determined using Tracer v1.7.1 (Rambaut et al. 2018). Independent runs were joined in LogCombiner v1.10.4 and the final tree produced in TreeAnnotator v1.10.4 (both in the BEAST package) and formatted in FigTree v1.4.4 (Rambaut 2012).

Species distributions and taxonomic review

Approximate species distributions were estimated based on both published and unpublished information (see Fig. 1). The area of occupancy (AOO) was calculated using ArcMap v10.7 (ESRI 2019) and hydrological network data from HydroAtlas v1.0 (Linke et al. 2019), as the total hydrological network area of lakes plus rivers larger than Strahler level three, a more realistic measure than a 2 × 2 km grid overlay (see Gomes-dos-Santos et al. 2019). Based on literature review, and present results, we comment on the biological status of each taxon, and review their taxonomic and conservation status based on IUCN criteria.

Historical biogeography and ancestral area reconstruction

To reconstruct the ancestral ranges of extant *Thymallus* diversity, we performed a Bayesian phylogeographic inference in discrete space using the Bayesian Stochastic Search Variable Selection (BSSVS; Lemey et al. 2009) implemented in BEAST. Considering both the current and paleohydrological drainage patterns, we considered seven biogeographic regions: Europe (west of Urals), Western Siberia (Kara Sea basin); East Siberia (north-flowing rivers, and Bering Sea); East Asia excluding the Amur (Okhotsk, Japan, Yellow and East China sea drainages); Amur River drainage; North America and the endorheic drainages of Western Mongolia. We used the same dataset, models and prior settings as in the dating analysis. BSSVS

parameters were as follows: symmetric substitution model, strict location-trait clock and an exponential distribution for the location rates prior (mean = 1, offset = 0).

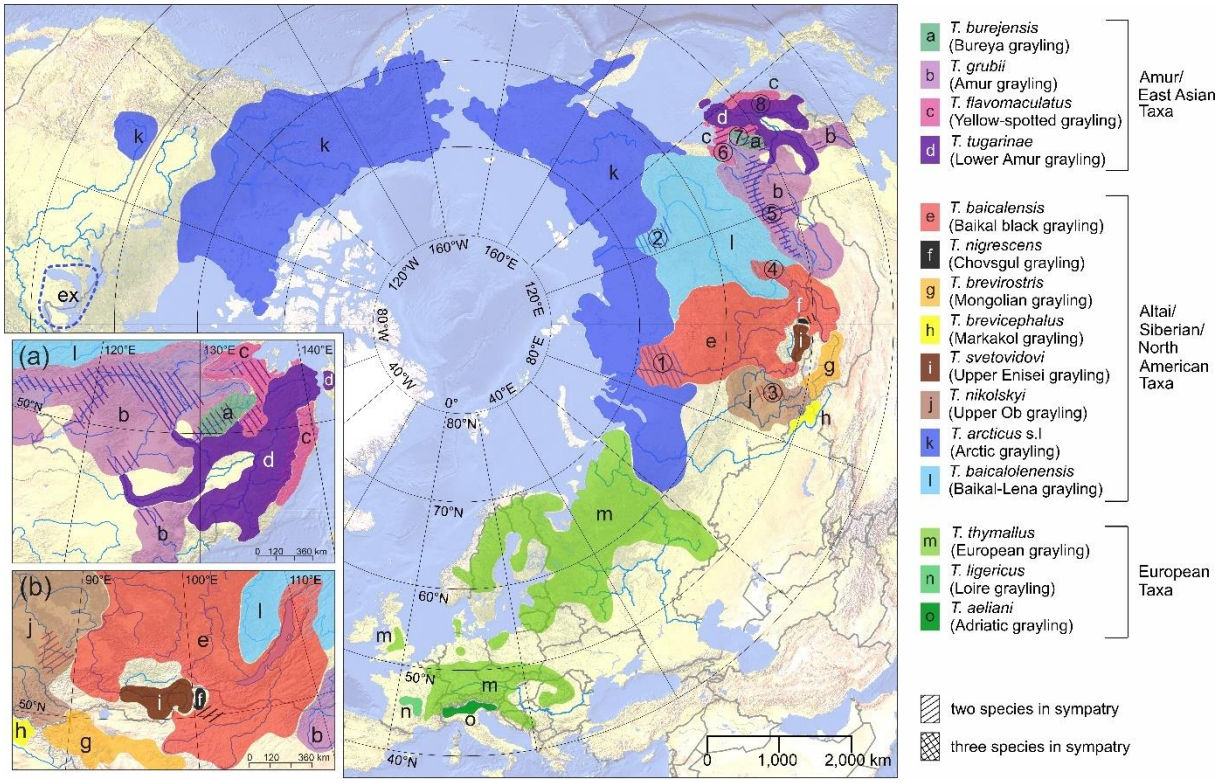



**Fig. 1** Map showing the global distribution range of *Thymallus* species. Information on sampling sites and species is given in Table S1. Numbers in the map refer to known contact zones of the following species: 1 = *T. arcticus* s.l. and *T. baicalensis* in the lower Enisei River; 2 = *T. arcticus* s.l. and *T. baicalolenensis* in the lower Lena River; 3 = *T. nikolskyi* and *T. baicalensis* in tributaries of the upper Ob River; 4 = *T. baicalolenensis* and *T. baicalensis* in tributaries of Lake Baikal; 5 = *T. grubii*, *T. tugarinae* and *T. baicalolenensis* (not all found in contact throughout this zone) in the upper and middle Amur drainage (including its tributaries such as the Ingoda, Argun and Zeya rivers); 6 = *T. flavomaculatus*, *T. tugarinae* and *T. baicalolenensis* in the Uda River drainage, 7 = *T. burejensis*, *T. grubii* and *T. baicalolenensis* in the upper Bureya River; 8 = *T. tugarinae* and *T. flavomaculatus* in tributaries of the lower Amur River.

## Results

### Mitogenome analyses

Low-coverage genome sequencing recovered mitogenomes for all 47 samples analysed. With on average 18 million genomic reads per sample, the number of mitochondrial reads ranged from almost 3 to more than 50 thousand per sample for an organelle coverage ranging from 25x to 390x (Table S2). Reads and circularized molecules were deposited on GenBank (MT062993-MT063053, BioProject: PRJNA604892) (Table S2). The mitogenomes included 13 protein-

coding genes, 22 tRNAs, 2 rRNAs and the CR (see Fig. S1, Electronic Supplementary Material). Average nucleotide composition and gene order were consistent across all individuals (Fig. S1).

Phylogenetic analyses and species delineation

ML and BI analyses showed the same topology; thus, only the Bayesian phylogenetic tree is represented (Fig. 2). The tree depicts three main groups of species, with the first split separating East Asian species (excluding *T. burejensis*) from all others, and the next split separates the three accepted European species from those in Siberia and the Altai, including *T. burejensis*. Pairwise divergence among a priori defined species ranged from 0.3% for *T. nigrescens* and *T. baicalensis*, to 5.5% between *T. tugarinae* and *T. aeliani* (Table S3). Net uncorrected p distances among all MOTUs revealed a bi-modal distribution, with a minor mode at 0.1–0.2% and a major mode centred on approximately 5% sequence divergence (Fig. S2). Congruence between well-supported clades, all delineation algorithms and a priori species assignment occurred for *T. baicalolenensis*, *T. burejensis*, *T. aeliani* and *T. ligericus*. *Thymallus baicalensis* was not monophyletic due to the placement of *T. nigrescens* within *T. baicalensis*. *Thymallus tugarinae* was also not completely congruent due to a divergent haplotype from the Kievka River, which drains to the Sea of Japan (KY078218).

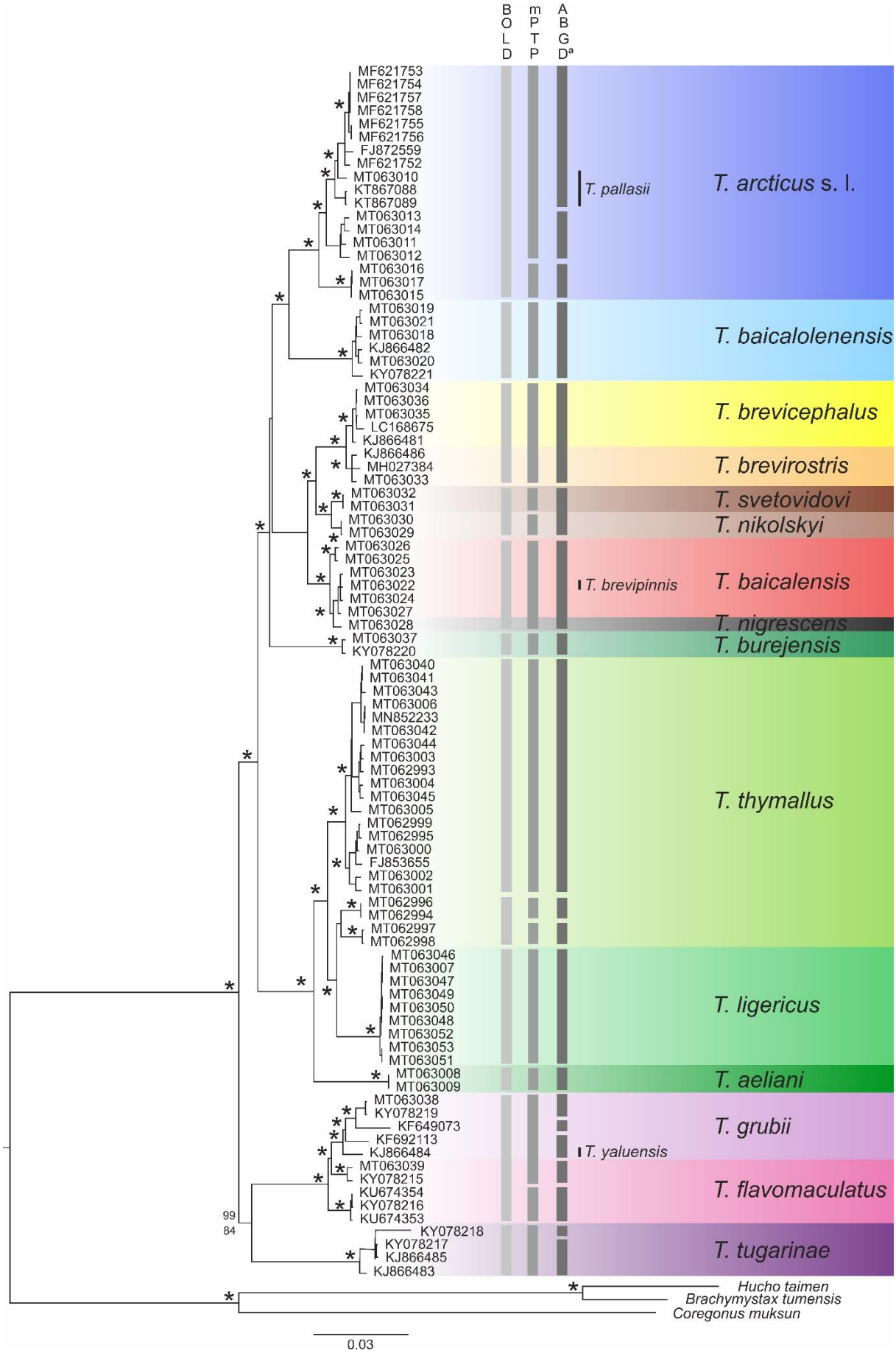
The most conservative delineation algorithm was BOLD, recognizing 12 MOTUs. The two ABGD schemes recognized 21 MOTUs, and the mPTP recognized 16 MOTUs. Paraphyletic relationships were inferred for *T. thymallus* and *T. flavomaculatus*. The former paraphyly results from a well-supported clade representing Alpine haplotypes from the upper Danube (Weiss et al. 2002, 2013) resolved as the sister clade (1.5% divergent) of *T. ligericus* and delineated as one (BOLD) or two (mPTP and ABGD) MOTUs. The node containing *T. grubii*, *T. flavomaculatus* and *T. yaluensis* represents the common ancestor of *T. flavomaculatus*. The earliest split and most divergent clade within *T. arcticus* s.l., delineated as a MOTU by mPTP and ABGD, was represented by haplotypes (MT063015-017) from north-eastern British Columbia in Canada (Table S1). No additional Arctic taxa (e.g. *T. pallasii*, *T. mertensii* or *T. signifer*) (see Appendix 1) could be recognized by MOTU delineation, though, as noted above, there is ambiguity concerning the assignment of these taxa to a location or sample in our data set. Phylogenetic analysis resolved the sister species *T. brevirostris* and *T. brevicephalus* (0.4% divergent), but not MOTU delineation. Similarly, phylogenetic analysis

resolved the sister species *T. nikolskyi* and *T. svetovidovi* (0.6% divergent), but only the mPTP algorithm delineated them as distinct MOTUs.

Time calibration and ancestral range reconstruction

The MRCA of Salmonidae was dated at 59.3 MY, at the end of Palaeocene (Fig. 3) while that for *Thymallus* was 9.11 MY, at the end of Miocene. The MRCA of three East Asian species, *T. tugarinae*, *T. flavomaculatus* and *T. grubii*, was estimated at 8.42 MY also in the late Miocene, similar to the split between European species and all others (7.54 MY), whereas the MRCA of all European species was estimated at 3.86 MY corresponding to the mid-Pliocene. The next oldest node (6.35 MY) represented the split between *T. burejensis* and remaining species. The estimated MRCA of *T. arcticus* s.l. and *T. baicalolenensis* 5.01 MY and that of all *T. arcticus* was 2.39 MY, in the Pliocene and early Pleistocene, respectively. The splits between two pairs of allopatric species in the Altai-Sayan mountain region [*T. brevirostris* and *T. brevicephalus* (0.85 MY), and *T. svetovidovi* and *T. nikolskyi* (0.77 MY)] were both dated to the mid-Pleistocene. The Amur River drainage was the most likely ancestral location of origin of the extant grayling diversity, with much higher support than every other option (Fig. 3). A single colonization of Europe from Western Siberia was predicted, but multiple exchanges may have occurred among the other considered regions, exemplified by the three-basin distribution of the young *T. baicalolenensis*, or two predicted ancestors of North American lineages over 1 MY apart.

**Fig. 2** Bayesian phylogenetic reconstruction based on whole mitogenome dataset (excluding the CR). Branch support values represent Bayesian posterior probabilities (above nodes) and maximum likelihood values (below nodes). An asterisk symbolizes that both values were ≥ 90%. To the right of the tree shown in grey bars are the results of three MOTU delineation methods. Each bar represents one MOTU: 12 for BOLD, 16 for mPTP and 21 for ABGDa (indicates both initial and recursive partition). Taxa with a smaller font positioned right of the MOTU bars represent names in use for which there is no biological support and thus may represent junior synonyms. Species are coloured according to Fig. 1.


BOLD
mPTP
ABGD[a]
MF621753
MF621754
MF621757
MF621758
MF621755
MF621756
FJ872559
MF621752
MT063010
KT867088
KT867089
MT063013
MT063014
MT063011
MT063012
MT063016
MT063017
MT063015
T. pallasii
T. arcticus s. l.
MT063019
MT063021
MT063018
KJ866482
MT063020
KY078221
T. baicalolenensis
MT063034
MT063036
MT063035
LC168675
KJ866481
T. brevicephalus
KJ866486
MH027384
MT063033
T. brevirostris
MT063032
MT063031
T. svetovidovi
MT063030
MT063029
T. nikolskyi
MT063026
MT063025
MT063023
MT063022
MT063024
MT063027
T. brevipinnis
T. baicalensis
MT063028
T. nigrescens
MT063037
KY078220
T. burejensis
MT063040
MT063041
MT063043
MT063006
MN852233
MT063042
MT063044
MT063003
MT062993
MT063004
MT063045
MT063005
MT062999
MT062995
MT063000
FJ853655
MT063002
MT063001
MT062996
MT062994
MT062997
MT062998
T. thymallus
MT063046
MT063007
MT063047
MT063049
MT063050
MT063048
MT063052
MT063053
MT063051
T. ligericus
MT063008
MT063009
T. aeliani
MT063038
KY078219
KF649073
KF692113
KJ866484
T. yaluensis
T. grubii
MT063039
KY078215
KU674354
KY078216
KU674353
T. flavomaculatus
99
84
KY078218
KY078217
KJ866485
KJ866483
T. tugarinae
Hucho taimen
Brachymystax tumensis
Coregonus muksun
0.03

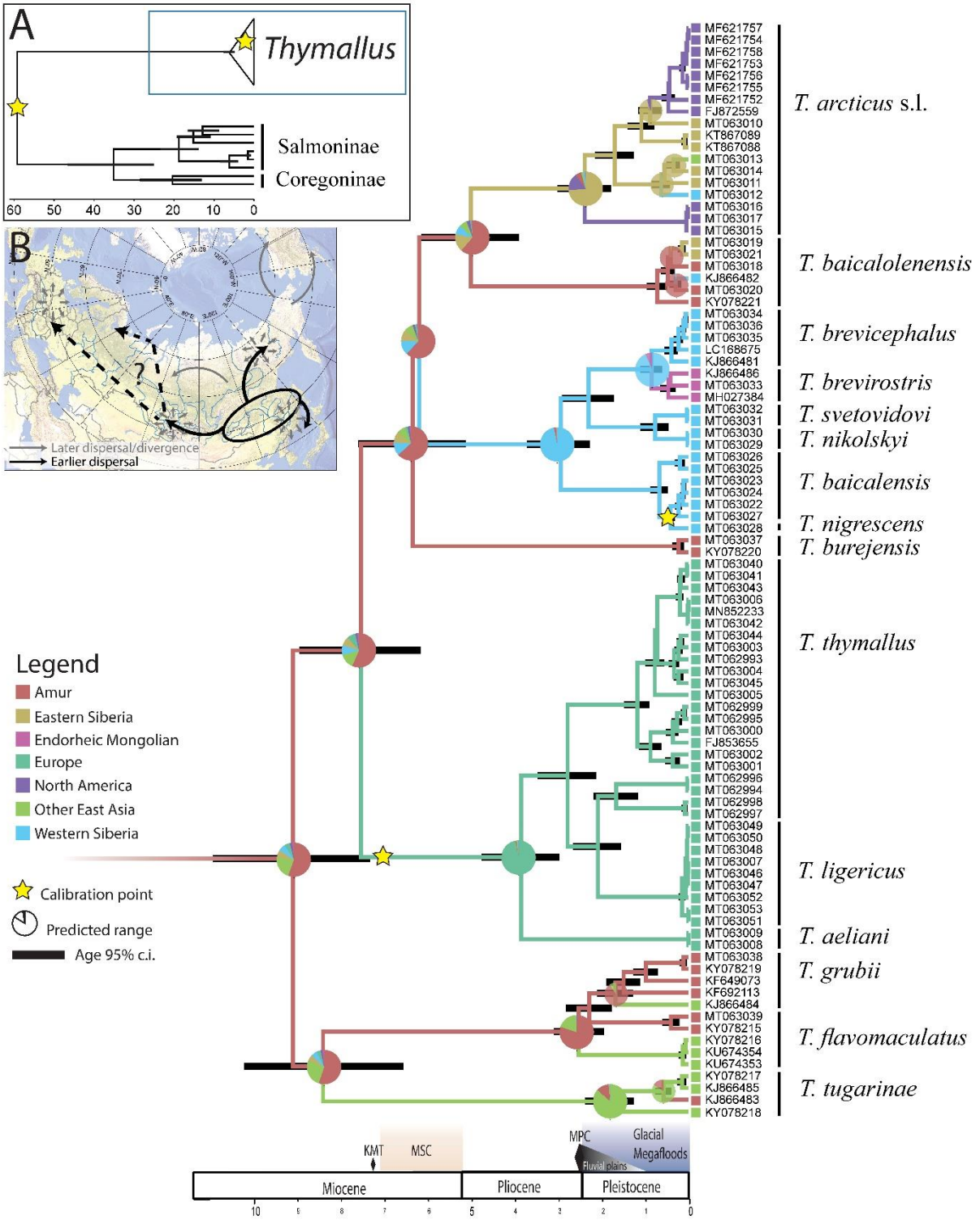



**Fig. 3** Time-calibrated phylogeny and biogeographic analysis of *Thymallus* spp. inferred from the whole mitogenome (excluding the CR). Three fossil calibration points (panel a) and a biogeography-based substitution rate were used (details in text). Ancestral ranges were reconstructed as a discrete trait following the phylogeographic diffusion BSSVS method; nodes and branches are colour-coded according to region. Hypothesised dispersal events are shown in panel b. Potentially relevant paleo-climatic events such as the Maximum Ponto-Caspian extent (MPC), the following Fluvial Plains Period and the Eurasian Glacial Cycles are represented at the bottom. MSC, Messinian salinity crisis; KMT, Khersonian-Meotian transition.

## Discussion

### Species delineation

Our results strongly support the biological validity of at least 13 grayling species globally, plus two additional species tentatively all but one supported by at least one MOTU delineation approach, the exception being *T. nigrescens*. Microsatellites and single mtDNA sequences allow no distinction between *T. nigrescens* and *T. baicalensis* (Koskinen et al. 2002; Kaus et al. 2019). Next-generation sequencing RAD data supported a 15–30 thousand year divergence between the two taxa (Roman et al. 2018). The taxon is morphologically distinct from *T. baicalensis* based on a small dorsal fin and a higher number of gill rakers (Knizhin and Weiss 2009; Olson et al. 2019), and thus, its occupation of a distinct ecological niche supports its status as a distinct species despite relatively low genetic divergence.

Mean pairwise distances among species (3.7%) conformed to minimum thresholds applied in other studies. The majority (> 90%) of vertebrate species differ from their nearest neighbour by > 2% (Avise et al. 1999) or > 3% (Hebert et al. 2004) using standard mtDNA gene markers. Such thresholds have been noted in many data sets involving fishes (e.g. Ward 2009; April et al. 2013). Our divergence values may be comparatively low as we use uncorrected p distances instead of model-corrected distances (see Srivathsan and Meier 2012), salmonid mtDNA exhibits a relatively slow substitution rate (ca. 1%/MY) (Smith 1992; Froufe et al. 2005) and our calculations include non-protein regions of the mitogenome (excl. the CR). Higher values are obtained when using protein-coding genes alone (Fig. S3). Besides the *baicalensis*/*nigrescens* species pair, two other pairwise distances in our data under these thresholds are between sister species pairs in the Altai-Sayan mountain region (*T. brevirostris* and *T. brevicephalus*; and *T. nikolskyi* and *T. svetovidovi*). These species pairs span allopatric drainages subject to catastrophic paleo-hydrological flooding (Weiss et al. 2020a), which may have played a role in their cross-basin colonization history.

Five additional valid *Thymallus* names (*T. yaluensis*, *T. brevipinnis*, *T. signifer*, *T. pallasii* and *T. mertensii*) are supported neither by MOTU delineation nor with existing morphological or genetic data (Fig. 2) (see Appendix I). Two phylogenetically well-supported clades, also supported by one or more delineation approaches, have no species assignment. The first of these makes *T. thymallus* paraphyletic and consists of upper Danubian haplotypes, forming a sister clade to *T. ligericus*. This relationship was not recovered in previous phylogenetic analyses based on mitochondrial CR (Weiss et al. 2002, 2013; Marić et al. 2014), underscoring the higher

resolution of the whole mitogenome and/or the importance of balanced taxon sampling, but also the need for additional investigation (see Appendix I for additional discussion). The second of these unassigned clades involves haplotypes (MT063015–017) of *T. arcticus* s.l., representing a lineage that is believed to derive from the Nahanni glacial refuge (Stamford and Taylor 2004). This mtDNA lineage was among three lineages described in North America (Stamford and Taylor 2004), whereby two colonization events into North America were proposed. The possibility that these haplotypes could represent a distinct taxon must be better explored with population genetic analysis.

The paraphyly of the *T. flavomaculatus* clade in our tree may be based on the complex phylogeographic structure of both *T. flavomaculatus* and *T. grubii* combined with our limited sampling. Both *T. flavomaculatus* and *T. tugarinae* are found in tributaries of the Amur River drainage, but also coastal drainages flowing directly into the seas of Japan and Okhotsk (Antonov and Knizhin 2014) promoting fragmentation and divergence (Appendix I).

## Biogeography

Species diversity and range size in *Thymallus* spp. reveal a latitudinal gradient with few species above the Arctic Circle and increasing diversity (Spearman's rank − 0.819, $P < 0.05$) and decreasing range size (Spearman's rank 0.785, $P < 0.001$) with decreasing latitude (Fig. 1). Ancestral trait reconstruction supports the present Amur River drainage as the place of origin for extant diversity (Fig. 3), with lineage sharing among coastal draining rivers reflecting drainage re-arrangements and dynamics of the Pleistocene (Grosswald 2009; Antonov 2012; Antonov and Mikheev 2016). Europe's colonization by *Thymallus* took place somewhere in the Pliocene, between the split from extant Asian lineages in the late Miocene, and the European lineage crown age in the mid-Pliocene. We note that the rough estimates of the timing of these events are affected by the fossil calibration point placed on the stem leading to the European radiation of the genus, derived from the 7.6 MY-old fossil remains of †*T. latisulcatus*. A previous analysis lacking this calibration point resulted in ages about 20% younger, but still falling in the same epochs described here, and provided even higher support for an Amur ancestral range (all nodes > 80%). Higher connectivity among Arctic Ocean drainages is illustrated by at least two colonisations of North America from Eastern Siberia and is analogous to Holarctic or near Holarctic distributions in *Pungitius* (Guo et al. 2019), *Esox* (Skog et al. 2014), and *Lota* (Van Houdt et al. 2005). Dispersal into Western Siberian rivers seems to have taken place multiple times, most recently from Eastern Siberia and the Amur, after a much more

ancient event. The suggested colonization of the endorheic basin in Mongolia from East Siberian basins evokes the still unconfirmed proposals of a Gobi-Amuran proto-glacial drainage system (Grosswald 1998). Coastal rivers of East Asia retained connectivity with the Amur during the Pleistocene, reflected by the absence of a geographically coherent monophyletic group. However, a Miocene-Pleistocene split and posterior isolation between the extant diversity in both basins cannot be excluded, due to the long un-split branches from Miocene to Pleistocene (Fig. 3). Overall, the biogeographic patterns of *Thymallus* are concordant with scenarios proposed for several taxa. Cobitid fishes reveal an early radiation in eastern Asia, and colonization of Europe from East Asia in the late Miocene (Šlechtová et al. 2008; Perdices et al. 2016). A colonization from northeast Asia to Europe is proposed for *Carassius* (Rylková et al. 2013), and similar colonization of *Rhodeus* (via vicariance) is inferred for the Pliocene (Bohlen et al. 2006), supported by the disjoint distribution and hypothesised vicariance of *Margaritifera* (Bolotov et al. 2016) a parasitic bivalve of *Rhodeus*. These studies, however, do not provide explicit paleo-hydrological pathways or events that may have facilitated these patterns. While some of these hypothesised colonization events are pre-Pleistocene, our knowledge at least of Pleistocene paleo-dynamics across Russia reveals numerous events of paleohydrological connectivity (via large paleo-lakes and mega floods) across Siberia and through the Ponto-Caspian region into Europe (Komatsu et al. 2016). These events have already served hypotheses concerning the phylogeographic structure of *Thymallus*, *Hucho* and *Brachymystax* across their Siberian range (Koskinen et al. 2002; Froufe et al. 2003a; Weiss et al. 2020a). Nonetheless, much work remains in matching the timing of specific events of dispersal and vicariance with specific paleo-events promoting cross-drainage connectivity and isolation.

The following section provides a very brief overview of the 15 focal taxa of our phylogenetic and MOTU analyses. Included is a general description of their geographic range and phylogenetic position as well as our opinion on each taxon's validity as a species (i.e. regarding its biology) and its suggested conservation status based on IUCN criteria. For more detailed comments on these taxa and those for which we do not currently find biological support, or cannot assign to a sample, see Appendix I.

**Amur and east Asian taxa**

***Thymallus tugarinae*** — Lower Amur grayling

Occurs in the middle to lower Amur River drainage and several coastal draining rivers including northwestern Sakhalin Island (Fig. 1). Highly divergent from its sister clade containing *T. grubii* (4.1%) and *T. flavomaculatus* (4.2%). Easily distinguished from sympatric *T. flavomaculatus* using meristic or morphological characters (Knizhin et al. 2007); reciprocally monophyletic with other Amur graylings (Fig. 2) (Froufe et al. 2003b, 2005); reproductive isolation supported (Froufe et al. 2003b). Shows significant phylogeographic structure across its range (Froufe et al. 2003b).

Taxonomic validity

A species reproductively isolated and easily distinguished from the sympatric *T. flavomaculatus* (see *T. flavomaculatus* below).

Conservation remarks

Area of occupancy exceeds 2000 sq. km. Suggested Global Status: Least Concern.

***Thymallus flavomaculatus*** — Yellow-spotted grayling

Occurs in some coastal rivers draining into the seas of Japan and Okhotsk, as well as some lower Amur River tributaries (Fig. 1); overlaps considerably with *T. tugarinae* (Froufe et al. 2003b; Antonov and Knizhin 2011). Easily distinguished from *T. grubii* and sympatric *T. tugarinae* based on a characteristic yellow-orange spot located in the posterior area of the dorsal fin (Knizhin et al. 2006a).

Taxonomic validity

The taxon was originally described as a subspecies of *T. grubii*, widely distributed in the Amur River drainage. Whether or not the taxon is treated as a species or a subspecies is beyond the scope of this manuscript. In our analysis, *T. flavomaculatus* is paraphyletic (see additional comments below for *T. grubii*).

Conservation remarks

Coastal populations are more threatened by anthropogenic changes and overfishing than interior populations. Some range fragmentation is present and the area of occupancy may be as little as 400 sq. km. Suggested Global Status: Near threatened.

***Thymallus grubii*** — Amur grayling

A small-sized grayling, easily diagnosed based on body and dorsal-fin colouration; occurs throughout the middle to upper Amur River drainage (Fig. 1). Significant phylogeographic structure is reported (Froufe et al. 2003b; Knizhin et al. 2004; Weiss et al. 2020b). Occurs in sympatry with *T. burejensis* and *T. baicalolenensis* in the upper Bureya River, where

reproductive isolation is strong but not complete (Weiss et al. 2020b), and in sympatry with *T. tugarinae* in the lower Zeya River and Ingoda River, and with *T. baicalolenensis* in the upper Zeya River and upper Ingoda River (Antonov and Mikheev 2016).

Taxonomic validity

A species diagnosable from all other grayling (Weiss et al. 2020b). Could be also treated as the nominal species of a three-taxon aggregate, consisting of *T.* (*grubii*) *grubii*, *T.* (*grubii*) *flavomaculatus* and *T.* (*grubii*) *yaluensis* (see Article 6.2 in the International Commission on Zoological Nomenclature 1999) (see also Appendix I).

Conservation remarks

Occupies a relatively large range and several relatively pristine river systems. Suggested Global Status: Least Concern.

***Thymallus burejensis*** — Bureya grayling

A robust-bodied grayling, endemic to the middle and upper reaches of the Bureya River. Occurs in sympatry in the upper Bureya River with *T. baicalolenensis* and *T. grubii*, and shows relatively strong (albeit not complete) reproductive isolation (Weiss et al. 2020b). Mitogenomic distances range from 3.0 to over 4.9% between *T. burejensis* and all other congeners.

Taxonomic validity

A species displaying significant reproductive isolation with two other grayling taxa.

Conservation remarks

The Bureya River is over 700 km in length but the mid-to lower reaches have been heavily impacted by hydropower, eliminating or fragmenting portions of the specie's range. Currently, its area of occupation does not exceed 100 sq. km. Suggested Global Status: Endangered.

**European taxa**

***Thymallus aeliani*** — Adriatic grayling

Occupies the middle to upper reaches of the Soca River in Slovenia and tributaries of the Po and Adige rivers in Italy (Fig. 1). Divergent from *T. thymallus* (2.7%) and *T. ligericus* (3.6%) within the clade of European grayling taxa (Fig. 2). Meraner et al. (2014) reported significant regional structure in the Adige River drainage.

Taxonomic validity

A species based on its deep divergence to all other grayling and allopatric distribution in Adriatic draining rivers.

Conservation remarks

River engineering measures, hydropower expansion and water pollution are among widespread threats that have reduced at least 50% of the species range; introgression with non-native lineages is a major threat to *T. aeliani* and there are few pure genetic populations left (Sušnik et al. 2004; Meraner et al. 2014). The area of aquatic occupancy may be as little as 100 sq. km. Suggested global status: Endangered.

***Thymallus thymallus*** — European grayling

Widely distributed (Fig. 1); until recently included all European stocks of grayling. They, along with *T. aeliani* and *T. ligericus*, are the only *Thymallus* species with a subterminal mouth. Significant phylogeographic structure throughout Western Europe (Weiss et al. 2002; Gum et al. 2009), and from the western Balkans and Caspian Sea catchment (Marić et al. 2012, 2014). The taxon is paraphyletic due to the systematic relationship to both *T. ligericus* and its sister clade of upper Danubian haplotypes.

Taxonomic validity

A species distinguished from all Asian grayling by a subterminal mouth and strict long-time allopatry to *T. ligericus* and *T. aeliani*.

Conservation remarks

Currently listed as a species of Least concern. Locally, and especially in the southern portions of its range, population declines or extinctions are widespread (see Weiss et al. 2013), leading to several endangered assignments at national levels. Suggested Global Status: Least Concern.

***Thymallus ligericus*** — Loire grayling

Recently described endemic of the upper Loire River drainage in France (Persat et al. 2019) (Fig. 1). Populations remain genetically pure despite 50 years of stocking with foreign strains (Persat et al. 2016), suggesting they either outcompete foreign lineages (i.e. *T. thymallus*) or display reproductive isolation. Morphologically distinguished from *T. thymallus* by a more pointed snout, more inferior mouth and profuse spotting (Persat et al. 2019). In our analysis, they appear as a shallow, monophyletic clade, 1.5% divergent from *T. thymallus* haplotypes from the upper Danube drainage and 2.2% divergent from all *T. thymallus* samples. The zoogeographic origins of this species in the Loire basin are unknown.

Taxonomic validity

A species based on its morphological and genetic distinction and long-term isolation (2 MY or more) from grayling of adjacent river drainages (Rhine and Rhône) (Persat et al. 2016).

Conservation remarks

Populations are in decline and there are concerns of decreasing water flows and rising water temperatures. Its area of occupancy may not exceed 35 sq. km, but at least six or more fragmented populations exist. Suggested Global Status: Vulnerable.

**Altai, Siberian and North American taxa**

***Thymallus arcticus*** s.l. — Arctic grayling

Occurs from just east of the Urals in Russia to Hudson Bay, Canada; a disjunct population in the Big Hole and Red Rock river drainages in Montana, USA (Fig. 1). In our analysis, *T. arcticus* s.l. is 2.6% divergent from its sister taxon *T. baicalolenensis*. Phylogeographic structure across the Arctic is weak; for example, the haplotype MT063012 near the type locality (Sob River, Ob) groups closely with haplotypes from the Lena River drainage and the Okhotsk Sea catchment in far eastern Russia. Haplotype MT063010 from the presumed type locality of *T. pallasii* (see Dyldin et al. 2017) (Appendix I) in eastern Siberia is intermediate between most North American haplotypes and those from Kamchatka, which some authors assign to *T. mertensii*.

Taxonomic validity

A species based on its clear genetic divergence to other taxa, morphological distinctiveness especially in the dorsal fin size and colouration (albeit with regional variation), and confirmed reproductive isolation to *T. baicalensis* (Weiss et al. 2007) and *T. baicalolenensis* (Weiss et al. 2006). See comments in Appendix I concerning potential recognition of additional taxa, herein treated as *T. arcticus* s.l.

Conservation remarks

The global population of *T. arcticus* s.l. is listed by the IUCN as a species of Least concern (LC). Numerous reports exist of population size declines for the species locally, both in North America and Russia. Suggested Global Status: Least Concern.

***Thymallus baicalensis*** — Baikal black grayling

Occurs throughout the Enisei River drainage including Lake Baikal and its major tributary the Selenga River (Fig. 1); also in some right-hand tributaries of the Ob River drainage (Mrassu and Kabyrza rivers) represented by haplotypes MT063026 and MT063025 (Fig. 2). Displays a

net divergence between 1.6 and 1.9% to four species comprising its sister clade (*T. nikolskyi*, *T. svetovidovi*, *T. brevicephalus* and *T. brevirostris*). Based on Koskinen et al. (2002), multiple samples throughout Lake Baikal were in Hardy-Weinberg Equilibrium reflecting a single taxon occupying Lake Baikal; this inference is also supported by morphological and genetic data from Knizhin et al. (2006b). All species delineation approaches allocated both *T. baicalensis* and *T. nigrescens* as a single MOTU.

Taxonomic validity

A species based on clear genetic divergence to other taxa, distinct dorsal-fin colouration and multiple contact zones with little to no gene flow with other species (*T. arcticus*, *T. baicalolenensis*) (see Knizhin et al. 2006b; Weiss et al. 2007). *Thymallus brevipinnis* is suggested to be a synonym of *T. baicalensis* (Appendix I).

Conservation remarks

Local population declines and extinctions reported; threats include hydropower development, overfishing and pollution. However, the species has a very large distribution range and occupies many habitats that are in pristine or near-pristine condition. Suggested Global Status: Least Concern.

***Thymallus baicalolenensis*** — Baikal-Lena grayling

Small-bodied grayling occurring throughout the Lena River drainage (Fig. 1). Also occurs in Lake Baikal tributaries, most notably the Barguzin River drainage, the Tiya River and Yakchinskie Lakes of the upper Angara River (Knizhin et al. 2006c, 2008b; Kirillov and Knizhin 2014). Also found in the Uda River drainage, Sea of Okhotsk (Antonov and Knizhin 2011) and upper Amur River drainage (Antonov and Knizhin 2011; Antonov and Mikheev 2016), including the upper Bureya River, together with *T. burejensis* and *T. grubii*, where it could be diagnosed with 100% accuracy based on morphological characters (Weiss et al. 2020b). Its morphological and genetic distinction from *T. arcticus* s.l. is shown in Weiss et al. (2006) (therein referred to as *T. a. lenensis*) and Koskinen et al. (2002) (therein referred to as *T. arcticus*, Lena basin). Both its body and dorsal-fin colourations are highly distinct (Knizhin et al. 2008b; Knizhin and Weiss 2009) from *T. arcticus* s.l. as well as all other members of the genus (Dyldin et al. 2017). *Thymallus baicalolenensis* is reciprocally monophyletic to *T. arcticus* s.l. with a net divergence of 2.6% (Fig. 2).

Taxonomic validity

A species showing relatively strong reproductive isolation to four species to which it comes into contact; *T. baicalensis*, *T. arcticus* s.l., *T. burejensis* and *T. grubii*.

Conservation remarks

Very large distribution area, occupies both large rivers and headwaters including small lakes. Found in numerous relatively remote and/or pristine systems. Suggested Global Status: Least Concern.

***Thymallus brevicephalus*** — Shorthead or Markakol grayling

Reported endemic to Lake Markakol (Dyldin et al. 2017); probably not limited to this lake. Haplotypes LC168675 and MT063035 (Fig. 2) stem from samples in the upper Irtysh River drainage; population genetic analysis shows close affinity with samples from the Kara-Kaba River (Weiss et al. 2020a) (Fig. 1). Both our mitogenome analysis and a population genetic analysis in Weiss et al. (2020a) show a very close (0.4%) sister clade relationship to *T. brevirostris* (Fig. 2). More data concerning morphology and ecology is needed.

Taxonomic validity

Although closely related genetically to *T. brevirostris*, viewed as species based on highly distinct morphology (short jaws, no dentation) and ecology (predominantly benthivore), as well as strict allopatric occurrence to *T. brevirostris*.

Conservation remarks

Populations within Lake Markakol are in serious decline due to overfishing (perhaps > 50% across recent decades, M. Baimukanov, pers. comm.); listed as endangered in Kazakhstan. Its strict area of aquatic occupancy (Markakol Lake) is < 700 sq. km, but its distribution is likely considerably larger. Suggested Global Status: Data deficient.

***Thymallus brevirostris*** — Mongolian grayling

Distributed across the large endorheic basin of Western Mongolia, extending into Tuva Republic, Russia (Fig. 1). Considered a large-growing, piscivorous grayling, with pronounced dentation on both jaws and vomer (Knizhin et al. 2008a). Displays considerable phenotypic variability concerning head and jaw size and dentation (Knizhin et al. 2008a; Weiss et al. 2020a). Displays a close (0.4%) sister clade relationship to *T. brevicephalus* from Lake Markakol (Fig. 2).

Taxonomic validity

Although closely related to *T. brevicephalus*, viewed as a species based on highly distinct morphology (large jaws, significant dentation), ecology (predominantly piscivorous) and strict allopatric occurrence to *T. brevicephalus* (see also Appendix I).

Conservation remarks

Overfishing (primarily illegal), hydropower development and spawning ground deterioration are a major concern. Suggested Global Status: Vulnerable.

***Thymallus nigrescens*** — Chovsgul grayling

Endemic to Lake Chovsgul in Mongolia. The single mitogenome (MT063028) in our analysis is 0.3% divergent from *T. baicalensis* (Fig. 2). Traditional genetic markers (microsatellites and single mtDNA sequences) allow no distinction between *T. nigrescens* and *T. baicalensis* (Koskinen et al. 2002; Kaus et al. 2019).

Taxonomic validity

Treated as a species despite its very close relationship to *T. baicalensis*. Morphologically distinct from *T. baicalensis* based on a small dorsal fin and a high number of gill rakers (Knizhin et al. 2008b; Olson et al. 2019), occupies a distinct allopatric ecological niche.

Conservation remarks

Listed as endangered in the Mongolia Red List (Ocock et al. 2006). Its habitat comprises the 2760 sq. km Lake Chovsgul. Substantial illegal harvest via gillnetting in the littoral zone has led to dramatic declines in population sizes (Free et al. 2015). Suggested Global Status: Vulnerable.

***Thymallus nikolskyi*** — Upper Ob grayling

Originally reported from the upper Ob River drainage; exact distribution is unclear. May occur together with *T. baicalensis* in the Mrassu and Kabyrza rivers of the Ob River drainage (Fig. 1). Our analysis reveals a close (0.6%) sister relationship to *T. svetovidovi* from the upper Enisei River drainage, and considerably more divergence from the two other taxa, *T. arcticus* s.l. and *T. brevicephalus*, in the Ob River drainage (3.0% and 1.7%, respectively). Population genetic analysis shows no gene flow between *T. nikolskyi* and *T. brevicephalus* (Weiss et al. 2020a).

Taxonomic validity

Treated tentatively as a species whereby the distributions and genetic relationships of all grayling in the Altai-Sayan mountain region require further investigation (Weiss et al. 2020a).

Conservation remarks

Insufficient data on its diagnosis and distribution. Suggested Global Status: Data deficient.

***Thymallus svetovidovi*** — Upper Enisei grayling

Recently described from the Sharga Gol River in Mongolia (Knizhin and Weiss 2009); occurs in headwater reaches of the Enisei River in Mongolia and possibly Tuva Republic (Fig. 1). Bright yellow caudal peduncle and fin is highly characteristic; a close (0.6%) sister taxon to *T. nikolskyi* from the upper Ob River drainage.

Taxonomic validity

A species based on a unique phenotype and genetic divergence from *T. baicalensis*. See Appendix I for comments on potential synonymies.

Conservation remarks

Known distribution range is limited (but uncertain), populations are reportedly dense and the river systems where this taxon is found are pristine. Thus, there are currently no threats to this taxon. Suggested Global Status: Data deficient.

**Authors' contributions** SW and EF conceived the study for this manuscript with input from DVG. GS, AGS, GD, CH, AA, HP and GE collected data and resources, and helped construct range distribution maps. Most raw data analysis was performed by GS, DVG, AGS and EF. SW wrote the first draft, and all authors contributed to writing.

**Funding** Open access funding provided by University of Graz. This work was supported by the Foundation for Science and Technology to GS (SFRH/BD/139069/2018) and to AGS (SFRH/BD/137935/2018), by Strategic Funding UIDB/04423/2020 and UIDP/04423/2020 and by the Institute of Biology at the University of Graz in the form of a Doc Academy Grant to GE under the supervision of SW. DVG is supported by project ConBiOmics (NORTE-01-0145-FEDER-030286, contract IA/2018–079). The 14 mitogenomes sequenced in France were done according to the partnership MNHN-AFB 2016–2018 and with the assistance of the 'Service de Systématique Moléculaire' of the Muséum national d'Histoire naturelle (CNRS UMS 2700).

**Data availability** Data generated or analysed during this study consist of mitogenome sequences. These sequences are deposited in GenBank under Bioproject number PRJNA604892; all GenBank accession numbers used in our analysis are also listed individually in the Electronic Supplementary File (Table S1).

**Supplementary Information**

The online version contains supplementary material available at https://doi.org/10.1007/s13127-020-00468-7

# Chapter 2

## General patterns of sexual dimorphism in graylings (*Thymallus*), with a comparison to other salmonid species

Gernot K. Englmaier, Alexander Antonov, Steven J. Weiss

**Citation:**
Englmaier GK, Antonov A, Weiss SJ (2022). General patterns of sexual dimorphism in graylings (*Thymallus*), with a comparison to other salmonid species. Reviews in Fish Biology and Fisheries, 32, 645–667. https://doi.org/10.1007/s11160-021-09694-4.

### Abstract

Among fishes, salmonids (family Salmonidae) have attracted a great deal of research attention focused on sexual dimorphism and associated selective forces. Most of this research has been directed toward anadromous and mostly semelparous salmon and trout (*Oncorhynchus*, *Salmo*), and comparatively little is known about intersexual variability in strictly iteroparous freshwater salmonids. We examined a comprehensive data set of 28 linear morphometric characters in 11 of 15 currently recognised species of grayling (Thymallinae, *Thymallus*), a genus consisting of iteroparous species only, to identify general patterns of intersexual morphological variability. Overall, we found that all grayling species show common sex-specific traits particularly relating to size dimensions of the dorsal, anal, pelvic and pectoral fins. Although the magnitude of sexual dimorphism differed among species, there was no significant phylogenetic signal associated with these differences across the genus. These results are discussed in terms of the assumed selection pressures driving sexual dimorphism in graylings and are compared to existing knowledge in Salmonidae as a whole where similarities and differences with both Salmoninae and Coregoninae exist. The present study provides the first detailed genus-wide comparison of sexually dimorphic phenotypic characters in graylings, and highlights the need for more large-scale comparative studies in multiple salmonid species to better understand general macroevolutionary trends among this important group of freshwater fishes.

**Keywords:** Salmonidae, Sexual selection, Dorsal fin, Anal fin, Morphology, Secondary sexual characters

## Introduction

In many animal taxa, a key aspect of intraspecific variability is associated with sexual dimorphism (Andersson 1994), the differences in physiology, morphology and behaviour of conspecific males and females (sensu Punzalan and Hosken 2010). By introducing the idea of sexual selection (Darwin 1871), Darwin set the stage for the general recognition of sex-specific roles in shaping organismal diversity. Although both empirical and theoretical studies suggest that the evolution and maintenance of sex-specific traits is more complex, also involving other selection mechanisms (e.g. Hedrick and Temeles 1989; Cooper 2010), the theory of sexual selection is still fundamental to a general understanding of intersexual variability (Clutton-Brock 2007).

The general mechanisms that drive the evolution of sexual dimorphism are well studied in several taxonomic groups including insects (Wilhelm et al. 2011), birds (Berns and Adams 2012), mammals (Swanson et al. 2013), reptiles (Agha et al. 2017), fishes (Oke et al. 2019) and amphibians (Pincheira-Donoso et al. 2021). Among fishes, salmonids (family Salmonidae) have become one of the most frequently studied groups of species used to address questions on the evolution of sexual dimorphism and associated selective forces (Fleming and Reynolds 2004). Salmonids are a diverse group of cold-water adapted fishes in the northern hemisphere and include salmon and trout (*Oncorhynchus*, *Salmo*), lenok (*Brachymystax*), taimen (*Hucho*), Sakhalin taimen (*Parahucho*), char (*Salvelinus*), whitefish and cisco (*Coregonus*), Round whitefish (*Prosopium*), inconnu (*Stenodus*), and grayling (*Thymallus*). Many salmonid species, particularly semelparous Pacific salmon (*Oncorhynchus*), undergo dramatic phenotypic change during the reproductive period, which includes the development of an elongated snout, enlarged teeth, hooked jaws, dorsal hump, elongated fins, thickened skin, and bright colouration (Fleming and Gross 1994; Quinn and Foote 1994). These exaggerated traits are usually male-biased (i.e. larger, thicker or more pronounced in males) and are assumed to have evolved as a consequence of sexual selection where males compete for fertilization opportunities (Fleming and Reynolds 2004). The presence and degree of sexual dimorphism in these traits differ between taxonomic groups and show great intraspecific variability, which often is habitat associated (Johnson et al. 2006; Oke et al. 2019).

Despite extensive research on sexually dimorphic characters in salmonids, the generality of these traits remains poorly investigated. So far, most attention has been given to large anadromous and semelparous species, with few studies addressing sexual dimorphism in iteroparous and/or freshwater salmonids. Graylings (subfamily Thymallinae) are freshwater

resident iteroparous species with a suite of distinctive morphological traits potentially relevant for the study of sexual dimorphism (Fig. 1). In comparison to other salmonids, graylings are easily characterised by their greatly enlarged dorsal fin, which often has a species-specific coloration pattern, that is not, at least not overtly, sex-specific (Knizhin 2009) and is known to undergo secondary sexual development (Ward 1951). While taxonomy and species level phylogeny in graylings are becoming well-resolved in recent years (Knizhin 2009; Weiss et al. 2021), the evolution of sexual dimorphism remains poorly studied, having only been addressed at all in a few species (e.g. Mikheev 2009). Increasing knowledge on sympatric occurrence of multiple grayling species (Shubin and Zakharov 1984; Weiss et al. 2007, 2020, 2021), with little evidence of hybridization and introgression (Froufe et al. 2003b; Weiss et al. 2007, 2020; Persat et al. 2016), has drawn increasing attention to elucidating the mechanisms that might support reproductive isolation.

Graylings are a monophyletic sister clade to Coregoninae (Campbell et al. 2020) and are widespread across most of Europe, Siberia, the Russian Far East and some parts of North America (Weiss et al. 2021). They are typical riverine fish, but also occur in many lacustrine habitats across their range, and spawn in spring or early summer after short or medium distanced potamodromous migrations (usually from lakes to rivers and within rivers). Graylings are gravel spawning salmonids, where both males and females are promiscuous with multiple spawning acts, usually involving different mates (Beauchamp 1990). In contrast to most other river spawning salmonids, dominant males occupy and defend spawning territories prior to the arrival of females (Fabricius and Gustafson 1955; Bishop 1971). Observations of frequent territorial contests and the generally increased aggressive behaviour during spawning season (Fabricius and Gustafson 1955) may suggest strong intrasexual (male-male) competition for territories and access to mates. Likewise, territorial contests as well as courtship and spawning include characteristic behaviours such as the specific display of the colourful dorsal and pelvic fins (Fabricius and Gustafson 1955; Kratt and Smith 1980). Intraspecific competition, however, is not restricted to the spawning season. Similar to other drift-feeding stream salmonids (Fausch and White 1981), feeding positions among graylings are established in dominance hierarchies (Hughes and Dill 1990; Hughes 1992) where the characteristic display of the dorsal and pelvic fins is an integral behavioural element (Fabricius and Gustafson 1955; Tack 1973).

Given this general behavioural framework and the assumption that selective forces shaping sexually dimorphic phenotypic characters are closely linked to the reproductive behaviour in salmonids (Fleming and Reynolds 2004), we hypothesize that the extent and direction of sexual dimorphism in graylings might be consistent across different species. Thus,

we analysed a comprehensive data set of linear morphometric traits to identify general trends of sexual dimorphism in graylings. Furthermore, we reviewed external morphometric characters subject to sexual dimorphism among salmonids (Salmoninae, Coregoninae, Thymallinae) in order to place our results in a broader phylogenetic context. Finally, by summarizing areas of potential future studies, we hope to foster cross-disciplinary research in ecology and evolution of graylings, which may aid future conservation and management efforts targeting this group of freshwater fishes.


**Fig. 1** Phenotypic comparison of female (above) and male (below) graylings in **A** *T. grubii*, during spawning season, both female and male from the Bureya River, Russia; **B** *T. baicalensis*, outside spawning season, both female and male from the Delger mörön River, Mongolia; **C** *T. flavomaculatus*, during spawning season, female from the Pody River and male from the Gobilly River, Russia; **D** *T. svetovidovi*, outside spawning season, both female and male from the Sharga Gol, Mongolia. Photos by A. Antonov (**A**, **C**) and C. Ratschan (**B**, **D**).

## Materials and methods

### Morphological data set

To test for general patterns of sexual dimorphism among graylings, we analysed a large morphological data set established over a period of more than 10 years and spanning > 1500 individual specimens from 11 species across the whole distribution range of the genus (Table 1). Many subsets of the data set have been used in a range of taxonomic, systematic, and evolutionary studies to date (e.g. Froufe et al. 2003b; Knizhin et al. 2004, 2006a, c, d, e, 2007, 2008a, b; Knizhin and Weiss 2009; Knizhin 2009; Weiss et al. 2006, 2020), but the data have never been investigated as a whole nor in the context of sexual dimorphism. Twenty-eight linear measurements following those introduced by Svetovidov (1936), Pravdin (1966), and Knizhin et al. (2004) were made point to point or as a projection to midline using a caliper to the nearest 0.1 mm (Fig. 2, Table S1). Measurements were taken from formalin (4%) preserved specimens.

Sex and stage of maturity were determined by visual examination of gonads following the classification of ovarian reproductive stages by Sakun and Butskaya (1968). Fishes of stage I (oogonia and oocytes did not yet start protoplasmic growth, immature condition) were excluded and only fishes between stage II (previtellogenic condition) and stage VI (postspawning condition, before returning to stage II) were included in the analysis. Most fishes were sampled after spawning season (July–October). Only few specimens of *T. arcticus* (n = 23), *T. baicalensis* (n = 60) and *T. thymallus* (n = 65) were taken before or during spawning season in spring or early summer.

**Table 1** Number of male and female specimens used in the analysis and mean fork length (Lsm ± SD) for each species.

| Species | Male | | Female | |
|---|---|---|---|---|
| | N | Lsm (mm) | N | Lsm (mm) |
| *T. arcticus* | 102 | 280.4±63.7 | 55 | 244.2±48.7 |
| *T. baicalensis* | 266 | 282.2±74.2 | 265 | 297.2±60.5 |
| *T. baicalolenensis* | 158 | 213.6±38.6 | 172 | 194.5±35.6 |
| *T. brevirostris* | 24 | 324.1±125.2 | 22 | 345.9±120.5 |
| *T. burejensis* | 32 | 300.5±66.5 | 35 | 288.7±59.1 |
| *T. flavomaculatus* | 25 | 233.2±29.8 | 23 | 221.3±33.5 |
| *T. grubii* | 75 | 199.0±40.8 | 50 | 173.4±27.0 |
| *T. nigrescens* | 15 | 280.9±25.0 | 15 | 268.7±20.5 |
| *T. svetovidovi* | 11 | 370.7±13.9 | 12 | 368.1±12.9 |
| *T. thymallus* | 66 | 330.5±44.6 | 58 | 300.3±37.4 |
| *T. tugarinae* | 32 | 215.9±21.0 | 26 | 204.6±24.9 |

Data transformation

In order to emphasize general trends in the data set as well as equalize variances among groups, we excluded measurements showing extreme values. These were defined as the deviation of 3 * IQR (inter quartile range) from the 25th (Q1 - 3 * IQR) and 75th (Q3 + 3 * IQR) percentile respectively, calculated using raw measurements relative to body length (referring to fork length). We excluded extreme values for each species and sex separately, rather than the entire data set, to retain the natural species-specific variability. In total, 52 specimens (3.3% of the entire data set) were excluded (19 males, 33 females). The final data set consisted of measurements for 1539 fish (806 males and 733 females) (Table 1). For a few individuals, some measurements were not obtained due to damage or poor preservation condition. To retain these specimens in the analyses, predicted values from linear regression models (per sex, species and trait) were used to substitute missing data. Such cases account for 0.6% of the entire data (0.7% of males and 0.6% of females).

All morphometric measurements were converted to their base 10 logarithm to linearize allometry and equalize variances (Sidlauskas et al. 2011). We evaluated potential allometric scaling among species by comparing species-specific slopes of reduced-major axis regression lines as outlined in Sidlauskas et al. (2011) and implemented in the R package 'smatr' (Warton et al. 2006) in R v.4.1.0 (R Core Team 2021). Some statistically significant differences among species were found (14 of 55 pairwise comparisons; Table S2). These differences, however, appeared related to sample size and body-size differences (Table 1), and thus we chose to apply a common slope in the following data transformation. To control for the effects of variation in body size and body size scaling, all morphometric traits were scaled to a common mean fork length using an allometric growth formula commonly applied in Salmonidae (Siwertsson et al. 2013; Jacobs et al. 2020): $\log_{10} Y_{std} = \log_{10} Y_{obs} + b * (\log_{10} L_{std} - \log_{10} L_{obs})$; where $Y_{std}$ is the corrected trait value, $Y_{obs}$ is the measured trait value, b is the slope of the regression of each ($\log_{10}$) trait against ($\log_{10}$) fork length, $L_{std}$ is the mean fork length of all specimens (mm), and $L_{obs}$ is the individual fork length. Terminology of these variables follows Siwertsson et al. (2013). The common slope b for each trait was derived from ANCOVA models using species and sex as factors while controlling for body length.

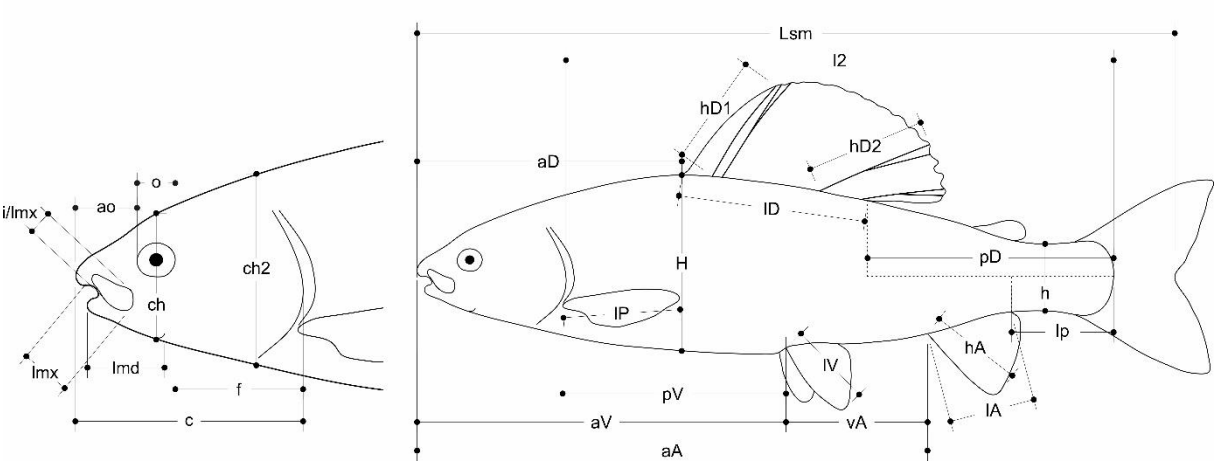



**Fig. 2** Schematic illustration of morphometric characters used in the present study. For a detailed description of each measurement see Table S1. Lsm, fork length; l2, trunk length; ao, snout length; o, horizontal eye diameter; f, postorbital length; c, head length; ch2, head depth at nape; ch, head depth through the eye; lmx, upper jaw length; i/lmx, upper jaw depth; lmd, lower jaw length; H, maximum body depth; h, minimal caudal peduncle depth; aD, predorsal length; pD, postdorsal length; aA, preanal length; aV, prepelvic length; lp, caudal peduncle length; pV, pectoral-pelvic distance; vA, pelvic-anal distance; lD, length of dorsal fin base; hD1, height of anterior part of dorsal fin; hD2, height of posterior part of dorsal fin; lA, length of anal fin base; hA, height of anal fin; lV, length of pelvic fin; lP, length of pectoral fin. Not illustrated are: k, forehead width; w, maximum width of body.

## Comparison of sexual dimorphism among graylings

We determined morphological characters contributing to the divergence between males and females, across the whole genus, using a two-way Analysis of Variance (ANOVA). ANOVA models (trait * sex + species + species * sex) were implemented in IBM SPSS Statistics v.26. A series of one-way ANOVA analyses with simple effects was additionally used to more precisely evaluate species-specific trends of sexual dimorphism. The assumptions of homoscedasticity and normal distribution of (unstandardized) residuals were not met in all cases. However, visual examination of residuals in histograms and normal Q–Q plots showed them to be approximately normally distributed. Overall, ANOVA analyses on large samples sizes (> 500 observations) are known to be robust against minor deviations of normality (e.g. Johnson 1998). To corroborate two-way ANOVA results and account for different sample sizes in sub-groups (i.e. species), we examined Welch's tests (unequal variance t-test) in cases where homogeneity of variances was not met (Ruxton 2006), and performed separate non-parametric Kruskal–Wallis tests to support one-way ANOVA results of species-specific trends where normality of residuals was violated.

## Phylogenetic comparative analysis

Because species are not independent and patterns of sexual dimorphism could have a phylogenetic component, we estimated the ancestral state of sexual dimorphism for each character and tested for a phylogenetic signal across species. The phylogenetic reconstruction is based on complete mitochondrial genomes of 15 *Thymallus* species (considering the current taxonomy presented in Weiss et al. (2021)) and two Coregonus species as outgroup with the following GenBank accession numbers: MT063012 (*T. arcticus*), MT063023 (*T. baicalensis*), MT063019 (*T. baicalolenensis*), MT063036 (*T. brevicephalus*), MT063033 (*T. brevirostris*), MT063037 (*T. burejensis*), MT063039 (*T. flavomaculatus*), MT063038 (*T. grubii*), MT063028 (*T. nigrescens*), MT063030 (*T. nikolskyi*), MT063004 (*T. thymallus*), KJ866485 (*T. tugarinae*), CM031715 (*C. clupeaformis*), and NC_025576 (*C. peled*). Alignment and analyses of mitochondrial data follow Weiss et al. (2021). All parts of the analysis were performed in PhyloSuite v.1.2.2 (Zhang et al. 2020) and IQ-TREE (Nguyen et al. 2015) was used for maximum likelihood (ML) analysis. The computed ML tree was visualised and edited with FigTree v.1.4.4 (Rambaut 2018) and CorelDRAW 2019.

For phylogenetic comparison, we excluded those species where morphological data were absent in our data set (*T. aeliani*, *T. brevicephalus*, *T. ligericus*, *T. nikolskyi*). We estimated and visualised the degree of sexual dimorphism in morphometric traits (based on mean pairwise differences (least square means) from linear two-way ANOVA models) at each node in the mtDNA phylogeny using the 'fastAnc' and 'contMap' functions in the R package 'phytools' (Revell 2012) in R v.4.1.0 (R Core Team 2021). To test for a phylogenetic signal of sexual dimorphism for each character, we estimated Pagel's k (Pagel 1999) and Blomberg's K (Blomberg et al. 2003) using the 'phylosig' function in 'phytools'.

## Literature review on sexually dimorphic traits in Salmonidae

To evaluate our observations in graylings in a broader phylogenetic context, we synthesized current knowledge of sexual dimorphism in external morphometric traits across Salmonidae. We included characters that are reversable and temporally linked to the breeding season as well as those that undergo a non-reversable change starting usually at the onset of sexual maturity. Sexual size dimorphism was not considered due to its high intraspecific variability (e.g. Jonsson and Jonsson 2015). Literature searches were performed in Web of Science, Scopus and Google Scholar (key words in different combinations: sexual dimorphism, secondary sexual

character(s), morphometric character(s), female, male, salmonids, *subfamily*, *genus*, *trait*). The results were augmented with additional studies, especially from Russia, that were not found with these search engines. A morphometric trait was considered sexually dimorphic if a statistical analysis was performed to test for intersexual variability or the morphological characters were described as explicit sexual dimorphism.

## Results

### Comparison of sexual dimorphism among graylings

The average fork length among all individuals was 260.5 ± 73.6 mm. Males were larger than females in the global data set (263.7 ± 73.7 mm vs. 256.9 ± 73.2 mm; two-way ANOVA, $F_{1,1517}$ = 8.208, $P$ = 0.004; Welch, $F_{1,\ 1513.03}$ = 3.938, $P$ = 0.047), but this pattern was not consistent across all species reflected in the significant species x sex interaction (Table 2). At the species-specific level, only *T. arcticus*, *T. baicalolenensis*, *T. grubii* and *T. thymallus* exhibited a significant male-biased size dimorphism (one-way ANOVA, F < 22.590, $P$ < 0.001; Kruskal Wallis H, $X^2$ < 20.770, $P$ < 0.001).

Two-way ANOVA analyses on morphometric characters showed that significant species x sex interactions were present in 21 out of 28 traits. Significant differences between the sexes were found in 18 characters, 12 of which remained significant after table-wide Bonferroni correction (Table 2). When species-specific pairwise differences were examined, predorsal length, length of dorsal-fin base, height of anterior part of dorsal fin, height of posterior part of dorsal fin, height of anal fin, and length of the pelvic and pectoral fins showed a uniform pattern of sexual dimorphism (Figs 3, S1), although these differences were not statistically significant for all species (Table S2). The magnitude of sexual dimorphism showed considerable variation in some characters, exemplified by a sevenfold difference in magnitude of the posterior part of dorsal fin between *T. baicalolenensis* (0.187, most dimorphic) and *T. burejensis* (0.033, least dimorphic).

In general, male graylings had significantly greater length and height dimensions of dorsal, pelvic and pectoral fins, a longer base of the anal fin, and a deeper caudal peduncle than conspecific females. In contrast, female-biased traits were related to the height of the anal fin, length dimensions of the abdomen, and a greater distance between the fins. Deviations from these general trends in *T. brevirostris*, *T. burejensis*, *T. flavomaculatus*, *T. svetovidovi* and *T.*

*tugarinae* (Fig. S1), were not statistically significant (one-way ANOVA, F < 0.809, P > 0.370; Kruskal Wallis H, $X^2$ < 0.510, *P* > 0.470).

While some attributes of the head and abdomen were non-significant in the global two-way ANOVA analyses, non-parametric Welch's tests suggested a significantly greater postorbital length (Welch, F1,1518.36 = 13.347, *P* < 0.001) and head length (Welch, F1, 1525.56 = 10.651, *P* = 0.001) in males, and a greater body width (Welch, F1, 1236.93 = 11.917, *P* < 0.001), postdorsal length (Welch, F1, 1526.86 = 48.073, *P* < 0.001) and preventral length (Welch, F1, 1536.71 = 27.415, *P* < 0.001) in females. These traits were found significant for sex-differentiation only in a few species (Table S2), although the direction of sexual dimorphism was similar for most species.

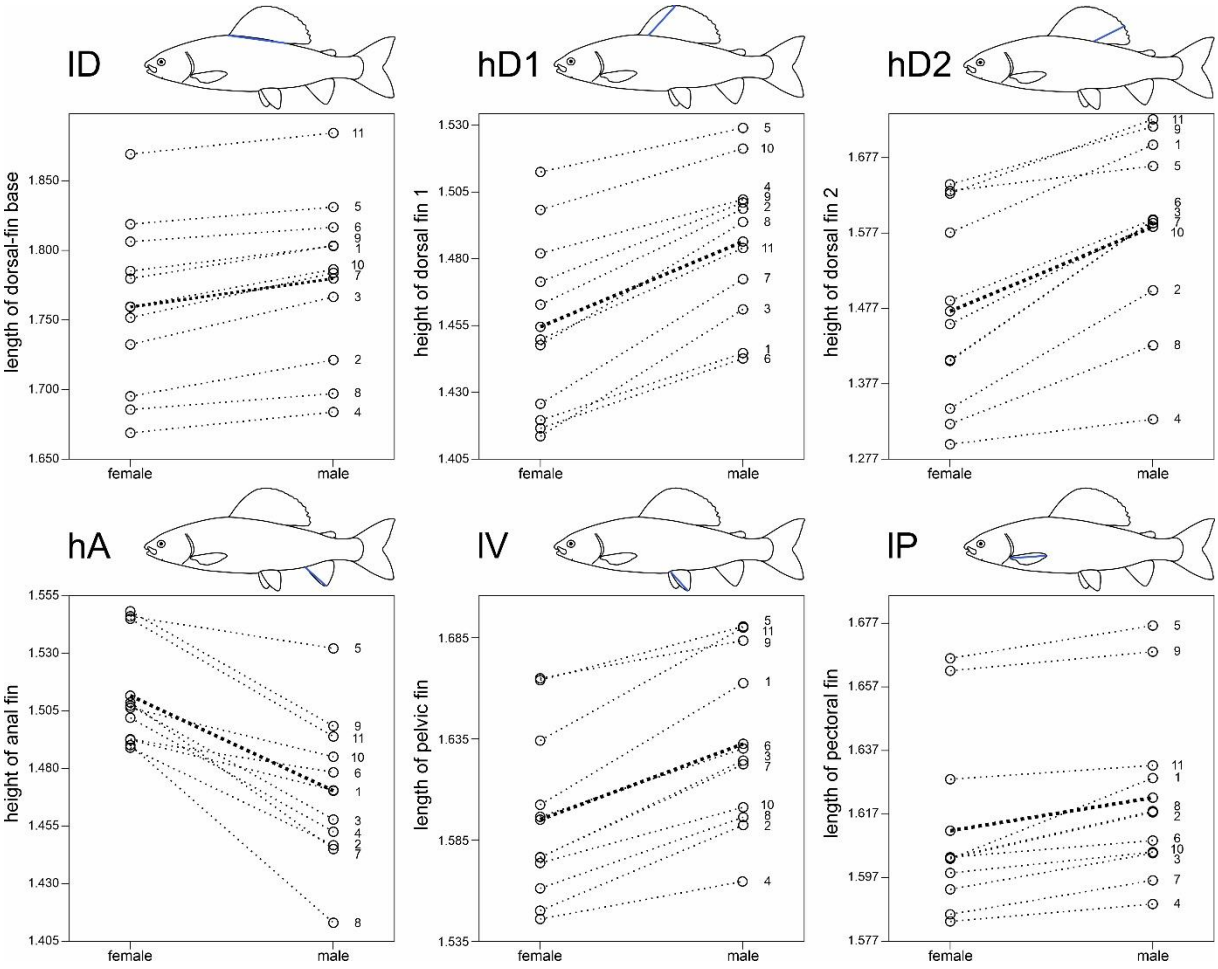



**Fig. 3** Comparison of pairwise differences (least square means) for the most uniform sexually dimorphic traits of the dorsal, anal, pelvic and pectoral fins (two-way ANOVA, significant after table-wide Bonferroni correction). Bold lines represent mean values for all species. Height of dorsal fin 1 refers to height of anterior part of dorsal fin and height of dorsal fin 2 refers to height of posterior part of dorsal fin. 1 = *T. arcticus*, 2 = *T. baicalensis*, 3 = *T. baicalolenensis*, 4 = *T. brevirostris*, 5 = *T. burejensis*, 6 = *T. flavomaculatus*, 7 = *T. grubii*, 8 = *T. nigrescens*, 9 = *T. svetovidovi*, 10 = *T. thymallus*, 11 = *T. tugarinae*.

**Table 2** Two-way ANOVA results of fork length and 28 standardized morphometric characters. Least square means (LSM) for each trait and sex are given with their standard errors.

| Variables | Sex | | | | | Species | | | Species x Sex | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | LSM male | LSM female | df | F | *P*-value | df | F | *P*-value | df | F | *P*-value |
| **Body length** | | | | | | | | | | | |
| Fork length (Lsm) | 2.424±0.005 | 2.404±0.005 | 1 | 8.208 | 0.004 | 10 | 115.861 | 0.000 | 10 | 5.661 | 0.000 |
| **Morphometric charaters** | | | | | | | | | | | |
| Trunk length (l2) | 2.306±0.000 | 2.306±0.001 | 1 | 0.004 | 0.953 | 10 | 26.847 | 0.000 | 10 | 2.897 | 0.001 |
| Snout length (ao) | 1.198±0.002 | 1.193±0.002 | 1 | 3.071 | 0.080 | 10 | 92.154 | 0.000 | 10 | 4.078 | 0.000 |
| Horizontal eye diameter (o) | 1.066±0.002 | 1.068±0.002 | 1 | 1.126 | 0.289 | 10 | 40.328 | 0.000 | 10 | 3.100 | 0.001 |
| Postorbital length (f) | 1.405±0.001 | 1.404±0.001 | 1 | 1.091 | 0.296 | 10 | 62.643 | 0.000 | 10 | 4.430 | 0.000 |
| Head length (c) | 1.701±0.001 | 1.698±0.001 | 1 | 5.135 | 0.024 | 10 | 76.975 | 0.000 | 10 | 4.682 | 0.000 |
| Head depth at nape (ch2) | 1.583±0.001 | 1.580±0.001 | 1 | 2.528 | 0.112 | 10 | 92.117 | 0.000 | 10 | 3.184 | 0.001 |
| Head depth through the eye (ch) | 1.425±0.002 | 1.419±0.002 | 1 | 6.571 | 0.011 | 10 | 65.819 | 0.000 | 10 | 3.365 | 0.000 |
| Forehead width (k) | 1.181±0.002 | 1.174±0.002 | 1 | 5.104 | 0.024 | 10 | 26.916 | 0.000 | 10 | 2.174 | 0.017 |
| Upper jaw length (lmx) | 1.164±0.002 | 1.167±0.002 | 1 | 1.168 | 0.280 | 10 | 72.336 | 0.000 | 10 | 5.296 | 0.000 |
| Upper jaw depth (i/lmx) | 0.698±0.003 | 0.694±0.003 | 1 | 1.545 | 0.214 | 10 | 29.787 | 0.000 | 10 | 2.888 | 0.001 |
| Lower jaw length (lmd) | 1.398±0.002 | 1.397±0.002 | 1 | 0.611 | 0.435 | 10 | 93.422 | 0.000 | 10 | 3.898 | 0.000 |
| Maximum body depth (H) | 1.739±0.002 | 1.740±0.002 | 1 | 0.750 | 0.387 | 10 | 128.123 | 0.000 | 10 | 0.681 | 0.743 |
| Minimal caudal peduncle depth (h)* | 1.269±0.001 | 1.262±0.001 | 1 | 12.646 | 0.000 | 10 | 201.091 | 0.000 | 10 | 0.892 | 0.540 |
| Maximum width of body (w) | 1.479±0.003 | 1.491±0.003 | 1 | 9.439 | 0.002 | 10 | 17.921 | 0.000 | 10 | 2.771 | 0.002 |
| Predorsal length (aD)* | 1.922±0.001 | 1.930±0.001 | 1 | 42.958 | 0.000 | 10 | 357.146 | 0.000 | 10 | 3.269 | 0.000 |
| Postdorsal length (pD) | 2.037±0.001 | 2.042±0.000 | 1 | 9.758 | 0.002 | 10 | 66.577 | 0.000 | 10 | 2.416 | 0.008 |
| Preanal length (aA)* | 2.262±0.000 | 2.266±0.001 | 1 | 36.540 | 0.000 | 10 | 22.651 | 0.000 | 10 | 2.949 | 0.001 |
| Prepelvic length (aV) | 2.076±0.001 | 2.079±0.001 | 1 | 8.548 | 0.004 | 10 | 38.705 | 0.000 | 10 | 2.179 | 0.017 |
| Caudal peduncle length (lp) | 1.634±0.001 | 1.635±0.001 | 1 | 0.113 | 0.737 | 10 | 53.547 | 0.000 | 10 | 0.518 | 0.878 |
| Pectoral-pelvic distance (pV)* | 1.864±0.001 | 1.872±0.001 | 1 | 24.976 | 0.000 | 10 | 18.483 | 0.000 | 10 | 1.213 | 0.277 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Pelvic-anal distance (vA)* | 1.819±0.001 | 1.825±0.001 | 1 | 10.487 | 0.001 | 10 | 18.789 | 0.000 | 10 | 1.445 | 0.155 |
| Length of dorsal fin base (lD)* | 1.780±0.002 | 1.759±0.002 | 1 | 69.945 | 0.000 | 10 | 316.737 | 0.000 | 10 | 1.966 | 0.033 |
| Height of anterior part of dorsal fin (hD1)* | 1.486±0.002 | 1.455±0.003 | 1 | 85.915 | 0.000 | 10 | 56.664 | 0.000 | 10 | 1.770 | 0.061 |
| Height of posterior part of dorsal fin (hD2)* | 1.585±0.005 | 1.473±0.005 | 1 | 257.779 | 0.000 | 10 | 147.709 | 0.000 | 10 | 8.347 | 0.000 |
| Length of anal fin base (lA)* | 1.391±0.002 | 1.369±0.002 | 1 | 57.600 | 0.000 | 10 | 53.979 | 0.000 | 10 | 2.858 | 0.002 |
| Height of anal fin (hA)* | 1.470±0.002 | 1.512±0.002 | 1 | 209.443 | 0.000 | 10 | 36.577 | 0.000 | 10 | 5.838 | 0.000 |
| Length of pelvic fin (lV)* | 1.633±0.002 | 1.595±0.002 | 1 | 150.800 | 0.000 | 10 | 93.433 | 0.000 | 10 | 2.478 | 0.006 |
| Length of pectoral fin (lP)* | 1.622±0.002 | 1.612±0.002 | 1 | 23.019 | 0.000 | 10 | 59.127 | 0.000 | 10 | 1.298 | 0.226 |

Traits that remained significant after table-wide Bonferroni correction (significance level at a = 0.05/28 = 0.0018) are marked with an *

Phylogenetic comparative analysis

Overall, there was no significant phylogenetic signal of sexual dimorphism among the traits (Pagel's $\lambda < 0.001$ ($P = 1$), Blomberg's K < 0.782 ($P > 0.160$)) (Table S3). Only the length of the pelvic fin had a low but non-significant phylogenetic signal (Pagel's $\lambda = 0.951$, $P = 0.543$; Blomberg's K = 0.788, $P = 0.087$). Thus, while ancestral state reconstruction showed that sexual dimorphism in the most general trends (identified by two-way ANOVA analyses) were also present in the most recent common ancestor (Fig. 4, Fig. S2), species-specific patterns were clearly not related to phylogeny. For example, closely related species often had a clearly different magnitude of sexual dimorphism (e.g. the sister species *T. flavomaculatus* and *T. grubii*, *T. arcticus* and *T. baicalolenensis*). This indicates that trends of sexual dimorphism are not more similar among closely related species than to distantly related relatives. While some species, such as *T. baicalensis* and *T. baicalolenensis* generally tend to have a high degree of intersexual variability, others such as *T. brevirostris* and *T. burejensis*, only show slight differences between the sexes or exhibit contrasting patterns to the general trends observed (Figs 3, S1).

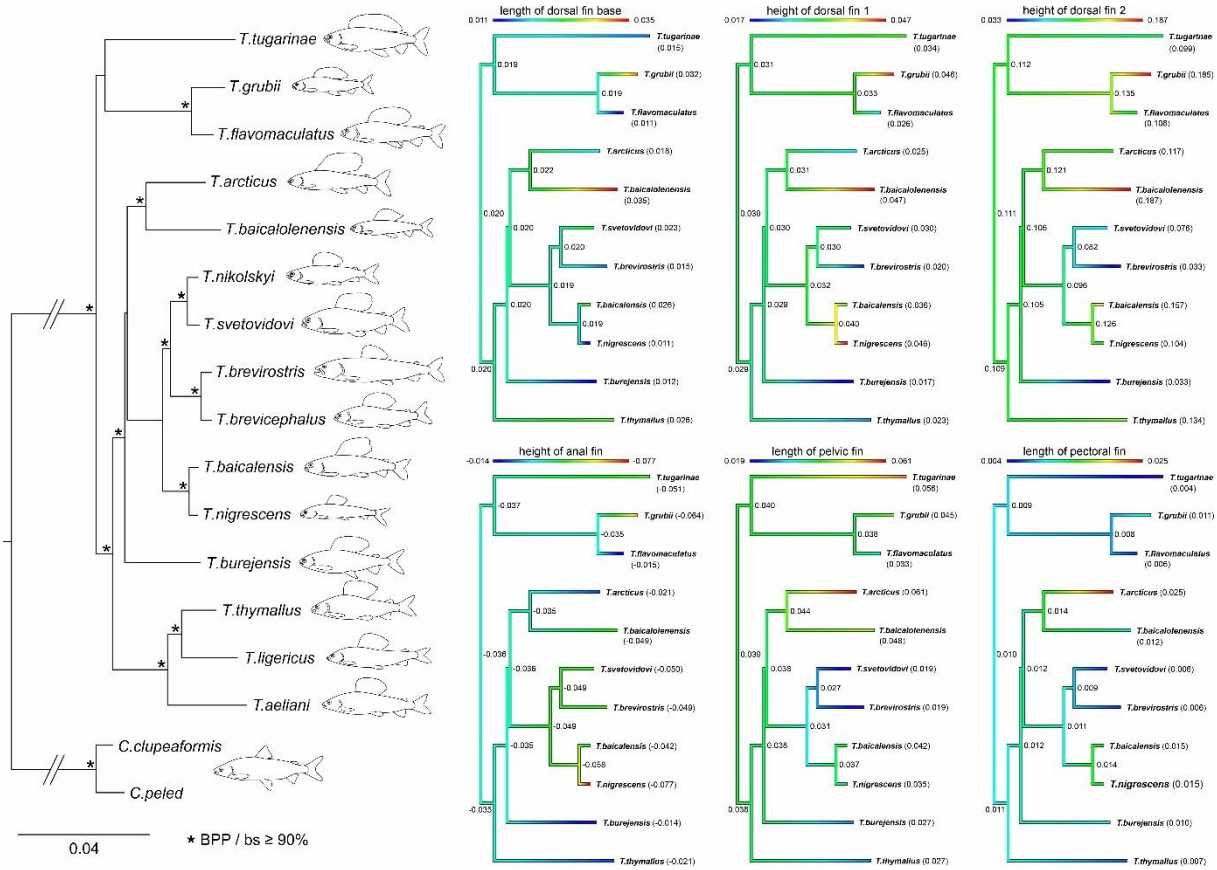


**Fig. 4** Phylogenetic reconstruction of *Thymallus* species based on whole mitochondrial genomes, and ancestral state reconstruction of the degree of sexual dimorphism in the most uniform sexually dimorphic traits of the dorsal, anal, pelvic and pectoral fins (red = most dimorphic, blue = least dimorphic). *Thymallus aeliani*, *T. brevicephalus*, *T. ligericus*, and *T. nikolskyi* were excluded from ancestral state reconstruction due to missing morphological data. Negative values indicate a female-biased trait.

## Review of sexual dimorphism in morphometric traits across Salmonidae

We reviewed 56 publications describing sexual dimorphism in morphometric traits among salmonid species (Table 3). The majority of these studies (n = 43) were based on the measurement of linear morphometric traits, while six used geometric morphometrics, and seven studies were descriptive and based on visual examination. A comparatively large number of these studies (n = 19) targeted anadromous (mostly semelparous) species of the genus *Oncorhynchus*. Overall, only a small percentage (11 %) of the large number of salmonid species (n = 247), listed as valid species in Fricke et al. (2021), has been explicitly investigated for sexual dimorphism in external morphometric traits. Although this problem may be overstated due to recent taxonomic inflation (e.g. Isaac et al. 2004), whole genera have apparently been ignored as no specific studies on sexual dimorphism in morphometric traits were found for the genera *Brachymystax*, *Hucho*, *Parahucho* and *Stenodus*. Among the species analysed, the most general traits of sexual dimorphism across different genera were:

### *Length of the jaws and snout*

The secondary sexual development of the male jaws (to a lesser extent also present in females), has been reported from several species in the subfamily Salmoninae. The transformation of the jaws (and the elongation of the snout) during the breeding period tends to be most characteristic for semelparous *Oncorhynchus* (upper jaw), and iteroparous *Salmo* and *Salvelinus* (lower jaw) (Table 3). A modification of the upper and/or lower jaw during the reproductive period has not been reported in either Coregoninae or Thymallinae.

### *Length and depth of the head*

The head tends to be generally more robust in male Salmoninae. In Coregoninae and Thymallinae, dimensions of the head appear to be more sexually monomorphic, though

observations by Nikulina and Polyaeva (2020) would suggest a larger head in female *Coregonus sardinella*.

*Length and height of the dorsal fin*

Sexual dimorphism in length and height dimensions of the dorsal fin and its base-length are usually male-biased and reported from species in *Oncorhynchus*, *Salmo*, *Coregonus*, *Prosopium* and *Thymallus*. Exceptions of female-biased dimensions in the dorsal fin may exist such as in *C. sardinella* (Table 3).

*Size of the adipose fin*

The adipose fin tends to be generally larger (height and length dimensions) in male *Oncorhynchus*, *Salmo*, and *Salvelinus* (Table 3), a trend that appears to be consistent across multiple *Salmo* species and not restricted to the spawning season (compare to data in Delling and Doardio 2005; Turan et al. 2011, 2012). Data of multiple *Coregonus* species in alpine lakes would suggest sexual monomorphism in this trait (Selz et al. 2020). Among Thymallinae, intersexual variability in the size of the adipose fin has not been investigated.

*Height of the anal fin*

The height of the anal fin was found to be female-biased in several species of *Oncorhynchus*, *Salmo* and *Thymallus* (Table 3). Morphological studies on European, Eurasian, and North African *Salmo* species suggest great interspecific variability in this character and a greater height of the anal fin in males of some species (e.g. Turan et al. 2011, 2012; Doadrio et al. 2015). Observations from *Coregonus* and *Prosopoum* may suggest a generally greater height of the anal fin in male Coregoninae (Table 3).

*Length of the paired pelvic and pectoral fins*

The pelvic and pectoral fins were found to be commonly longer in males of species in *Oncorhynchus*, *Salmo*, *Salvelinus*, *Coregonus*, *Prosopium* and *Thymallus*; a trend that appears most consistent in the subfamilies Coregoninae and Thymallinae (Table 3).

*Size of breeding tubercles*

Breeding tubercles are known only from iteroparous species. They are most characteristic for Coregoninae but were also found in individual species of Thymallinae (*T. arcticus*; Kratt and Smith (1978)) and Salmoninae (*Salvelinus namaycush*; Muir et al. (2012)). Size and abundance of breeding tubercles were found to be usually male-biased.

**Table 3** Sexually dimorphic external morphometric traits in Salmonidae.

| Subfamily | Genus | Species | Parity | Male-biased | Female-biased | Reference |
|---|---|---|---|---|---|---|
| Salmoninae | | | | | | |
| | | *gorbuscha* (Pink salmon) | semelparous | upper jaw length*, head length*, head width, body depth, caudal peduncle depth*, dorsal fin base length*, adipose fin length*, anal fin base length* | anal fin height | Davidson (1935); Beacham and Murray (1983, 1985, 1986); Beacham et al. (1988); Zhivotovsky and Kim (2015) |
| | | *keta* (Chum salmon) | semelparous | snout length*, upper jaw length*, head length*, postorbital head length*, body depth*, caudal peduncle depth*, prepelvic length*, dorsal fin height*, dorsal fin base length*, adipose fin height*, adipose fin length* | horizontal eye diameter*, anal fin height*, anal fin base length* | Beacham and Murray (1983, 1985, 1987); Beacham (1984); Myoung et al. (1993)[a] |
| | *Oncorhynchus* | *kisutch* (Coho salmon) | semelparous | snout length*, upper jaw length*, tooth length, body depth*, dorsal fin height*, adipose fin length*, pelvic fin length*, pectoral fin lenght* | caudal peduncle depth*, anal fin height*, anal fin base length* | Shapovalov and Taft (1954); Beacham and Murray (1983, 1986); Fleming and Gross (1994) |
| | | *mykiss* (Steelhead trout) | iteroparous | jaw length, tooth length, body depth | | Shapovalov and Taft (1954) |
| | | *nerka* (Sockeye salmon) | semelparous | snout length, upper jaw length*, tooth length, body depth, caudal peduncle depth, adipose fin length* | | Beacham and Murray (1983, 1986); Quinn and Foote (1994); Hendry and Berg (1999); Johnson et al. (2006); Oke et al. (2019) |
| | | *nerka* (Kokanee salmon) | semelparous | snout length*, jaw length*, tooth length, body depth*, caudal fin height, pelvic fin length*, pectoral fin length* | anal fin height*, anal fin base length* | Ricker (1938), Winans et al. (2003); Thorn and Morbey (2016) |

| | | | | | | |
|---|---|---|---|---|---|---|
| | *Salmo* | *tshawytscha* (Chinook salmon) | semelparous | snout length*, upper jaw length*, head length*, adipose fin length*, adipose fin height* | | Beacham and Murray (1983, 1986); Merz and Merz (2004) |
| | | *fahrettini* | iteroparous | upper jaw length, mouth gape length, adipose fin base length | | Turan et al. (2020) |
| | | *kottelati* | iteroparous | upper jaw length, mouth gape length, mouth gape width, head length | | Turan et al. (2014) |
| | | *salar* (Atlantic salmon) | iteroparous | jaw length, adipose fin size* | | Tchernavin (1944), Næsje et al. (1988); Järvi (1990) |
| | | *trutta* (Brown trout) | iteroparous | upper jaw length, head length*, body depth, dorsal fin height, adipose fin length | abdomen length*, predorsal length, pectoral-pelvic distance | Reyes-Gavilán et al. (1997); Monet et al. (2006) |
| | | *alpinus* (Arctic char) | iteroparous | mouth size*, head length*, head depth*, body depth*, pectoral fin length* | | Janhunen et al. (2009) |
| | | *confluentus* (Bull trout) | iteroparous | head length, adipose fin height | | McPhail and Murray (1979, seen in McPhail and Baxter (1996)); Nitychoruk et al. (2013) |
| | *Salvelinus* | *fontinalis* (Brook trout) | iteroparous | snout length, lower jaw length, mouth width, head length, head depth, pelvic fin length, pectoral fin length | body width | Willson (1997); Proulx and Magnan (2004); Kazyak et al. (2013) |
| | | *malma* (Dolly Varden trout) | iteroparous | snout length, body depth, adipose fin height | | McPhail and Murray (1979, seen in Beacham and Murray (1983)); Yamamoto et al. (2017) |
| Coregoninae | *Coregonus* | *artedi* (Cisco) | iteroparous | dorsal fin length, pectoral fin length, anal fin length | | Jacobson et al. (2020) |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | *clupeaformis* (Lake whitefish) | iteroparous | upper jaw length*, pelvic fin length*, pectoral fin length* | | Casselman and Schulte-Hostedde (2004) |
| | | *lavaretus* (Lavaret) | iteroparous | caudal peduncle width*, anal fin base length*, pelvic fin length*, pectoral fin length* | predorsal length*, prepelvic length*, pectoral-pelvic distance*, body with | Heese (1987) |
| | | *peled* (Peled) | iteroparous | horizontal eye diameter*, minimum body depth*, dorsal fin height*, dorsal fin length*, anal fin height* | pectoral-pelvic distance* | Mamcarz and Nowak (1986)[b] |
| | | *sardinella* (Least cisco) | iteroparous | interorbital width*, prepectoral length* | head length*, head width*, head depth at nape*, body depth*, caudal peduncle length*, dorsal fin base length*, dorsal fin height*, anal fin base length*, pectoral fin base length* | Nikulina and Polyaeva (2020) |
| | | *zugensis* (Albeli) | iteroparous | size of breeding tubercles* | | Wedekind et al. (2008) |
| | *Prosopium* | *coulteri* (Pygmy whitefish) | iteroparous | dorsal fin height*, anal fin height*, pelvic fin length*, pectoral fin length* | | McCart (1965) |
| | | *cylindraceum* (Round whitefish) | iteroparous | size of breeding tubercles | abdomen length | Normandeau (1963) |
| Thymallinae | *Thymallus* | *arcticus* (Arctic grayling) | iteroparous | postorbital length*, head length*, head depth at eye*, body depth*, postdorsal distance*, dorsal fin height*, dorsal fin base length*, anal fin base length*, pelvic fin length*, pectoral fin length* | horizontal eye diameter*, predorsal length, preanal length*, pectoral-anal distance*, anal fin height* | Rawson (1950); Ward (1951); Bishop (1967, 1971); Tack (1973); Ridder (1989); Zinovjev and Bogdanov (2012); Romanov (2016) |

| | | | | | |
|---|---|---|---|---|---|
| | *flavomaculatus* (Yellow-spotted grayling) | iteroparous | lower jaw length, dorsal fin height, dorsal fin base length | anal fin height | Semenchenko (2005)[c]; Knizhin et al. (2006a) |
| | *thymallus* (European grayling) | iteroparous | dorsal fin height*, dorsal fin base length*, anal fin base length*, pelvic fin length*, pectoral fin length* | anal fin height | Magreiter (1951); Persat (1977); Zinovjev (2012); Kucheruk et al. (2015) |
| | tugarinae (Lower-Amur grayling) | iteroparous | dorsal fin height, anal fin base length, pelvic fin length, pectoral fin length | lower jaw length, preanal length, pectoral-pelvic distance, anal fin height | Mikheev (2009) |

Traits reported significant at P\0.05 are marked with an *

[a] Myoung et al. (1993) reported a significantly longer postorbital head length in females while Beacham and Murray (1987) reported the character to be male-biased

[b] Sexual dimorphism in fishes of age 4+ is reported

[c] Therein reported as *T. arcticus grubii* (Samarga River, Prymorsky Territory)

## Discussion

### Sexual dimorphism in graylings

Our analyses revealed a suite of morphometric characters that display significant sexual dimorphism across most grayling species, particularly relating to size dimensions of the dorsal, anal, pelvic and pectoral fins. The differences were pronounced, despite the variation in magnitude and the significant species x sex interaction in our global analysis. Previous studies suggest that the differentiation between the sexes in these characters starts with the beginning of maturity (Ward 1951; Tack 1973; Kratt and Smith 1979) and is then permanently present in adult fishes outside and during the breeding season. Before sexual maturity, these characters may essentially follow similar growth trajectories in males and females (Kratt and Smith 1979). Yet, our data of fishes at reproductive stage I (immature condition; see Sakun and Butskaya (1968)) indicate a largely similar, but often non-significant pattern of sexual dimorphism for the fins (ANCOVA, $P > 0.05$; n = 285, across 6 species; data not shown).

The general predictions on the different reproductive strategies and energy investments suggest differential selection acting on the sexes (Fleming and Gross 1994), and thus the existence of sexual dimorphism in specific morphometric traits in order to increase reproductive success. In male graylings, these predictions are consistent with the male-biased length and height dimensions of the dorsal, pelvic and pectoral fins. The display of the dorsal and pelvic fins is an integral behavioural element in male-male competition and territorial behaviour during spawning season (Fabricius and Gustafson 1955). Sex-specific differences in these characters have already been described for *T. arcticus* (Romanov 2016), *T. flavomaculatus* (Semenchenko 2005), *T. tugarinae* (Mikheev 2009) and *T. thymallus* (Persat 1977; Zinovjev 2012; Kucheruk et al. 2015). The pectoral fin is known to support swimming stability and manoeuvrability in many fish groups (Bone and Moore 2008) and may be favourable in male-male competition and mate acquisition. However, the most noticeable morphological trait among graylings—the large colourful dorsal fin—may not have evolved as a direct consequence of sex-specific selection as the trait is permanently present in both sexes. Thus, the initial driver of this accentuated character, which does not occur in any other salmonid fish, is most probably rooted in natural selection perhaps in the form of intraspecific competition for position in the typically drift-feeding graylings (Fabricius and Gustafson 1955; Hughes and Dill 1990).

In female graylings, the commonly observed female-biased dimorphism in size among other salmonid species (e.g. Tamate and Maekawa 2004; Morbey 2018) is not supported in our analysis, although length dimensions of the abdomen and distances between fins were generally longer in females. However, these differences may be a secondary effect of the position of the fins perhaps in relation to the extended dorsal and anal fin bases in males. The strongest pattern of female-biased dimorphism was evident for the height of the anal fin. Kratt and Smith (1979) reported that the anal fin is used in lateral display, but the lack of iridescent colouration compared to other fins (see Fig. 1) may indicate a reduced visual function. It is more likely that, similar to other gravel spawning salmonids, the anal fin holds a female-specific (mechanistic) function in reproduction (see Thorn and Morbey 2016), related to the female ‘‘probing’’ behaviour (Groot 1996; Esteve 2005) or oviposition. Compared to other gravel spawning salmonid species, however, female graylings do not construct spawning redds or actively cover the eggs with substrate after fertilization (Fabricius and Gustafson 1955). Instead, the eggs are buried into the substrate by the characteristic spawning behaviour, whereby the caudal region of the female is forced into the porous gravel substrate by vigorous quivering of both sexes and tail flapping of the male (Kratt and Smith 1980). This grayling-specific behaviour may also be key to understanding the observed sexual dimorphism in the caudal region, which includes a deeper caudal peduncle and an elongated length of the anal fin base in males.

An interesting finding of our study is that the extent of sex-specific differences in the general traits is not equal across the genus. The fact that these differences do not have a significant phylogenetic signal would seem to support a differential strength of intrasexual selection across species. Among the few species that showed weaker or a lack of pronounced sexual dimorphism, *T. brevirostris* from the species poor (i.e. impoverished ichthyofauna; Kottelat (2006)) Altai region of Western Mongolia stands out. This species has comparatively small trait sizes for the fins and generally shows only weak differences in morphometric characters between the sexes. Besides *T. nigrescens*, it is the only species in our data set where in both sexes the greatest height of the dorsal fin is in its anterior part, which does not only affect the shape of the fin (see Knizhin et al. 2008a: Fig. 3), but likely also its display function as the extended posterior part is usually the most colourful region. In contrast, *T. burejensis*, endemic to the Bureya River, a tributary of the Amur River in the Russian Far East, is characterized by having large dorsal, anal, pelvic and pectoral fins in both sexes but significant sexual dimorphism only in the pelvic fin. *Thymallus burejensis* further contrasts with *T. brevirostris* as it occurs in sympatry with up to three other grayling species (*T. baicalolenensis*, *T. grubii* and *T. tugarinae*; see Antonov (2004); Knizhin et al. (2004)). This raises the

possibility that the unique combination of sex-specific attributes reflects past competition (sensu Connell 1980) among sympatric species, supporting species recognition (i.e. distinction between con- and heterospecific individuals) and thus may help to avoid hybridization.

## Review of sexual dimorphism in morphometric traits across Salmonidae

Previous studies have shown that sexual dimorphism in salmonids is primarily driven by breeding competition (Fleming and Reynolds 2004). Other factors such as life history tactics (e.g. early maturation) (Koseki and Maekawa 2000) or habitat characteristics (Oke et al. 2019) can have a profound effect on the development and/or expression of sexually dimorphic traits, which underlines the intraspecific (among-population) variability and facultative nature of sexual dimorphism in this group. By reviewing the collective evidence of the factors responsible for trait-specific sexual dimorphism, patterns emerge supporting links between reproductive behaviour or particular environmental conditions and specific morphometric character development associated with sex. The reduced taxonomic coverage among existing studies on sexual dimorphism in salmonids limits evolutionary interpretations, but our review nonetheless expands support for a number of assumptions that have been made for single taxa or genera in the past.

Sexual dimorphism is present in all subfamilies of salmonids but the specific traits or patterns of expression differ among groups. In Thymallinae, intersexual variability in morphometric traits is primarily associated with length and height attributes of the fins, although these characters are subject to sexual dimorphism in Coregoninae and Salmoninae as well. The height or length of the dorsal, adipose, pelvic and pectoral fins are male-biased in multiple genera/species, and were found to play a role in behaviours related to display (e.g. Fabricius and Gustafson 1955; Esteve et al. 2009a; Muir et al. 2012) and female choice (Järvi 1990) but are apparently not so energetically costly as to make them facultative and dependent on life history. Interestingly, the height of the anal fin is the only fin-specific character that shows frequent female-biased dimorphism in gravel spawning Thymallinae and Salmoninae but appears to be monomorphic in the open substrate spawning Coregoninae. Thus, this female-biased trait may have evolved in response to selection pressure on females, related to a mechanistic or sensory function involving oviposition or selection of a suitable spawning habitat (Thorn and Morbey 2016) in order to increase offspring survival. However, the height and shape (see Gruchy and Vladykov 1968) of the anal fin have not yet been investigated in

several species, and data on *Brachymystax* suggest that both sharp- and blunt-snouted lenok are sexually monomorphic in the height of the anal fin (Alekseyev S., unpublished data).

The characteristic transformation of the jaws and snout is limited to Salmoninae, but absent from the basal genera *Brachymystax* and *Hucho* (Esteve and McLennan 2008; Esteve et al. 2009b), and apparently reduced or absent in *Parahucho* (Esteve et al. 2009a). Thus, these traits may represent a derived set of characters possibly related to the intense breeding competition in Salmoninae. Interestingly, the greatest expression of sexual dimorphism in jaws and snout is exhibited in anadromous and primarily semelparous species (absent in precocious parr; Koseki and Maekawa (2000)), life history strategies that presumably allow more energy to be invested in such asymmetric growth (Fleming and Reynolds 2004). This contrasts somewhat with anadromous *O. mykiss* (Steelhead trout); a species that shows comparatively little dimorphism in the jaws. However, although the species is iteroparous only a rather small percentage of anadromous individuals manage to breed a second time (10% (0.6–31.3%) in Fleming (1998); 2.4% in Christie et al. (2018)). The fact that another iteroparous anadromous salmonid, *Salmo salar*, exhibits pronounced sexual dimorphism in the lower jaw, as well as a low frequency of repeat spawning with a mean of 11% (0.7–42.5%; Fleming (1998)), suggests a mechanistic relationship between anadromous behaviour or (facultative) semelparity and the development of sexual dimorphism at the level of the individual as opposed to a fixed population or species-specific trait.

Compared to Salmoninae and Thymallinae, species in Coregoninae generally exhibit a low degree of sexual dimorphism (Willson 1997). This may be rooted in the fact that Coregoninae contrast sharply in a range of reproductive behaviours compared to other salmonids. For example, Coregoninae are open substrate spawners, exhibit reduced intrasexual competition for access to mates and commonly spawn at night (Fabricius and Lindroth 1954; Karjalainen and Marjomäki 2018), all behavioural traits that are not common in Salmoninae and Thymallinae (Fabricius and Gustafson 1955; Esteve 2005). Thus, spawning behaviour in Coregoninae may favour a different set of (non-visual) signals such as the development of breeding tubercles, which are commonly male biased in size and abundance (Willson 1997). The lacustrine open substrate spawning *Salvelinus namaycush* (Lake char) with well-described male-biased tubercles (Muir et al. 2012), would support this hypothesis, but there is too little information on tubercles in salmonids in general to draw further conclusions on their prominence and relation to reproductive behaviours and sexual dimorphism.

Overall, it would be revealing to test the effects of environmental conditions vs. common ancestry on trait evolution across a broader phylogenetic range of salmonids with particular

focus on species showing diverging life-history tactics or a behavioural repertoire that contrasts with closely related congeners such as observed in *S. namaycush* (Muir et al. (2012).

Limitations and future research perspectives

While we think that the general trends of sexual dimorphism in graylings reported in this analysis are robust, a few comments on potential caveats and data limitations are warranted. Multiple populations across the range of some species (e.g. *T. arcticus*, *T. baicalensis*), contrast with single populations or low sample sizes of some others (e.g. *T. nigrescens*, *T. svetovidovi*). Thus, some species-specific results may not capture the natural range of variability that is present. Furthermore, the size range analysed does not capture the entire range of sexually mature fish, especially for those species with lower sample sizes. This may be important because the assumed allometric growth component of morphometric traits is likely to lead to a higher degree of dimorphism with age, a variable that we could not assess directly due to the limitation of insufficient sample sizes across multiple age-classes. Lastly, our analyses rely on linear morphometric characters only, without evaluating dichromatism or shape dimorphism, and thus the full scope of sexual dimorphism is assumed to be underestimated. We therefore recognize specific areas of research that could further clarify the patterns and hypotheses concerning the evolution of sexual dimorphism in graylings.

a. The allometric growth component in morphometric traits could be investigated directly for both males and females, perhaps most simply using hatchery-reared populations and the measurement of individuals across their entire life-cycle (following an ontogenetic approach). These studies could provide direct evidence for the secondary sexual development of specific morphological traits and may also include shape dimorphism (geometric morphometrics) to capture a broader extent of intersexual morphological variability. Moreover, such studies could clarify to what extent sexual dimorphism is related to age and growth.

b. The large colourful dorsal fin in graylings has long been suggested to play a key role in reproductive behaviour (Fabricius and Gustafson 1955) and its species-specific colouration pattern (Knizhin 2009) raises the question of species recognition. The recognition of potential conspecific mates may be more important for those species living in sympatry with congeners, as this could help maintain reproductive isolation. It

is, however, unclear if and to what extent graylings actively choose mates based on visual traits and more generally what mechanisms may underlie species recognition and mate choice in graylings. In other salmonid species, mate choice appears to be quite common where several sexually dimorphic traits such as the adipose fin in *Oncorhynchus* and *Salmo* (Beacham and Murray 1983; Järvi 1990), the kype in *Salmo* (Perry et al. 2019) or the breeding tubercles in *Coregonus* (Wedekind et al. 2008) are thought to be, at least partly, subject to female choice and/or serve a function in displaying status.

c. The clarification of which morphological characters are more driven by natural vs. sexual selection as well as the relevance of intra- vs. interspecific competition can be investigated by research on contact zones, such as in the Amur drainage, where at least three if not four grayling species can be found in sympatry (Antonov 2004; Knizhin et al. 2004; Weiss et al. 2020b, 2021). In these zones, our sample sizes were limited for some species, and the spectrum of investigated characters could be expanded to include both coloration and shape, if not also differential gene expression that may be mechanistically driving these patterns. Such studies would also benefit from more detailed behavioural data, especially on the spatiotemporal distribution of spawning, as well as both the accentuation and display of body and fin colouration during the entire reproduction period. Likewise, genome-wide sequencing and expression studies can help identify sex-biased gene expression (e.g. Sharma et al. 2014) and alleles (Mohammed et al. 2019), which will foster the understanding of the mechanisms that maintain reproductive isolation.

d. In addition to contact zones, several widespread species, such as *T. arcticus*, *T. baicalensis*, *T. baicalolenensis* and *T. grubii*, which are found in diverse habitats either in sympatry with other grayling species or alone (Weiss et al. 2021), could be investigated specifically for potential morphological (sexually dimorphic) patterns that change based on habitat or the presence or absence of congeners. Thus far, Weiss et al. (2020) touched on the morphological differences within and between populations of *T. baicalolenensis* across three different major drainage systems (Amur, Lena, Yenisei), and the likelihood that these differences are driven by different selection mechanisms, potentially including ecological niche partitioning in one drainage, but interspecific competition or its avoidance in another drainage.

More generally, some of our observations can be extended to other genera in Salmonidae, where relatively little is known about general patterns of sexual dimorphism across multiple species and different habitats. We assume that both population and species-specific patterns of sexual dimorphism are relevant for long-term population viability. Thus, we should recognize that our increased knowledge of sexual dimorphism and the underlying evolutionary processes clearly contribute to the growing consensus that the management of salmonid fish populations should aim to avoid using artificial rearing and stocking to supplement populations and above all, inter-basin transfers, whether involving conspecific or congeneric material (Laikre et al. 2010; Rand et al. 2012; Weiss et al. 2013). In areas where such crossbasin transfers have already occurred, and extensive hybridization exists between divergent lineages, such as between *T. thymallus* and *T. aeliani* in much of the original range of the latter species (Meraner et al. 2014), it would be revealing to examine sexual dimorphism in populations showing introgression and see whether or not the patterns of trait divergence conform to the general patterns of sexual dimorphism observed in the present study.

**Acknowledgements**

The data set presented herein was established by the late I. Knizhin and represents the most comprehensive and detailed collection of morphological information on grayling (*Thymallus*) species to date. We express our deep gratitude for the help in collecting materials to the inspectors of the Bureinsky Reserve V. Shichanin, P. Sarychev and G. Taranik. We also thank C. Ratschan (EZB) for providing photographs used in the present study, S. Alekseyev (Russian Academy of Sciences) for sharing information on *Brachymystax*, and K. Sefc (University of Graz) for providing constructive comments on an earlier version of the manuscript. The present work was supported by a grant from the Doctoral Academy Graz, Ecology and Evolution in Changing Environments (EECE) of the University Graz to GKE.

**Author contributions**

Conceived and coordinated the study: GKE, SJW. Analysed the data: GKE. Contributed data and field observations: AA. Wrote the first draft of the paper: GKE, SJW. All three authors contributed equally to the improvement of the manuscript.

**Funding**

Open access funding provided by University of Graz. Doctoral Academy Graz—Ecology and Evolution in Changing Environments (EECE-consortium) of the University of Graz.

**Data availability**

Data are available from the corresponding author upon reasonable request.

**Supplementary Information**

Supplementary data to this article can be found online at https://link.springer.com/article/10.1007/s11160-021-09694-4#Sec23

# Chapter 3

## Evaluating a species phylogeny using ddRAD SNPs: Cyto-nuclear discordance and introgression in the salmonid genus *Thymallus* (Salmonidae)

Secci-Petretto, G., Englmaier, G. K., Weiss, S. J., Antonov, A., Persat, H., Denys, G., Schenekar, T., Romanov, V. I., Taylor, E. B., Froufe, E.

**Citation:**

Secci-Petretto G, Englmaier GK, Weiss SJ, Antonov A, Persat H, Denys G, Schenekar T, Romanov VI, Taylor EB, Froufe E (2023). Evaluating a species phylogeny using ddRAD SNPs: cytonuclear discordance and introgression in the salmonid genus *Thymallus* (Salmonidae). Molecular Phylogenetics and Evolution, 178, 107654. https://doi.org/10.1016/j.ympev.2022.107654

### Abstract

Hybridization and introgression are very common among freshwater fishes due to the dynamic nature of hydrological landscapes. Cyclic patterns of allopatry and secondary contact provide numerous opportunities for interspecific gene flow, which can lead to discordant paths of evolution for mitochondrial and nuclear genomes. Here, we used double digest restriction-site associated DNA sequencing (ddRADseq) to obtain a genome-wide single nucleotide polymorphism (SNP) dataset comprehensive for all *Thymallus* (Salmonidae) species to infer phylogenetic relationships and evaluate potential recent and historical gene flow among species. The newly obtained nuclear phylogeny was largely concordant with a previously published mitogenome-based topology but revealed a few cyto-nuclear discordances. These incongruencies primarily involved the placement of internal nodes rather than the resolution of species, except for one European species where anthropogenic stock transfers are thought to be responsible for the observed pattern. The analysis of four contact zones where multiple species are found revealed a few cases of mitochondrial capture and limited signals of nuclear introgression. Interestingly, the mechanisms restricting interspecific gene flow might be distinct; while in zones of secondary contact, small-scale physical habitat separation appeared as a limiting factor, biologically based reinforcement mechanisms are presumed to be operative in areas where species presumably evolved in sympatry. Signals of historical introgression were largely congruent with the routes of species dispersal previously inferred from mitogenome

data. Overall, the ddRADseq dataset provided a robust phylogenetic reconstruction of the genus *Thymallus* including new insights into historical hybridization and introgression, opening up new questions concerning their evolutionary history.

**Keywords:** Cyto-nuclear discordance, Introgression, Grayling, ddRADseq, Phylogenetics

## 1. Introduction

Cyto-nuclear discordance is a common pattern in genetic studies of both plants and animals and is found at various levels of analyses ranging from population genetic to large-scale phylogenetic studies (Asmussen et al., 1987; Scribner and Avise, 1994, Avise, 1995, Toews and Brelsford, 2012). It occurs when nuclear and cytoplasmic genetic markers show contrasting patterns of diversification and thus discordant phylogenies. If this visualized discordance is based simply on incomplete lineage sorting (ILS) then no real evolutionary discordance has taken place. However, both recent or historical hybridization and introgression coupled with selection and drift can result in organelle and nuclear genomes having different evolutionary histories. Indeed, both ILS and true evolutionary discordance can also take place within the same organismal group (Rose et al., 2021). Thus, the evolutionary history of populations both within and between closely related species cannot be reconstructed with confidence based solely on organelle markers, such as mitochondrial DNA (mtDNA) (Ballard and Whitlock, 2004, De Ré et al., 2017, Wallis et al., 2017) making the incorporation of nuclear DNA (nDNA) data more or less obligatory for robust phylogenetic and/or phylogeographic inferences. Nonetheless, mtDNA barcoding, at least in some groups such as birds (Kerr et al., 2007) or mussels (Froufe et al., 2016; Lopes-Lima et al., 2019), reveals high rates of concordance with existing taxonomic schemes, and many unique properties of mtDNA (as well as plastid molecules in plants) support their continued use as complementary to, rather than inconsistent with nDNA-based inferences. This is especially clear when an organelle-based phylogeny reflects historical gene flow among taxa or lineages, whereas the nDNA-based phylogeny reveals primarily reciprocally monophyletic relationships (e.g., Rose et al., 2021).

The advent of high-throughput sequencing techniques together with their increasing affordability has elevated the power to detect hybridization and introgression, even in non-model organisms. For instance, genome-wide single nucleotide polymorphisms (SNPs) can reveal gene flow stemming from recent or ancient introgression in a variety of organisms (e.g., Guo et al., 2019, Paetzold et al., 2019, Hughes et al., 2020). Hybridization and introgression,

the primary drivers of cyto-nuclear discordance, may be more prevalent in freshwater environments, due to, for example, the high frequency of hybridization in fishes (Hubbs, 1955, Scribner et al., 2000) coupled with paleo-hydromorphological processes that lead to shifting or cyclic phases of allopatry and contact among lineages (Wallis et al., 2017, Mendes et al., 2021). Such landscape dynamics should provide more opportunities for hybridization and introgression, promoting reticulate evolution at some time in the past, despite restricted gene flow in the present (e.g., Sušnik et al., 2006, MacGuigan and Near, 2019, Mason et al., 2019).

Large-scale paleo-hydromorphological dynamics are relevant for several groups of freshwater fishes throughout the world and promote hybridization, cyto-nuclear discordance and cryptic diversity (Goodier et al., 2011, Bangs et al., 2020, Weiss et al., 2021, Cambell et al., 2022), with genome-wide datasets playing an increasingly important role in elucidating the processes behind such complex evolutionary histories. The application of SNP data in the widely distributed freshwater genus *Pungitius* (sticklebacks), for example, underscored the importance of generating a robust nDNA-based phylogeny and diminished the importance of several events of mitochondrial capture by revealing ancient hybridization yet limited nuclear gene flow, a phenomenon shown across a wide range of taxa (Good et al., 2015, Rose et al., 2021). The Eurasian range of *Pungitius* extends across some of the world's largest temperate river systems (e.g., Ob, Enisei, Lena, and Amur), whose complex paleo-hydromorphological history is far from fully understood (e.g., Komatsu et al., 2016). Phylogeographic structure among freshwater fishes across this range reveals drainage-specific patterns with some notable exceptions that likely involve Pleistocene or pre-Pleistocene dynamics (Froufe et al., 2003a, Froufe et al., 2005, Bohlen et al., 2006). One genus of salmonid fishes, *Thymallus* (grayling) is widely distributed across this region and has been the recent focus of considerable study related to phylogeographic structure and paleo-hydromorphological processes.

Recently, Weiss et al. (2021) provided a mitogenome-based phylogenetic and biogeographic analysis of grayling supporting the existence of 13–15 species with high resolution. The majority of this diversity was found in boreal regions of Eurasia, particularly in the Amur River drainage. Studies of salmonid fishes, including grayling, in these regions, support repeated cross-drainage colonization leading to contact among previously allopatric lineages (Weiss et al., 2002, Weiss et al., 2020a, Weiss et al., 2020b, Froufe et al., 2003a). Interspecific hybridization is reported in grayling, but extensive introgression has thus far been rare and limited to Europe, primarily due to anthropogenic processes such as stocking programs (Weiss et al., 2013). In Asia, multiple areas of both sympatric and parapatric distribution among grayling species are described, especially in the Amur drainage, where up to five species are

present (Antonov and Mikheev, 2016, Weiss et al., 2021). However, many conclusions have been based on mtDNA datasets or local population genetic studies with a limited number of genetic markers. The combination of a widely distributed genus, large-scale paleo-hydrological dynamics, and a mitogenome phylogeny together with a range of background data at the population level provide an excellent foundation for evaluating cyto-nuclear discordance.

This study aims to provide broadly applicable inferences concerning cyto-nuclear discordance and events of historic and recent introgression using *Thymallus* as a model. To accomplish this, a genome-wide SNP dataset including all currently known *Thymallus* species was generated through double digest restriction-site associated DNA sequencing (ddRADseq) to: 1) infer a nuclear-based phylogeny of the genus and compare it to an existing mtDNA-based topology; 2) evaluate potential hybridization and introgression among sympatric species, in selected contact zones; and 3) infer historical introgression, aiming to differentiate between recent and more ancient events, using both genome-wide SNPs and mitochondrial data.

## 2. Materials & methods

### 2.1. Taxon sampling

The choice of samples was based on the two major objectives of the study – the genome-wide nDNA phylogeny, and the evaluation of admixture and historical introgression within four contact zones where multiple species are found. Thus, samples for the phylogenetic analysis were based on the taxon list discussed in Weiss et al. (2021), with two to 35 individuals per taxon (Table 1 and Supplementary Table S1). Larger sample numbers were used to target more broadly distributed taxa, which may contain strong phylogeographic structure or cryptic diversity. For the contact zone analyses, between six and 21 individuals were sequenced per taxon, bringing the total number of analyzed samples to 201 (Table 1 and Supplementary Table S1). The four chosen contact zones are: 1) the Kantaiskoye Lake system (lower Enisei) with *T. arcticus* (Pallas, 1776) and *T. baicalensis* Dybowski, 1874 (Weiss et al., 2007, Romanov, 2020); 2) the Lake Baikal basin, with *T. baicalensis* and *T. baicalolenensis* Matveev, Samusenok, Pronin & Tel'pukhovsky, 2005 (Knizhin et al., 2006b); 3) the upper Bureya River drainage, a tributary of the Amur River, with sympatric *T. baicalolenensis*, *T. burejensis* Antonov, 2004, and *T. grubii* Dybowski, 1869 (Weiss et al., 2020b); and 4) the Anui River drainage, a tributary of the lower Amur River, where *T. flavomaculatus* Knizhin, Antonov & Weiss, 2006 and *T.*

*tugarinae* Knizhin, Antonov, Safronov & Weiss, 2007 are found in sympatry (Froufe et al., 2003b).

**Table 1** List of all samples used in the present work, showing the species names, sampling locations, drainage systems, coordinates (WGS84) of sampling locations and number of individuals (N) analyzed from each location. Asterisks indicate admixed individuals identified with fastSTRUCTURE. Further details are given in Supplementary Table S1.

| Species | Location | Drainage | Coordinates | N |
|---|---|---|---|---|
| *T. aeliani* | Adige | Adige | 46.616389, 11.184167 | 1 |
| | Piave | Piave | 46.394156, 12.357128 | 1 |
| | Rienz | Isarco-Adige | 46.759745, 12.051320 | 1 |
| *T. arcticus* | Edyngde | Khantaiskoye-Enisei | 68.266667, 91.266667 | 4 |
| | Gogochenda | Khantaiskoye-Enisei | 68.283333, 91.116667 | 8 |
| | Ulya | Ulya | 58.58, 141.24 | 1 |
| | Maly Lagorta | Voikar-Ob | 66.353854, 63.574339 | 1 |
| | Kolyma | Kolyma | 66.166667, 151.05 | 1 |
| | Tuser | Lena | 72.127778, 126.872778 | 1 |
| | Urak | Urak | 59.709769, 141.932469 | 1 |
| | Big Hole | Jefferson-Missouri-Mississippi | 45.616667, -113.45 | 1 |
| | Adsett | Nahanni-Mackenzie | 58.100036, -122.7704346 | 2 |
| | Tuchodi | Muskwa- Fort Nelson-Liard-Mackenzie | 58.363889, -123.623611 | 1 |
| *T. baicalensis* | Gogochenda | Khantaiskoye-Enisei | 68.283333, 91.116667 | 7 |
| | Nerotkar | Khantaiskoye-Enisei | 68.416667, 91.216667 | 5 |
| | Barguzin | Baikal-Angara-Enisei | 54.418256, 110.628797 | 3 |
| | Dagary Bay | Baikal-Angara-Enisei | 55.783333, 109.75 | 2 |
| | Selenga | Baikal-Angara-Enisei | 49.36308, 103.56167 | 1 |
| | Mrassu | Tom-Ob | 52.711944, 88.5975 | 2 |
| | Khoboi Cape | Baikal-Angara-Enisei | 53.4, 107.783333 | 2 |
| | Tompa | Baikal-Angara-Enisei | 55.482806, 109.960539 | 2 |
| | Ushkaniy Islands | Baikal-Angara-Enisei | 53.666667, 108.616667 | 2 |
| *T. baicalolenensis* | Bureya | Amur | 51.916667, 134.883333 | 12 |
| | Tiya | Baikal-Angara-Enisei | 56.052175, 109.405697 | 5 |
| | Barguzin | Baikal-Angara-Enisei | 54.555367, 111.399297 | 7 |
| | Verkhnyaya Angara | Baikal-Angara-Enisei | 55.864289, 108.896714 | 1 |
| | Ikcheya | Baikal-Angara-Enisei | 56.083333, 110.766667 | 2 |
| | Yakchiy | Baikal-Angara-Enisei | 56.069319, 110.794269 | 4 |
| | Sobopol | Lena | 67.0525, 128.013889 | 1 |
| | Okonon | Zeya-Amur | 55.677778, 130.166944 | 1 |
| | Goudzhekit | Tiya-Baikal-Angara-Enisei | 55.730314, 108.890081 | 2 |
| *T. brevicephalus* | Kaldzhir | Irtysh-Ob | 48.416389, 85.183056 | 2 |
| | Kara-Kaba | Irtysh-Ob | 48.8, 86.516667 | 1 |
| | Urunkhaika | Irtysh-Ob | 48.766667, 86.116667 | 1 |
| *T. brevirostris* | Khukh | Zavkhan | 47.516667, 98.45 | 1 |
| *T. burejensis* | Bureya | Amur | 51.715717, 134.309278 | 8 |
| *T. flavomaculatus* | Urmi | Tunguska-Amur | 48.713056, 134.266944 | 2 |
| | Gobili | Anui-Amur | 49.25, 138.316667 | 6 |
| | Koppi | Tatar Strait | 48.615906, 139.590178 | 1 |
| *T. grubii* | Umalta Makit | Bureya-Amur | 51.679767, 134.239539 | 2 |

| | | | | |
|---|---|---|---|---|
| | Bureya | Amur | 51.715717, 134.309278 | 5 |
| | Onon | Amur | 48.663056, 110.425 | 3 |
| *T. ligericus* | Alagnon | Loire | 45.106389, 2.896667 | 1 |
| | Sioule | Loire | 46.025833, 2.904167 | 1 |
| | | Eg-Selenga-Baikal- | 51.45, 100.65 | 2 |
| *T. nigrescens* | Chovsgul | Angara-Enisei | | |
| *T. nikolskyi* | Biya | Ob | 51.82, 87.154722 | 1 |
| | Anuy | Ob | 51.691417, 84.239306 | 1 |
| *T. svetovidovi* | Belin | Enisei | 52.19117, 98.6648 | 1 |
| | Belin | Enisei | 51.64529, 98.15946 | 1 |
| *T. thymallus* | Sieg | Rhine | 50.782222, 7.213333 | 1 |
| | Lafnitz | Raab-Danube | 47.245742, 16.084938 | 2 |
| | Saglbach | Inn-Danube | 47.311667´, 11.106111 | 1 |
| | Heiterwanger | Lech-Danube | 47.456944, 10.776667 | 1 |
| | Ain | Rhone | 46.7451317, 5.809925 | 1 |
| | Ure | Ouse | 54.305556, -1.939722 | 1 |
| | Padje | Kalixälven | 67.703038, 18.600424 | 1 |
| | Kaitumjaure | | | |
| | Mindyak | Ural | 53.984444, 58.799444 | 1 |
| *T. tugarinae* | Anui | Amur | 49.283333, 137.916667 | 1 |
| | Gobili | Anui-Amur | 49.294994, 138.520103 | 3 |
| *T. baicalolenensis* | Verkhnyaya | Baikal-Angara-Enisei | 55.864289, 108.896714 | 1 |
| x *T. baicalensis* | Angara | | | |
| *T. baicalolenensis* | | Baikal-Angara-Enisei | 54.555367, 111.399297 | 2 |
| x *T. baicalensis* | Barguzin | | | |
| *T. baicalolenensis* | | Amur | 51.715717, 134.309278 | 2 |
| x *T. burejensis* | Bureya | | | |
| *B. lenok* | Urunkhaika | Irtysh-Ob | 48.766667, 86.116667 | 1 |
| *C. renka* | Achensee | Seeache-Isar-Danube | 47.4755843, 11.7061217 | 1 |
| *H. hucho* | Pielach | Danube | 48.1993185, 15.4345248 | 1 |

2.2. DNA extraction, ddRADseq library preparation and sequencing

Total genomic DNA was extracted from fin clips or liver tissue using a high-salt ammonium acetate protocol (Winstanley and Rapley, 2003). DNA quality was checked using gel electrophoresis and quantified with an Implen Nanophotometer. So-called ddRADseq libraries, each containing 48 individuals, were prepared following the protocol of Peterson et al. (2012) with modifications described in Lecaudey et al. (2018) and Schedel (2020). For each individual, 400–500 ng of genomic DNA was digested using the restriction enzymes SphI and ApoI (New England Biolabs). Digested fragments were cleaned with a DCC-5 kit (Zymo Research) and adapters [unique P1 and common P2 from Peterson et al. (2012)] were ligated to the enzyme cut sites. Fragments > 250 bp were selected from each individual using 1.8X AMPure XP beads (Beckman Coulter) and quantified using an Implen Nanophotometer. Samples were subsequently pooled in equimolar proportions (100 ng DNA per sample), purified with a DCC-5 kit and size-selected (300 bp, tight range) on a BluePippin (Sage Science). For library

amplification we used a Phusion High-Fidelity PCR kit (New England Biolabs) with Illumina indexed primers [PCR1 and PCR2 (Idx1-Idx6)], according to Peterson et al. (2012). PCRs were done in 50 µl volumes with each reaction containing 16 µl ddH2O, 10 µl 5x Buffer HF, 1 µl dNTPs (10 mM), 5 µl Primer 1 (4 µM), 5 µl Primer 2 (4 µM), 0.5 µl Phusion Taq, and 12.5 µl template. Cycling conditions consisted of an initial denaturation at 98 °C for 30 sec, followed by 13 or 14 cycles at 98 °C for 40 sec, annealing at 65 °C for 30 sec and 72 °C for 30 sec and final extension at 72 °C for 7 min. PCR products were pooled and purified with a DCC-5 kit, followed by a final 1.8X AMPure XP size-selection (>250 bp) to remove primer dimers. The fragment size distribution of libraries was determined on a TapeStation 2200 (Agilent) and concentration was quantified using qPCR (NEBNext Library Quant Kit for Illumina). Genomic libraries were single-end (100 bp) sequenced on an Illumina HiSeq V4 (two libraries of 48 individuals each, pooled on a single lane) or a NovaSeq SP (four libraries of 48 individuals each, pooled onto two lanes) at the Vienna BioCenter Core Facilities (VBCF, Austria). To facilitate appropriate cluster detection, runs were spiked with 5 % PhiX (Illumina HiSeq V4) or 15 % PhiX (Illumina NovaSeq SP).

### 2.3. Raw data analyses

The raw sequencing data was sorted with samtools v.1.9 (Li et al., 2009) and quality was checked with FastQC v.0.11.8 (Andrews, 2010). The sequencing recovered an average of $4.62 \times 10^6$ reads per sample, with a minimum of $2.12 \times 10^6$ and a maximum of $19.12 \times 10^6$. Raw reads were deposited on NCBI under BioProject PRJNA842553 (Supplementary Table S1).

To determine homology and identify orthologous sites, the data was assembled in ipyrad v.0.9.81 (Eaton and Overcast, 2020) using a reference-based approach with the most recent available genome of *T. thymallus* (Linnaeus, 1758) (GCA_004348285.1; Sävilammi et al., 2019) and default parameters for raw read filtering. Quality filtered reads were mapped to chromosomes and unplaced scaffolds (CM014990.1–CM015040.1) were ignored. The mean coverage per sample ranged from 3x to 77x whereby only samples with > 5x coverage were retained resulting in downstream analysis of 145 (72 %) samples (Supplementary Table S1). The mean percentage of loci mapping to the reference genome was 97 % and ranged from 92 % to 98 % with the largest variance occurring among samples within and not between species, underscoring that sample quality and not interspecific divergence was the key factor influencing locus recovery. The minimum number of samples per locus (parameter 21) was set to 75 % for

each data set, and the assembly steps were replicated once for the phylogenetic dataset, comprising all species (18,667 loci, containing 212,145 SNPs), and once separately for each of the four contact zone datasets (13,799 to 27,457 SNPs).

### 2.4. Nuclear phylogeny and cophylogenetic analyses

The phylogenetic dataset included 128 individuals of *Thymallus* and three specimens from other Salmonidae subfamilies as outgroups (*Brachymystax lenok* Pallas, 1773, *Hucho hucho* Linnaeus, 1758, and *Coregonus renke* Schrank, 1783) (Supplementary Table S1). Individuals identified as admixed in downstream contact zone analyses (see below) were not included in the phylogenetic analysis. Maximum Likelihood (ML) phylogenetic inference was carried out in IQ-TREE v.1.6.12 (Nguyen et al., 2015) using a concatenated alignment of all ddRADseq loci. Using ModelFinder (Kalyaanamoorthy et al., 2017) and a Bayesian information Criterion (BIC), the evolutionary model TVM + F + R2 was chosen as the best fit for the dataset. The phylogeny was estimated using this model and 1,000 ultrafast bootstrap replicates (Hoang et al., 2018).

To evaluate the congruence between the nuclear-based phylogeny and the mitogenome phylogeny presented in Weiss et al. (2021), comparative cophylogenetic analyses were performed. To make the processes computationally efficient, both nuclear and mitogenome datasets were reduced to one individual per species (Supplementary Table S2). When available, the same individuals were kept in both trees. To test for congruence of mitogenome and nuclear phylogenies, the following tests were performed in IQ-TREE v.1.6.12 (Nguyen et al., 2015) using either the mitogenome or the nuclear alignments and a concatenated file with the mitogenome and nuclear ML trees: 1) Bootstrap Proportion (BP) (Kishino et al., 1990); 2) Kishino-Hasegawa (Kishino and Hasegawa, 1989); 3) Shimodaira-Hasegawa (Shimodaira and Hasegawa, 1999); 4) Approximately Unbiased (AU) (Shimodaira, 2002); and 5) Expected Likelihood Weights (ELW) (Strimmer and Rambaut, 2002). The tests were performed with 1,000 replicates using the Resampling Estimated Log-Likelihoods (RELL) method (Kishino et al., 1990). To identify taxa contributing to the potential topological incongruence, PACo analysis v.0.4.2 (Balbuena et al., 2013) was run in R v.4.0.2 (R Core Team, 2021) with 100,000 permutations, following the pipeline from Pérez-Escobar et al. (2016). For this analysis, the reduced ML trees were used as input. Additionally, MrBayes v.3.2.7a (Ronquist and Huelsenbeck, 2003) was run for the reduced datasets, using the best substitution models previously obtained in ModelFinder for the nuclear data and PartitionFinder2 v.2.1.2 (Lanfear

et al., 2012) for the mitogenome data (Weiss et al., 2021). Two runs were performed (107 generations, sampling every 100 generations), with default chains. The set of trees obtained from the first run for the mitogenome and ddRADseq data were used to infer the potential outlier associations in the PACo pipeline.

### 2.5. Potential LORe regions influence

Because ohnologs with lineage-specific delayed rediploidization (LORe regions) may influence phylogenetic inferences in salmonids (Lien et al., 2016, Robertson et al., 2017, Gundappa et al., 2022), we estimated the percentage of loci in LORe regions. The LORe regions identified in the *T. thymallus* genome by Gundappa et al. (2022) were used to build a reference database. The demultiplexed reads of each sample were mapped against this reference database with the program BBMap v.37.77 (Bushnell, 2014). The percentage of reads per individual, mapped to LORe regions was between 4.1 % and 8.5 % with an average of 6.2 % (Supplementary Table S3). The ML topology inferred on the assembly obtained after LORe region removal did not show any changes to our original analysis with LORe regions retained (data not shown) and thus all analyses were carried out with the full dataset.

### 2.6. Nuclear admixture analysis in the contact zones

Evaluation of potential admixture among sympatric species was carried out with fastSTRUCTURE v.1.0 (Raj et al., 2014) using a flat beta-prior with population-specific allele frequencies at each locus (known as a “simple prior”). For each dataset, values of K ranging from 1 to 4 (number of species plus 1) were evaluated, each with 10 replicates. Results obtained in fastSTRUCTURE were then imported into the webserver StructureSelector (Li and Liu, 2018), which calculates the most likely K and helps visualize the standard estimators used to choose the best K and plots the results with the integration of the program Clumpack (Kopelman et al., 2015). To depict the genetic relationships of individuals within each contact zone, Principal Components Analyses (PCA) were carried out using the R package adegenet v.2.1.3 (Jombart, 2008). Genetic differentiation between species in sympatry was evaluated using pairwise $F_{ST}$ values, estimated with the R package hierfstat v.0.5–7 (Goudet, 2005). Both analyses were performed on the reference-based assemblies using a single random SNP per locus.

## 2.7. Evaluating historical nuclear introgression

To evaluate potential events of historical introgression, Patterson's D (ABBA-BABA test) and the $f_4$-ratio (Patterson et al., 2012) along with the *f*-branch metric (Malinsky et al., 2018) were calculated with the software package Dsuite v.0.4 (Malinsky et al., 2021). The *f*-branch metric was used to assign signals of admixture to specific branches on a given phylogeny, based on the $f_4$-ratio results (Malinsky et al., 2021). Patterson's D and related statists are widely used to detect introgression and estimate admixture proportions ($f_4$-ratio). In short: given four taxa with the phylogenetic relationships (((P1,P2),P3),O) and considering "A" as the ancestral allele and "B" as the derived allele, the occurrence of two particular allelic patterns ABBA and BABA among the chosen taxa is calculated. These patterns should occur in equal frequencies when no introgression in present (due to ILS). In the presence of gene flow, an excess of the ABBA pattern (positive D-statistic) indicates gene flow between species P2 and P3, while an excess of the BABA pattern (negative D-statistic) indicates gene flow between species P1 and P3 (Durand et al., 2011, Patterson et al., 2012, Malinsky et al., 2021).

The analyses were performed on the phylogenetic dataset comprising 18,667 loci and using the ddRADseq-based topology. Potential hybrids identified with fastSTRUCTURE were excluded because they would bias the evaluation by introducing signals of recent introgression. Z-scores and associated *p*-values were calculated to assess the significance of the results.

## 2.8. mtDNA

To identify potential mitochondrial capture events, a fragment of 300 bp of the COI gene was amplified and sequenced for all contact zone samples for which mitochondrial data was not previously analyzed in Weiss et al. (2021). PCRs were performed using a pair of primers designed for metazoans mlCOIintF (Leray et al., 2013) and dgHCO2198 (Meyer, 2003). PCR conditions were as follows: 94 °C for 3 min, 94 °C for 30 sec, 54 °C for 40 sec, 72 °C for 1 min, and 72 °C for 10 min (34 cycles). PCR products were sent for sequencing to Macrogen (Spain) with the same primers. The quality of the sequences was evaluated in Chromas v.2.6.5 (https://www.technelysium.com.au/chromas.html) and each sequence was BLAST against the NCBI database for species identification. All new COI sequences are available on GenBank under the accession numbers ON711084–ON711175 (Supplementary Table S1).

## 3. Results

### 3.1. Nuclear phylogeny and cophylogenetic analyses

The first split in the ML phylogeny divided *T. flavomaculatus*, *T. grubii*, and *T. tugarinae* (all endemic to the Amur and coastal drainages in the Russian Far East) from other species in the genus (Fig. 1). A fourth species, endemic to the Amur drainage, *T. burejensis*, clustered outside this group forming a well-supported clade [bootstrap value (bs) 100], but with an unsupported internal node (bs 68). The remaining species formed two well-supported (bs 96, 100) groups.

One of these contained the widely distributed sister taxa *T. arcticus* and *T. baicalolenensis*. Both species revealed high levels of intraspecific substructure across distinct drainage systems. The well-supported clades within *T. baicalolenensis* (bs 100) corresponded to isolated populations in the Amur, Lena and Enisei rivers. Similarly, strong geographic structure within *T. arcticus* showed the presence of a "western" and "eastern" subclade, both being well-supported (bs 100); the "eastern" subclade was made up of populations from North America and the Russian Far East, and the "western" subclade included samples from the Ob, Enisei and Lena rivers.

The other major group consisted of species inhabiting drainage systems in Europe, Siberia and the Altai-Sayan Mountain region. Among the European species, only *T. ligericus* Persat, Weiss, Froufe, Secci-Petretto & Denys, 2019 was recovered as monophyletic (bs 100). The samples identified as *T. aeliani* Valenciennes, 1848 (SRR19545555, Piave River, and SRR19545556 and SRR19545501, both from the Adige River drainage) clustered together with samples of *T. thymallus* from different localities in the Danube River drainage north and south of the Alps (Fig. 1). Considerable phylogeographic structure was observed in the clade of taxa from Siberia and adjacent drainage systems. This well-supported group (bs 100) included species from Central- and Southern Siberia (*T. baicalensis*, *T. nigrescens* Dorogostaisky, 1923, *T. svetovidovi* Knizhin & Weiss, 2009, *T. nikolskyi* Kaschenko, 1899), Western Mongolia (*T. brevirostris* Kessler, 1879) and the Irtysh River drainage in Northern Kazakhstan (*T. brevicephalus* Mitrofanov, 1971). Except for *T. baicalensis*, clustering together with *T. nigrescens* from Chovsgul Lake (upper Selenga River), all these species formed well-supported monophyletic (bs 100) assemblages.

For the cophylogenetic analyses, both mitogenome and nuclear datasets included a total of 18 individuals (Supplementary Table S2). All tests performed in IQ-TREE rejected the hypothesis of congruence between the two topologies (Supplementary Table S4). The analysis

of incongruence performed with PACo, identified six out of the 15 *Thymallus* taxa with potential nuclear-mitochondrial outlier associations (Fig. 2), all with their entire 95 % confidence range exceeding the program's threshold for congruence (i.e., values of normalized squared residual above 0.6). These six taxa were *T. thymallus*, *T. ligericus*, *T. aeliani*, *T. baicalolenensis*, *T. arcticus*, and *T. burejensis*.

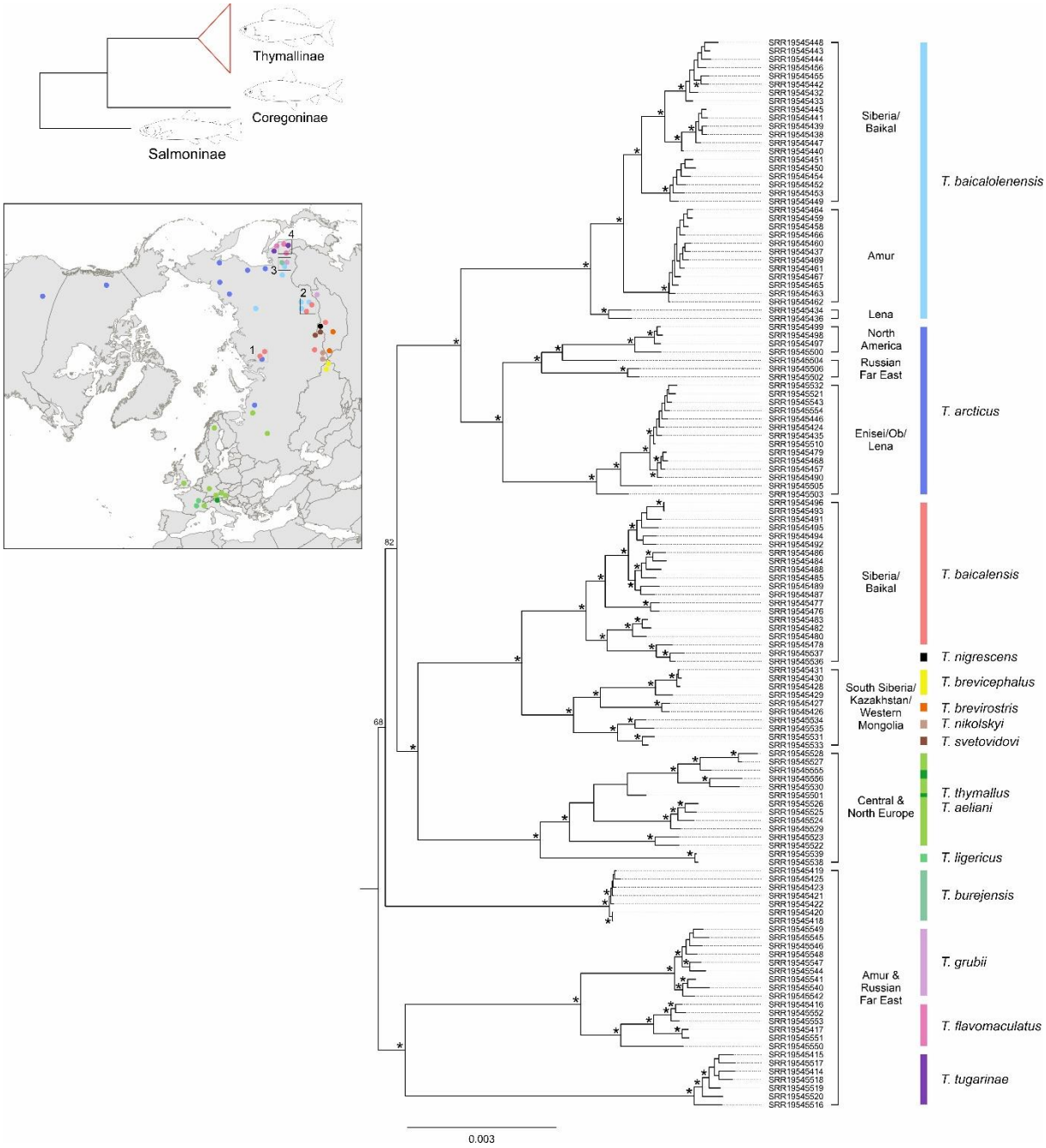



**Fig. 1.** Maximum Likelihood phylogeny inferred from genome-wide ddRADseq data for the genus *Thymallus*, including major drainages/geographic areas for each group of samples; asterisks indicate bootstrap values > 95 %. Outgroup taxa and phylogenetic placement in Salmonidae are depicted in the top left. Sampling localities are shown in the insert map together with the four contact zones analyzed: 1 = Kantaiskoye Lake system (*T. arcticus* and *T. baicalensis*), 2 = Lake Baikal basin (*T. baicalensis* and

*T. baicalolenensis*), 3 = upper Bureya River drainage (*T. baicalolenensis*, *T. burejensis* and *T. grubii*), 4 = Anui River drainage (*T. flavomaculatus* and *T. tugarinae*).

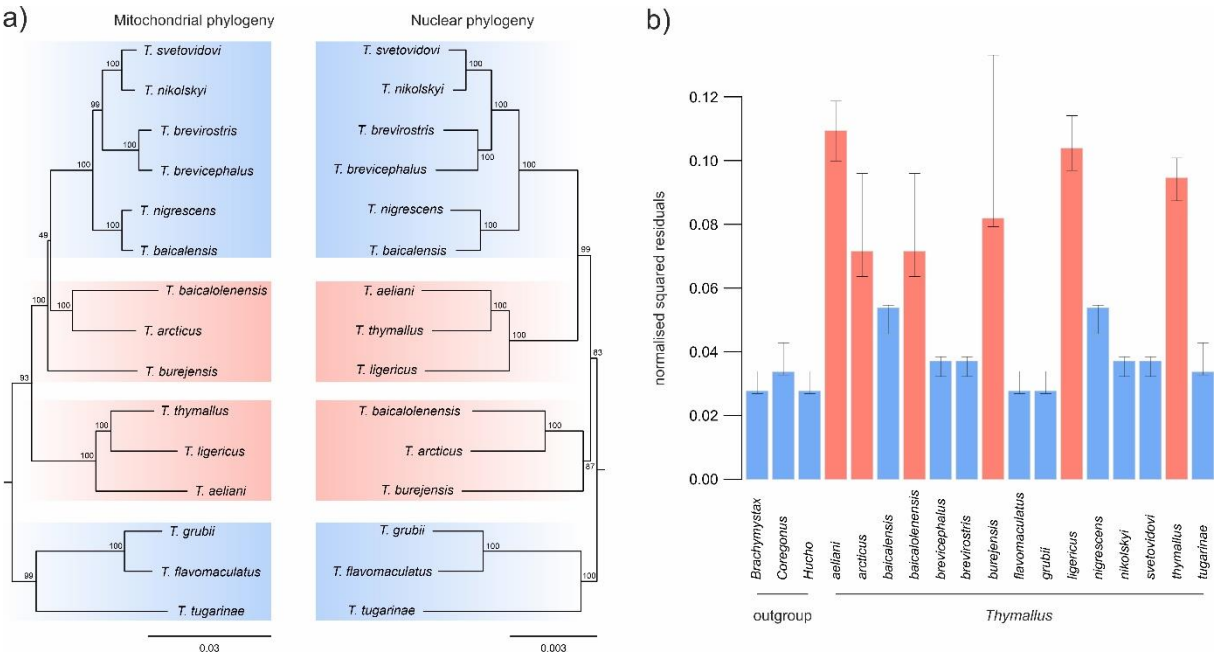



**Fig. 2.** Comparison between mitochondrial and nuclear phylogenies for *Thymallus* spp. a) ML phylogenies using a single individual per species. b) values of normalized squared residuals inferred with PACo for each species. Outliers with normalized square residuals above the program's standard threshold (0.6) are marked in red, mito-nuclear incongruencies are marked in blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### 3.2. Admixture analyses in the contact zones

In the Kantaiskoye Lake system, fastSTRUCTURE and PCA provided clear evidence for the distinction of *T. arcticus* and *T. baicalensis* (Fig. 3a). No clear signal of strong nuclear admixture was detected and pairwise $F_{ST}$ values (0.8) showed high genetic differentiation between the two species. Only one individual of *T. baicalensis* (SRR19545493, Gogochenda River) showed weak admixture with *T. arcticus* (8 % membership probability according to fastSTRUCTURE, K = 2). However, mitochondrial COI data revealed three cases of mitochondrial capture – two individuals (SRR19545435 and SRR19545543, Gogochenda River) with nDNA from *T. arcticus* and mtDNA from *T. baicalensis*, and one individual (SRR19545496, Gogochenda River) with nDNA from *T. baicalensis* and mtDNA from *T. arcticus* (Fig. 3a).

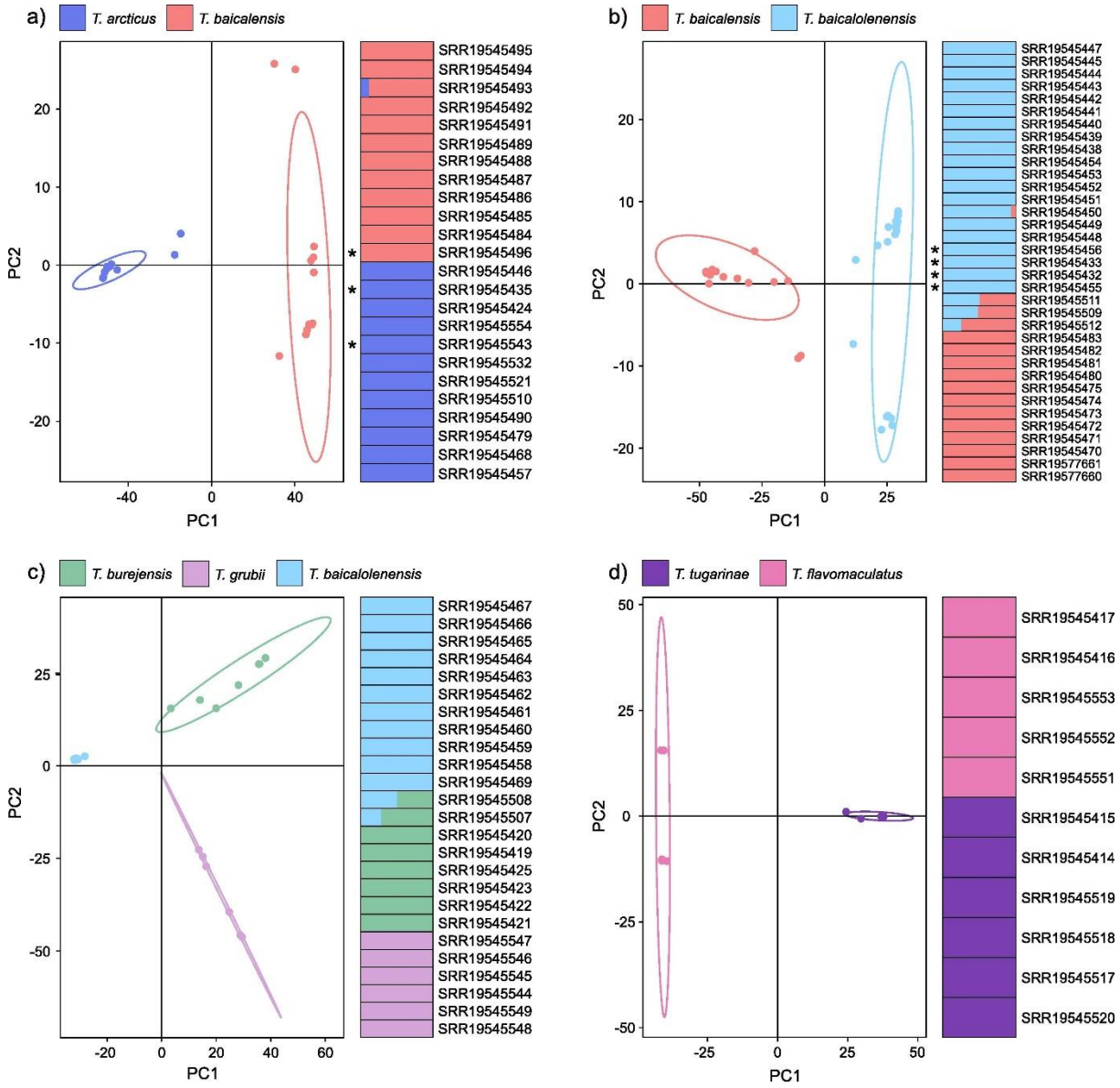
In the Lake Baikal basin where *T. baicalensis* and *T. baicalolenensis* co-occur in several tributaries, fastSTRUCTURE and PCA recovered two genetically distinct groups,

corresponding to the two species (Fig. 3b). Pairwise $F_{ST}$ values suggested high genetic differentiation ($F_{ST}$ = 0.7) between these species. Among the 37 individuals analyzed, three showed signals of admixture, two of which were presumably F1 hybrids (SRR19545511 and SRR19545509, both from the Barguzin River) and the third an F2 backcross (SRR19545512, Verkhnyaya Angara River). Assignment of hybrid generations (F1, F2) was based on membership probabilities to each group (at K = 2), with SRR19545511 (50 % *T. baicalensis*, 50 % *T. baicalolenensis*) and SRR19545509 (53 % *T. baicalensis*, 47 % *T. baicalolenensis*) showing almost equal proportions and SRR19545512 having a higher membership probability to *T. baicalensis* (77 %). Sequences of the mitochondrial COI fragment indicated that *T. baicalensis* is the maternal lineage for two individuals (SRR19545511 and SRR19545512), and *T. baicalolenensis* for the other individual (SRR19545509). Furthermore, putative mitochondrial capture events were identified in four individuals from the Tiya River that exhibited nDNA from *T. baicalolenensis* and mtDNA from *T. baicalensis* (Fig. 3b).

The three species in the upper Bureya River were genetically distinct (Fig. 3c) and pairwise $F_{ST}$ values ranged from 0.79 between *T. burejensis* and *T. grubii* to 0.86 between *T. baicalolenensis* and *T. grubii*. Two individuals showed admixture between *T. baicalolenensis* and *T. burejensis*; one an F1 hybrid (SRR19545508, 50 % *T. baicalolenensis*, 50 % *T. burejensis*) and the other a potential backcross (SRR19545507, 27 % *T. baicalolenensis*, 73 % *T. burejensis*). Based on the mitochondrial COI haplotypes, both admixed individuals stem from a female *T. burejensis* and a male *T. baicalolenensis*. For the remaining individuals, mtDNA was congruent with nDNA.

The analysis of *T. flavomaculatus* and *T. tugarinae* in the contact zone of the lower Amur River revealed no signal of admixture or mitochondrial capture. Both fastSTRUCTURE and PCA showed two distinct groups with a high level of genetic differentiation ($F_{ST}$ = 0.84).

**Fig. 3.** PCA and fastSTRUCTURE results for a) *T. arcticus* and *T. baicalensis* from the Kantaiskoye Lake system (based on 27,457 SNPs); b) *T. baicalensis* and *T. baicalolenensis* from the Lake Baikal basin (based on 16,093 SNPs); c) *T. baicalolenensis*, *T. burejensis* and *T. grubii* from the upper Bureya River drainage (based on 17,806 SNPs); d) *T. flavomaculatus* and *T. tugarinae* from the Anui River drainage (based on 13,799 SNPs). Asterisks indicate cases of mitochondrial capture.




3.3. Evaluating historical introgression

A total of 286 trio combinations were evaluated in the Dsuite analysis. Of those, 74 had a Z-score > 3 indicating a significant ($p < 0.002$) deviation from zero and providing support for gene flow in several comparisons. The *f*-branch method (based on the $f_4$-ratio results) highlighted 18 potential introgression events, but admixture proportions above 5 % were only found in seven cases (Fig. 4, Supplementary Table S5). The strongest signal of introgression (admixture proportion of 15 %) was evident between *T. arcticus* and the ancestor (internal branch in Fig. 4) of the species assemblage in Siberia and adjacent regions (*T. baicalensis*, *T. nikolskyi*, *T. svetovidovi*, *T. brevirostris*, and *T. brevicephalus*). Elevated values for the *f*-branch (7 % and 9 %) also suggested gene flow from *T. burejensis* and *T. baicalolenensis* into this

ancestral lineage. Admixture proportions were high (12 %) between *T. baicalensis* and the common ancestor of *T. svetovidovi* (upper Enisei) and *T. nikolskyi* (upper Ob) which may indicate that introgression took place before the split of *T. nikolskyi* and *T. svetovidovi*. Additionally, proposed events of gene flow had lower admixture proportions (<10 %) but supported introgression between *T. nikolskyi* and *T. brevicephalus* (allopatric species in the Ob drainage), and between *T. baicalensis* and *T. arcticus* (allopatric species with a geographically limited contact zone in the lower Enisei). *T. burejensis*, the only species with an unsupported internal node in the ML phylogeny, showed a weak signal (4 %) of gene flow with the ancestral branch of *T. arcticus* and *T. baicalolenensis*.

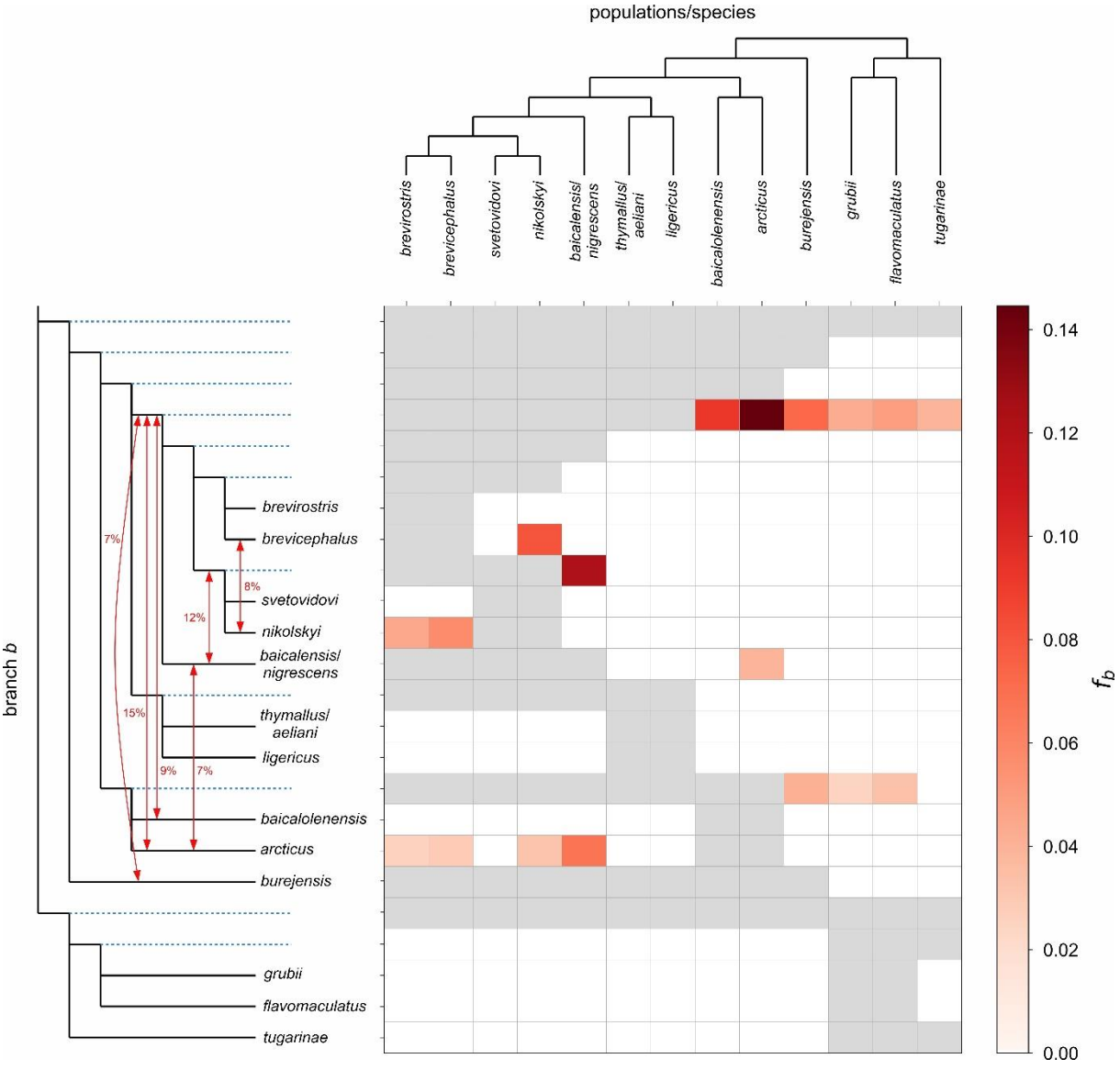



**Fig. 4.** Results of Dsuite *f*-branch statistics for *Thymallus* spp. The phylogenetic tree used is shown along the x- and y-axes (the y-axis showing an expanded form of the tree where both terminal- and internal

branches point to a corresponding row). Grey squares represent trio combinations that cannot be tested for due to topological constraints.

## 4. Discussion

The newly generated ddRADseq data provided the first high-resolution nuclear phylogeny of the genus *Thymallus*. Although species relationships were largely congruent with the previously published mtDNA-based phylogeny (Weiss et al., 2021), some important cyto-nuclear discordances underscored the known pitfalls of relying solely on organelle data for phylogenetic inference. Evaluation of hybridization and introgression in Asian contact zones of grayling revealed limited admixture among species, which may contrast somewhat with the expectations of gene flow in landscapes with high levels of paleo-hydromorphological dynamics, raising questions about the relevant mechanisms responsible for restricted gene flow. Despite signals of limited contemporary gene flow, several events of mitochondrial capture reflect historical events of hybridization that may not have been so clear based on ddRADseq data alone, highlighting the complimentary efficacy of using both organelle and nuclear gene analyses for evaluating the evolutionary history of organisms.

### 4.1. Nuclear phylogeny and cyto-nuclear discordance

All species that were resolved with the mtDNA-based analysis of Weiss et al. (2021) were recovered as monophyletic groups based on ddRADseq data, underscoring that the species-level resolution of the two analyses was relatively similar. The exception, reflecting cyto-nuclear discordance, is *T. aeliani*, which appeared scattered across minor clades of *T. thymallus* in the ddRADseq phylogeny, resulting in the paraphyly of both taxa. This is congruent with signals of admixture with stocked lineages previously reported using microsatellites (Sušnik et al., 2004, Meraner et al., 2014). Based on mtDNA data, *T. aeliani* is monophyletic and 1.8 % divergent from *T. thymallus* (Weiss et al., 2021). While the individuals bearing Adriatic mtDNA haplotypes were chosen to represent the Adriatic clade in the mitogenome phylogeny (as defined in Weiss et al., 2002, Meraner and Gandolfi, 2012), we could not retrieve a non-introgressed nuclear genome for *T. aeliani* due to the difficulty of finding a population with no record of stocking or admixture. This is generally problematic for analyzing phylogenetic relationships of species subject to decades of human-mediated admixture, common for salmonids throughout Europe (Sušnik et al., 2004, Weiss et al., 2013, Meraner et al., 2014), and

thus we hypothesize that the discordance observed for *T. aeliani* is based on anthropogenic introduction and subsequent introgression with exotic lineages.

Other instances of cyto-nuclear discordance in *Thymallus* involve the position of species or the placement of internal nodes in our phylogenies rather than the resolution of the species themselves.

First, the ddRADseq data support *T. ligericus*, endemic to the Loire River drainage (Persat et al., 2019) as the most divergent lineage among European taxa, while mtDNA-based analyses (Marić et al., 2014, Weiss et al., 2021) depict *T. ligericus* as a sister clade to the north and south Alpine Danubian lineages of *T. thymallus*. Interestingly, these relationships based on nuclear as opposed to mtDNA data are not congruent with any known phylogeographic patterns from other similarly distributed fishes (Costedoat and Gilles, 2009), including the co-occurring genera *Phoxinus* (Denys et al., 2020b, Palandačić et al., 2022) and *Salmo* (Hashemzadeh Segherloo et al., 2021). Much of our current knowledge of freshwater phylogeography in Europe, however, is still exclusively based on mtDNA. A wider geographic sampling in combination with nuclear genomic markers is needed to explain the colonization history of *Thymallus* in the Loire River.

Second, apart from the European taxa, strong cyto-nuclear discordance was evident in the placement of *T. burejensis*. The species is endemic to a single Amur River tributary system (Bureya River) but is not genetically related to the other Amur drainage species (*T. grubii*, *T. flavomaculatus*, *T. tugarinae*), showing a closer relationship to the clade containing *T. arcticus* and *T. baicalolenensis*. This pattern of poor support together with a short internal branch may suggest ancient incomplete lineage sorting, which is frequently observed in rapidly radiating species (Takahashi et al., 2001; Koblmüller et al., 2010). An alternative hypothesis considering ancient hybridization cannot be supported with the present ddRADseq dataset (D-statistics).

Third, cyto-nuclear discordance was also noted at one of the most internal nodes in the phylogeny. Based on the ddRADseq data, European grayling species form a strongly supported sister clade to species from Siberia and adjacent drainage systems, whereas the mitogenome-based analysis groups European species with all Asian species, except the three Amur taxa. This is noteworthy as the nDNA-based phylogeny is congruent with the hypothesized colonization routes for *Thymallus* from Asia to Europe (Weiss et al., 2021).

The examples of cyto-nuclear discordance observed in *Thymallus* were all visually obvious and reflected in the cophylogenetic analyses, however, they provide little insight into the mechanisms responsible. Cyto-nuclear discordance could be the result of either recent introgression, as seen for the relationships within European taxa, or historical introgression as

is likely to be the case for the different positions of the whole European clade in the mitogenome and nuclear phylogenies. Thus, instances of cyto-nuclear discordance, whether recent or more historical, are likely at least affected by, if not caused by, hybridization and introgression, the most common mechanism promoting cyto-nuclear discordance in freshwater fishes (Wallis et al., 2017).

Finally, one instance of a lack of species-level resolution exhibited no signal of hybridization nor cyto-nuclear discordance. Neither the mitochondrial nor the ddRADseq-based phylogenetic analyses supported the monophyly of *T. baicalensis* due to the position of *T. nigrescens* placed within *T. baicalensis* (Weiss et al., 2021; and Fig. 1). Nonetheless, the two species are morphologically distinct (Olson et al., 2019) and *T. nigrescens* forms a monophyletic group within the diverse assemblage of *T. baicalensis*. A recent study based on RADseq data and limited to samples from the Selenga River drainage revealed two different genetic clusters, corresponding to the two taxa, indicating that the paraphyly of *T. baicalensis* is most likely the result of very recent divergence of *T. nigrescens* (Roman et al., 2018).

### 4.2. Contact zones of *Thymallus* in Asia

The ddRADseq-based analysis of four contact zones revealed only very rare cases of contemporary hybridization, concordant with previous population genetic analyses using microsatellites (Froufe et al., 2003b, Weiss et al., 2007, 2020b), supporting relatively strong levels of reproductive isolation among grayling species, especially those in Asia. This may underlie why incidences of cyto-nuclear discordance were limited in our phylogenetic inference, but it begs the question of what mechanisms are operative in maintaining species integrity and if any patterns can be recognized between species that do exhibit some hybridization and introgression and those that do not.

The presence of morphological and genetic constraints, as well as environmental features delimiting contact or ecological niches often explain restricted gene flow between species (e.g., Feller et al., 2020, Kautt et al., 2020). Thus, it is important to consider to what extent physical, ecological or fine-scale niche segregation exists when evaluating reproductive isolation. For two contact zones in our analysis (the Kantaiskoye Lake system, and the Lake Baikal basin), some level of physical habitat separation may be playing a role in restricting interspecific gene flow. A recent study on the distribution of *T. arcticus* and *T. baicalensis* in the Kantaiskoye Lake system showed that these species may be physically delimited by the presence of waterfalls (Romanov, 2020). Our data are somewhat congruent with this picture, as no

contemporary gene flow is evident. However, historically it is clear that these taxa were in contact, as instances of mitochondrial capture are evident and ancient introgression at low levels, was supported by D-statistics. Similarly, in the Lake Baikal basin, both *T. baicalensis* and *T. baicalolenensis* occur in tributaries such as the Barguzin River (Knizhin et al., 2006b, 2008b). Knizhin et al. (2008b) showed that while *T. baicalensis* is primarily distributed in the lower reaches of these rivers, *T. baicalolenensis* is more restricted to their headwaters and isolated lakes in the mountain ranges. Thus, *T. baicalolenensis* never seems to reach Lake Baikal itself, despite the possibility of downstream migration, suggesting that competition or competitive exclusion from *T. baicalensis* or other species native to the lake may be occurring; see Alekseyev et al. (2021) for a similar distribution pattern observed in *Salvelinus alpinus* (Linnaeus, 1758). In the Barguzin River itself, as well as the Verkhnyaya Angara River, a northern tributary of Lake Baikal, we found signals of limited contemporary hybridization between the two species (F1 or F2 individuals). Knizhin et al. (2008b) noted that the migration of *T. baicalensis* into the upper reaches of these rivers may occur during spawning migration. Interestingly, in one additional northern tributary (Tiya River) with headwaters isolated from Lake Baikal, mitochondrial capture was evident but no sign of contemporary hybridization. Such examples of mitochondrial capture, presumably reflecting more ancient hybridization yet limited nuclear introgression, have been reported for other freshwater fishes such as *Pungitius* (Guo et al., 2019), but are also known from a wide range of organisms such as mammals (Good et al., 2015) and birds (Joseph, 2021).

In summary, both species pairs (*T. arcticus*/*T. baicalensis* and *T. baicalensis*/*T. baicalolenensis*) occur in the same drainage systems, but physical habitat separation may at least partly explain the pattern of reproductive isolation. Thus, although these paleo-hydrologically dynamic systems (e.g., Komatsu et al., 2016, Margold et al., 2018) repeatedly shaped periods of allopatry and contact among species (Weiss et al., 2020a), small-scale physical habitat structure can remain an important factor in restricting interspecific gene flow.

When physical separation is not operative, habitat or trophic niche segregation can also restrict gene flow and help maintain reproductive isolation. Both mechanisms are very common among freshwater fishes, well-documented in salmonid fishes (e.g., May-Mcnally et al., 2015, Olson et al., 2019, Jacobs et al., 2020), but occur predominately in lacustrine habitats (Greenwood, 1984, Simonsen et al., 2017, Dias et al., 2022) and generally seem to be rare in rivers (but see Levin et al., 2021). In our contact zone analyses within the Amur River drainage (Bureya and Anui rivers), the relevant species live in sympatry in multiple circumscribed areas (Knizhin et al., 2004; Antonov and Mikheev, 2016), but trophic niche segregation has not yet

been reported. The Amur River drainage is likely the ancestral range of the genus *Thymallus* (Weiss et al., 2021) and three endemic species (*T. grubii*, *T. flavomaculatus*, *T. tugarinae*) have presumably evolved in sympatry or parapatry. During sampling in these regions, primarily done by angling, it was not uncommon to catch different species on successive casts from the same pools, underlining their sympatric existence at a small physical habitat scale. In the Anui River, there was no evidence of contemporary introgression between *T. tugarinae* and *T. flavomaculatus*, two species that are distantly related but have either evolved in sympatry or have been in sympatry for 100 s of thousands if not several million years (Froufe et al., 2003b). In the Bureya River, there were two instances of recent hybridization but no introgression in the samples analyzed, and the revealed hybrids all involved *T. baicalolenensis*, a species that most likely evolved in the Lena River drainage and not the Amur (Knizhin et al., 2006b, Weiss et al., 2021). The absence of trophic niche segregation, however, would suggest that other mechanisms must be operative in maintaining reproductive isolation. Although exclusively based on conjecture, the leading hypothesis to date has been that external phenotype, such as the color pattern of the dorsal fins, which is very prominent and well-differentiated among grayling species, may be supporting species recognition and mate choice (Englmaier et al., 2022), and thus possibly serving as a reinforcement mechanism enhancing reproductive isolation in sympatry (Rundle and Schluter, 1998, Pfennig, 2016).

### 4.3. Possible scenarios of historical introgression in *Thymallus*

Despite the lack of contemporary introgression in Asian *Thymallus* species, inference of historical gene flow may help to understand current phylogenetic relationships and distribution ranges. Based on D-statistics, introgression between *T. arcticus* and the ancestor of species from Siberia and adjacent drainage systems may relate to ancient routes of colonization and subsequent species radiation leading to the diversity found in this area (Weiss et al., 2020a). Especially *T. arcticus* and *T. baicalensis* may have come into more frequent contact during the Pleistocene glacial periods, as repeated phases of hydrological connectivity are evident from proglacial lakes in the lower Enisei River drainage (Margold et al., 2018). The weaker signal of ancient introgression between *T. baicalolenensis* and the same ancestor of species from Siberia likely results from the shared ancestry between *T. baicalolenensis* and *T. arcticus*, producing such correlated *f*-branch signals (Malinsky et al., 2021).

Evidence of historical introgression was also found among species that are currently not in contact but inhabit a region (Altai-Sayan Mountain region) subject to repeated paleo-floods

and drainage re-arrangements (Weiss et al., 2020a, and references therein). For example, the closely related *T. nikolskyi* (upper Ob River) and *T. brevicephalus* (upper Irtysh River, Markakol Lake) inhabit different areas of the Ob River drainage but revealed a significant signal of introgression. The distribution ranges of these two species are not clearly defined (see also Weiss et al., 2020a), and it is unclear whether they are currently strictly allopatric, parapatric, or perhaps found in sympatry in a yet-to-be-disclosed area. In the Altai Mountains just northeast of Markakol Lake, the headwater tributary systems feeding the lake come into very close contact (a few km) with tributaries of the Ob River, near the locality where *T. nikolskyi* is found. This area may have served as a possible corridor between the Ob and Irtysh rivers, which today appear to be rather isolated reflected by the distinct freshwater ecoregions (Abell et al., 2008). A similar pattern of past introgression was detected between *T. baicalensis* and the common ancestor of *T. svetovidovi* and *T. nikolskyi*, which can also be related to opportunities for secondary contact between currently isolated drainage systems in the Altai-Sayan Mountain region (Weiss et al., 2020a). While *T. svetovidovi* occurs exclusively in the Enisei River drainage and *T. nikolskyi* is restricted to the Ob River drainage, *T. baicalensis* has a very broad distribution spanning from Lake Baikal and its tributaries to the lower Enisei River and even tributaries of the Ob River (Weiss et al., 2021). These results underscore the impact of paleo-environmental perturbations on distribution ranges and the evolutionary history, concordant with other such studies in the region (Bochkarev et al., 2018, Guo et al., 2019).

## 5. Conclusion

Although cyto-nuclear discordance seems to be common in the phylogenetic analyses of freshwater fishes, it appeared relatively limited across the genus *Thymallus*. The discordance that was found did not involve the resolution of species, except for *T. aeliani* where anthropogenic stock transfers are the likely cause of interspecific admixture and the lack of monophyly in our ddRADseq-based phylogeny. Where cyto-nuclear discordance was found, hybridization was seen as the most likely causative factor, affecting node placement only. Signals of hybridization and introgression were more common among species thought to have evolved in allopatry subsequently forming geographically limited zones of secondary contact, as opposed to species that presumably evolved in sympatry. Despite significant paleo-hydrological dynamics fostering cyclical events of isolation and contact among lineages, small-scale physical habitat separation likely plays a significant role in restricting interspecific gene flow. Where such separation is not evident, and species have been living for significant periods

in sympatry, other biologically based reinforcement mechanisms such as species recognition and mate choice are likely the major factor maintaining reproductive isolation. The continued use of organelle-based phylogenetic inference can be seen as an important complementary source of information, supporting ddRADseq and other nuclear genomic levels of phylogenetic analysis as specific events of historical hybridization and introgression can be verified and described in terms of the relevant donor species. Genome-wide nDNA markers, however, clearly provide the most reliable and robust phylogenies and are invaluable in describing historical introgression, including putative gene flow from or to ancestral species.

**Acknowledgements**

The authors are thankful to the editor and the reviewers for their constructive suggestions, which played a fundamental role in the improvement of this manuscript.

**CRediT authorship contribution statement**

Giulia Secci-Petretto: Methodology, Software, Formal analysis, Investigation, Writing – original draft. Gernot K. Englmaier: Methodology, Software, Formal analysis, Investigation, Writing – review & editing. Steven J. Weiss: Conceptualization, Methodology, Resources, Writing – original draft, Supervision. Alexander Antonov: Resources, Writing – review & editing. Henri Persat: Resources, Writing – review & editing. Gael P.J. Denys: Resources, Writing – review & editing. Tamara Schenekar: Methodology, Writing – review & editing. Vladimir I. Romanov: Resources, Writing – review & editing. Eric B. Taylor: Resources, Writing – review & editing. Elsa Froufe: Conceptualization, Methodology, Writing – review & editing, Supervision.

**Funding**

This work was supported by the Foundation for Science and Technology to GS (SFRH/BD/139069/2018) and EF (CEECIND/00627/2017), by Strategic Funding UIDB/04423/2020 and UIDP/04423/2020; and by a grant from the Doctoral Academy Graz, Ecology and Evolution in Changing Environments (EECE), to GKE.

**Supplementary Information**

Supplementary data to this article can be found online at
https://doi.org/10.1016/j.ympev.2022.107654

# Chapter 4

## Phylogeography and biogeographic origins of graylings (*Thymallus*, Salmonidae) in Europe: new insights from nuclear genomic SNP data

Gernot K. Englmaier[1,2]*, Nuria Viñuela Rodríguez[3], Jernej Bravničar[4], Lukas Zangl[1,5], Henri Persat[6], Saša Marić[7], Clemens Ratschan[8], Bo Delling[9], Duarte V. Gonçalves[10], Giulia Secci-Petretto[10], Elsa Froufe[10], Steven J. Weiss[1]

[1] Institute of Biology, University of Graz, Universitätsplatz 2, 8010 Graz, Austria
[2] Present address: Academy of Sciences of the Czech Republic, Institute of Vertebrate Biology, Květná 8, 60365 Brno, Czech Republic
[3] Department of Ecology, Faculty of Science, Charles University, Viničná 7, CZ-12844 Prague 2, Czech Republic
[4] Department of Animal Science, Biotechnical Faculty, University of Ljubljana, Groblje 3, 1230 Domžale, Slovenia
[5] Universalmuseum Joanneum, Studienzentrum Naturkunde, Weinzöttlstraße 16, 8045 Graz, Austria
[6] Société Française d'Ichthyologie, Muséum National d'Histoire Naturelle Paris, France, 57 rue Cuvier CP26, 75005 Paris, France
[7] Institute of Zoology, Faculty of Biology, University of Belgrade, Studentski Trg 16, 11001 Belgrade, Serbia
[8] 3ezb–TB Zauner GmbH, Marktstraße 35, 4090 Engelhartszell, Austria
[9] Department of Zoology, Swedish Museum of Natural History, P.O. Box 50007, SE-104 05 Stockholm, Sweden
[10] CIIMAR/CIMAR, Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Terminal de Cruzeiros do Porto de Leixões, Avenida General Norton de Matos, S/N, 4450-208 Matosinhos, Portugal

* **Corresponding author:** gernot.englmaier@uni-graz.at

### Abstract

Rivers and lake systems in Europe have undergone frequent paleoenvironmental changes, and thus provide a unique framework for understanding the evolutionary dynamics driving genetic structure of their fauna on a macro-biogeographic scale. Repeated shifting of hydrological isolation and connectivity have promoted faunal exchange between adjacent catchments, shaping fine-scaled phylogeographic structure. Here, the first large-scale phylogeographic study of *Thymallus* (Salmonidae) in Europe using genome-wide single nucleotide polymorphisms is presented. These data were compared with complete mitochondrial genomes and a meta-analysis of existing mtDNA control region data. The results corroborated previously

recognized genetic diversity but revealed strong mito-nuclear discordance suggesting a colonisation that dissents previous hypotheses based on mtDNA alone, and supports frequent events of secondary contact, introgression and mitochondrial capture. Early diversification probably originated from the Danube River with independent dispersal to adjacent drainage basins. Divergence time estimates within extant lineages coincide with the Pleistocene, suggesting multiple spatially separated glacial refugia. Splits between major phylogeographic lineages, however, appear considerably older dating back to the early Pliocene (the late Miocene from mtDNA data). These results challenge earlier mtDNA-based hypotheses on the role of late Pleistocene glacial cycles as a general driving force for lineage differentiation.

**Keywords:** biogeography, ddRADseq, freshwater fish, glacial refugia, mito-nuclear discordance, palaeohydrology

## 1. Introduction

Climatic oscillations in the Pleistocene (2.6–0.01 Ma) were major drivers for the evolution and distribution of biota in the temperate and arctic regions of Europe, as alternating glacial/interglacial periods dictated the retraction and expansion of distribution ranges (Hewitt, 1999, 2004). However, the evolutionary history of most organism groups cannot be explained by Pleistocene climatic changes alone. Fossil records support the presence of many present-day taxonomic groups in Europe since the late Miocene (11.6–5.3 Ma) and Pliocene (5.3–2.6 Ma) (e.g., Kovács et al., 2020; NOW Community, 2022). At this time, both the landscape topography and major drainage systems were largely established (Popov et al., 2004), allowing widespread colonisation. Drastic climatic changes during the Pleistocene led to frequent and significant changes in range boundaries, with regression or expansion of distribution ranges following suitable climatic conditions (Hewitt, 2004). These periodic variations introduced considerable complexity in distribution ranges and refugial areas (e.g., Weiss and Ferrand, 2002; Sommer and Zachos, 2009; Hantemirova et al., 2017; Horreo and Fitze, 2018; van Rensburg et al., 2021), but also led to a noticeable coincidence in re-colonisation patterns and zones of secondary contact among European taxa (Hewitt, 2004).

Freshwater fishes are of particular interest because their (re-)colonisation pathways are usually dependent on waterway connectivity and thus may reflect the evolution of drainage systems at the catchment scale (Van Steenberge et al., 2020). Changes in size and flow direction of drainage systems are linked to processes of long-term landscape evolution involving tectonic

events, hydrology, hydrogeomorphology, erosion, and climate (Coulthard and Van De Wiel, 2012), triggering temporal connectivity between drainage systems and allowing the unidirectional or bidirectional exchange of biota (Bishop, 1995). Allopatric differentiation among distinct drainage (sub-)systems is well documented in the European ichthyofauna (e.g., Bănărescu, 1992) and described for widely distributed genera such as *Phoxinus* (Palandačić et al., 2022), *Squalius* (Seifertová et al., 2012), *Telestes* (Gilles et al., 2010), or *Thymallus* (Marić et al., 2014). Not all European fish species, however, show such clear patterns of lineage diversification among contemporary drainage networks and are thus characterised by rather low geographical variability (e.g., *Cobitis taenia* (Culling et al., 2006); *Vimba vimba* (Hänfling et al., 2009)). Drainages with dynamically shifting phases of isolation and connectivity reflect historical complexity promoting the expectation of high intra- and interspecific lineage diversity. Such palaeohydrological dynamics are evident in the headwaters of the Danube, Elbe, Rhine, and Rhône rivers (Berger et al., 2005) with evidence for cross-drainage distribution of mitochondrial DNA (mtDNA) haplotypes in *Cottus gobio* (Englbrecht et al., 2000; Šlechtova et al., 2004), *Phoxinus csikii* (Palandačić et al., 2017), *Thymallus thymallus* (Weiss et al., 2002), and *Salmo trutta* (Lerceteau-Köhler et al., 2013).

The genus *Thymallus* (graylings) has been frequently shown to be a suitable model to study the interplay between paleohydrology and lineage differentiation on a macro-biogeographic scale (e.g., Stamford and Taylor, 2004; Weiss et al., 2021). On the European continent, the phylogeographic structure of *Thymallus* has been intensively studied over the past 20 years, initially relating to two species – *T. thymallus* (European grayling) and *T. arcticus* (Arctic grayling); whereby *T. arcticus* inhabits only few river systems in the extreme north-east of Europe (Shubin and Zakharov, 1984; Koskinen et al., 2000) with its primary distribution area in Siberia and North America (Weiss et al., 2021).

Within the *T. thymallus* clade, Weiss et al. (2002) and subsequent phylogenetic studies identified eight divergent mtDNA lineages – Adriatic, Loire, Northern Alps (Danube), Southern Alps (Danube), Western Balkans (Danube), Scandinavian, Caspian, and a "Mixed Central European" clade with haplotypes from the Rhône, Rhine, and Elbe rivers as well as drainage systems in England, Denmark and the Baltic region (see Gum et al., 2009; Marić et al., 2012, 2014). The deep divergence across these lineages, with a pairwise difference of up to 3.8% (based on mtDNA), and the presence of distinct groups of haplotypes in large drainage systems such as the Danube suggested a complex scenario of expansions from different refugial areas during the Pleistocene cold periods (Weiss et al., 2002). Divergence time estimates (Weiss et al., 2021) and fossil data (Rückert-Ülkümen and Kaya, 1993; Rückert-Ülkümen et al., 2006),

however, indicated that the evolutionary history of *Thymallus* in Europe may be considerably older than the Pleistocene, potentially predating the Messinian Salinity Crisis (MSC) in the late Miocene. Based on these observations, morphological and genetic studies (Persat et al., 2019; Bravničar et al., 2020) supported the distinction of two additional species in Europe – *T. ligericus* (Loire grayling) and *T. aeliani* (Adriatic grayling), rendering *T. thymallus* as paraphyletic (Weiss et al., 2021).

A limitation of these earlier studies was the strong reliance on mitochondrial genetic markers or limited geographic sampling (Secci-Petretto et al., 2023). Despite the advantages of mtDNA, examples of mito-nuclear discordance are widespread, including also salmonid fishes (e.g., Yamamoto et al., 2006) and often lead to incorrect or incomplete assumptions on phylogenetic relationships (Toews and Brelsford, 2012; Wallis et al., 2017). Thus, the present study aims to 1) re-evaluate the phylogeographic structure of European *Thymallus* spp. (*T. thymallus*, *T. aeliani*, *T. ligericus*) using nuclear genomic markers obtained from double digest restriction-site associated DNA sequencing (ddRADseq) and compare them with complete mitochondrial genomes (hereafter 'mitogenomes') and a meta-analysis of existing mtDNA control region (CR) data; and 2) unravel the colonisation history of *Thymallus* in Europe in the context of major paleohydrological scenarios.

## 2. Materials and Methods

### 2.1. ddRADseq

#### 2.1.1. Sampling, library preparation and sequencing

Samples ($N$ = 217) include all three European species and all major European drainage systems representing previously known mitochondrial lineages (Figure 1a, Supporting Information 1, Table S1). Due to difficulty obtaining samples from unmanaged Adriatic drainage populations, *T. aeliani* is presented by only two samples from one locality. Tissue samples (fin clips or muscle tissue) were preserved in 95% ethanol and genomic DNA was extracted using a high salt (ammonium acetate) protocol (Sambrook et al., 1989). Some samples stem from either extractions (stored at -20°C) or tissues (stored at either -20°C or room temperature) up to 20 years old. DNA integrity was assessed on a 1% Agarose gel, and quantified using a NanoDrop Spectrometer.

Double digest restriction-site associated DNA (ddRAD) library preparation followed Peterson et al. (2012) with modifications described in Secci-Petretto et al. (2023). In brief, 400–500ng of genomic DNA were used for enzyme digestion using the restriction enzymes *SphI* and *ApoI* (New England Biolabs). Each ddRADseq library contained 48 individual samples and a 300 bp size selection (tight range, 1.5% cartridge) on a BluePippin (Sage Science). Phusion PCR amplification was done with custom primers (PCR1 and PCR2.1–PCR2.6, Illumina indexed primers following Peterson et al. (2012)) and the following cycling conditions: initial denaturation at 98°C for 30 sec, followed by 13 or 14 cycles at 98°C for 40 sec, 68°C for 30 sec and 72°C for 30 sec, and a final extension at 72°C for 7 min. Pooled samples (48 per index) were sequenced on Illumina HiSeq 2500 (a single library per lane) or Illumina NovaSeq (six libraries pooled on two lanes) platforms using the 100 bp single-end read mode by the NGS Facility at Vienna BioCenter Core Facilities (VBCF, Austria).

### 2.1.2. Data processing and assembly

Quality control of raw reads was performed using FastQC v.0.11.8 (Andrews, 2010). Raw reads were then demultiplexed, quality-filtered (with default parameters), and assembled using ipyrad v.0.9.81 (Eaton and Overcast, 2020) with trimming the first five base pairs of each read (*parameter 25*). Raw ddRADseq reads have been deposited at NCBI's Sequence Read Archive (BioProject: XXXXX, Table S1).

A reference-based approach was applied to determine homology and identify orthologous sites using the latest reference genome of *T. thymallus* (ASM434828v1; Sävilammi et al., 2019). Quality-filtered reads were mapped to chromosomes (CM014990.1–CM015040.1) only, ignoring unplaced scaffolds. The percentage of mapped reads ranged from 95.8% to 98.4% per individual with a mean coverage across all loci between 1.8x and 79.8x. Samples with a mean coverage < 5.0x were excluded ($N$ = 39), retaining 178 individuals for analysis. The minimum number of samples per locus (*parameter 21*) was set to 85% ($N$ = 151) to create a stringent dataset, resulting in a total of 19,574 loci and 103,943 single nucleotide polymorphisms (SNPs). Other parameters were left at default. Assemblies with a lower minimum number of samples per locus resulted in higher numbers of polymorphic loci and SNPs (75%: 26,692 loci, 143,309 SNPs; 65%: 33,971 loci, 184,323 SNPs).

Additionally, a de novo (reference-independent) approach was applied in ipyrad to confirm that the reference-based assembly (using one species) does not deliver biased results. Orthologous sites were identified with default parameters and a clustering threshold of 0.85.

Three datasets were constructed with a minimum number of samples per locus (*parameter 21*) using 85%, 75%, and 65%. The resulting matrices contained 11,894, 16,504, and 21,163 loci with 55,632, 79,113, and 103,426 SNPs respectively.

### 2.1.3. Genetic structure

To determine genetic clusters and visualize inter-cluster differentiation without prior information on the underlying genetic structure, a Principal Component Analysis (PCA) was carried out on SNP genotypes using the R packages adegenet v.2.2.5 (Jombart, 2008; Jombart and Ahmed, 2011) and ade4 v.1.7-18 (Thioulouse et al., 1997; Chessel et al., 2004; Dray and Dufour, 2007). The analysis was run on both the reference-based SNP matrix with 19,574 loci and the de novo assembly with 11,894 loci (outgroup taxa excluded), using a single random SNP per locus.

To provide an alternative, genetic model-based evaluation of individual admixture and genetic structure, fastSTRUCTURE v.1.0 (Raj et al., 2014) was used on the same assemblies as for PCA. Values of *K* ranging from 1–15 were tested on the complete dataset, comprising all individuals, using a simple prior and 10 replicates each. StructureSelector (Li and Liu, 2018) was used to evaluate the optimal number of clusters. Separate fastSTRUCTURE analyses were run for each of the drainage specific clusters identified, following a hierarchical approach.

### 2.1.4. Phylogenetic analysis

A maximum likelihood (ML) approach implemented in IQtree v.2.1.4 (Minh et al., 2020) was used to infer evolutionary relationships among *Thymallus* throughout Europe based on a concatenated alignment of all ddRADseq loci. Two Asian taxa, *T. flavomaculatus* and *T. tugarinae*, were used as outgroups. The best-fit nucleotide substitution model (TVM+F+R2) was identified using ModelFinder (Kalyaanamoorthy et al., 2017) implemented in IQtree with three independent runs and based on the BIC. IQtree was run on both the reference-based (19,574 loci, 26,692 loci, 33,971 loci) and the de novo (11,894 loci, 16,504 loci, 21,163 loci) assemblies. ML phylogenies were estimated with 1,000 ultrafast bootstrap replicates, applying the *-bnni* flag to reduce the impact of overestimating branch support (Hoang et al., 2018), as well as with 100 standard nonparametric bootstrap replicates.

### 2.1.5. Historical introgression

To identify signals of past gene flow between distinct phylogenetic lineages, Patterson's *D* and $f_4$-ratio statistics (estimating the admixture proportion) (Patterson et al., 2012), and the *f*-branch metric (Malinsky et al., 2018), were calculated with Dsuite v.0.4 (Malinsky et al., 2021). Analyses were run with both the reference-based (19,574 loci) and the de novo assemblies (11,894 loci) using the topology inferred from ML analysis. Individuals identified as hybrids by fastSTRUCTURE ($N$ = 32) were excluded from the analyses. Samples from the Sesia River (> 94% membership at $K$ = 6 in fastSTRUCTURE), and the Drava and Mur rivers (> 96% membership at $K$ = 6 in fastSTRUCTURE) were kept as they represent the purest samples from the Adriatic and the Danubian Southern Alps lineages in the dataset, respectively. Assignment of groups (i.e., lineages) followed fastSTRUCTURE analysis; outgroups remained the same as those used for phylogenetic inference.

2.1.6. Divergence time estimation

A time-calibrated phylogeny was inferred under the multi-species-coalescent model and a strict molecular clock implemented in SNAPP v.1.0 (Bryant et al., 2012), an add-on package for BEAST2 v.2.6.6 (Bouckaert et al., 2014). Group assignment followed the delimitation of distinct genetic clusters in fastSTRUCTURE. Samples identified as hybrids ($N$ = 32) were not included, except for specimens from the Sesia, Drava and Mur rivers (see above). Due to the high computational demands in SNAPP analyses with large sample sizes (Stange et al., 2018), the number of analysed individuals was limited to two per group because the smallest sample size per genetic cluster was two (Adriatic). To account for within-lineage variability, SNAPP analyses were performed on 10 random subsets using the reference-based assembly with 19,574 loci. Outgroups were the same as in ML analysis. The datasets were reduced to sets of 13,859–16,858 bi-allelic SNPs (following: https://github.com/mmatschiner/tutorials/tree/master/divergence_time_estimation_with_snp_data) with a single SNP per locus; sites closer than 300 bp from one another (size selection in ddRADseq libraries was 300 bp) were excluded. XML files were generated using a RUBY script "snapp_prep.rb" (https://github.com/mmatschiner/snapp_prep). For each of the 10 datasets, two independent runs were performed with a chain length of 5 million MCMC iterations. Convergence (estimated sample size > 200) was checked with Tracer v.1.7.2 (https://github.com/beast-dev/tracer) and trees were visualised with DensiTree v.2.2.7 (Bouckaert, 2010), discarding the first 10% as burn-in. A maximum clade credibility tree

(MCC) was summarized with TreeAnnotator v.2.6.7 (Drummond et al., 2012) and visualised in FigTree v.1.4.4 (Rambaut, 2018).

### 2.1.7. Calibration points

The fossil record of *Thymallus* in Europe is scarce and most fossil remains are dated to the late Pleistocene and early Holocene (e.g., 127–10 ka in Böhme and Ilg (2003); 43–27 ka in Conard et al. (2013); approx. 100–10 ka in Stefaniak et al. (2020)). Divergence time estimates based on mitogenomes (Weiss et al., 2021) and the oldest fossil data (Rückert-Ülkümen and Kaya, 1993) both predate the Pleistocene. The earliest suggested evidence of the genus in Europe dates back to the uppermost Miocene (Rückert-Ülkümen and Kaya, 1993) but cannot be dated with certainty and is thus not used in this study (see Supporting Information 2).

The lack of unequivocally assignable fossil records over long periods means that alternative calibration points (paleogeographical data) must be sought. In most cases, geomorphological settings in tectonically active regions allow identification of time slots for possible freshwater pathways between different catchments. Here, two calibration points were used. First, the estimated last direct connection between the Rhône and Danube catchments (Berger et al., 2005, map 19, age classification: 4.2–2.9 Ma) was used to set the minimum time of the most recent common ancestor (MRCA) of European *Thymallus* spp. (using a uniform distribution), dating the split between the ancestral Rhône/Loire and the Danube clades. The chronostratigraphic unit of interest is the stratum "Sundgau gravel" (local term: "Sundgauschotter"). Second, a calibration point was placed at 1.8–0.5 Ma (uniform distribution) to date the split between *Thymallus* in the Rhône and Loire rivers. Based on the poorly documented sand and gravel deposits of the fluvio-lacustrine formation of Guye (local formation name in Middle-Eastern France), interconnections between these catchments seem likely, for example via the watershed at Montchanin (Dept. Saône-et-Loire) during the Pleistocene (Straffin et al., 1999, fig 3b; also refer to explanations of geological maps of France 1:50,000, e.g., Bonvalot et al., 1984; BRGM, 1999; Donzeau et al., 2001).

### 2.1.8. Historical biogeography

Ancestral ranges were estimated with the R package BioGeoBEARS (Matzke, 2013) under all available models (DEC and DEC+J, DIVA-like and DIVA-like+J, BAYAREA-like and BAYAREA-like+J). The best fitting model was selected using the Akaike information criterion

(AIC). Eight geographic ranges were defined based on the present and past hydrographic networks as well as the current distribution of *Thymallus* in Europe, which largely reflect the freshwater ecoregions (Abell et al., 2008): a) Northern Adriatic basin, b) Black Sea basin (Danube River), c) Loire River, d) Rhône River, e) North Sea and Southern Baltic Sea basins, f) Scandinavia (most of the Northern Baltic Sea basin), g) Barents Sea basin (Pechora River), and h) the Caspian Sea basin (Volga and Ural rivers). Extant *Thymallus* lineages inhabit a maximum of two of these geographic ranges; dispersal between nonadjacent areas was restricted.

## 2.2. Mitochondrial DNA

### 2.2.1. Mitochondrial CR

The complete dataset comprised 1,221 sequences retrieved from GenBank and 1,179 unpublished sequences stemming from contract reports (133 localities from 11 river drainages). For the latter, the complete mtDNA CR was amplified using the primers LRBT-25 and LRBT-1195, following Weiss et al. (2013) (Supporting Information 2). Newly obtained haplotypes were deposited on GenBank under accession numbers XXXXX–XXXXX.

Sequences retrieved from GenBank have been used in a range of phylogenetic studies (e.g., Weiss et al., 2002; Duftner et al., 2005; Marić et al., 2011, 2012, 2014; Meraner and Gandolfi, 2012; Persat et al., 2016). Specimens originating from hatcheries and stocked rivers (outside of the native range) were not considered, but samples from a Finnish hatchery representing stocks from the Oura River (Weiss et al., 2002) were included. Three haplotypes (AY841359.1, JX961600.1, MT062993.1) were discarded due to the presence of an 82 bp repeat (Sušnik et al., 2006; Meraner et al., 2013) found in diverse salmonids. The poly-T region was removed from all sequences due to its high variability (unreliability of alignments), which is likely phylogenetically uninformative (see discussion in Sušnik et al., 2006).

Sequences were aligned with ClustalW, implemented in BioEdit v.7.2.6 (Hall, 1999). Haplotypes were sorted with DnaSP v.6.12.01 (Rozas et al., 2017). The final alignment had a length of 959 bp (trimmed to the length of the shortest sequences). Haplotypes were then classified into 12 distinct groups, based on previously published data (see above): 1) Adriatic (Northern Adriatic basin), 2) Danube – Northern Alps (northern Alpine drainages), 3) Danube – Southern Alps (southern Alpine drainages), 4) Danube – Lafnitz (Lafnitz River, Austria), 5) Danube – Western Balkans (Western Balkan drainages), 6) Danube – Tisza (Tisza River,

Romanian Carpathians), 7) Loire (Atlantic Loire River), 8) Rhône (Mediterranean Rhône River), 9) North Sea basin – Rhine (Rhine and Maas rivers, and drainages from England), 10) North Sea basin – Elbe (Elbe River), 11) Scandinavia, and 12) North-Eastern Europe (Caspian (Volga and Ural rivers) and Barents Sea (Pechora River) basins). The haplotype network was constructed with TCS v.1.21 (Clement et al., 2000) following statistical parsimony (Templeton et al., 1992) with a 95% connection limit, and visualized in tcsBU (TCS Beautifier; dos Santos et al., 2016) and Inkscape v.0.92.3 (Inkscape Project, https://inkscape.org). MEGA v.7.0 (Kumar et al., 2016) was used to calculate pairwise distances, as well as uncorrected *p*-distances within and between the groups described above.

### 2.2.2. Mitogenomes

Ten mitogenomes of *Thymallus* from Central-, Northern-, and South-Eastern Europe were sequenced on an Illumina Platform at Novogene (UK, Europe) using a paired-end read mode (150 bp) with approx. 2 Gb per individual. For the assembly, the custom-made bioinformatic pipeline MitoComp (https://github.com/SamLMG/MitoComp) was used. This pipeline includes quality trimming (trimmomatic, Bolger et al., 2014), subsampling (seqtk, Li, 2012) and interleaving (BBMap, Bushnell, 2014) of raw short reads followed by the assembly of mitogenomes from genomic reads using five different third-party assembly tools (Mitoflex, GetOrganelle, NOVOplasty, Norgal and MITObim – see Li et al., 2021; Jin et al., 2020; Dierckxsens et al., 2017; Al-Nakeeb et al., 2017; Hahn et al., 2013) and subsequent annotation. The pipeline was run omitting the subsampling option, utilizing all five different assemblers and a COI sequence of *T. thymallus* (KM373646.1; Knebelsbeger et al., 2015) with 10 rounds for GetOrganelle (preset) followed by annotation (see Supporting Information 2 for details). Assemblies for subsequent analyses were chosen based on annotation completeness (number of recovered genes, tRNAs and rRNAs) and overall assembly length. Sequences were deposited on GenBank under accession numbers XXXXX–XXXXX (Table S2).

These 10 new mitogenomes were combined with 17 existing mitogenomes from GenBank, chosen to cover all previously known lineages of European *Thymallus* spp., together with *T. flavomaculatus* and *T. tugarinae* as outgroups. Alignment and phylogenetic reconstruction followed Weiss et al. (2021) but were implemented in PhyloSuite v.1.2.2 (Zhang et al., 2020). ML analysis was performed with IQtree v.2.1.4 (Minh et al., 2020) using 1,000 ultrafast bootstrap replicates (Hoang et al., 2018) under the nucleotide substitution model GTR+I+G. For Bayesian Inference (BI), MrBayes v.3.2.6 (Ronquist et al., 2012) was used, with

two independent Markov Chain Monte Carlo searches (MCMC) run simultaneously for 5 million generations, sampling trees every 100 generations. The first 25% were discarded as burn-in. Parameter convergence and stationarity were assessed with Tracer v.1.7.2 (Rambaut and Drummond, 2007). The computed ML and BI trees were visualised with FigTree v.1.4.4 (Rambaut, 2018). MEGA v.7.0 (Kumar et al., 2016) was used to calculate uncorrected *p*-distances between groups.

## 3. Results

### 3.1. Genetic structure and admixture

#### 3.1.1. Nuclear genomic SNPs

The analyses of nuclear genomic SNPs supported the recognition of several genetically distinct groups within European *Thymallus* spp. (Figure 1) with multiple cases of genetic clusters shared across drainages, especially in the Upper Rhine and Danube rivers. Highest genetic differentiation was centred around the southern margin of the distribution range, associated with the European Alps.

Strong geographic structure with nine distinct groups was evident along the first seven principal components (PCs). The first two axes (8.6% and 7.3% of the total variance) distinguished three genetic clusters: 1) the two French river drainages, Rhône and Loire, 2) the North Sea basin and river systems in Scandinavia and North-Eastern Europe, and 3) the Danube River drainage including the Northern Adriatic basin (Figure 1b). Along PC3, samples from Scandinavia and North-Eastern Europe were separated from the North Sea basin (Figure 1b); PC4 differentiated four distinct genetic groups within the Danube River drainage, corresponding to major drainage subsystems: 1) Northern Alps, 2) Southern Alps, 3) Western Balkans including the Northern Adriatic basin, and 4) the Tisza River; PC5 separated samples from the Tisza River from other Danubian samples; PC6 distinguished samples from the Rhône and Loire rivers; and PC7 separated the Adriatic samples from all other groups.

A similar pattern was recovered with fastSTRUCTURE (Figure 1c). Based on the marginal likelihood and the median and maximum number of clusters (MedMeaK, MedMedK, MaxMeaK, MaxMedK), the optimal number of clusters for the global dataset was $K = 5$ (marginal likelihood) or $K = 6$ (both median and maximum), corresponding to: 1) Rhône and Loire rivers, 2) Scandinavia and North-Eastern Europe, 3) the North Sea basin, 4) Danubian

Northern Alps, 5) Danubian Tisza River, and 6) a homogenous group of Danubian (Southern Alps, Western Balkans) and Adriatic samples. Hierarchical partitioning (2nd and 3rd row of Figure 1c) of these assemblages showed additional geographic structure similar to results obtained with PCA, adding fine-scale genetic differentiation in Scandinavia and North-Eastern Europe as well as between the Upper- and Lower Rhine River.

The fastSTRUCTURE analysis also revealed limited degrees of admixture between some regions such as the North Sea basin and the Danubian Northern Alps, and between the North Sea basin and the Rhône River (Figure 1c). Lower levels of admixture were found between the North Sea basin and Scandinavia, as well as between Scandinavia and North-Eastern Europe represented by two individuals from the western headwaters of the Volga River.

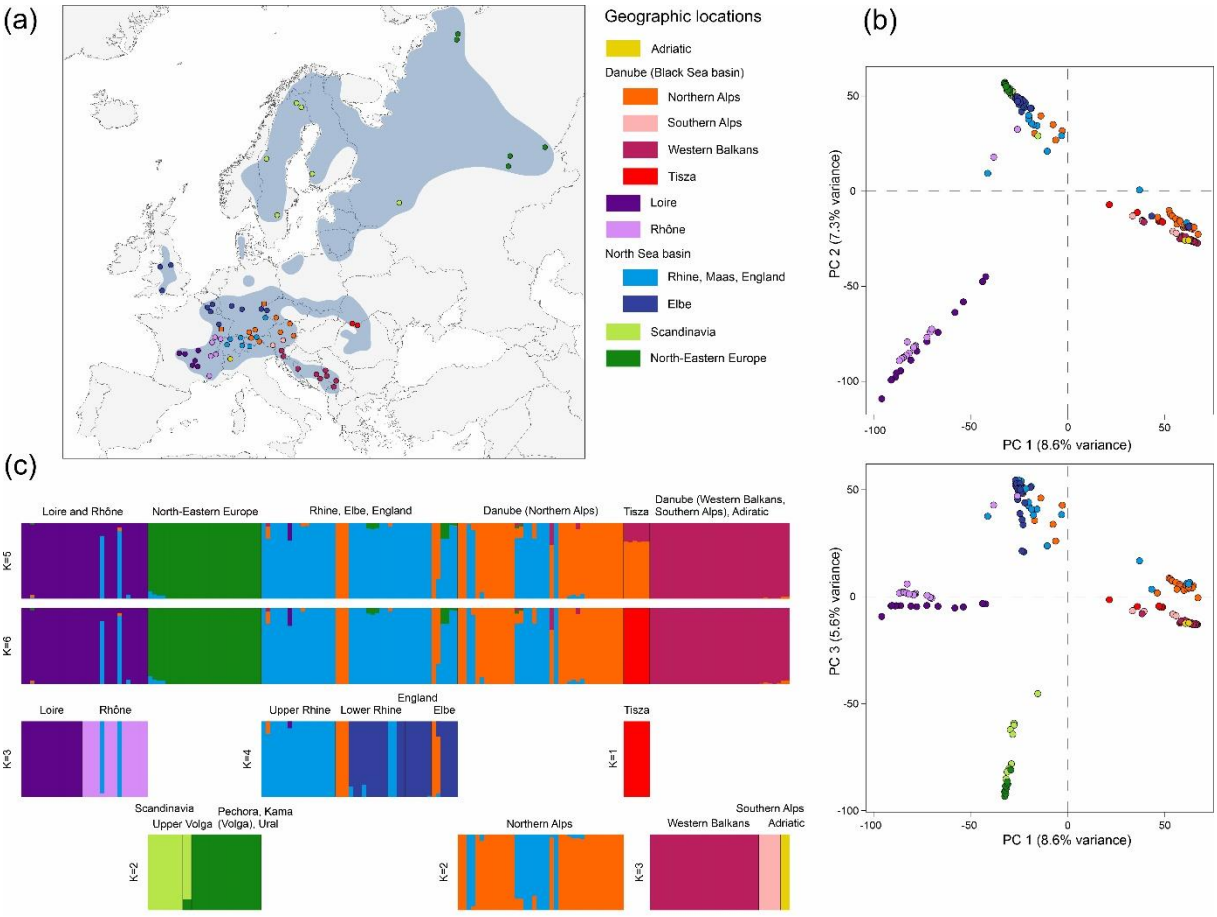



**Figure 1.** Sampling locations and genetic structure (based on nuclear genomic SNPs) of the three species of *Thymallus* in Europe: *T. aeliani* (Adriatic grayling, in yellow), *T. ligericus* (Loire grayling, in purple), and *T. thymallus* (European grayling, all remaining locations). **(a)** Schematic native range of European *Thymallus* spp. (blue background), modified from Freyhof (2011) and Persat et al. (2016), and sampling locations; circles = supposedly native populations, squares = known stocked populations. **(b)** Principal Component Analysis showing PC1 vs. PC2 and PC1 vs. PC3. **(c)** fastSTRUCTURE analysis of the global dataset (all individuals), showing $K = 5$ and $K = 6$ (optimal number of clusters based on marginal likelihood, and median and maximum respectively). The 2nd and 3rd row of structure graphs, are achieved

by hierarchical analyses, and show additional substructure within previously identified drainage specific clusters.

### 3.1.2. Mitochondrial CR

Among the 2,400 sequences analysed, 125 distinct haplotypes were found (Table S3). While the overall distribution of haplotypes reflected the structure of predefined haplogroups and was congruent with the nuclear genomic SNPs, 35 haplotypes (28%) were shared between groups (Figure 2). The Adriatic cluster was the most variable with a mean within-group divergence of 1.9%. Low within-group variability was found in the North-Eastern European (0.3%) and Scandinavian (0.3%) haplogroups.

Despite the obvious divergence (beyond the 95% connection limit) of Loire drainage haplotypes, the Adriatic cluster was most divergent overall, with mean uncorrelated *p*-distances ranging from 1.9% to 2.5% (Table S4). One haplotype (34 in Figure 2) found in the Loire drainage (upper Allier River) was clearly foreign, occurring also in the Rhine, Rhône and Lafnitz rivers. Haplotypes from the Danube River drainage were found in most Adriatic localities, but no Adriatic clade haplotype was found outside of the Northern Adriatic basin. Most samples from the Po River drainage formed a subcluster (haplotypes 3–7) within the Adriatic haplogroup and only haplotype 8 was shared between the Po and Adige rivers and represents the neotype of *T. aeliani* from Lake Maggiore (MT762347).

Danubian haplotypes were strongly structured following major drainage sub-systems (Figure 2). Samples from the Romanian Carpathians (Tisza River) revealed a single, previously unknown haplotype (116), the most divergent within all Danubian haplogroups; mean distances ranged from 1.5% (Northern Alps) to 1.9% (Western Balkans). A high proportion of Southern Alps haplotypes was detected in the Northern Adriatic basin, and Northern Alps haplotypes appeared widely distributed in the Northern Adriatic and Southern Alps regions, as well as in the Rhine and Rhône rivers. Two haplotypes from the Lafnitz River, a tributary of the Raab River drainage in Austria, grouped close to the Western Balkans (85, 112) but haplotypes associated with the Northern Alps (42, 49, 58, 121), and the Rhône River and North Sea basin (28, 34) were also found in the Lafnitz River.

Samples from the North Sea basin were distinct from Scandinavia and North-Eastern Europe, but grouped with haplotypes of the Rhône River. Samples from Scandinavia and those originating from the Arctic Ocean (Pechora River) and Caspian Sea drainages (Volga and Ural rivers) were well-differentiated from other *Thymallus* in Europe. North-Eastern European haplotypes were exclusively private (haplotypes 44–48, 114, 124); Scandinavian haplotypes

were mostly private (22–24, 26, 27, 39), but some were also found in the Rhine (101) and Danube drainages (21, 113). The haplotype (25) from the Oura hatchery was shared with individuals from the Danubian Northern- and Southern Alps, as well as the Rhine River.

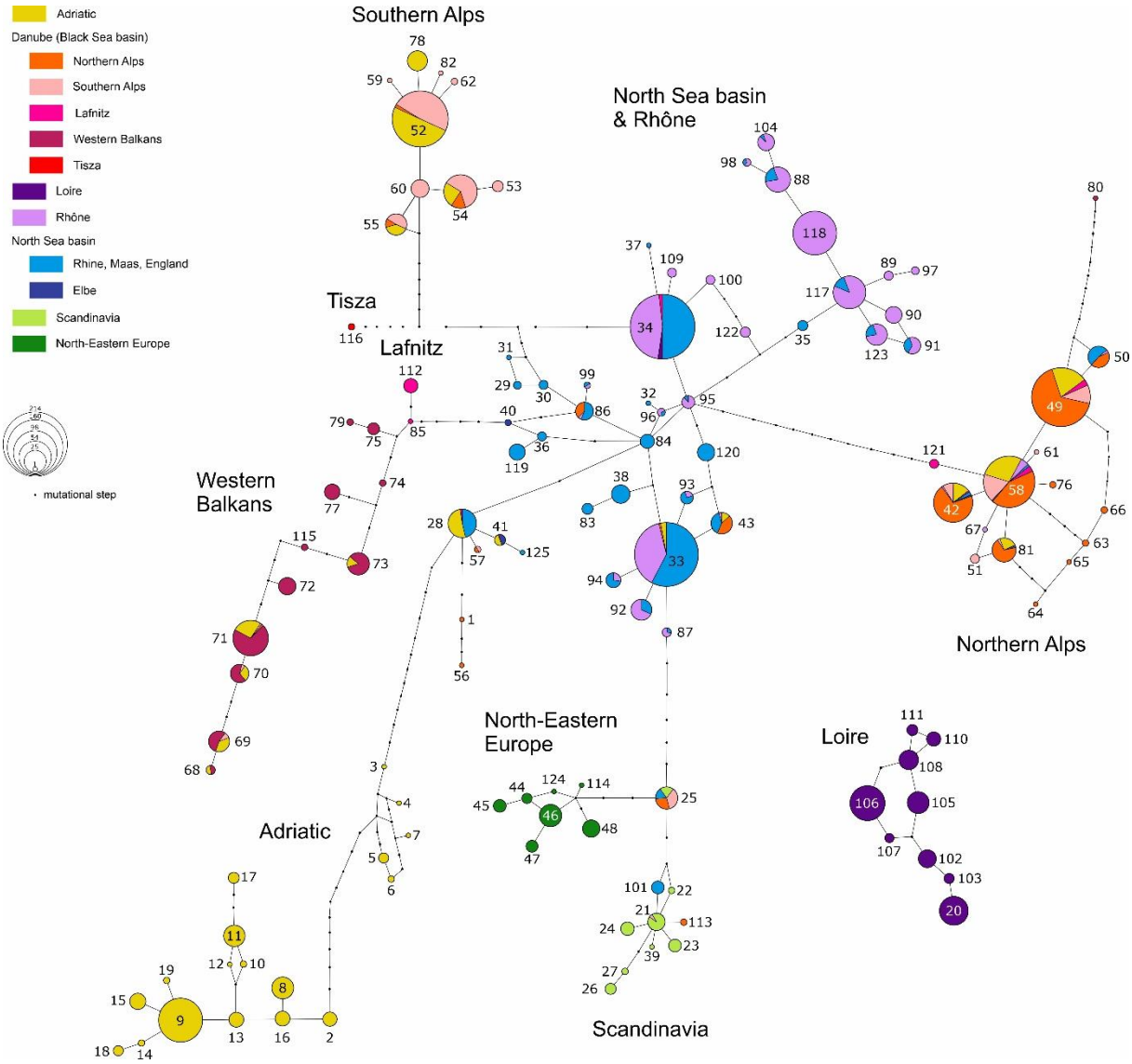



**Figure 2.** A statistical parsimony haplotype network (95% connection limit) of European *Thymallus* spp., based on 2,400 mtDNA CR sequences (959 bp). Haplotype size reflects frequency in the data set, and colours correspond to 12 pre-defined geographic regions.

## 3.2. Phylogenetic relationships

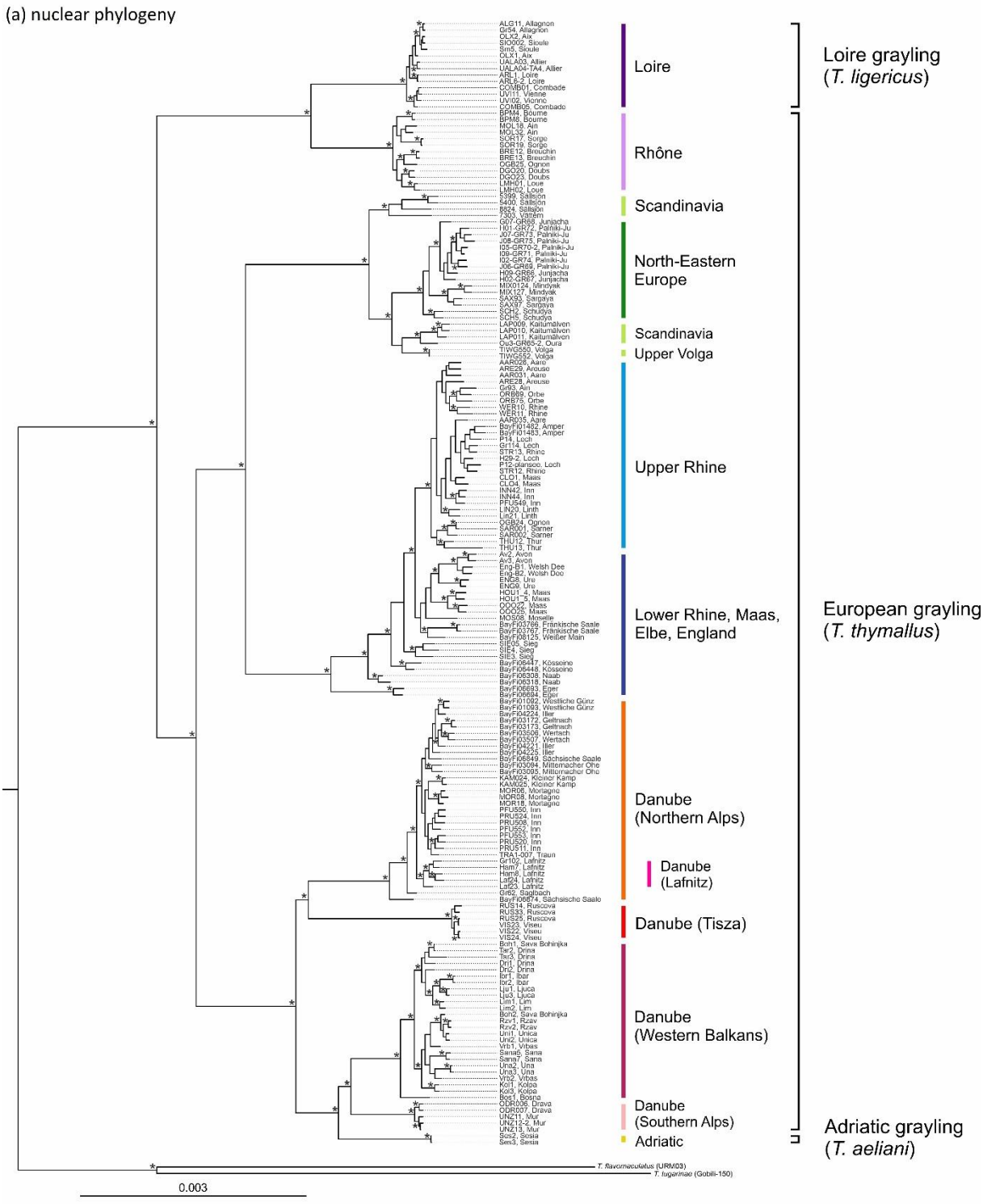
### 3.2.1. Nuclear genomic SNPs

ML analysis of 19,574 concatenated nuclear loci resolved three well-supported clades although not corresponding to the three species of *Thymallus* in Europe (Figure 3a). Additional ML

analyses for different reference-based and de novo assemblies revealed the same topology (data not shown). Similar to PCA, the three major clades were: I) the two French river drainages, Rhône (*T. thymallus*) and Loire (*T. ligericus*), II) the North Sea basin and river systems in Scandinavia and North-Eastern Europe (all *T. thymallus*), and III) the Danube River drainage (*T. thymallus*) including the Adriatic basin (*T. aeliani*). Substructure within these clades largely followed the demarcation of distinct drainage (sub-)systems. Within clade II, samples from the North Sea basin were distinct from North-Eastern Europe and Scandinavia. The Upper- and Lower Rhine were distinct, the latter also including *Thymallus* from the Maas and Elbe rivers, and drainage systems in England. Samples from Central Sweden (5399, 5400, 8824, 7303) formed a monophyletic assemblage distinct from samples from Northern Sweden (LAP009, LAP010, LAP011) and the Upper Volga (TIWG550, TIWG552). Samples from the Kama and Ural rivers (Caspian Sea) group with those from the Pechora River (Barents Sea). Within clade III, four well-supported Danubian clades were inferred. Samples from the Northern Alps formed a monophyletic group with those from the Tisza River and samples from the Southern Alps grouped with those from the Western Balkans. The two samples from the Adriatic basin (*T. aeliani*) were recovered as a sister clade to the Danubian Southern Alps and Western Balkans.

### 3.2.2. Mitogenomes

Both the ML and BI analyses recovered the same, well-supported topology (Figure 3b). The most divergent lineage was found in the Northern Adriatic basin (*T. aeliani*), with a pairwise distance to all other groups ranging from 3.1% to 3.3%. Mitogenomes from the Loire River (*T. ligericus*) appeared as sister clade to three Danubian lineages but were distinct from the neighbouring Rhône River (mean *p*-distance 2.4%). The greatest lineage diversity was found in the Danube drainage (five distinct mtDNA lineages) although these lineages did not form a monophyletic assemblage; samples from the Western Balkans and the Lafnitz River formed a well-supported sister clade to *Thymallus* from Scandinavia and North-Eastern Europe (Figure 3b). Samples from the Danubian Alpine drainages (Northern-/Southern Alps) grouped with those from the Tisza River. Pairwise distances among the Danubian lineages ranged from 0.9% to 2.0%. *Thymallus* from the North Sea basin grouped with samples from the Rhône River, forming a sister clade to samples from Scandinavia and North-Eastern Europe and the Danubian Western Balkans (including the Lafnitz River).


(a) nuclear phylogeny
Loire
Rhône
Scandinavia
North-Eastern Europe
Scandinavia
Upper Volga
Upper Rhine
Lower Rhine, Maas, Elbe, England
Danube (Northern Alps)
Danube (Lafnitz)
Danube (Tisza)
Danube (Western Balkans)
Danube (Southern Alps)
Adriatic
Loire grayling (T. ligericus)
European grayling (T. thymallus)
Adriatic grayling (T. aeliani)
T. flavomaculatus (URM03)
T. lugarinae (Gobili-150)
0.003

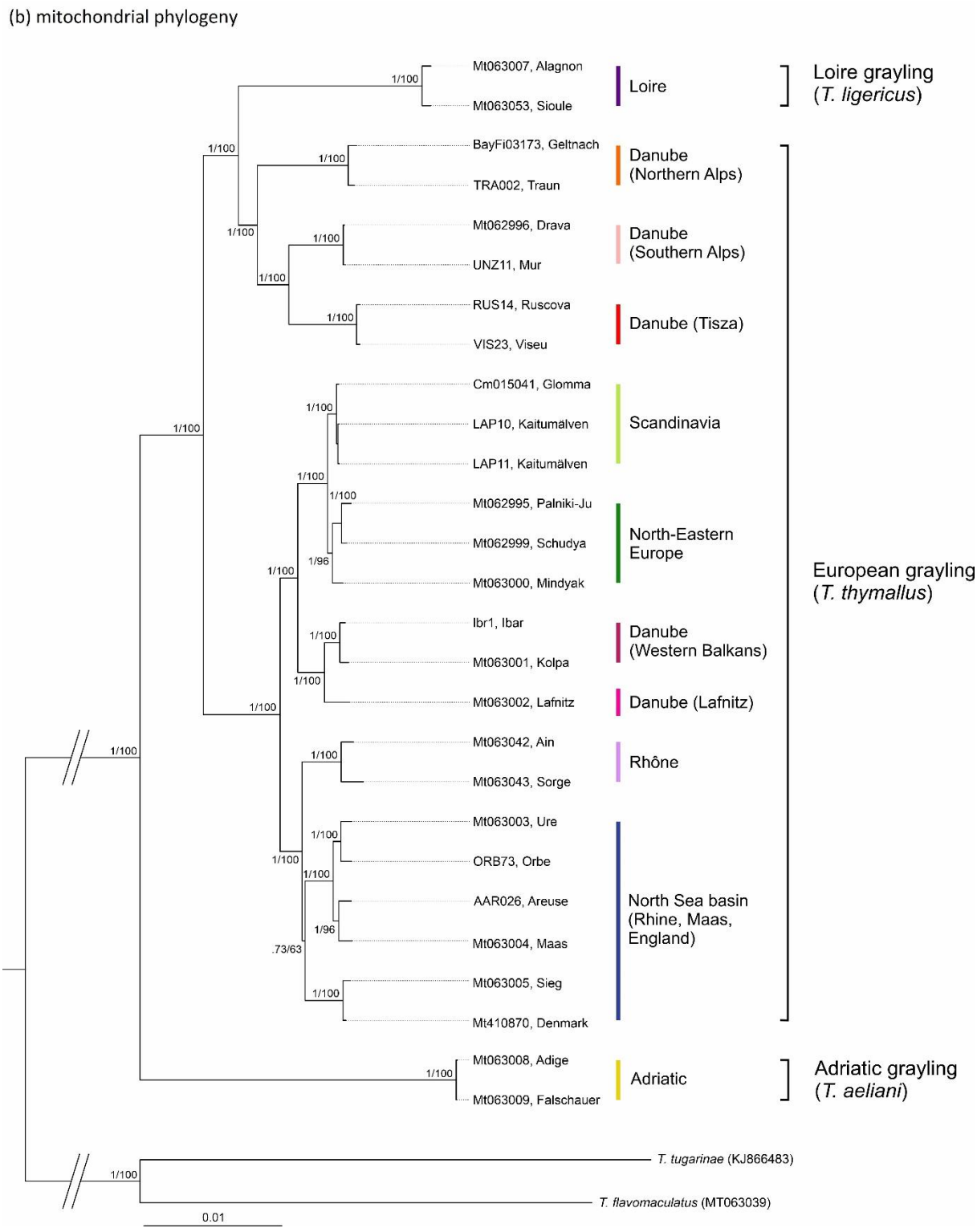



**Figure 3.** Phylogenetic inference in European *Thymallus* spp. **(a)** Maximum likelihood (ML) phylogeny based on a concatenated alignment of 19,574 polymorphic nuclear loci and 178 individuals; asterisks represent bootstrap values (bs) > 95. **(b)** Phylogenetic relationships based on 27 mitogenomes; numbers above branches represent Bayesian posterior probability/bs. Colour coding as in Figure 1; sample information in Table S1.

## 3.3. Scenarios of historical introgression

Both the reference-based and the de novo assemblies revealed a similar pattern of introgression among distinct lineages. For the reference-based assembly, the Dsuite analysis estimated admixture proportions ($f_4$-ratio statistic) above 5% in 53 trios. The visualisation of the *f*-branch metric (based on the nuclear tree topology) highlighted 36 potential introgression events (Figure 4), but admixture proportions above 5% were found in only 10 cases. The strongest signals of introgression were evident between lineages of the Danube and Rhine rivers, but evidence for excess allele sharing was also found between *Thymallus* in the Rhine, Rhône and Loire rivers.

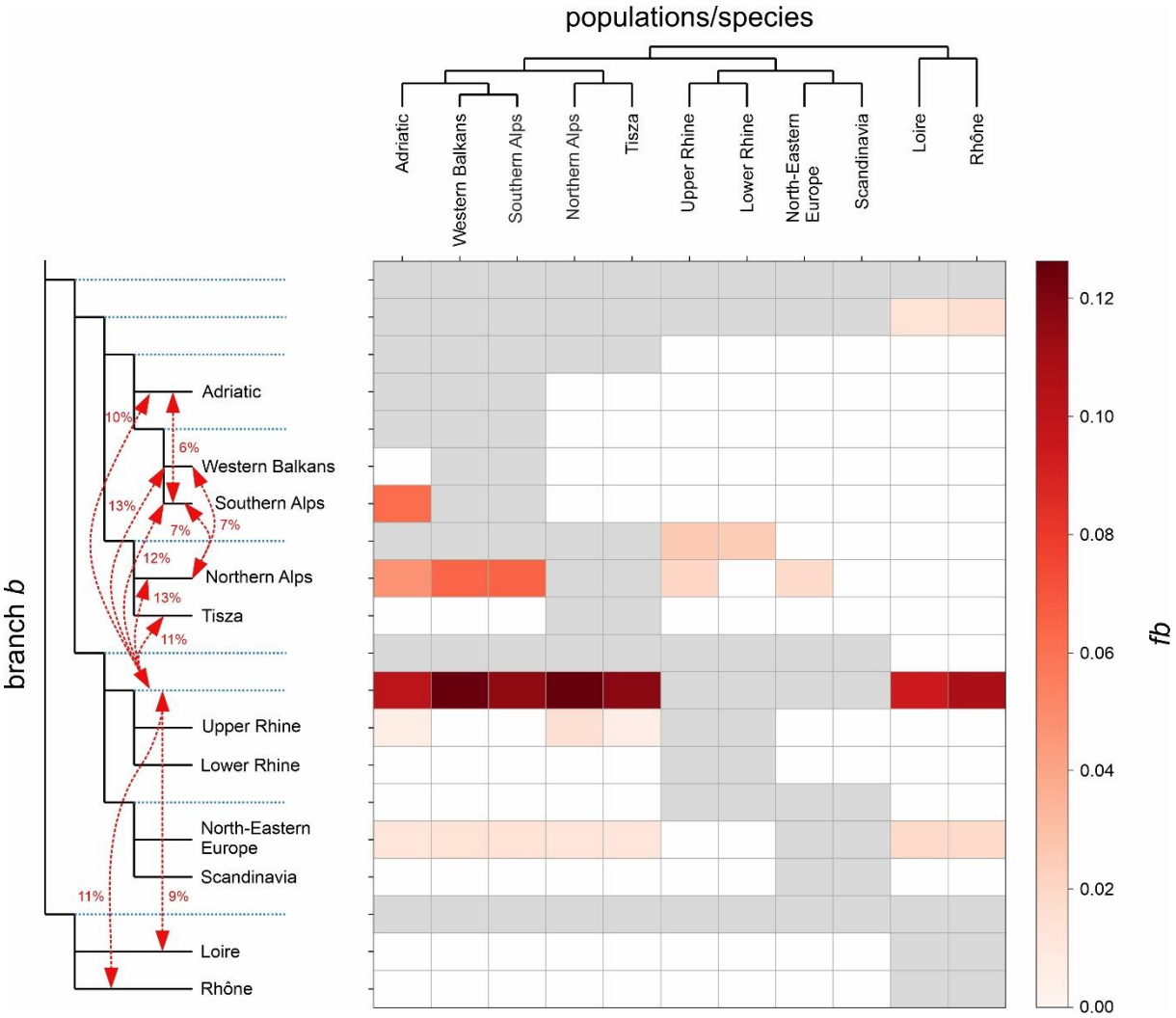



**Figure 4.** Results of the *f*-branch metric based on the nuclear topology. Along the y-axis, each branch (also ancestral branches, dotted lines) points to a row with corresponding proportions of excess allele sharing between the specific branch and the population/species (x-axis). Red arrows highlight admixture proportions > 5%.

### 3.4. Divergence-time and ancestral area estimation

All time-constrained SNAPP phylogenies identified the same three divergent phylogeographic clades as seen with the ML analyses (Figure 5). In contrast to the ML inferences, however, the SNAPP analyses showed that the placement of the Danubian Northern Alps and Tisza clades is uncertain and their sister relationship is not strongly supported (nine datasets, Bayesian posterior probability (BPP) ranging from 0.70 to 0.99). An alternative topology inferred from one out of 10 datasets suggested an early split of the Tisza from all other Danubian lineages (Figure S1) and thus questions the reliability of time estimates for that particular node.

Using two geomorphological age constraints, the oldest split among European *Thymallus* spp. was estimated at 3.5 Ma (95% highest probability density (HPD) interval: 2.9–4.1 Ma), representing the divergence between the Rhône and Loire (clade I) and all others (Figure 5). Based on the best-fit DIVALIKE model (AIC = 68.04), this split likely occurred between the Upper Danube and Rhône rivers (Figure 5: b1) followed by the colonisation of the Loire River (from the Rhône River, Figure 5: c2) at 1.4 Ma (95% HPD interval: 1.2–1.7 Ma). The MRCA of extant Danubian lineages was dated at 2.2 Ma (95% HPD interval: 1.8–2.6 Ma), and the MRCA of the North Sea basin, Scandinavian and North-Eastern European lineages was at 2.6 Ma (95% HPD interval: 2.2–3.1 Ma), both falling within the Pliocene–Pleistocene transition. The colonisation of Central- and Northern Europe probably originated from the Danube River drainage (Figure 5: b2, b3) with subsequent divergence into two geographically distinct lineages. Differentiation between the Scandinavian and North-Eastern European lineage was estimated at 0.5 Ma (95% HPD interval: 0.4–0.6 Ma), predating the last glacial maximum (LGM) at the Pleistocene–Holocene transition.

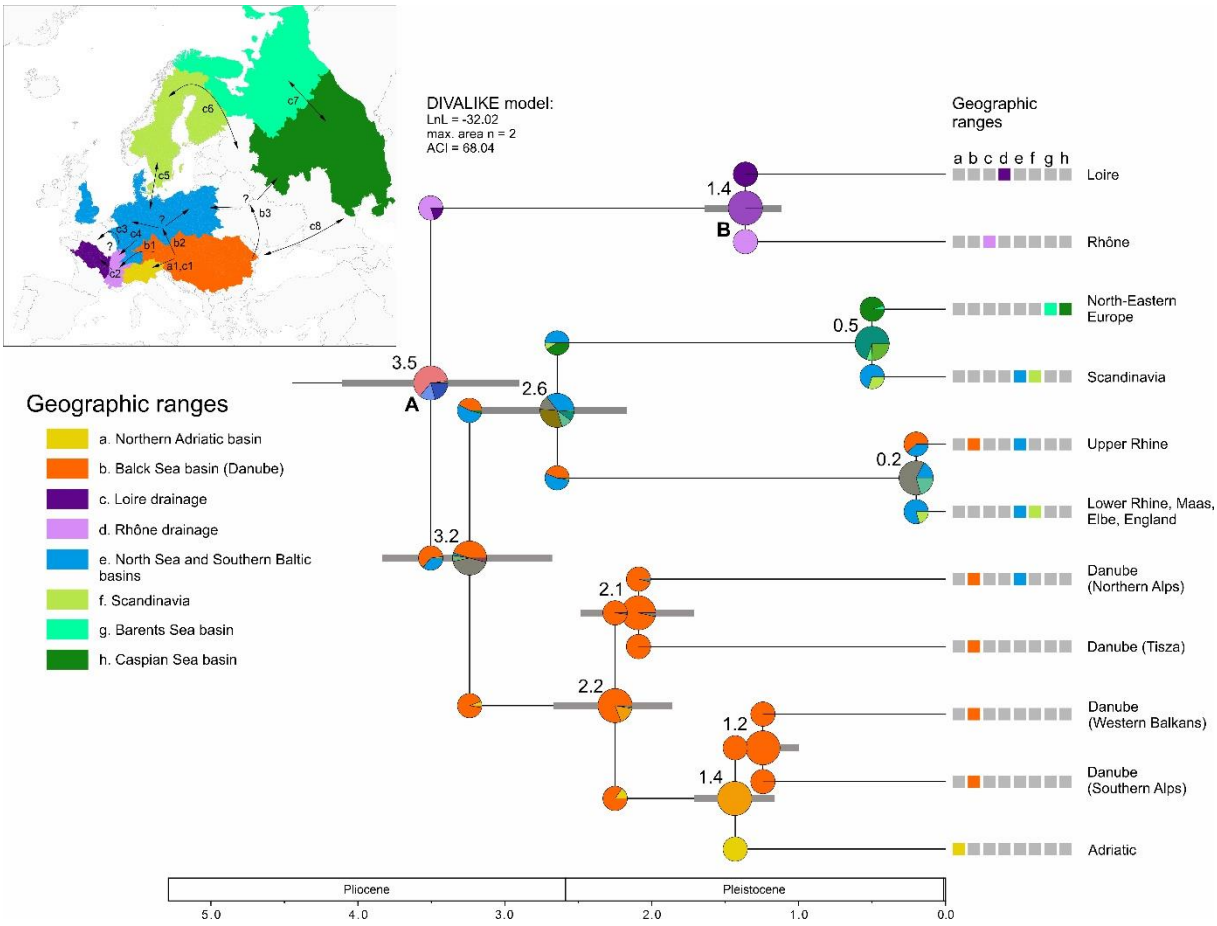



**Figure 5.** Time-calibrated SNAPP phylogeny (maximum-clade-credibility tree) of major lineages/species of *Thymallus* in Europe, and biogeographical reconstruction from BioGeoBEARS under the DIVALIKE model. The SNAPP analysis is based on two biogeographical calibration points (A, B). Divergence times in million years (median values) refer to one of nine topologically identical analyses (see text for additional information); see Figure S1 for an alternative topology, suggesting an early split of the Tisza from all other Danubian lineages. Grey bars indicate the 95% highest posterior density intervals. The inserted map shows biogeographical ranges and possible migration pathways of *Thymallus* in the Miocene (arrow a), the Pliocene (arrows b), and the Pleistocene (arrows c). Biogeographical ranges referring to the presence/absence of *Thymallus* (compare to Figure 1) and the availability of samples.

## 4. Discussion

### 4.1. Lineage diversity and mito-nuclear discordance

A notable finding of this study is the discordance between the robust phylogenetic inference based on nuclear genomic SNPs compared to mitogenomes. This discordance relates directly to the phylogenetic placement of the three European species, as well as the hypothesised colonisation history of *Thymallus* in Europe. While both nuclear genomic SNPs and mtDNA sequences support high lineage diversity overall, lineage-specific or species-specific diversity

differs markedly between the two sets of markers. Such phylogenetic discordance can be caused by multiple mechanisms (e.g., Antunes and Ramos, 2005; Koblmüller et al., 2010; Bonnet et al., 2017; Guo et al., 2019), whereby historical introgression was found to be particularly common among freshwater fishes (Wallis et al., 2017). Some of these mechanisms imply secondary contact of previously isolated lineages, underscoring historical complexity of river drainage evolution, and/or other dispersal mechanisms such as human-mediated translocations.

The sister clade structure of *T. thymallus* from the Rhône drainage and *T. ligericus* from the Loire drainage was not anticipated due to the distant relationship of their respective mitochondrial lineages, as well as the physical isolation of these two drainage systems. Interestingly, while the mitochondrial lineages, including many specific mtDNA CR haplotypes, are shared between the Rhône and Rhine drainages, the *T. ligericus* mtDNA lineage is private to the Loire River. This appears to reflect the complete replacement of the ancestral mtDNA in *Thymallus* from the Rhône drainage with mtDNA from the Rhine lineage.

While the actual genetic relationship based on nuclear genomic SNPs between *Thymallus* in the Rhine and Rhône rivers is rather distinct (except for two samples from the Ain and Ognon rivers, Figure 1c), a relatively strong signal of historical introgression is supported from an ancestral Rhine lineage into the Rhône (as well as the Loire). Interestingly, the Rhine and Rhône rivers share 17 out of 40 mtDNA CR haplotypes. Although recent stocking activities in this region (Persat, 1996; Weiss et al., 2002; Cattanéo et al., 2011), in theory, could explain some haplotype sharing, there are several reasons to reject an anthropogenic cause for this pattern. The shared distribution of many tip haplotypes as well as two more ancestral haplotypes (33, 34) are more easily explained by ancestral sharing. Shared tip- or derived haplotypes can be explained by the fact that the temporal resolution of mtDNA CR haplotypes in salmonids is roughly 100,000 years (based on a 2% per Ma divergence rate), and thus likely does not reflect identical haplotypes at the mitogenome-level. This is additionally congruent with the absence of a general pattern of hybridization observed between these drainages. Lastly, the management of *Thymallus* in the Rhône drainage has been extremely limited compared to other salmonid species, such as *S. trutta* (Cucherousset et al., 2020), or the management of *Thymallus* in the Austrian Alps (Weiss et al., 2013). Thus, the signal of historical introgression between the Rhône and Rhine clades supports a natural scenario of secondary contact followed by mitochondrial capture and low levels of nuclear introgression.

*Thymallus* in Central- and Northern Europe are characterised by wider geographical ranges with less drainage-specific substructure, and a deep divergence between two phylogeographic lineages (North Sea basin vs. Scandinavia/North-Eastern Europe). This

pattern is largely congruent with mtDNA data (Weiss et al., 2021) and implies distinct Pleistocene glacial refugia and postglacial re-colonisation pathways (Gum et al., 2009). In Southern- and Central Sweden both lineages presumably form a zone of contact, an area that is generally known as a complex hybrid zone for many European taxa (Hewitt, 2004). Repeated contact with phylogeographic clades of adjacent drainage systems is suggested by signals of historical introgression. This is not only supported for the Rhine, Rhône, and Loire rivers, but also for the Volga and Danube rivers, where mtDNA haplotypes from the Danubian Western Balkans and the Lafnitz River form a sister clade to the Scandinavian/North-Eastern European lineage (Marić et al., 2014; Weiss et al., 2021). This also represents a major discordance within Danubian *Thymallus*, which shows strong genetic structure by major sub-drainages with both nuclear genomic SNPs and mtDNA data. However, with the nuclear genomic markers, the Western Balkan haplotypes form a sister clade to the Southern Alps, and the Lafnitz River haplotypes fall within the Northern Alps. Speculating on the possible evolutionary mechanisms for mito-nuclear discordances within the Danube River drainage presents additional difficulties because 1) the lineages in question are more closely related to each other, and 2) the paleohydrological development of distinct sub-drainages is complex and far from clear (see below). Additionally, frequent sharing of mtDNA CR haplotypes, especially in the Alpine region, can be explained by decades of human-mediated stocking (Weiss et al., 2013), where source populations originated from within the Danube River drainage, but also from the North Sea basin and Scandinavia.

The problem of human-mediated stocking also creates significant challenges for understanding the evolutionary history of *T. aeliani* (Adriatic clade). Secci-Petretto et al. (2023) have shown that *T. aeliani* samples from the Adige and Piave rivers were nested within Danubian samples and concluded that anthropogenic introgression was the most likely mechanism responsible for the observed mito-nuclear discordance, as *Thymallus* in the eastern part of the Northern Adriatic basin are known to be heavily managed and introgression with allochthonous lineages widespread (Meraner et al., 2014). In this study, samples were specifically selected from the putatively non-managed Sesia River (Po River drainage), nevertheless only two individuals were obtained, causing potential sample bias in downstream analyses. The placement of these samples in the SNP-based phylogeny is still discordant with the mitogenome analysis, where *T. aeliani* haplotypes represent the first split among all European lineages; but they do form a clade distinct from the Danubian Western Balkans and Southern Alps clades, a result not obtained by Secci-Petretto et al. (2023). While it may be tempting to hypothesize that human-mediated stocking is responsible, there is no obvious sign

of recent nuclear introgression. Elevated *D* and $f_4$-ratio results rather indicate continuous/high level historical introgression. Interestingly, even in rivers where most individuals appear to be of admixed origin, divergent Adriatic mtDNA haplotypes are predominant and not outcompeted by foreign mtDNA strains (Sušnik et al., 2004; Meraner and Gandolfi, 2012). Such patterns may result from genotype-by-environment interactions as recently shown for *S. trutta* (Folio et al., 2021), supporting a scenario with selection for specific (local) mitochondrial variants, despite nuclear introgression (Toews and Brelsford, 2012, and references therein). While this picture of mito-nuclear discordance in *T. aeliani* seems to involve both natural introgression and more recent stocking activities, as well as some type of natural selection, the limited number of samples results in a degree of uncertainty concerning these inferences but cast considerable doubt on the notion that any non-introgressed populations of *T. aeliani* can be found.

### 4.2. Divergence-time estimation and biogeography

The phylogenetic inferences based on nuclear genomic data have important implications for the biogeography of *Thymallus* in Europe and allow novel insights into a dynamic colonisation history. In the following section, the findings are discussed from a biogeographic perspective with emphasis on possible palaeohydrological scenarios.

*(a) European colonisation:* Fossil evidence (Rückert-Ülkümen and Kaya, 1993) and divergence times estimated from mitogenomes (Weiss et al., 2021; but see Supporting Information 2 and Figure S2) suggest an arrival of *Thymallus* in Europe in the late Miocene. The split of Asian and European *Thymallus* spp. (7.8–10.8 Ma, Figure S2) largely overlaps with the split of northern *Pungitius pungitius*/*P. laevis* and southern *P. platygaster*/*P. hellenicus* (Guo et al., 2019) and coincides with the oldest fossil record of *Hucho* in Europe (Kovalchuk, 2015). The colonisation of *Thymallus* in Europe may stem from a connection of the paleo-Ob River and the Ponto-Caspian basin (Knizhin 2009) and likely followed river systems along the northern side of the Paratethys region (see maps in Popov et al., 2004; and Krijgsman et al., 2010), which may have allowed passage for freshwater fishes from Central Asia and Siberia to the paleo Black-Sea basin over long periods (Artamonova et al., 2021).

*(b) The Northern Adriatic basin:* Based on mitochondrial data, the earliest split among European *Thymallus* spp. coincides with the hypothesised colonisation of the Northern Adriatic basin (Figure 5**:** a1). This makes a colonisation of the Danube system in the late Miocene or

early Pliocene plausible. Especially the Lago Mare stage (5.5–5.3 Ma) at the end of the MSC is considered an important biogeographical event for freshwater diversity in the Mediterranean region (Dubut et al., 2012). Although connections between the Danube drainage and the Northern Adriatic region are supported by phylogeographic patterns in other taxa (e.g., *C. gobio* (Šlechtova et al., 2004); *Squalius cephalus* (Seifertová et al., 2012)), as well as *Thymallus* in the late Pleistocene (Meraner et al., 2014), the pattern of high levels of historical introgression and replacement of the original genetic structure of *Thymallus* in the Adriatic basin by peri-alpine lineages suggests previously undetected and continuous colonisation/contact taking place during Quaternary glaciation cycles (Figure 5: c1).

*(c) Rhône and Loire rivers:* The presence of mitochondrial haplotypes in the Loire River that could not have come from either the Rhône or the Rhine rivers, and the mito-nuclear discordance of the Rhône clade require a differentiated view on the colonisation history of *Thymallus* in this region. The paleohydrology in the headwaters of the Danube and Rhône (Aare and Doubs) rivers, and the later established Rhine River provide evidence for a dynamic drainage evolution with repeated phases of isolation and connectivity (Giamboni et al., 2004; Berger et al., 2005). The last direct connection between the Danube and the Rhône rivers, via the Aare River, is documented by the geological evidence of the "Sundgau gravel" (4.2–2.9 Ma, Berger et al., 2005) and possibly enabled an initial colonisation of *Thymallus* and other freshwater fishes (Olivier et al., 2009) into the Rhône River (Figure 5: b1). Faunal exchange in later (Pleistocene) periods could have only occurred indirectly via the Rhine system.

Paleohydrological corridors between the Rhône and Loire rivers in the Pleistocene are suggested from the geologically poorly documented “Formations fluvio-lacustres de la Guye” (BRGM, 1999). These sediments are widespread north/north-east of the Massif Central and indicate dynamic fluvial systems with shallow watersheds in the vicinity of backwater lakes in the Saône River valley, which were regularly formed by the advances of the Rhône glacier into the foreland during the Pleistocene cold phases (Persat et al., 2020). Such geomorphological conditions may have allowed repeated faunal exchange (Figure 5: c2). Alternatively, a passage via the Rhine/Maas and Seine rivers into the Lower Loire River (Weiss et al., 2002) (Figure 5: c3) may be hypothesised as faunal similarities between the Lower Rhine and Loire rivers are evident for several freshwater fishes (Bănărescu, 1992). This, however, is contradicted by the absence of *Thymallus* in the Seine River (Weiss et al., 2002), and although the Loire and Seine rivers were likely connected at the Pliocene/Pleistocene transition (Persat et al., 2020) connectivity in later stages is unknown from the sedimentological record.

The Rhône/Loire split was presumably followed by a secondary colonisation of *Thymallus* from the Rhine River into the Rhône River (Figure 5: c4) dated at 0.7–1.2 Ma from mitogenomes (Figure S2). Later stages of connectivity between the Rhine and Rhône rivers are plausible and supported by shared ichthyofaunal elements (Olivier et al., 2009).

*(d) North Sea basin and Scandinavia/North-Eastern Europe:* Based on the assumption of an early (Miocene) Danubian colonisation, range expansion of *Thymallus* towards the north/north-east seems possible in both the upper and lower reaches of the Danube River. In the upper reaches (Figure 5: b2), dispersal to the North Sea and Baltic Sea basins is plausible via river capture events (e.g., Domokos et al., 2000). Ichthyofaunal similarities between these basins are documented for several species (e.g., Bănărescu, 1992; Englbrecht et al., 2000; Lercetau-Köhler et al., 2013), but are largely associated with the late Pleistocene glacial/interglacial stages. An alternative route along the Lower Danube and the shallow Dnieper Depression via the "Baltic Gate" (Figure 5: b3) may have allowed repeated dispersal opportunities across lakes, marshes and cross-basin river systems in the paleo-Pripyet region over longer geological periods (Meulenkamp and Sissingh, 2003; Popov et al., 2004). This area denotes the lowest point of the main European watershed and would have provided hydrological connections to both the west (paleo-Weichsel system) and the east/north-east (paleo-Volga system), possibly explaining the early split of *Thymallus* in the North Sea basin and Scandinavia/North-Eastern Europe. The clear genetic differentiation between these lineages implies distinct refugia during the Pleistocene and independent re-colonisation of previously glaciated areas (Figure 5: c5, c6, c7). Southern- and Central Sweden was likely re-colonised from refugial areas in both the North Sea basin and North-Eastern Europe, following a common biogeographic pattern for many European biota (Hewitt, 1999).

*(e) The Danube drainage and its subsystems:* The Danube is one of the biogeographically most complex river systems in Europe (Bănărescu, 1992). This may be associated with its geomorphological setting – “within-drainage differentiation”, and the complex colonisation history – “between-drainage similarity”. Divergence times estimated from both mtDNA and ddRADseq suggest that major splits within Danubian *Thymallus* predate the environmental perturbations of the late Pleistocene, implying mechanisms other than glacial/interglacial periods driving lineage differentiation. The sequence of multiple basins (e.g., Vienna, Pannonian, and Dacian basins), separated by reaches of steeper gradients (e.g., Hainburg Gate, Dunakanyar, Iron Gates), promotes differentiation of large drainage subsystems within the

Danube (Domokos et al., 2000; fig. 1 in Krézsek and Olariu, 2021), possibly shaping the observed phylogeographic patterns.

The Black Sea region and the Danube River are generally considered to be important migration corridors and Pleistocene refugial areas for many biota (Hewitt, 2004). This would explain faunal similarities between the Danube River and other European drainage systems with regard to widely distributed species (e.g., Englbrecht et al., 2000; Takács et al., 2022). Earlier studies (Marić et al., 2014; Weiss et al., 2021) have already shown the similarity of *Thymallus* in the Western Balkans and the Lafnitz River with the Scandinavian/North-Eastern European clade based on mtDNA. Mitochondrial sequences date the split between these lineages in the Pleistocene (mitogenomes: 0.9–1.4 Ma, Figure S2; mtDNA CR: 1.5–0.6 Ma, Marić et al., 2014), which coincides with the onset of the Pleistocene cold periods and suggests migration corridors between the Caspian and Black Sea regions (Figure 5: c8).

## 5. Conclusion

This study generates the most comprehensive analysis of the phylogeographic structure and biogeographic origins of European *Thymallus* spp. to date, reflecting very high genetic diversity as well as novel insights into colonisation patterns following known paleohydrological events and cross-drainage colonisation corridors. Mito-nuclear discordances underscore the necessity of applying genome-wide markers for elucidating the evolutionary history of widespread taxa. However, it is also demonstrated how information on the distribution of mitochondrial lineages can complement, rather than contradict inferences based on nuclear genomic markers. Generally, the phylogenetic structure of mitochondrial lineages reflects more ancient patterns of allopatry that cannot necessarily be uncovered by contemporary genetic structure with nuclear genomic SNPs. The potential to draw complimentary inferences from mtDNA and nuclear genomic data is likely universal. However, it may be particularly salient in steno-thermic aquatic organisms in regions marked by glaciations, where the cyclic nature of allopatry and secondary contact generate complex phylogeographic patterns that are reflected differently across nuclear vs. mitochondrial genomes.

Lastly, the results underscore the distinctiveness of *T. ligericus* in the Loire River drainage, and the questionable genetic integrity of the critically endangered *T. aeliani*. Yet, for both species, their SNP-supported phylogenetic position also draws attention to the paraphyly of *T. thymallus*.

**Acknowledgements**

This work was funded by a grant from the Doctoral Academy Graz, Ecology and Evolution in Changing Environments (EECE), at the University of Graz, to GKE. We thank Stefania Trasforini and Cesare Puzzi (Graia, Italy), Ulrich Schliewen (ZSM, Germany), Martin Schletterer (TIWAG, Austria), Omar Bašić (Montenegro), Rifat Tahirić and Vlado Kerkez (Bosnia and Herzegovina), and Ramo Berović and Milan Lugić (Serbia) for assistance with samples. T. Schenekar (University of Graz) provided valuable comments and support to the study. Fieldwork in Serbia was supported by the Ministry of Natural Resources, Mining and Spatial Planning (permit number 324-04-154/2013-02). Slovenian samples were collected based on permit 3420-30/2017/5 issued by the Ministry of Agriculture, Forestry and Food. Samples from Switzerland were collected within the project "Genetics of grayling populations in Switzerland" (permit number 00.0214.PZ/I284-1246), a cooperation between the Federal Office for the Environment (OFEV, represented by D. Hefti) and Haute école du paysage, d'ingénierie et d'architecture de Genève (Hepia, represented by F. Cattanéo).

**Authors' contributions**

Gernot K. Englmaier: Conceptualization, Methodology, Software, Formal analysis, Investigation, Visualization, Writing- Original draft preparation. Nuria Viñuela Rodríguez: Methodology, Software, Formal analysis, Visualization. Jernej Bravničar: Methodology, Investigation, Resources, Writing - Review & Editing. Lukas Zangl: Methodology, Software, Investigation, Writing - Review & Editing. Henri Persat: Resources, Writing - Review & Editing. Saša Marić: Resources, Writing - Review & Editing. Clemens Ratschan: Resources, Writing - Review & Editing. Bo Delling: Resources, Writing - Review & Editing. Duarte V. Gonçalves: Formal analysis. Giulia Secci-Petretto: Formal analysis. Elsa Froufe: Writing - Review & Editing. Steven J. Weiss: Conceptualization, Methodology, Resources, Writing- Reviewing and Editing, Supervision.

**Data availability**

Data will become available after (open access) publication of the manuscript. In the meantime data are available upon request from the corresponding author.

## Supporting Information 1

**Table S1.** Information of the *Thymallus* samples used for ddRADseq analyses. Included is the original sample code, accession number, and sample site location (including drainage, basin, country and geo-coordinates).

| Species | Sample code | Accession number | Sampling date | Location | Drainage Progression | Basin | Country | Coordinates |
|---|---|---|---|---|---|---|---|---|
| *T. flavomaculatus* | URM03 | SRR19545551 | N/A | Urmi River (Khabarovsk Krai) | Urmi/Tunguska/Amur | Sea of Okhotsk (North Pacific Ocean) | Russia | 48.713056, 134.266944 |
| *T. tugarinae* | Gobili-150 | SRR19545519 | N/A | Gobili River | Gobili/Anyui/Amur | Sea of Okhotsk (North Pacific Ocean) | Russia | 49.294994, 138.520103 |
| *T. aeliani* | Ses2 | submitted | 2018 | Sesia River | Sesia/Po | Adriatic Sea | Italy | 45.772898, 8.097358 |
| *T. aeliani* | Ses3 | submitted | 2018 | Sesia River | Sesia/Po | Adriatic Sea | Italy | 45.772898, 8.097358 |
| *T. ligericus* | ARL1 | submitted | 2011 | Loire River at Arlempdes | Loire | Atlantic Ocean | France | 44.889444, 3.927778 |
| *T. ligericus* | ARL6_2 | submitted | 2011 | Loire River at Arlempdes | Loire | Atlantic Ocean | France | 44.889444, 3.927778 |
| *T. ligericus* | OLX1 | submitted | 2004 | Aix River at Saint-Germain-Laval | Aix/Loire | Atlantic Ocean | France | 45.830556, 4.011667 |
| *T. ligericus* | OLX2 | submitted | 2004 | Aix River at Saint-Germain-Laval | Aix/Loire | Atlantic Ocean | France | 45.830556, 4.011667 |
| *T. ligericus* | ALG11 | submitted | N/A | Allagnon River at Chapelle d'Alagnon | Allagnon/Allier/Loire | Atlantic Ocean | France | 45.106389, 2.896667 |
| *T. ligericus* | Alg06(GR54) | SRR19545539 | N/A | Allagnon River at Chapelle d'Alagnon | Allagnon/Allier/Loire | Atlantic Ocean | France | 45.106389, 2.896667 |
| *T. ligericus* | SIO002 | submitted | N/A | Sioule River near Menat | Sioule/Allier/Loire | Atlantic Ocean | France | 46.025833, 2.904167 |
| *T. ligericus* | SM5 | SRR19545538 | N/A | Sioule River near Menat | Sioule/Allier/Loire | Atlantic Ocean | France | 46.025833, 2.904167 |
| *T. ligericus* | UALA04_TA4 | submitted | 2007 | Allier River | Allier/Loire | Atlantic Ocean | France | 44.792203, 3.788725 |
| *T. ligericus* | UALA03 | submitted | 2007 | Allier River | Allier/Loire | Atlantic Ocean | France | 44.792203, 3.788725 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *T. ligericus* | COMB01 | submitted | 2011 | Combade River at Chateauneuf | Combade/Vienne/Loire | Atlantic Ocean | France | 45.712778, 1.628056 |
| *T. ligericus* | COMB05 | submitted | 2011 | Combade River at Chateauneuf | Combade/Vienne/Loire | Atlantic Ocean | France | 45.712778, 1.628056 |
| *T. ligericus* | UVI11 | submitted | 2011 | Vienne River at Bazenant | Vienne/Loire | Atlantic Ocean | France | 45.764167, 1.619444 |
| *T. ligericus* | UVI02 | submitted | 2011 | Vienne River at Bazenant | Vienne/Loire | Atlantic Ocean | France | 45.764167, 1.619444 |
| *T. thymallus* | MIX127 | submitted | N/A | Mindyak River | Mindyak/Ural | Caspian Sea | Russia | 53.984444, 58.799444 |
| *T. thymallus* | MIX0124 | SRR19545522 | N/A | Mindyak river | Mindyak/Ural | Caspian Sea | Russia | 53.984444, 58.799444 |
| *T. thymallus* | SAX97 | submitted | N/A | Sargaya River | Sargaya/Ufa/Kama/Volga | Caspian Sea | Russia | 56.260886, 57.712611 |
| *T. thymallus* | SAX93 | submitted | N/A | Sargaya River | Sargaya/Ufa/Kama/Volga | Caspian Sea | Russia | 56.260886, 57.712611 |
| *T. thymallus* | SCH2 | submitted | N/A | Schudya River | Kama/Volga | Caspian Sea | Russia | N/A |
| *T. thymallus* | SCH5 | submitted | N/A | Schudya River | Kama/Volga | Caspian Sea | Russia | N/A |
| *T. thymallus* | TIWG550 | submitted | N/A | Volga River at Staritsa | Volga | Caspian Sea | Russia | 56.51225, 34.925889 |
| *T. thymallus* | TIWG552 | submitted | N/A | Volga River at Staritsa | Volga | Caspian Sea | Russia | 56.51225, 34.925889 |
| *T. thymallus* | J06(GR69) | submitted | 2015 | Palniki-Ju River | Palniki-Ju/Junjacha/Usa/Pechora | Arctic Ocean | Russia | 66.523955, 62.782175 |
| *T. thymallus* | I09(GR71) | submitted | 2015 | Palniki-Ju River | Palniki-Ju/Junjacha/Usa/Pechora | Arctic Ocean | Russia | 66.523955, 62.782175 |
| *T. thymallus* | H01(GR72) | submitted | 2015 | Palniki-Ju River | Palniki-Ju/Junjacha/Usa/Pechora | Arctic Ocean | Russia | 66.523955, 62.782175 |
| *T. thymallus* | J07(GR73) | submitted | 2015 | Palniki-Ju River | Palniki-Ju/Junjacha/Usa/Pechora | Arctic Ocean | Russia | 66.523955, 62.782175 |
| *T. thymallus* | I02(GR74) | submitted | 2015 | Palniki-Ju River | Palniki-Ju/Junjacha/Usa/Pechora | Arctic Ocean | Russia | 66.523955, 62.782175 |
| *T. thymallus* | J08(GR75) | submitted | 2015 | Palniki-Ju River | Palniki-Ju/Junjacha/Usa/Pechora | Arctic Ocean | Russia | 66.523955, 62.782175 |
| *T. thymallus* | I05(GR70)_2 | submitted | 2015 | Palniki-Ju River | Palniki-Ju/Junjacha/Usa/Pechora | Arctic Ocean | Russia | 66.523955, 62.782175 |
| *T. thymallus* | H09(GR66) | submitted | 2015 | Junjacha River | Junjacha/Usa/Pechora | Arctic Ocean | Russia | 66.620283, 62.943028 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *T. thymallus* | H02(GR67) | submitted | 2015 | Junjacha River | Junjacha/Usa/Pechora | Arctic Ocean | Russia | 66.620283, 62.943028 |
| *T. thymallus* | G07(GR68) | submitted | 2015 | Junjacha River | Junjacha/Usa/Pechora | Arctic Ocean | Russia | 66.620283, 62.943028 |
| *T. thymallus* | Ou3(GR65)_2 | submitted | 2001 | Oura (Aura) River | Oura (Aura) | Baltic Sea | Finland | 61.833333, 21.333333 |
| *T. thymallus* | LAP009 | submitted | 2020 | Lake Padje Kaitumjaure | Kaitumälven/Kalixälven | Gulf of Bothnia | Sweden | 67.741117, 18.290243 |
| *T. thymallus* | LAP010 | SRR19545523 | 2020 | Lake Padje Kaitumjaure | Kaitumälven/Kalixälven | Gulf of Bothnia | Sweden | 67.741117, 18.290243 |
| *T. thymallus* | LAP011 | submitted | 2020 | Lake Padje/Kaska Kaitumjaure | Kaitumälven/Kalixälven | Gulf of Bothnia | Sweden | 67.651413, 18.721269 |
| *T. thymallus* | 8824 (NRM 56971) | submitted | 2007 | Lake Sällsjön | Storbodströmmen/Indalsälven | Gulf of Bothnia | Sweden | 63.210054, 13.6983803 |
| *T. thymallus* | 5400 (NRM 56971) | submitted | 2007 | Lake Sällsjön | Storbodströmmen/Indalsälven | Gulf of Bothnia | Sweden | 63.210054, 13.6983803 |
| *T. thymallus* | 5399 (NRM 56971) | submitted | 2007 | Lake Sällsjön | Storbodströmmen/Indalsälven | Gulf of Bothnia | Sweden | 63.210054, 13.6983803 |
| *T. thymallus* | 7303 (NRM 43982) | submitted | 1999 | Lake Vättern, Ålebäcken at Sverkerskapellet, Östergötlands län | Motala ström | Baltic Sea | Sweden | 58.2899598, 14.6402213 |
| *T. thymallus* | Eng_B1 | submitted | N/A | Welsh Dee | Dee | Irish Sea | England | N/A |
| *T. thymallus* | Eng_B2 | submitted | N/A | Welsh Dee | Dee | Irish Sea | England | N/A |
| *T. thymallus* | Av2 | submitted | N/A | River Avon at Hampshire | Avon | Atlantic Ocean | England | N/A |
| *T. thymallus* | Av3 | submitted | N/A | River Avon at Hampshire | Avon | Atlantic Ocean | England | N/A |
| *T. thymallus* | ENG8 | SRR19545524 | 2000 | Ure River | Ure/Ouse/Humber | North Sea | England | 54.305556, -1.939722 |
| *T. thymallus* | ENG9 | submitted | 2000 | Ure River | Ure/Ouse/Humber | North Sea | England | 54.305556, -1.939722 |
| *T. thymallus* | ZSM-PIS-035907 \| BayFi06447 | submitted | 2007 | Kösseine stream in Marktredwitz | Kösseine/Röslau/Eger/Elbe | North Sea | Germany | 50.002967, 12.0963586 |
| *T. thymallus* | ZSM-PIS-035907 \| BayFi06448 | submitted | 2007 | Kösseine stream in Marktredwitz | Kösseine/Röslau/Eger/Elbe | North Sea | Germany | 50.002967, 12.0963586 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *T. thymallus* | ZSM-PIS-036203 \| BayFi06693 | submitted | 2007 | Eger stream in Marktleuthen, NE of Weissenstadt | Eger/Elbe | North Sea | Germany | 50.091956, 11.9099323 |
| *T. thymallus* | ZSM-PIS-036203 \| BayFi06694 | submitted | 2007 | Eger stream in Marktleuthen, NE of Weissenstadt | Eger/Elbe | North Sea | Germany | 50.091956, 11.9099323 |
| *T. thymallus* | ZSM-PIS-036238 \| BayFi06849 | submitted | 2007 | Sächsische Saale stream N of Förbau, E of Münchberg | Sächsische Saale/Elbe | North Sea | Germany | 50.215521, 11.9231713 |
| *T. thymallus* | ZSM-PIS-036238 \| BayFi06874 | submitted | 2007 | Sächsische Saale stream N of Förbau, E of Münchberg | Sächsische Saale/Elbe | North Sea | Germany | 50.215521, 11.9231713 |
| *T. thymallus* | THU12 | submitted | 2009 | Thur River | Thur/Rhine | North Sea | Switzerland | 47.392027, 9.0788703 |
| *T. thymallus* | THU13 | submitted | 2009 | Thur River | Thur/Rhine | North Sea | Switzerland | 47.392027, 9.0788703 |
| *T. thymallus* | LIN20 | submitted | 2009 | Linth River (Linthkanal) | Linth/Limmat/Aare/Rhein | North Sea | Switzerland | 47.166667, 9 |
| *T. thymallus* | LIN21 | submitted | 2009 | Linth River (Linthkanal) | Linth/Limmat/Aare/Rhein | North Sea | Switzerland | 47.166667, 9 |
| *T. thymallus* | ORB69 | submitted | 2009 | Orbe River | Orbe/Thielle/Zihlkanal/Aare/Rhein | North Sea | Switzerland | 46.7, 6.35 |
| *T. thymallus* | ORB75 | submitted | 2009 | Orbe River | Orbe/Thielle/Zihlkanal/Aare/Rhein | North Sea | Switzerland | 46.7, 6.35 |
| *T. thymallus* | SAR001 | submitted | 2009 | Sarner Aa River | Sarber Aa/Reuss/Aare/Rhein | North Sea | Switzerland | 46.899754, 8.2446948 |
| *T. thymallus* | SAR002 | submitted | 2009 | Sarner Aa River | Sarber Aa/Reuss/Aare/Rhein | North Sea | Switzerland | 46.899754, 8.2446948 |
| *T. thymallus* | WER10 | submitted | 2009 | Werdenberger Binnenkanal (Alpine Rhine) | Werdenberger Binnenkanal/Rhine | North Sea | Switzerland | 47.271786, 9.5256863 |
| *T. thymallus* | WER11 | submitted | 2009 | Werdenberger Binnenkanal (Alpine Rhine) | Werdenberger Binnenkanal/Rhine | North Sea | Switzerland | 47.271786, 9.5256863 |
| *T. thymallus* | ARE28 | submitted | 2009 | Areuse River | Areuse/Zihlkanal/Aare/Rhein | North Sea | Switzerland | 46.949263, 6.7491353 |
| *T. thymallus* | ARE29 | submitted | 2009 | Areuse River | Areuse/Zihlkanal/Aare/Rhein | North Sea | Switzerland | 46.949263, 6.7491353 |
| *T. thymallus* | STR12 | submitted | 2009 | Rhine River at Stein am Rhein | Rhine | North Sea | Switzerland | 47.69069, 8.7496053 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *T. thymallus* | STR13 | submitted | 2009 | Rhine River at Stein am Rhein | Rhine | North Sea | Switzerland | 47.69069, 8.7496053 |
| *T. thymallus* | AAR035 | submitted | 2009 | Aare River at Aaran | Aare/Rhine | North Sea | Switzerland | 47.39525, 8.04131 |
| *T. thymallus* | AAR026 | submitted | 2009 | Aare River at Aaran | Aare/Rhine | North Sea | Switzerland | 47.39525, 8.04131 |
| *T. thymallus* | AAR031 | submitted | 2009 | Aare River at Aaran | Aare/Rhine | North Sea | Switzerland | 47.39525, 8.04131 |
| *T. thymallus* | MOR06 | submitted | 2010 | Mortagne River at Autrey | Mortagne/Meurthe/Moselle/Rhine | North Sea | France | 48.2823501, 6.6930834 |
| *T. thymallus* | MOR08 | submitted | 2010 | Mortagne River at Autrey | Mortagne/Meurthe/Moselle/Rhine | North Sea | France | 48.2823501, 6.6930834 |
| *T. thymallus* | MOR18 | submitted | 2010 | Mortagne River at Autrey | Mortagne/Meurthe/Moselle/Rhine | North Sea | France | 48.2823501, 6.6930834 |
| *T. thymallus* | MOS08 | submitted | 2011 | Moselle River at Rupt | Moselle/Rhine | North Sea | France | 48.2007498, 6.4414482 |
| *T. thymallus* | ZSM-PIS-033673 \| BayFi03766 | submitted | 2005 | Thulba stream at Oberthulba | Fränkische Saale/Main/Rhine | North Sea | German | 50.204219, 9.9642813 |
| *T. thymallus* | ZSM-PIS-033673 \| BayFi03767 | submitted | 2005 | Thulba stream at Oberthulba | Fränkische Saale/Main/Rhine | North Sea | German | 50.204219, 9.9642813 |
| *T. thymallus* | ZSM-PIS-037937 \| BayFi08125 | submitted | 2008 | Weißer Main River E of Kulmbach | Weißer Main/Main/Rhine | North Sea | German | 50.109899, 11.4824263 |
| *T. thymallus* | SIE05 | SRR19545529 | 2012 | Sieg River near Bonn | Sieg/Rhine | North Sea | German | 50.782222, 7.213333 |
| *T. thymallus* | SIE3 | submitted | 2012 | Sieg River near Bonn | Sieg/Rhine | North Sea | German | 50.782222, 7.213333 |
| *T. thymallus* | SIE4 | submitted | 2012 | Sieg River near Bonn | Sieg/Rhine | North Sea | German | 50.782222, 7.213333 |
| *T. thymallus* | HOU1_4 | submitted | 2013 | Houille River at Givet | Houille/Maas/Hollands Diep | North Sea | Belgium | 50.127778, 4.8425 |
| *T. thymallus* | HOU1_5 | submitted | 2013 | Houille River at Givet | Houille/Maas/Hollands Diep | North Sea | Belgium | 50.127778, 4.8425 |
| *T. thymallus* | CLO4 | submitted | 2013 | Crusnes River at Longuyon | Crusnes/Chiers/Maas/Hollands Diep | North Sea | France | 49.440215, 5.6062393 |
| *T. thymallus* | CLO1 | submitted | 2013 | Crusnes River at Longuyon | Crusnes/Chiers/Maas/Hollands Diep | North Sea | France | 49.440215, 5.6062393 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *T. thymallus* | OOO22 | submitted | 1999 | Ourthe Occidentale River | Ourthe Occidentale/Maas/Hollands Diep | North Sea | Belgium | 50.347333, 5.4382143 |
| *T. thymallus* | OOO25 | submitted | 1999 | Ourthe Occidentale River | Ourthe Occidentale/Maas/Hollands Diep | North Sea | Belgium | 50.347333, 5.4382143 |
| *T. thymallus* | DGO20 | submitted | 2009 | Doubs River at Gogniat | Doubs/Saône/Rhône | Mediterranean Sea | Switzerland | 47.340266, 7.156468 |
| *T. thymallus* | DGO23 | submitted | 2009 | Doubs River at Gogniat | Doubs/Saône/Rhône | Mediterranean Sea | Switzerland | 47.340266, 7.156468 |
| *T. thymallus* | LMH01 | submitted | 2012 | Loue River at Mouthier-Haute-Pierre | Loue/Doubs/Saône/Rhône | Mediterranean Sea | France | 47.0373921, 6.2745824 |
| *T. thymallus* | LMH02 | submitted | 2012 | Loue River at Mouthier-Haute-Pierre | Loue/Doubs/Saône/Rhône | Mediterranean Sea | France | 47.0373921, 6.2745824 |
| *T. thymallus* | OGB24 | submitted | 2012 | Ognon River at Belonchamp | Ognon/Saône/Rhône | Mediterranean Sea | France | 47.7711, 6.6163153 |
| *T. thymallus* | OGB25 | submitted | 2012 | Ognon River at Belonchamp | Ognon/Saône/Rhône | Mediterranean Sea | France | 47.7711, 6.6163153 |
| *T. thymallus* | BRE12 | submitted | 2012 | Breuchin River | Breuchin/Saône/Rhône | Mediterranean Sea | France | 47.834866, 6.4805503 |
| *T. thymallus* | BRE13 | submitted | 2012 | Breuchin River | Breuchin/Saône/Rhône | Mediterranean Sea | France | 47.834866, 6.4805503 |
| *T. thymallus* | CRO01 (GR93) | SRR19545525 | N/A | Ain River at Crotenay | Ain/Rhône | Mediterranean Sea | France | 46.7451317, 5.809925 |
| *T. thymallus* | MOL32 (aka AIN) pos. 17 | submitted | 2009 | Ain River at Mollon | Ain/Rhône | Mediterranean Sea | France | 45.9441907, 5.2475483 |
| *T. thymallus* | MOL18 | submitted | 2009 | Ain River at Mollon | Ain/Rhône | Mediterranean Sea | France | 45.9441907, 5.2475483 |
| *T. thymallus* | BPM4 | submitted | 2012 | Bourne River at Pont de Manne | Bourne/Isère/Rhône | Mediterranean Sea | France | 45.060348, 5.2854893 |
| *T. thymallus* | BPM8 | submitted | 2012 | Bourne River at Pont de Manne | Bourne/Isère/Rhône | Mediterranean Sea | France | 45.060348, 5.2854893 |
| *T. thymallus* | SOR17 | submitted | 2011 | Sorge River in L'Isle-sur-la-Sorgue | Sorgue/Rhône | Mediterranean Sea | France | 43.921389, 5.051111 |
| *T. thymallus* | SOR19 | submitted | 2011 | Sorge River in L'Isle-sur-la-Sorgue | Sorgue/Rhône | Mediterranean Sea | France | 43.921389, 5.051111 |
| *T. thymallus* | ZSM-PIS-031530 \| BayFi01092 | submitted | 2004 | Westliche Günz stream upstream of Westerheim, near Memmingen | Westliche Günz/Günz/Danube | Black Sea | Germany | 48.03806, 10.2842863 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *T. thymallus* | ZSM-PIS-031530 \| BayFi01093 | submitted | 2004 | Westliche Günz stream upstream of Westerheim, near Memmingen | Westliche Günz/Günz/Danube | Black Sea | Germany | 48.03806, 10.2842863 |
| *T. thymallus* | ZSM-PIS-031699 \| BayFi01482 | submitted | 2004 | Amper stream 3 km downriver of Lake Ammersee, near Stegen | Amper/Isar/Danube | Black Sea | Germany | 48.103719, 11.1343753 |
| *T. thymallus* | ZSM-PIS-031699 \| BayFi01483 | submitted | 2004 | Amper stream 3 km downriver of Lake Ammersee, near Stegen | Amper/Isar/Danube | Black Sea | Germany | 48.103719, 11.1343753 |
| *T. thymallus* | ZSM-PIS-033488 \| BayFi03094 | submitted | 2005 | Mitternacher Ohe stream, Zehrer Mühle at Schönberg | Mitternacher Ohe/Ilz/Danube | Black Sea | Germany | 48.858097, 13.3043153 |
| *T. thymallus* | ZSM-PIS-033488 \| BayFi03095 | submitted | 2005 | Mitternacher Ohe stream, Zehrer Mühle at Schönberg | Mitternacher Ohe/Ilz/Danube | Black Sea | Germany | 48.858097, 13.3043153 |
| *T. thymallus* | ZSM-PIS-033503 \| BayFi03173 | submitted | 2005 | Geltnach stream below Bertoldshofen near Marktoberdorf | Geltnach/Wertach/Lech/Danube | Black Sea | Germany | 47.775609, 10.6576623 |
| *T. thymallus* | ZSM-PIS-033503 \| BayFi03172 | submitted | 2005 | Geltnach stream below Bertoldshofen near Marktoberdorf | Geltnach/Wertach/Lech/Danube | Black Sea | Germany | 47.775609, 10.6576623 |
| *T. thymallus* | ZSM-PIS-033627 \| BayFi03506 | submitted | 2005 | Wertach stream below Maria Rain, NW of Nesselwang | Wertach/Lech/Danube | Black Sea | Germany | 47.63266, 10.4933373 |
| *T. thymallus* | ZSM-PIS-033627 \| BayFi03507 | submitted | 2005 | Wertach stream below Maria Rain, NW of Nesselwang | Wertach/Lech/Danube | Black Sea | Germany | 47.63266, 10.4933373 |
| *T. thymallus* | ZSM-PIS-033852 \| BayFi04221 | submitted | 2005 | Iller River upstream of Sonthofen (4212) | Iller/Danube | Black Sea | Germany | 47.539163, 10.2635893 |
| *T. thymallus* | ZSM-PIS-033747 \| BayFi04224 | submitted | 2005 | Iller River at Blaichach, Swabia (4220-21) | Iller/Danube | Black Sea | Germany | 47.539163, 10.2635893 |
| *T. thymallus* | ZSM-PIS-033747 \| BayFi04225 | submitted | 2005 | Iller River at Blaichach, Swabia (4220-21) | Iller/Danube | Black Sea | Germany | 47.539163, 10.2635893 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *T. thymallus* | ZSM-PIS-035826 \| BayFi06308 | submitted | 2007 | Haidenaab stream near Hütten, SE of Grafenwöhr | Naab/Danube | Black Sea | Germany | 49.688163, 11.9608793 |
| *T. thymallus* | ZSM-PIS-035826 \| BayFi06318 | submitted | 2007 | Haidenaab stream near Hütten, SE of Grafenwöhr | Naab/Danube | Black Sea | Germany | 49.688163, 11.9608793 |
| *T. thymallus* | 33(GR114) | SRR19545526 | 2010 | Lake Heiterwanger | Archbach/Lech/Danube | Black Sea | Austria | 47.456944, 10.776667 |
| *T. thymallus* | H29_2 | submitted | 2010 | Lake Heiterwanger | Archbach/Lech/Danube | Black Sea | Austria | 47.456944, 10.776667 |
| *T. thymallus* | P12_plansee | submitted | 2010 | Lake Plansee | Archbach/Lech/Danube | Black Sea | Austria | 47.46925, 10.800661 |
| *T. thymallus* | P14 | submitted | 2010 | Lake Plansee | Archbach/Lech/Danube | Black Sea | Austria | 47.46925, 10.800661 |
| *T. thymallus* | 18 (GR62) | submitted | 2014 | Saglbach stream | Saglbach/Inn/Danube | Black Sea | Austria | 47.311667, 11.106111 |
| *T. thymallus* | INN42 | submitted | 2009 | Inn River in Oberengadin, Samedan | Inn/Danube | Black Sea | Switzerland | 46.5320427, 9.87425 |
| *T. thymallus* | INN44 | submitted | 2009 | Inn River in Oberengadin, Samedan | Inn/Danube | Black Sea | Switzerland | 46.5320427, 9.87425 |
| *T. thymallus* | PFU549 | submitted | 2018 | Inn River at Pfunds | Inn/Danube | Black Sea | Austria | 46.999375, 10.589803 |
| *T. thymallus* | PFU550 | submitted | 2018 | Inn River at Pfunds | Inn/Danube | Black Sea | Austria | 46.999375, 10.589803 |
| *T. thymallus* | PFU552 | submitted | 2018 | Inn River at Pfunds | Inn/Danube | Black Sea | Austria | 46.999375, 10.589803 |
| *T. thymallus* | PFU553 | submitted | 2018 | Inn River at Pfunds | Inn/Danube | Black Sea | Austria | 46.999375, 10.589803 |
| *T. thymallus* | PRU508 | submitted | 2018 | Inn River at Prutz | Inn/Danube | Black Sea | Austria | 47.082703, 10.668367 |
| *T. thymallus* | PRU511 | submitted | 2018 | Inn River at Prutz | Inn/Danube | Black Sea | Austria | 47.082703, 10.668367 |
| *T. thymallus* | PRU520 | submitted | 2018 | Inn River at Prutz | Inn/Danube | Black Sea | Austria | 47.082703, 10.668367 |
| *T. thymallus* | PRU524 | submitted | 2018 | Inn River at Prutz | Inn/Danube | Black Sea | Austria | 47.082703, 10.668367 |
| *T. thymallus* | TRA1_007 | submitted | N/A | Traun River | Traun/Danube | Black Sea | Austria | 48.081514, 13.856544 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *T. thymallus* | KAM024 | submitted | 2018 | Kleiner Kamp stream at Rappottenstein | Kleiner Kamp/Kamp/Danube | Black Sea | Austria | 48.5064535, 15.0750933 |
| *T. thymallus* | KAM025 | submitted | 2018 | Kleiner Kamp stream at Rappottenstein | Kleiner Kamp/Kamp/Danube | Black Sea | Austria | 48.5064535, 15.0750933 |
| *T. thymallus* | 119o (GR102) | SRR19545528 | 2012 | Lafnitz River | Lafnitz/Raab/Dabube | Black Sea | Austria | 47.305, 16.063056 |
| *T. thymallus* | Laf23 | submitted | 2010 | Lafnitz River between Lafnitz und Rohrbach | Lafnitz/Raab/Dabube | Black Sea | Austria | 47.3855895, 15.9997946 |
| *T. thymallus* | Laf24 | submitted | 2010 | Lafnitz River between Lafnitz und Rohrbach | Lafnitz/Raab/Dabube | Black Sea | Austria | 47.3855895, 15.9997946 |
| *T. thymallus* | Ham7 | submitted | 2006 | Lafnitz River at Hammerwald | Lafnitz/Raab/Dabube | Black Sea | Austria | 47.333333, 16.033333 |
| *T. thymallus* | Ham8 | submitted | 2006 | Lafnitz River at Hammerwald | Lafnitz/Raab/Dabube | Black Sea | Austria | 47.333333, 16.033333 |
| *T. thymallus* | ODR006 | SRR19545530 | 2013 | Upper Drau River | Drau/Danube | Black Sea | Austria | 46.774381, 13.344769 |
| *T. thymallus* | ODR007 | submitted | 2013 | Upper Drau River | Drau/Danube | Black Sea | Austria | 46.774381, 13.344769 |
| *T. thymallus* | UNZ11 | submitted | 2004 | Mur River at Unzmarkt | Mur/Drau/Danube | Black Sea | Austria | 47.2, 14.416667 |
| *T. thymallus* | UNZ13 | submitted | 2004 | Mur River at Unzmarkt | Mur/Drau/Danube | Black Sea | Austria | 47.2, 14.416667 |
| *T. thymallus* | UNZ12_2 | submitted | 2004 | Mur River at Unzmarkt | Mur/Drau/Danube | Black Sea | Austria | 47.2, 14.416667 |
| *T. thymallus* | Boh1 | submitted | 2017 | Sava Bohinjka stream | Sava Bohinjka/Sava/Danube | Black Sea | Slovenia | 46.334056, 14.061694 |
| *T. thymallus* | Boh2 | submitted | 2017 | Sava Bohinjka stream | Sava Bohinjka/Sava/Danube | Black Sea | Slovenia | 46.334056, 14.061694 |
| *T. thymallus* | Uni1 | submitted | 2017 | Unica River | Unica/Ljubljanica/Sava/Danube | Black Sea | Slovenia | 45.839528, 14.274222 |
| *T. thymallus* | Uni2 | submitted | 2017 | Unica River | Unica/Ljubljanica/Sava/Danube | Black Sea | Slovenia | 45.839528, 14.274222 |
| *T. thymallus* | Kol1 | submitted | 2015 | Kolpa River | Kolpa/Sava/Danube | Black Sea | Slovenia | 45.52575, 14.700778 |
| *T. thymallus* | Kol3 | submitted | 2015 | Kolpa River | Kolpa/Sava/Danube | Black Sea | Slovenia | 45.52575, 14.700778 |
| *T. thymallus* | Una2 | submitted | 2016 | Una River | Una/Sava/Danube | Black Sea | Bosnia and Herzegovina | 44.532222, 16.114444 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *T. thymallus* | Una3 | submitted | 2016 | Una River | Una/Sava/Danube | Black Sea | Bosnia and Herzegovina | 44.532222, 16.114444 |
| *T. thymallus* | Ibr1 | submitted | 2008 | Ibar River | Ibar/West Morava/Great Morava/Danube | Black Sea | Serbia | 42.911507, 20.345006 |
| *T. thymallus* | Ibr2 | submitted | 2008 | Ibar River | Ibar/West Morava/Great Morava/Danube | Black Sea | Serbia | 42.911507, 20.345006 |
| *T. thymallus* | Rzv1 | submitted | 2008 | Rzav River (stocked from Slovenia) | Rzav/Golijska Moravica/West Morava/Great Morava/Danube | Black Sea | Serbia | 43.773694, 19.933306 |
| *T. thymallus* | Rzv2 | submitted | 2008 | Rzav River (stocked from Slovenia) | Rzav/Golijska Moravica/West Morava/Great Morava/Danube | Black Sea | Serbia | 43.773694, 19.933306 |
| *T. thymallus* | Lju1 | submitted | 2008 | Ljuča River | Ljuča/Lim/Drina/Sava/Danube | Black Sea | Montenegro | 42.575, 19.890889 |
| *T. thymallus* | Lju3 | submitted | 2008 | Ljuča River | Ljuča/Lim/Drina/Sava/Danube | Black Sea | Montenegro | 42.575, 19.890889 |
| *T. thymallus* | Dri1 | submitted | 2008 | Drina River | Drina/Sava/Danube | Black Sea | Serbia | 43.957805, 19.431628 |
| *T. thymallus* | Dri2 | submitted | 2008 | Drina River | Drina/Sava/Danube | Black Sea | Serbia | 43.957805, 19.431628 |
| *T. thymallus* | Lim1 | submitted | 2008 | Lim River | Lim/Drina/Sava/Danube | Black Sea | Serbia | 43.596963, 19.381493 |
| *T. thymallus* | Lim2 | submitted | 2008 | Lim River | Lim/Drina/Sava/Danube | Black Sea | Serbia | 43.596963, 19.381493 |
| *T. thymallus* | Tar2 | submitted | 2003 | Drina River | Drina/Sava/Danube | Black Sea | Bosnia and Herzegovina | 43.400389, 18.777806 |
| *T. thymallus* | Tar3 | submitted | 2003 | Drina River | Drina/Sava/Danube | Black Sea | Bosnia and Herzegovina | 43.400389, 18.777806 |
| *T. thymallus* | Bos1 | submitted | 2004 | Željeznica River | Bosna/Sava/Danube | Black Sea | Bosnia and Herzegovina | 43.73225, 18.402167 |
| *T. thymallus* | Vrb1 | submitted | 2020 | Vrbas River | Vrbas/Sava/Danube | Black Sea | Bosnia and Herzegovina | 44.614700, 17.148969 |
| *T. thymallus* | Vrb2 | submitted | 2020 | Vrbas River | Vrbas/Sava/Danube | Black Sea | Bosnia and Herzegovina | 44.614700, 17.148969 |
| *T. thymallus* | Sana5 | submitted | 2020 | Sana River | Sana/Una/Sava/Danube | Black Sea | Bosnia and Herzegovina | 44.479238, 16.814854 |
| *T. thymallus* | Sana7 | submitted | 2020 | Sana River | Sana/Una/Sava/Danube | Black Sea | Bosnia and Herzegovina | 44.479238, 16.814854 |
| *T. thymallus* | RUS33 | submitted | 2000 | Ruscova River | Ruscova/Vișeu/Tisza/Danube | Black Sea | Romania | 47.790499, 24.2830173 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *T. thymallus* | RUS14 | submitted | 2000 | Ruscova River | Ruscova/Vișeu/Tisza/Danube | Black Sea | Romania | 47.790499, 24.2830173 |
| *T. thymallus* | RUS25 | submitted | 2000 | Ruscova River | Ruscova/Vișeu/Tisza/Danube | Black Sea | Romania | 47.790499, 24.2830173 |
| *T. thymallus* | VIS22 | submitted | 2000 | Vișeu River | Vișeu/Tisza/Danube | Black Sea | Romania | 47.907501, 24.1440773 |
| *T. thymallus* | VIS23 | submitted | 2000 | Vișeu River | Vișeu/Tisza/Danube | Black Sea | Romania | 47.907501, 24.1440773 |
| *T. thymallus* | VIS24 | submitted | 2000 | Vișeu River | Vișeu/Tisza/Danube | Black Sea | Romania | 47.907501, 24.1440773 |

**Table S2.** Information of the *Thymallus* samples used for mitogenome analyses. Included is the original sample code, accession number, and sample site location (including drainage, basin, country and geo-coordinates).

| **Species** | **Sample code** | **Accession number** | **Sampling date** | **Location** | **Drainage Progression** | **Basin** | **Country** | **Coordinates** |
|---|---|---|---|---|---|---|---|---|
| *T. flavomaculatus* | URM003 | MT063039 | N/A | Urmi River (Khabarovsk Krai) | Urmi/Tunguska/Amur | Sea of Okhotsk (North Pacific Ocean) | Russia | 48.713056, 134.266944 |
| *T. tugarinae* | N/A | KJ866483 | N/A | Huma River | Huma/Amur | Sea of Okhotsk (North Pacific Ocean) | China | 52.3, 124.7 |
| *T. aeliani* | FAL02 | MT063009 | 2014 | Falschauer River | Falschauer/Adige | Adriatic Sea | Italy | 46.627938, 11.171812 |
| *T. aeliani* | ETS01 | MT063008 | 2014 | Adige River (south of Meran) | Adige | Adriatic Sea | Italy | 46.616389, 11.184167 |
| *T. ligericus* | S20 | MT063053 | 2017 | Sioule River (Châteauneuf-les-Bains) | Sioule/ Allier/Loire | Atlantic Ocean | France | 46.022778, 2.904167 |
| *T. ligericus* | Alg06 | MT063007 | N/A | Alagnon River (Chapelle d'Alagnon) | Alagnon/Allier/Loire | Atlantic Ocean | France | 46.106389, 2.896667 |
| *T. thymallus* | UNZ11 | submitted | 2004 | Mur River at Unzmarkt | Mur/Drau/Danube | Black Sea | Austria | 47.2, 14.416667 |
| *T. thymallus* | ODR15 | MT062996 | 2013 | Upper Drau River | Drau/Danube | Black Sea | Austria | 46.774381, 13.344769 |
| *T. thymallus* | TRA002 | submitted | N/A | Traun River | Traun/Danube | Black Sea | Austria | 48.081514, 13.856544 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *T. thymallus* | ZSM-PIS-033503 \| BayFi03173 | submitted | 2005 | Geltnach stream below Bertoldshofen near Marktoberdorf | Geltnach/Wertach/Lech/Danube | Black Sea | Austria | 47.775609, 10.6576623 |
| *T. thymallus* | VIS23 | submitted | 2000 | Vișeu River | Vișeu/Tisza/Danube | Black Sea | Romania | 47.907501, 24.1440773 |
| *T. thymallus* | RUS14 | submitted | 2000 | Ruscova River | Ruscova/Vișeu/Tisza/Danube | Black Sea | Romania | 47.790499, 24.2830173 |
| *T. thymallus* | TTM2012 | CM015041 | N/A | Glomma River | Glomma | Oslofjord (Baltic Sea) | Norway | 61.420000, 11.090000 |
| *T. thymallus* | LAP010 | submitted | 2020 | Lake Padje Kaitumjaure | Kaitumälven/Kalixälven | Gulf of Bothnia | Sweden | 67.741117, 18.290243 |
| *T. thymallus* | LAP011 | submitted | 2020 | Lake Padje/Kaska Kaitumjaure | Kaitumälven/Kalixälven | Gulf of Bothnia | Sweden | 67.651413, 18.721269 |
| *T. thymallus* | Thy23 | MT063000 | N/A | Mindyak River | Mindyak/Ural | Caspian Sea | Russia | 53.984444, 58.799444 |
| *T. thymallus* | SCH001 | MT062999 | N/A | Schudya River | Schudya/Kama/Volga | Caspian Sea | Russia | N/A |
| *T. thymallus* | I05 | MT062995 | 2015 | Palniki-Ju | Palniki-Ju/Junjacha/Usa/Pechora | Barents Sea (Arctic Ocean) | Russia | 66.523955, 62.782175 |
| *T. thymallus* | L12-13 | MT063002 | 2012 | Lafnitz River | Lafnitz/Raab/Danube | Black Sea | Austria | 47.305, 16.063056 |
| *T. thymallus* | KOLPA25 | MT063001 | 2016 | Kolpa River (Osilnica) | Kolpa/Sava/Danube | Black Sea | Slovenia | 45.52575, 14.700778 |
| *T. thymallus* | Ibr1 | submitted | 2008 | Ibar River | Ibar/West Morava/Great Morava/Danube | Black Sea | Serbia | 42.911507, 20.345006 |
| *T. thymallus* | DM364 | MT410870 | N/A | N/A | N/A | N/A | Denmark | 56.019094, 10.449945 |
| *T. thymallus* | Thy27 | MT063005 | 2012 | Sieg River (Buisdorff) | Sieg/Rhine | North Sea | Germany | 50.782222, 7.213333 |
| *T. thymallus* | GR113 | MT063004 | 2013 | Houille River (Givet) | Houille/Maas/Rhine | North Sea | France | 50.127778, 4.8425 |
| *T. thymallus* | Thy26 | MT063003 | 2000 | Ure River | Ure/Ouse | North Sea | England | 54.305556, -1.939722 |
| *T. thymallus* | AAR026 | submitted | 2009 | Aare River at Aaran | Aare/Rhine | North Sea | Switzerland | 47.39525, 8.04131 |
| *T. thymallus* | ORB69 | submitted | 2009 | Orbe River | Orbe/Thielle/Zihlkanal/Aare/Rhein | North Sea | Switzerland | 46.7, 6.35 |
| *T. thymallus* | OMB14 | MT063043 | 2011 | Sorge River in L'Isle-sur-la-Sorgue | Sorgue/Rhône | Mediterranean Sea | France | 43.921389, 5.051111 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *T. thymallus* | OMB10 | MT063042 | 2015 | Ain River (Marigny) | Ain/Rhône | Mediterranean Sea | France | 46.686389, 5.764722 |

**Table S3.** List of CR haplotypes used in this study (959 bp). Shown are the species, the haplotype number from Figure 2, the original haplotype number from the source publication or report, the GenBank accession number, the source references, and additional remarks regarding haplotype nomenclature. Note that original published haplotype nomenclature corresponds to a longer sequence and thus variation in regions not included here can lead to the collapsing of some haplotypes (see Remarks). Additionally, an overlap in the naming of haplotypes in Marić et al. (2012) and Meraner & Gandolfi (2012) resulted in duplicate haplotype names Ad7, Da31, Da32 and Da33. Original codes were preserved for haplotypes published by Meraner & Gandolfi (2012), and haplotypes in Marić et al. (2012) were referred to as Ad7cs, Da30cs, Da31cs, Da32cs and Da33cs.

| Species | Number in Figure 2 | Haplotype | Accession | Reference | Remarks |
|---|---|---|---|---|---|
| *T. aeliani* | 2 | **Ad1** | AF522419.1 | Weiss et al. 2002 | |
| *T. aeliani* | 3 | **Ad2** | AF522420.1 | Weiss et al. 2002 | |
| *T. aeliani* | 4 | **Ad3** | AF522421.1 | Weiss et al. 2002 | |
| *T. aeliani* | 5 | **Ad4** | AF522422.1 | Weiss et al. 2002 | |
| *T. aeliani* | 6 | **Ad5** | AF522423.1 | Weiss et al. 2002 | |
| *T. aeliani* | 7 | **Ad6** | AF522424.1 | Weiss et al. 2002 | |
| *T. aeliani* | 8 | **Ad7** | JN796420.1 | Meraner & Gandolfi, 2012 | includes MT762347 (neotype) |
| *T. aeliani* | 9 | **Ad9** | JN796422.1 | Meraner & Gandolfi, 2012 | includes haplotypes Ad10, Ad11, Ad7cs and samples ETS01 and FAL02 |
| *T. aeliani* | 9 | Ad7cs | JX099344.1 | Marić et al. unpublished | collapses to Ad9 |
| *T. aeliani* | 9 | Ad10 | JN796423.1 | Meraner & Gandolfi, 2012 | collapses to Ad9 |
| *T. aeliani* | 9 | Ad11 | JN796424.1 | Meraner & Gandolfi, 2012 | collapses to Ad9 |
| *T. aeliani* | 10 | **Ad8** | JN796421.1 | Meraner & Gandolfi, 2012 | includes Ad22 |
| *T. aeliani* | 10 | Ad22 | JN796435.1 | Meraner & Gandolfi, 2012 | collapses to Ad8 |
| *T. aeliani* | 11 | **Ad12** | JN796425.1 | Meraner & Gandolfi, 2012 | includes haplotype Ad15 |
| *T. aeliani* | 11 | Ad15 | JN796428.1 | Meraner & Gandolfi, 2012 | collapses to Ad12 |
| *T. aeliani* | 12 | **Ad13** | JN796426.1 | Meraner & Gandolfi, 2012 | |
| *T. aeliani* | 13 | **Ad14** | JN796427.1 | Meraner & Gandolfi, 2012 | |
| *T. aeliani* | 14 | **Ad16** | JN796429.1 | Meraner & Gandolfi, 2012 | |
| *T. aeliani* | 15 | **Ad17** | JN796430.1 | Meraner & Gandolfi, 2012 | |
| *T. aeliani* | 16 | **Ad18** | JN796431.1 | Meraner & Gandolfi, 2012 | includes sample Ses189 |

| | | | | | |
|---|---|---|---|---|---|
| *T. aeliani* | 16 | Ses189 | | unpublished | collapses to Ad18 |
| *T. aeliani* | 17 | **Ad19** | JN796432.1 | Meraner & Gandolfi, 2012 | |
| *T. aeliani* | 18 | **Ad20** | JN796433.1 | Meraner & Gandolfi, 2012 | |
| *T. aeliani* | 19 | **Ad21** | JN796434.1 | Meraner & Gandolfi, 2012 | |
| *T. ligericus* | 20 | **At1** | AF522425.1 | Weiss et al. 2002 | includes samples OMB3, OMB4 and OMB5 |
| *T. ligericus* | 102 | Ola23 | | unpublished | collapses to Ht20 |
| *T. ligericus* | 103 | **Ht21** | | unpublished | includes samples ACO02, Alg06 and F1 |
| *T. ligericus* | 105 | Olx2 | | unpublished | collapses to Ht23 |
| *T. ligericus* | 106 | **Ht24** | | unpublished | |
| *T. ligericus* | 107 | **Ht25** | | unpublished | |
| *T. ligericus* | 108 | **Ht26** | | unpublished | |
| *T. ligericus* | 110 | **Ht29** | | unpublished | |
| *T. ligericus* | 111 | **Ht30** | | unpublished | |
| *T. thymallus* | 1 | **A45** | AY594181.1 | Duftner et al. 2005 | |
| *T. thymallus* | 21 | **At2** | AF522426.1 | Weiss et al. 2002 | includes samples LAP10 and LAP11 |
| *T. thymallus* | 22 | **At3** | AF522427.1 | Weiss et al. 2002 | |
| *T. thymallus* | 23 | **At4** | AF522428.1 | Weiss et al. 2002 | |
| *T. thymallus* | 24 | **At5** | AF522429.1 | Weiss et al. 2002 | |
| *T. thymallus* | 25 | **At6** | AF522430.1 | Weiss et al. 2002 | |
| *T. thymallus* | 26 | **At7** | AF522431.1 | Weiss et al. 2002 | |
| *T. thymallus* | 27 | **At8** | AF522432.1 | Weiss et al. 2002 | |
| *T. thymallus* | 28 | **At9** | AF522433.1 | Weiss et al. 2002 | includes haplotype At28 and samples FEI3_45 and RBA19 |
| *T. thymallus* | 28 | At28 | JN796437.1 | Meraner & Gandolfi, 2012 | collapses to At9 |
| *T. thymallus* | 29 | **At10** | AF522434.1 | Weiss et al. 2002 | |
| *T. thymallus* | 30 | **At11** | AF522435.1 | Weiss et al. 2002 | |
| *T. thymallus* | 31 | **At12** | AF522436.1 | Weiss et al. 2002 | |
| *T. thymallus* | 32 | **At13** | AF522437.1 | Weiss et al. 2002 | |
| *T. thymallus* | 33 | **At14** | AF522438.1 | Weiss et al. 2002 | includes haplotype At20 and samples AAR023, OMB16 |
| *T. thymallus* | 33 | At20 | AF522444.1 | Weiss et al. 2002 | collapses to At14 |
| *T. thymallus* | 34 | **At15** | AF522439.1 | Weiss et al. 2002 | includes samples OMB15 and ORB73 |
| *T. thymallus* | 35 | **At16** | AF522440.1 | Weiss et al. 2002 | includes At17 |

| | | | | | |
|---|---|---|---|---|---|
| *T. thymallus* | 35 | At17 | AF522441.1 | Weiss et al. 2002 | collapses to At16 |
| *T. thymallus* | 36 | **At18** | AF522442.1 | Weiss et al. 2002 | |
| *T. thymallus* | 37 | **At19** | AF522443.1 | Weiss et al. 2002 | |
| *T. thymallus* | 38 | **At21** | AF522445.1 | Weiss et al. 2002 | |
| *T. thymallus* | 39 | **AT22** | AY841355.1 | Gum et al. 2005 | |
| *T. thymallus* | 40 | **AT23** | AY841356.1 | Gum et al. 2005 | |
| *T. thymallus* | 41 | **AT24** | AY841357.1 | Gum et al. 2005 | |
| *T. thymallus* | 42 | **Da4** | AF522398.1 | Weiss et al. 2002 | includes haplotype At25 and sample Muerz |
| *T. thymallus* | 42 | AT25 | AY841360.1 | Gum et al. 2005 | collapses to Da4 |
| *T. thymallus* | 42 | DA22 | AY841358.1 | Gum et al. 2005 | collapses to Da4 |
| *T. thymallus* | - | AT26 | AY841359.1 | Gum et al. 2005 | removed (330 bp repeat region) |
| *T. thymallus* | 43 | **At27** | JN796436.1 | Meraner & Gandolfi, 2012 | collapses to Ht3 |
| *T. thymallus* | 43 | Ht3 | | unpublished | collapses to At27 |
| *T. thymallus* | - | At29 | JX961600.1 | Meraner et al. 2013 | removed (330 bp repeat region) |
| *T. thymallus* | 44 | **Ca1** | JX144730.1 | Marić et al. 2014 | |
| *T. thymallus* | 45 | **Ca2** | JX144731.1 | Marić et al. 2014 | |
| *T. thymallus* | 46 | **Ca3** | JX144732.1 | Marić et al. 2014 | includes sample Thy23 |
| *T. thymallus* | 47 | **Ca4** | KF280207.1 | Marić et al. 2014 | |
| *T. thymallus* | 48 | **Ca5** | KF280208.1 | Marić et al. 2014 | |
| *T. thymallus* | 49 | **Da1** | AF522395.1 | Weiss et al. 2002 | includes haplotype A86 and sample TRA002 |
| *T. thymallus* | 49 | A86 | AY594182.1 | Duftner et al. 2005 | collapses to Da1 |
| *T. thymallus* | 50 | **Da1a / Da2** | AF522396.1 | Weiss et al. 2002 | |
| *T. thymallus* | 51 | **Da3** | AF522397.1 | Weiss et al. 2002 | |
| *T. thymallus* | 52 | **Da5** | AF522399.1 | Weiss et al. 2002 | |
| *T. thymallus* | 53 | **Da6** | AF522400.1 | Weiss et al. 2002 | |
| *T. thymallus* | 54 | **Da7** | AF522401.1 | Weiss et al. 2002 | includes samples ODR15, SAG31 and UNZ11 |
| *T. thymallus* | 55 | **Da8** | AF522402.1 | Weiss et al. 2002 | includes Da16 and Da31 |
| *T. thymallus* | 55 | Da16 | AF522410.1 | Weiss et al. 2002 | collapses to Da8 |
| *T. thymallus* | 55 | Da31 | JN796439.1 | Meraner & Gandolfi, 2012 | collapses to Da8 |
| *T. thymallus* | 56 | **Da9** | AF522403.1 | Weiss et al. 2002 | |

| | | | | | |
|---|---|---|---|---|---|
| *T. thymallus* | 57 | **Da10** | AF522404.1 | Weiss et al. 2002 | |
| *T. thymallus* | 58 | **Da11** | AF522405.1 | Weiss et al. 2002 | includes Da33 and samples BayFi03173and BayFi06874 |
| *T. thymallus* | 58 | Da33 | JN796441.1 | Meraner & Gandolfi, 2012 | collapses to Da11 |
| *T. thymallus* | 59 | **Da12** | AF522406.1 | Weiss et al. 2002 | |
| *T. thymallus* | 60 | **Da13** | AF522407.1 | Weiss et al. 2002 | |
| *T. thymallus* | 61 | **Da14** | AF522408.1 | Weiss et al. 2002 | |
| *T. thymallus* | 62 | **Da15** | AF522409.1 | Weiss et al. 2002 | |
| *T. thymallus* | 63 | **Da17** | AF522411.1 | Weiss et al. 2002 | |
| *T. thymallus* | 64 | **Da18** | AF522412.1 | Weiss et al. 2002 | |
| *T. thymallus* | 65 | **Da19** | AF522413.1 | Weiss et al. 2002 | |
| *T. thymallus* | 66 | **Da20** | AF522414.1 | Weiss et al. 2002 | |
| *T. thymallus* | 67 | **Da21** | AF522415.1 | Weiss et al. 2002 | |
| *T. thymallus* | 68 | **Da22** | AF522416.1 | Weiss et al. 2002 | |
| *T. thymallus* | 69 | **Da23** | AF522417.1 | Weiss et al. 2002 | collapses to Da23cs |
| *T. thymallus* | 69 | Da23cs | JX099336.1 | Marić et al. unpublished | collapses to Da23 |
| *T. thymallus* | 70 | **Da24** | AF522418.1 | Weiss et al. 2002 | |
| *T. thymallus* | 71 | **Da25cs** | JX099337.1 | Marić et al. unpublished | same as Da25; collapses to Da34 |
| *T. thymallus* | 71 | Da34 | JN796442.1 | Meraner & Gandolfi, 2012 | collapses to Da25cs |
| *T. thymallus* | 72 | **Da26cs** | JX099338.1 | Marić et al. unpublished | |
| *T. thymallus* | 73 | **Da27cs** | JX099339.1 | Marić et al. unpublished | |
| *T. thymallus* | 74 | **Da28cs** | JX099340.1 | Marić et al. unpublished | |
| *T. thymallus* | 75 | **Da29cs** | JX099341.1 | Marić et al. unpublished | |
| *T. thymallus* | 76 | **Da30** | JN796438.1 | Meraner & Gandolfi, 2012 | |
| *T. thymallus* | 77 | **Da30cs** | JX099342.1 | Marić et al. unpublished | |
| *T. thymallus* | 77 | **Da31cs** | JX099343.1 | Marić et al. unpublished | collapses to Da30cs |
| *T. thymallus* | 78 | **Da32** | JN796440.1 | Meraner & Gandolfi, 2012 | |
| *T. thymallus* | 79 | **Da32cs** | JX099345.1 | Marić et al. unpublished | |
| *T. thymallus* | 80 | **Da33cs** | JX099346.1 | Marić et al. unpublished | |
| *T. thymallus* | 81 | **Da35** | JN796443.1 | Meraner & Gandolfi, 2012 | includes sample BayFi06849 |
| *T. thymallus* | 82 | **Da36** | JX524179.1 | Meraner et al. 2013 | |
| *T. thymallus* | 83 | **Eng10** | | unpublished | |

| | | | | | |
|---|---|---|---|---|---|
| *T. thymallus* | 84 | **Eng12** | | unpublished | includes sample EngA and Thy26 |
| *T. thymallus* | 85 | **FEI3_48** | | unpublished | |
| *T. thymallus* | 86 | **Ht1** | | unpublished | |
| *T. thymallus* | 87 | **Ht2** | | unpublished | |
| *T. thymallus* | 88 | **Ht4** | | unpublished | includes sample OMB14 |
| *T. thymallus* | 89 | **Ht5** | | unpublished | |
| *T. thymallus* | 90 | **Ht6** | | unpublished | |
| *T. thymallus* | 91 | **Ht7** | | unpublished | |
| *T. thymallus* | 92 | **Ht8** | | unpublished | |
| *T. thymallus* | 93 | **Ht9** | | unpublished | |
| *T. thymallus* | 94 | **Ht10** | | unpublished | includes Ht18 |
| *T. thymallus* | 94 | Ht18 | | unpublished | collapses to Ht10 |
| *T. thymallus* | 95 | **Ht11** | | unpublished | |
| *T. thymallus* | 96 | **Ht12** | | unpublished | |
| *T. thymallus* | 97 | **Ht13** | | unpublished | |
| *T. thymallus* | 98 | **Ht14** | | unpublished | |
| *T. thymallus* | 99 | **Ht15** | | unpublished | |
| *T. thymallus* | 100 | **Ht16** | | unpublished | |
| *T. thymallus* | 101 | **Ht17** | | unpublished | |
| *T. thymallus* | 101 | Ht19 | | unpublished | collapses to Ht17 |
| *T. thymallus* | 102 | **Ht20** | | unpublished | includes Ola23 and sample OMB17 |
| *T. thymallus* | 104 | **Ht22** | | unpublished | |
| *T. thymallus* | 105 | **Ht23** | | unpublished | includes Olx2 and samples OMB8 and S20 |
| *T. thymallus* | 109 | **Ht28** | | unpublished | |
| *T. thymallus* | 112 | **La14** | | unpublished | |
| *T. thymallus* | 113 | **LEO03** | | unpublished | |
| *T. thymallus* | 114 | **I05** | | unpublished | |
| *T. thymallus* | 115 | **Ibr1** | | unpublished | includes sample KOLPA25 |
| *T. thymallus* | 116 | **RUS14** | | unpublished | includes sample VIS23 |
| *T. thymallus* | 117 | **Rh4** | AF522449.1 | Weiss et al. 2002 | includes Rh7 and samples Mes41 and OMB10 |
| *T. thymallus* | 117 | Rh7 | AF522452.1 | Weiss et al. 2002 | collapses to Rh4 |

| | | | | | |
|---|---|---|---|---|---|
| *T. thymallus* | 118 | **Rh1** | AF522446.1 | Weiss et al. 2002 | includes Rh2 and Rh3 and samples OMB08 and OMB09 |
| *T. thymallus* | 118 | Rh2 | AF522447.1 | Weiss et al. 2002 | collapses to Rh1 |
| *T. thymallus* | 118 | Rh3 | AF522448.1 | Weiss et al. 2002 | collapses to Rh1 |
| *T. thymallus* | 119 | **OEN33** | | unpublished | |
| *T. thymallus* | 120 | **OEN34** | | unpublished | includes sample GR113 |
| *T. thymallus* | 121 | **RAA01** | | unpublished | |
| *T. thymallus* | 122 | **Rh5** | AF522450.1 | Weiss et al. 2002 | |
| *T. thymallus* | 123 | **Rh6** | AF522451.1 | Weiss et al. 2002 | |
| *T. thymallus* | 124 | **SCH001** | | unpublished | |
| *T. thymallus* | 125 | **Thy27** | | unpublished | |

**Table S4.** Estimates of Evolutionary Divergence over Sequence Pairs between Groups (based on mtDNA control region sequences).

| Group 1 | Group 2 | Dist | Std. Err |
|---|---|---|---|
| Adriatic | Loire | 0.0249 | 0.0035 |
| Adriatic | Scandinavia | 0.0246 | 0.0039 |
| North-Eastern Europe | Tisza | 0.0235 | 0.0047 |
| Adriatic | North-Eastern Europe | 0.0232 | 0.0039 |
| Scandinavia | Tisza | 0.0232 | 0.0045 |
| Adriatic | Tisza | 0.0223 | 0.0035 |
| Loire | North-Eastern Europe | 0.0215 | 0.0044 |
| Loire | Scandinavia | 0.0214 | 0.0043 |
| Adriatic | Western Balkans | 0.0211 | 0.0033 |
| Adriatic | Southern Alps | 0.0204 | 0.0028 |
| Northern Alps | Adriatic | 0.0202 | 0.0028 |
| Adriatic | Elbe | 0.0200 | 0.0029 |
| Adriatic | Lafnitz | 0.0200 | 0.0029 |
| Adriatic | Rhône | 0.0194 | 0.0029 |
| Loire | Western Balkans | 0.0193 | 0.0039 |
| Adriatic | Rhine | 0.0192 | 0.0029 |
| Loire | Southern Alps | 0.0192 | 0.0032 |
| Loire | Tisza | 0.0190 | 0.0041 |
| Scandinavia | Elbe | 0.0189 | 0.0034 |
| Western Balkans | Tisza | 0.0185 | 0.0038 |
| Loire | Lafnitz | 0.0182 | 0.0032 |
| Loire | Elbe | 0.0180 | 0.0031 |
| Northern Alps | Loire | 0.0180 | 0.0030 |
| Scandinavia | Southern Alps | 0.0176 | 0.0031 |
| Elbe | North-Eastern Europe | 0.0175 | 0.0034 |
| Loire | Rhône | 0.0174 | 0.0033 |
| Southern Alps | North-Eastern Europe | 0.0173 | 0.0033 |
| Northern Alps | Scandinavia | 0.0173 | 0.0030 |
| Scandinavia | Lafnitz | 0.0172 | 0.0031 |
| Northern Alps | North-Eastern Europe | 0.0170 | 0.0032 |
| Loire | Rhine | 0.0165 | 0.0033 |
| Lafnitz | North-Eastern Europe | 0.0163 | 0.0033 |
| Elbe | Tisza | 0.0157 | 0.0030 |
| Lafnitz | Tisza | 0.0156 | 0.0029 |
| Southern Alps | Tisza | 0.0156 | 0.0028 |
| Scandinavia | Rhône | 0.0155 | 0.0032 |
| Northern Alps | Tisza | 0.0154 | 0.0028 |
| Rhine | Tisza | 0.0145 | 0.0031 |
| Rhône | Tisza | 0.0144 | 0.0031 |
| Scandinavia | Rhine | 0.0143 | 0.0030 |
| Rhône | North-Eastern Europe | 0.0142 | 0.0032 |
| Southern Alps | Western Balkans | 0.0137 | 0.0023 |
| Northern Alps | Western Balkans | 0.0131 | 0.0023 |
| Elbe | Western Balkans | 0.0131 | 0.0025 |

| | | | |
|---|---|---|---|
| Southern Alps | Rhône | 0.0131 | 0.0021 |
| Scandinavia | Western Balkans | 0.0130 | 0.0029 |
| Rhine | North-Eastern Europe | 0.0129 | 0.0030 |
| Southern Alps | Rhine | 0.0129 | 0.0021 |
| North-Eastern Europe | Western Balkans | 0.0125 | 0.0030 |
| Northern Alps | Southern Alps | 0.0121 | 0.0018 |
| Southern Alps | Lafnitz | 0.0120 | 0.0018 |
| Southern Alps | Elbe | 0.0120 | 0.0019 |
| Northern Alps | Rhône | 0.0117 | 0.0018 |
| Rhône | Western Balkans | 0.0115 | 0.0024 |
| Northern Alps | Rhine | 0.0114 | 0.0018 |
| Lafnitz | Western Balkans | 0.0114 | 0.0021 |
| Rhine | Western Balkans | 0.0109 | 0.0024 |
| Rhône | Elbe | 0.0109 | 0.0019 |
| Rhine | Elbe | 0.0108 | 0.0018 |
| Northern Alps | Lafnitz | 0.0107 | 0.0016 |
| Lafnitz | Rhône | 0.0105 | 0.0017 |
| Rhine | Lafnitz | 0.0104 | 0.0018 |
| Northern Alps | Elbe | 0.0104 | 0.0016 |
| Lafnitz | Elbe | 0.0092 | 0.0017 |
| Scandinavia | North-Eastern Europe | 0.0082 | 0.0024 |
| Rhine | Rhône | 0.0079 | 0.0014 |

**Supporting Information 2**

**Comments on fluvial deposits described in Rückert-Ülkümen and Kaya (1993), p. 62**

The authors characterized the fossiliferous bed as fluviatile and possibly resedimented (Rückert-Ülkümen & Kaya, 1993, p. 62). Fluvial deposits of this type were possible in several regression periods from the late Khersonian (approx. 8 Ma, see Krézsek & Olariu, 2021) up to the onset of the Messinian Salinity Crisis in the Eastern Paratethys (5.95 Ma, see Krijgsman et al., 2010). "Sarmatium-Pannonium" in the sense of Rückert-Ülkümen & Kaya (1993) are chronostatigraphical units of the Central Paratethys and may cover several regional strata in the Eastern Paratethys, ranging from Volhynian even up to the base of Pontian (Odessian in the sense of Krijgsman et al. (2010)), including Khersonian and Maeotian (13.7–6.8 Ma according to Papp (1985); up to 6.05 Ma according to Krijgsman et al. (2010)).

**mtDNA control region (CR) sequencing: technical details and information on sample choice**

While the entire CR (approx. 1,082 bp) was targeted, difficulties involving an 82 bp repeat region in the second half (3'-end) were encountered. To overcome this problem, the first half of the CR (5'-end) was amplified with LRBT-25 and INT-5` (5'-ATATAAGAGAACGCCCGGCT-3'), and the second half (3'-end) with a newly designed primer CRint3F (5`-GAAACCACTCACTGAAAGCCG-3') and HN20 (5'-GTGTTATGCTTTAGTTAAGC-3') (Bernatchez & Danzmann, 1993). In some cases, three different PCR product were generated, the first involving LRBT-25 and INT-5`, a second involving the newly designed primers CRint3F and CRI_int1R (5`-ACTTCCTGGTTTAGGGGTTTGAC-3`) and a third using INT-3` (5'-TCCTTGTTTTCCTGTCAAACC-3') and CRII_int2R (5`-TGAGTTTCCTTGGGGGTGTG-3`). This difficulty has been reported in several other species of salmonid genera, such as *Coregonus* and *Hucho*, and was also mentioned for *Thymallus* in Weiss et al. (2002) and Gum et. al (2005). Populations containing an unknown number of the 82 bp repeat corresponded primarily to the Danube River and the Rhine River drainage of Switzerland.

**Mitogenome assembly: technical details**

The five assemblers did not work equally well for all samples. While MITObim, GetOrganelle and Norgal produced assemblies for every sample, NOVOplasty failed for one (Ibr1) and Mitoflex for three samples (Ibr1, ORB73 and RUS14). Annotation results varied between the assemblers, i.e., the number of recovered genes, rRNAs and tRNAs as well as the overall length (number of bp). The most comprehensive and homogenous results for the dataset were produced by MITObim and NOVOplasty. Assemblies with NOVOplasty generally showed lower deviation in overall length and were thus selected for downstream analyses. Only for Ibr1 the MITObim assembly was chosen, as this sample did not produce a NOVOplasty assembly. For all ten samples, 13 out of 13 mitochondrial protein coding genes, two out of two rRNAs and 22 out of 22 tRNAs were recovered. Genome length ranged from 16,657 to 16,712 bp.

**The time calibrated phylogeny in Weiss et al. (2021)**

The time calibrated phylogeny in Weiss et al. (2021) contains a misinterpretation of Wilson & Li (1999) in relation to the classification of †*Eosalmo driftwoodensis* (basal to the subfamily Salmoninae vs. basal to the family Salmonidae). The calibration scheme included a first calibration point set at 50 Ma for the most recent common ancestor (MRCA) of Salmonidae, using the fossil †*Eosalmo driftwoodensis*, another point set at 0.13 Ma for the expansion of *T. baicalensis* into Lake Baikal, and a third point set at 7.6 Ma as stem of the European lineages, based on the fossil remains of †*Thymallus latisulcatus*. However, upon revisiting the bibliography in support of the calibration point of *Eosalmo*, the fossil prior would be better placed as stem only of the subfamily Salmoninae instead of the MRCA of all Salmonidae (Wilson & Li, 1999). In order to verify the implications of the replacement of this contentious fossil calibration point, the analyses was re-run as reported in Weiss et al. (2021). Most of the topology remained the same, with the exception of the subfamily grouping (Figure S2). In this new topology, Salmoninae appeared as the basal lineage to Thymallinae + Coregoninae, in agreement with (Macqueen & Johnston, 2014). Some node ages also became slightly older which was particularly important for the split of Asian and European *Thymallus* (7.8–10.8 Ma vs. 6.2–9.0 Ma) and thus affecting the time of colonisation of Europe. Most significantly, the age of the first split among European *Thymallus* became older (4.0–6.0 Ma vs. 3.0–4.8 Ma) possibly suggesting a pre-Messinian colonisation of the Danube River region.

**References in Supporting Information 2**

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(5), 1002–1014. https://doi.org/10.1093/oxfordjournals.molbev.a040062

Duftner N, Koblmüller S, Weiss S, Medgyesy N, Sturmbauer C (2005). The impact of stocking on the genetic structure of European grayling *Thymallus thymallus*, Salmonidae) in two alpine rivers. Hydrobiologia, 542, 121–129. https://doi.org/10.1007/s10750-004-4951-7

Gum B, Gross R, Kuehn R (2005). Mitochondrial and nuclear DNA phylogeography of European grayling (*Thymallus thymallus*): evidence for secondary contact zones in central Europe. Molecular Ecology, 14, 1707–1725. https://doi.org/10.1111/j.1365-294X.2005.02520.x

Krézsek C, Olariu C (2021). Filling of sedimentary basins and the birth of large rivers: The lower Danube network in the Dacian Basin, Romania. Global and Planetary Change, 197, 103391. https://doi.org/10.1016/j.gloplacha.2020.103391

Krijgsman W, Stoica M, Vasiliev I, Popov VV (2010). Rise and fall of the Paratethys Sea during the Messinian Salinity Crisis. Earth and Planetary Science Letters, 290, 183–191. https://doi.org/10.1016/j.epsl.2009.12.020

Macqueen DJ, Johnston IA (2014). A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. Proceedings of the Royal Society B: Biological Sciences, 281, 20132881. https://doi.org/10.1098/rspb.2013.2881

Rückert-Ülkümen N, Kaya O (1993). Neue Beiträge zur Tertiär-Stratigraphie und Otolithenfauna der Umgebung von Istanbul (Küçükçekmece-Büyükçekmece See), Türkei. Mitteilungen der Bayerischen Staatssammlung für Paläontologie und Historische Geologie, 33, 51–89 [in German].

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 mitochondrial DNA control region. Molecular Ecology, 11(8), 1393–1407. https://doi.org/10.1046/j.1365-294X.2002.01544.x

Weiss S, Gonçalves DV, Secci-Petretto G, Englmaier GK, Gomes-Dos-Santos A, Denys GPJ, Persat H, Antonov A, Hahn C, Taylor EB, Froufe E (2021). Global systematic diversity, range distributions, conservation and taxonomic assessments of graylings (Teleostei: Salmonidae; *Thymallus* spp.). Organisms Diversity & Evolution, 21, 25–42. https://doi.org/10.1007/s13127-020-00468-7

Wilson MVH, Li G-Q (1999). Osteology and systematic position of the Eocene salmonid †*Eosalmo driftwoodensis* Wilson from western North America. Zoological Journal of the Linnean Society, 125(3), 279–311, https://doi.org/10.1111/j.1096-3642.1999.tb00594.x

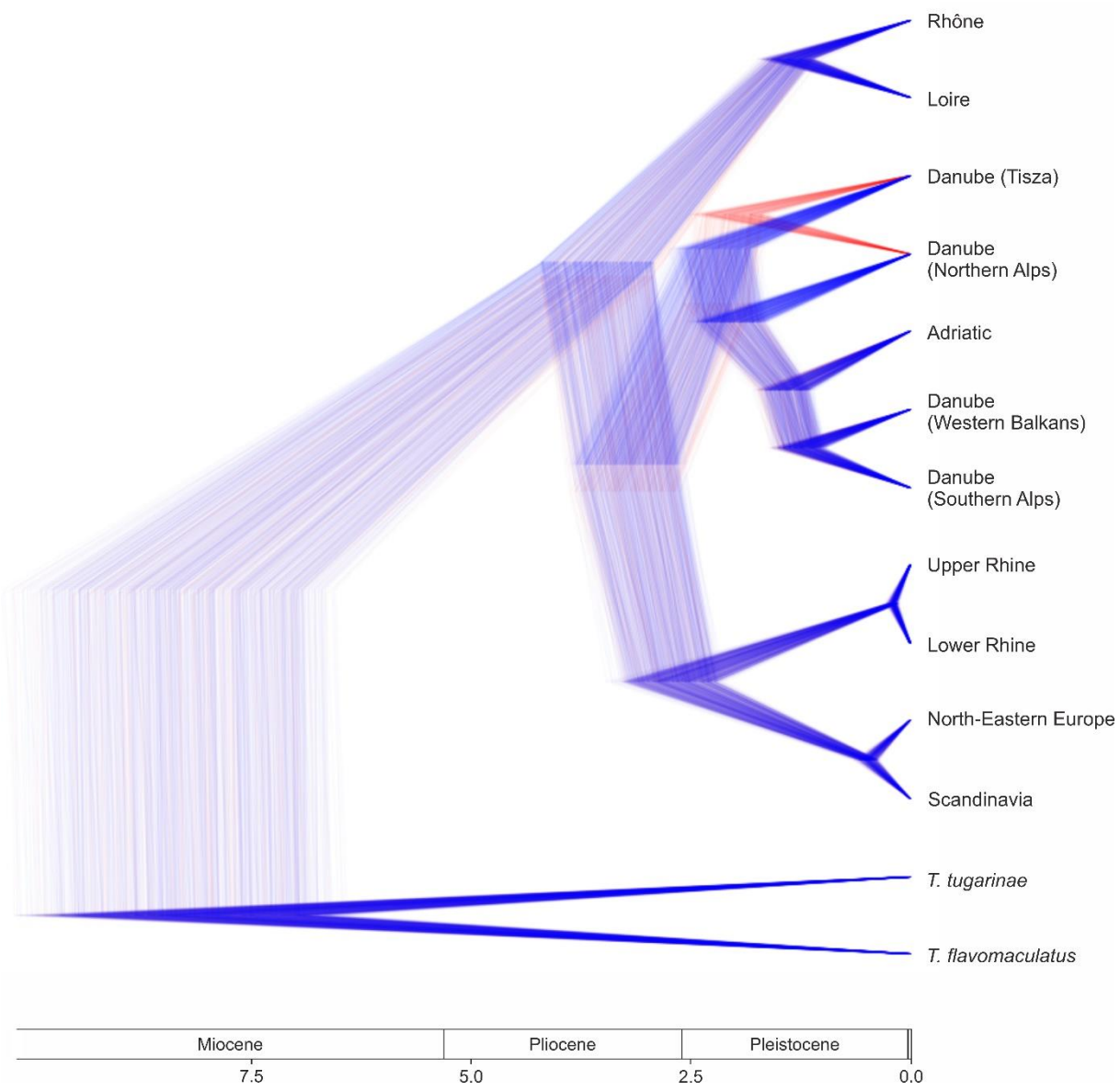



**Figure S1.** Time-calibrated SNAPP phylogeny showing the alternative topology with an early split of the Tisza from all other Danubian lineages. The blue colour indicating the most common topology; the red colour indicating the second most common topology.

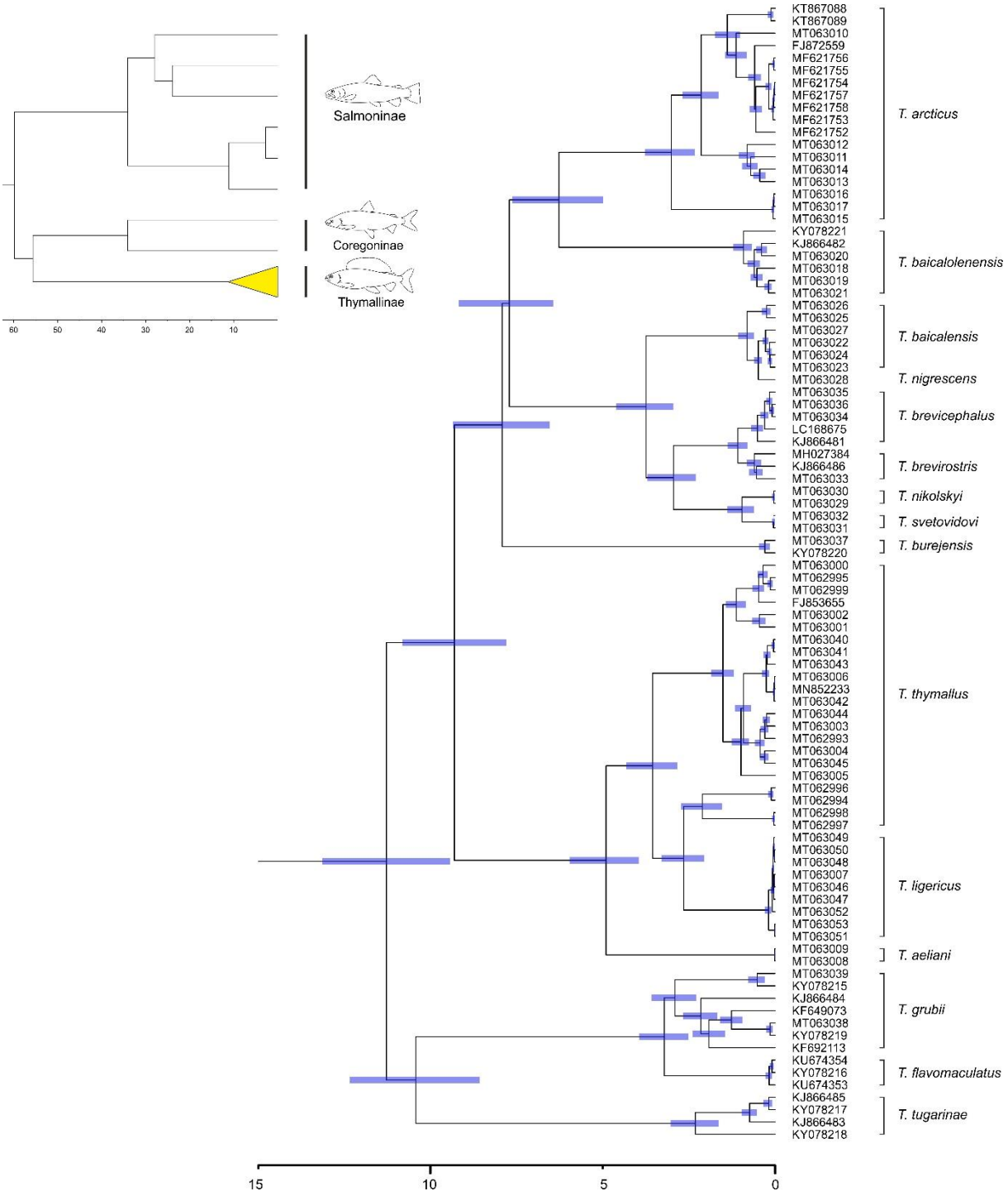



**Figure S2.** Time-calibrated mitogenome phylogeny corrected from Weiss et al. (2021) with the fossil calibration point of †*Eosalmo driftwoodensis* set as stem of the subfamily Salmoninae.

## References

Abell R, Thieme ML, Revenga C, Bryer M, Kottelat M, Bogutskaya N, Coad B, Mandrak N, Balderas SC, Bussing W, Stiassny MLJ, Skelton P, Allen GR, Unmack P, Naseka A, Ng R, Sindorf N, Robertson J, Armijo E, Higgins JV, Heibel TJ, Wikramanayake E, Olson D, López HL, Reis RE, Lundberg JG, Sabaj Pérez MH, Petry P (2008). Freshwater Ecoregions of the World: A New Map of Biogeographic Units for Freshwater Biodiversity Conservation. BioScience, 58(5), 403–414. https://doi.org/10.1641/B580507

Agha M, Ennen JR, Nowakowski AJ, Lovich JE, Sweat SC, Todd BD (2017). Macroecological patterns of sexual size dimorphism in turtles of the world. Journal of Evolutionary Biology, 31, 336–345. https://doi.org/10.1111/jeb.13223

Alekseyev SS, Samusenok VP, Matveev AN, Andreev RS, Samusenok IV, Vokin AI, Yur'ev AL (2021). Finding of Arctic Charr *Salvelinus alpinus* (Salmonidae) and Sculpin *Cottus* cf. *poecilopus* (Cottidae) in the Verkhnyaya Angara basin Near the Divide with the Lena River System: New Data About the Invasion Routes of the Representatives of Lena Ichthyofauna. Journal of Ichthyology, 61, 396–406. https://doi.org/10.1134/S0032945221030024

Allendorf FW, Thorgaard GH (1984). Tetraploidy and the Evolution of Salmonid Fishes. In Evolutionary Genetics of Fishes (pp. 1–53). Springer US. https://doi.org/10.1007/978-1-4684-4652-4_1

Al-Nakeeb K, Petersen TN, Sicheritz-Pontén T (2017). Norgal: extraction and de novo assembly of mitochondrial DNA from whole-genome sequencing data. BMC Bioinformatics, 18(1), 1–7. https://doi.org/10.1186/s12859-017-1927-y

Andersson M (1994). Sexual Selection. Princeton University Press, New Jersey. https://doi.org/10.2307/j.ctvs32s1x

Andrews S (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Antonov AL (2004). A new species of grayling *Thymallus burejensis* sp. nov. (Thymallidae) from the Amur Basin. Journal of Ichthyology, 44, 401–411.

Antonov AL (2009). Ichthyofauna of the upper river basin Maya R (southwestern Priokhotye). All-Russian Conference "Readings in Memory of Academician K.V. Simakova". Magadan, November 25-27, 2009. Magadan: Publishing House of the North-East Scientific Center of the Far Eastern Branch of the Russian Academy of Sciences, 153–154. [In Russian]

Antonov AL (2012). Influence of hydrological network reconstruction on formation of habitats for some fish species in the Amur River basin and adjacent territories. Izvestia Russian Geographical Society, 144(6), 30–37. [In Russian]

Antonov AL, Knizhin IB (2011). Distribution, ecological features, and possible trends in development of habitats of graylings (Thymallidae) in the Amur basin. Herald North-Eastern Scientific Center of the Far Eastern Branch of the Russian Academcy of Sciences, 1, 41–48. [In Russian]

Antonov AL, Knizhin IB (2014). Graylings (Thymallidae) form the Amur basin: history of studies and modern concepts about diversity. Amur Zoological Journal, 6(3), 298–307. [In Russian]

Antonov AL, Mikheev IE (2016). New Data about Distribution of Lower Amur Grayling *Thymallus tugarinae* and Baikal-Lena Grayling *Thymallus* sp. (Thymallidae) in the Basins of the Upper Amur River and Zeya River. Journal of Ichthyology, 56(2), 208–216. https://doi.org/10.1134/S0032945216020016

April J, Hanner RH, Mayden RL, Bernatchez L (2013). Metabolic Rate and Climatic Fluctuations Shape Continental Wide Pattern of Genetic Divergence and Biodiversity in Fishes. PLOS ONE, 8(7). https://doi.org/10.1371/journal.pone.0070296

Artamonova VS, Bolotov IN, Vinarski MV, Makhrov AA (2021). Fresh- and Brackish-Water Cold-Tolerant Species of Southern Europe: Migrants from the Paratethys That Colonized the Arctic. Water, 13, 1161. https://doi.org/10.3390/w13091161

Arzhannikov SG, Ivanov AV, Arzhannikova AV, Demonterova EI, Jansen JD, Preusser F, Kamenetsky MB (2018). Catastrophic events in the Quaternary outflow history of Lake Baikal. Earth-Science Reviews, 177, 76–113. https://doi.org/10.1016/J.EARSCIREV.2017.11.011

Asmussen MA, Arnold J, Avise JC (1987). Definition and properties of disequilibrium statistics for associations between nuclear and cytoplasmic genotypes. Genetics, 115(4), 755–768. https://doi.org/10.1093/genetics/115.4.755

Avise JC (1995). Mitochondrial DNA Polymorphism and a Connection Between Genetics and Demography of Relevance to Conservation. Conservation Biology, 9, 686–690. https://doi.org/10.1046/j.1523-1739.1995.09030686.x

Avise JC, Walker D, Burns JM, Janzen DH, Hallwachs W (1999). Species realities and numbers in sexual vertebrates: perspectives from an asexually transmitted genome. Proceedings of the National Academy of Sciences of the United States of America, 96(3), 992–995. https://doi.org/10.1073/pnas.96.3.992

Balakirev ES, Romanov NS, Mikheev PB, Ayala FJ (2016). Complete mitochondrial genome of Siberian taimen, *Hucho taimen* not introgressed by the lenok subspecies, *Brachymystax*

*lenok* and *B. lenok tsinlingensis*. Mitochondrial DNA, 27(2), 815–816. https://doi.org/10.3109/19401736.2014.919455

Balakirev ES, Romanov NS, Ayala FJ (2017). Complete mitochondrial genome of the Kamchatka grayling *Thymallus mertensii* (Salmoniformes, Salmonidae). Mitochondrial DNA Part A: DNA Mapping, Sequencing, and Analysis, 28(1), 135–136. https://doi.org/10.3109/19401736.2015.1111358

Balbuena JA, Míguez-Lozano R, Blasco-Costa I (2013). PACo: A Novel Procrustes Application to Cophylogenetic Analysis. PLOS ONE, 8, e61048. https://doi.org/10.1371/JOURNAL.PONE.0061048

Ballard JWO, Whitlock MC (2004). The incomplete natural history of mitochondria. Molecular Ecology, 13, 729–744. https://doi.org/10.1046/j.1365-294X.2003.02063.x.

Bănărescu P (1992). Zoogeography of Fresh Waters. Volume 2. Distribution and Dispersal of Freshwater Animals in North America and Eurasia. Aula-Verlag Wiesbaden, pp. 520–1091.

Bangs MR, Douglas MR, Brunner PC, Douglaus MC (2020). Reticulate evolution as a management challenge: Patterns of admixture with phylogenetic distance in endemic fishes of western North America. Evolutionary Applications, 13(6), 1400–1419.

Beacham TD (1984). Age and Morphology of Chum Salmon in Southern British Columbia. Transactions of the American Fisheries Society, 113, 727–736. https://doi.org/10.1577/1548-8659(1984)113<727:AAMOCS>2.0.CO;2

Beacham TD, Murray CB (1983). Sexual Dimorphism in the Adipose Fin of Pacific Salmon (*Oncorhynchus*). Canadian Journal of Fisheries and Aquatic Sciences, 40, 2019–2024. https://doi.org/10.1139/f83-231

Beacham TD, Murray CB (1985). Variation in Length and Body Depth of Pink Salmon (*Oncorhynchus gorbuscha*) and Chum Salmon (*O. keta*) in Southern British Columbia. Canadian Journal of Fisheries and Aquatic Sciences, 42, 312–319. https://doi.org/10.1139/f85-040

Beacham TD, Murray CB (1986). Sexual Dimorphism in Length of Upper Jaw and Adipose Fin of Immature and Maturing Pacific Salmon (*Oncorhynchus*). Aquaculture, 58, 269–276. https://doi.org/10.1016/0044-8486(86)90092-X

Beacham TD, Murray CB (1987). Adaptive variation in body sine, age, morphology, egg size, and developmental biology of chum salmon (*Oncorhynchers keta*) in British Columbia. Canadian Journal of Fisheries and Aquatic Sciences, 44, 244–261. https://doi.org/10.1139/f87-034

Beacham TD, Withler RE, Murray CB, Barner LW (1988). Variation in Body Size, Morphology, Egg Size, and Biochemical Genetics of Pink Salmon in British Columbia. Transactions of the American Fisheries Society, 117, 109–126. https://doi.org/10.1577/1548-8659(1988)117<0109:VIBSME>2.3.CO;2

Beauchamp DA (1990). Movements, Habitat Use, and Spawning Strategies of Arctic Grayling in a Subalpine Lake Tributary. Northwest Science, 64(4), 195–207.

Berbel-Filho WM, Ramos TPA, Jacobina UP, Maia DJG, Torres RA, Lima SMQ (2018). Updated checklist and DNA barcode-based species delimitations reveal taxonomic uncertainties among freshwater fishes from the mid-north-eastern Caatinga ecoregion, north-eastern Brazil. Journal of Fish Biology, 93(2), 311–323. https://doi.org/10.1111/jfb.13758

Berger J-P, Reichenbacher B, Becker D, Grimm M, Grimm K, Picot L, Storni A, Pirkenseer C, Derer C, Schaefer A (2005). Paleogeography of the Upper Rhine Graben (URG) and the Swiss Molasse Basin (SMB) from Eocene to Pliocene. International Journal of Earth Sciences, 94, 697–710. https://doi.org/10.1007/s00531-005-0475-2

Berns CM, Adams DC (2012). Becoming Different But Staying Alike: Patterns of Sexual Size and Shape Dimorphism in Bills of Hummingbirds. Evolutionary Biology, 40, 246–260. https://doi.org/10.1007/s11692-012-9206-3

Betancur-R R, Wiley O, Arratia G, Acero A, Bailly N, Miya M, Lecointre G, Ortí G (2017). Phylogenetic classification of bony fishes. BMC Evolutionary Biology, 17, 162. https://doi.org/10.1186/s12862-017-0958-3

Bishop FG (1967). The biology of the Arctic grayling, *Thymallus arcticus* (Pallas), in Great Slave Lake. Master thesis, Department of Zoology, University of Alberta, Edmonton.

Bishop FG (1971). Observations on Spawning Habits and Fecundity of the Arctic Grayling. The Progressive Fish-Culturist, 33, 12–19. https://doi.org/10.1577/1548-8640(1971)33[12:OOSHAF]2.0.CO;2

Bishop P (1995). Drainage rearrangement by river capture, beheading and diversion. Progress in Physical Geography, 19(4), 449–473. https://doi.org/10.1177/030913339501900402

Blomberg SP, Garland Jr T, Ives AR (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution, 57, 717–745. https://doi.org/10.1111/j.0014-3820.2003.tb00285.x

Bochkarev NA, Zuykova EI, Pestryakova LA, Zakharov ES, Romanov VI, Sokolov VV, Politov DV (2018). Siberian Whitefish (*Coregonus lavaretus pidschian*, Coregonidae) from the Anabar River: Morphogenetic Structure of the Population. Russian Journal of Genetics, 54 (9), 1078–1088.

Bohlen J, Šlechtová V, Bogutskaya N, Freyhof J (2006). Across Siberia and over Europe: Phylogenetic relationships of the freshwater fish genus *Rhodeus* in Europe and the phylogenetic position of *R. sericeus* from the River Amur. Molecular Phylogenetics and Evolution, 40(3), 856–865. https://doi.org/10.1016/j.ympev.2006.04.020

Böhme M, Ilg A (2003). fosFARbase, https://www.fosfarbase.org/ [accessed July 2022].

Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bolotov IN, Vikhrev IV, Bespalaya YV, Gofarov MY, Kondakov AV, Konopleva ES, Bolotov NN, Lyubas AA (2016). Multi-locus fossil-calibrated phylogeny, biogeography and a subgeneric revision of the Margaritiferidae (Mollusca: Bivalvia: Unionoida). Molecular Phylogenetics and Evolution, 103, 104–121. https://doi.org/10.1016/j.ympev.2016.07.020

Bonvalot J, Fleury R, Gélard JP (1984). Notice explicative de la feuille Chalon-sur-Saône à 1/50 000. Bureau de Recherches Géologique National Orléans. France. [in French]

Bone Q, Moore RH (2008). Biology of Fishes. Third edition. Taylor & Francis Group, New York, pp. 478.

Boo GH, Hughey JR (2019). Phylogenomics and multigene phylogenies decipher two new cryptic marine algae from California, *Gelidium gabrielsonii* and *G. kathyanniae* (Gelidiales, Rhodophyta). Journal of Phycology, 55(1), 160–172. https://doi.org/10.1111/jpy.12802

Bouckaert RR (2010). DensiTree: making sense of sets of phylogenetic trees. Bioinformatics, 26(10), 1372–1373, https://doi.org/10.1093/bioinformatics/btq110

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLOS Computational Biolog, 10(4), e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Bravničar N, Palandačić A, Sušnik Bajec S, Snoj A (2020). Neotype designation for *Thymallus aeliani* Valenciennes, 1848 from a museum topotype specimen and its affiliation with Adriatic grayling on the basis of mitochondrial DNA. ZooKeys, 999, 165–178. https://doi.org/10.3897/zookeys.999.56636

BRGM (1999). Carte géologique de la France: Montceau-les-Mines à 1/50 000. Bureau de Recherches Géologique National Orléans, 578. France. [in French]

Brown KH, Drew RE, Weber LA, Thorgaard GH (2006). Intraspecific variation in the rainbow trout mitochondrial DNA genome. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, 1(2), 219–226. https://doi.org/10.1016/J.CBD.2005.11.004

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A (2012). Inferring Species Trees Directly from Biallelic Genetic Markers: Bypassing Gene Trees in a Full Coalescent Analysis. Molecular Biology and Evolution, 29(8), 1917–1932. https://doi.org/10.1093/molbev/mss086

Bushnell B (2014). BBMap: A fast, Accurate, Splice-Aware Aligner (No. LBNL-7065E). Lawrence Berkeley National Lab. Berkeley, CA, United States.

Cambell CSM, Dutoit L, King TM, Craw D, Burridge CP, Wallis GP, Waters JM (2022). Genome-wide analysis resolves the radiation of New Zealand's freshwater *Galaxias vulgaris* complex and reveals a candidate species obscured by mitochondrial capture. Diversity and Distribution, 28(10), 2255–2267. https://doi.org/10.1111/ddi.13629

Campbell MA, Buser TJ, Alfaro ME, López JA (2020). Addressing incomplete lineage sorting and paralogy in the inference of uncertain salmonid phylogenetic relationships. PeerJ, 8, e9389. https://doi.org/10.7717/peerj.9389

Casselman SJ, Schulte-Hostedde AI (2004). Reproductive roles predict sexual dimorphism in internal and external morphology of lake whitefish, *Coregonus clupeaformis*. Ecology of Freshwater Fish, 13, 217–222. https://doi.org/10.1111/j.1600-0633.2004.00053.x

Cattanéo F, Winkler K, Grimardias D, Persat H, Weiss S (2011). Caractérisation génétique des populations d'ombre commun (*Thymallus thymallus* L.) de Suisse et France transfrontalière, Applications à la conservation et a la gestion de l'espèce. Hesso Genève (Haute École spécialisée de Suisse occidentale), pp. 76.

Christie M, McNicklec G, Frenche R, Blouin M (2018). Life history variation is maintained by fitness trade-offs and negative frequency-dependent selection. PNAS, 115, 4441–4446. https://doi.org/10.1073/pnas.1801779115

Clement M, Posada D, Crandall KA (2000). TCS: a computer program to estimate gene genealogies. Molecular Ecology, 9, 1657–1659. https://doi.org/10.1046/j.1365-294x.2000.01020.x

Clutton-Brock T (2007). Sexual Selection in Males and Females. Science, 318, 1882–1885. https://doi.org/10.1126/science.1133311

Chessel D, Dufour AB, Thioulouse J (2004). The ade4 package - I: One-table methods. R News, 4(1), 5–10. https://cran.r-project.org/doc/Rnews/

Conard NJ, Kitagawa K, Krönneck P, Böhme M, Münzel SC (2013). The Importance of Fish, Fowl and Small Mammals in the Paleolithic Diet of the Swabian Jura, Southwestern Germany. In JD Clark, JD Speth (Eds.), Zooarchaeology and Modern Human Origins: Human Hunting Behavior during the Late Pleistocene. Vertebrate Paleobiology and Paleoanthropology (pp. 173–190). Springer, London. https://doi.org/10.1007/978-94-007-6766-9_11

Connell JH (1980). Diversity and the Coevolution of Competitors, or the Ghost of Competition Past. Oikos, 35, 131–138. https://doi.org/10.2307/3544421

Cooper IA (2010). Ecology of Sexual Dimorphism and Clinal Variation of Coloration in a Damselfly. American Naturalist,176, 566–572. https://doi.org/10.1086/656491

Corral-Lou A, Perea S, Aparicio E, Doadrio I (2019). Phylogeography and species delineation of the genus *Phoxinus* Rafinesque, 1820 (Actinopterygii: Leuciscidae) in the Iberian Peninsula. Journal of Zoological Systematics and Evolutionary Research, 57(4), 926–941. https://doi.org/10.1111/jzs.12320

Costedoat C, Gilles A (2009). Quaternary Pattern of Freshwater Fishes in Europe: Comparative Phylogeography and Conservation Perspective. Open Conservation Biology Journal, 3, 36–48. https://doi.org/10.2174/1874839200903010036

Coulthard TJ, Van De Wiel MJ (2012). Modelling river history and evolution. Philosophical Transactions of the Royal Society A, 370, 2123–2142. https://doi.org/10.1098/rsta.2011.0597

Crête-Lafrenière A, Weir LK, Bernatchez L (2012). Framing the Salmonidae Family Phylogenetic Portrait: A More Complete Picture from Increased Taxon Sampling. PLOS ONE, 7(10), e46662. https://doi.org/10.1371/journal.pone.0046662

Cucherousset J, Lassus R, Riepe C, Millet P, Santoul F, Arlinghaus R, Buoro M. (2020). Quantitative estimates of freshwater fish stocking practices by recreational angling clubs in France. Fisheries Management and Ecology, 28, 295–304. https://doi.org/10.1111/fme.12471

Culling MA, Janko K, Boron A, Vasil'ev VP, Côté IM, Hewitt GM (2006). European colonization by the spined loach (*Cobitis taenia*) from Ponto-Caspian refugia based on mitochondrial DNA variation. Molecular Ecology, 15(1), 173–190. https://doi.org/10.1111/j.1365-294X.2005.02790.x

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(11), 1740–1754. https://doi.org/10.1139/e81-159

Darwin CR (1871). The descent of man and selection in relation to sex. John Murray, London.

Davidson FA (1935). The development of the secondary sexual characters in the pink salmon (*Oncorhynchus gorbuscha*). Journal of Morphology, 57, 169–183. https://doi.org/10.1002/jmor.1050570110

Dawnay N, Dawnay L, Hughes RN, Cove R, Taylor MI (2011). Substantial genetic structure among stocked and native populations of the European grayling (*Thymallus thymallus*, Salmonidae) in the United Kingdom. Conservation Genetics, 12(3), 731–744. https://doi.org/10.1007/s10592-010-0179-4

Delling B Doardio I (2005). Systematics of the trouts endemic to Moroccan lakes, with description of a new species (Teleostei: Salmonidae). Ichthyological Exploration of Freshwaters, 16, 49–64.

Denys GPJ, Secci-Petretto G, Gomez dos Santos A (2020a). The complete mitochondrial genome of *Thymallus thymallus* (Linnaeus, 1758) (Actinopterygii, Salmonidae) obtained by long range PCRs and double multiplexing. Cybium, 44(2), 91–94.

Denys GP, Dettai A, Persat H, Daszkiewicz P, Hautecoeur M, Keith P (2020b). Revision of *Phoxinus* in France with the description of two new species (Teleostei, Leuciscidae). Cybium, 44, 205–237. https://doi.org/10.26028/cybium/2020-443-003

Dierckxsens N, Mardulyn P, Smits G (2017). NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Research, 45(4), e18. https://doi.org/10.1093/nar/gkw955

de Queiroz K (2007). Species Concepts and Species Delimitation. Systematic Biology, 56(6), 879–886. https://doi.org/10.1080/10635150701701083

De Ré FC, Robe LJ, Wallau GL, Loreto ELS (2017). Inferring the phylogenetic position of the *Drosophila flavopilosa* group: Incongruence within and between mitochondrial and nuclear multilocus datasets. Journal of Zoological Systematics and Evolutionary Research, 55, 208–221. https://doi.org/10.1111/jzs.12170

Dias RM, Tófoli RM, Barbosa da Silva JC, Gomes LC, Agostinho AA (2022). Effects of habitat complexity on trophic interactions of three congeneric fish species. Aquatic Ecology, 56, 877–889. https://doi.org/10.1007/s10452-022-09954-w

Doadrio I, Perea S, Yahyaoui A (2015). Two new species of Atlantic trout (Actinopterygii, Salmonidae) from Morocco. Graellsia, 71, e031. http://dx.doi.org/10.3989/graellsia.2015.v71.142

Domokos M, Neppel F, Somogyi S (2000). Palägeografische Geschicht der Donau und ihres Einzugsgebiets. Hydrologie und Wasserbewirtschaftung, 44(4), 172–183. [in German]

Donzeau M, Chèvremont P, Marteau P (2001). Notice explicative de la feuille Paray-le-Monial à 1/50 000. Bureau de Recherches Géologique National Orléans, 600. France. [in French]

Dray S, Dufour A (2007). The ade4 Package: Implementing the Duality Diagram for Ecologists. Journal of Statistical Software, 22(4), 1–20. https://doi.org/10.18637/jss.v022.i04

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006). Relaxed Phylogenetics and Dating with Confidence. PLOS Biology, 4(5), e88. https://doi.org/10.1371/journal.pbio.0040088

Drummond AJ, Rambaut A (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology, 7(1), 214. https://doi.org/10.1186/1471-2148-7-214

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution, 29(8), 1969–1973. https://doi.org/10.1093/molbev/mss075

Dubut V, Fouquet A, Voisin A, Costedoat C, Chappaz R, Gilles A (2012). From Late Miocene to Holocene: Processes of Differentiation within the *Telestes* Genus (Actinopterygii: Cyprinidae). PLOS ONE, 7(3), e34423. https://doi.org/10.1371/journal.pone.0034423

Duftner N, Koblmüller S, Weiss S, Medgyesy N, Sturmbauer C (2005). The impact of stocking on the genetic structure of European grayling *Thymallus thymallus*, Salmonidae) in two alpine rivers. Hydrobiologia, 542, 121–129. https://doi.org/10.1007/s10750-004-4951-7

Durand EY, Patterson N, Reich D, Slatkin M (2011). Testing for ancient admixture between closely related populations. Molecular Biology and Evolution, 28, 2239–2252. https://doi.org/10.1093/molbev/msr048

Dyldin Y, Hanel V, Romanov V, Plesnik J (2017). A review of the fish genus *Thymallus* (Pisces: Salmoniformes, Salmonidae, Thymallinae) with taxonomic notes. Bulletin Lampetra, ZO ČSOP Vlašim, 8, 103–126.

Eaton DAR, Overcast I (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. Bioinformatics, 36(8) 2592–2594. https://doi.org/10.1093/bioinformatics/btz966

Englbrecht CC, Freyhof J, Nolte A, Rassmann K, Schliewen U, Tautz D (2000). Phylogeography of the bullhead *Cottus gobio* (Pisces: Teleostei: Cottidae) suggests a pre-Pleistocene origin of the major central European populations. Molecular Ecology, 9(6), 709–722. https://doi.org/10.1046/j.1365-294x.2000.00912.x

Englmaier GK, Antonov A, Weiss SJ (2022). General patterns of sexual dimorphism in graylings (*Thymallus*), with a comparison to other salmonid species. Reviews in Fish Biology and Fisheries, 32, 645–667. https://doi.org/10.1007/s11160-021-09694-4

ESRI (2019). ArcGIS Desktop: Release 10.7. Redlands, CA: Environmental Systems Research Institute.

Esteve M (2005). Observations of spawning behaviour in Salmoninae: *Salmo, Oncorhynchus* and *Salvelinus*. Reviews in Fish Biology and Fisheries, 15, 1–21. https://doi.org/10.1007/s11160-005-7434-7

Esteve M, McLennan DA (2008). Spawning Behavior of Lenok, *Brachymystax lenok* (Salmoniformes) from the Uur River, Northern Mongolia. Journal of Ichthyology, 48, 1031–1036. https://doi.org/10.1134/S0032945208110040

Esteve M, McLennan DA, Kawahara M (2009a). Spawning behaviour of Sakhalin taimen, *Parahucho perryi*, from northern Hokkaido, Japan. Environmental Biology of Fishes, 85, 265–273. https://doi.org/10.1007/s10641-009-9495-2

Esteve M, Gilroy D, McLennan DA (2009b). Spawning behaviour of taimen (*Hucho taimen*) from the Uur River, Northern Mongolia. Environmental Biology of Fishes, 84, 185–189. https://doi.org/10.1007/s10641-008-9407-x

Fabricius E, Gustafson K-J (1955). Observations on the spawning behaviour of the grayling, *Thymallus thymallus* (L.). Report / Institute of Freshwater Research Drottningholm, 36, 75–103.

Fausch KD, White RJ (1981). Competition between brook trout (*Salvelinus fontinalis)* and brown trout (*S. trutta*) for positions in a Michigan stream. Canadian Journal of Fisheries and Aquatic Sciences, 38, 1220–1227. https://doi.org/10.1139/f81-164

Feller AF, Haesler MP, Peichel CL, Seehausen O (2020). Genetic architecture of a key reproductive isolation trait differs between sympatric and non-sympatric sister species of Lake Victoria cichlids. Proceedings of the Royal Society B: Biological Sciences, 287(1924), 20200270. https://doi.org/10.1098/rspb.2020.0270

Fleming IA (1998). Pattern and variability in the breeding system of Atlantic salmon (*Salmo salar*), with comparisons to other salmonids. Canadian Journal of Fisheries and Aquatic Sciences, 55, 59–76. https://doi.org/10.1139/d98-009

Fleming IA, Gross MR (1994). Breeding Competition in a Pacific Salmon (Coho: *Oncorhynchus kisutch*): Measures of Natural and Sexual Selection. Evolution, 48, 637–657. https://doi.org/10.2307/2410475

Fleming IA, Reynolds JD (2004). Salmonid breeding systems. In: Hendry AP, Stearns SC (eds) Evolution illuminated: *Salmon* and their relatives. Oxford University Press, Oxford, pp 264–294.

Folio DM, Gil J, Caudron A, Labonne J (2021). Genotype-by-environment interactions drive the maintenance of genetic variation in a *Salmo trutta* L. hybrid zone. Evolutionary Applications, 14(11), 2698–2711. https://doi.org/10.1111/eva.13307

Free CM, Jensen OP, Mendsaikhan B (2015). A mixed-method approach for quantifying illegal fishing and its impact on an endangered fish species. PLOS ONE, 10(12), 1–17. https://doi.org/10.1371/journal.pone.0143960

Freyhof J (2011). *Thymallus thymallus*. The IUCN Red List of Threatened Species 2011: e.T21875A9333742. https://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T21875A9333742.en. [accessed August 2022]

Fricke R, Eschmeyer WN, Van der Laan R (eds) (2020). Eschmeyer's catalog of fishes: genera, species, references. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp [electronic version accessed 15 February 2020]

Fricke R, Eschmeyer WN, Van der Laan R (eds) (2021). Eschmeyer's catalog of fishes: genera, species, references. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp [electronic version accessed 15 September 2021]

Fricke R, Eschmeyer WN, Van der Laan R (eds) (2022). Eschmeyer's catalog of fishes: genera, species, references. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp [electronic version accessed 23 December 2022]

Froufe E, Alekseyev S, Knizhin I, Alexandrino P, Weiss S (2003a). Comparative phylogeography of salmonid fishes (Salmonidae) reveals late to post-Pleistocene exchange between three now-disjunct river basins in Siberia. Diversity and Distributions, 9(4), 269–282. https://doi.org/10.1046/j.1472-4642.2003.00024.x

Froufe E, Knizhin I, Koskinen MT, Primmer CR, Weiss S (2003b). Identification of reproductively isolated lineages of Amur grayling (*Thymallus grubii* Dybowski 1869): Concordance between phenotypic and genetic variation. Molecular Ecology, 12, 2345–2355. https://doi.org/10.1046/j.1365-294X.2003.01901.x

Froufe E, Knizhin I, Weiss S (2005). Phylogenetic analysis of the genus *Thymallus* (grayling) based on mtDNA control region and ATPase 6 genes, with inferences on control region

constraints and broad-scale Eurasian phylogeography. Molecular Phylogenetics and Evolution, 34, 106–117. https://doi.org/10.1016/j.ympev.2004.09.009

Froufe E, Prié V, Faria J, Ghamizi M, Gonçalves DV, Gürlek ME, Karaouzas I, Kebapçi Ü, Şereflişan H, Sobral C, Sousa R, Teixeira A, Varandas S, Zogaris S, Lopes-Lima M (2016). Phylogeny, phylogeography, and evolution in the Mediterranean region: News from a freshwater mussel (*Potomida*, Unionida). Molecular Phylogenetics and Evolution, 100, 322–332. https://doi.org/10.1016/j.ympev.2016.04.030

Gernhard T (2008). The conditioned reconstructed process. Journal of Theoretical Biology, 253(4), 769–778. https://doi.org/10.1016/J.JTBI.2008.04.005

Giamboni M, Ustaszewski K, Schmid SM, Schumacher ME, Wetzel A (2004). Plio-Pleistocene transpressional reactivation of Paleozoic and Paleogene structures in the Rhine-Bresse transform zone (northern Switzerland and eastern France). International Journal of Earth Sciences, 93, 207–223. https://doi.org/10.1007/s00531-003-0375-2

Gilles A, Costedoat C, Barascud B, Voisin A, Banarescu P, Bianco PG, Economidis PS, Marić D, Chappaz R (2010). Speciation pattern of *Telestes souffia* complex (Teleostei, Cyprinidae) in Europe using morphological and molecular markers. Zoologica Scripta, 39(3), 225–242. https://doi.org/10.1111/j.1463-6409.2010.00417.x

Goble CW, Auer NA, Huckins CJ, Danhoff BM, Holtgren JM, Ogren SA (2018). Fish distributions and habitat associations in Manistee River, Michigan, Tributaries: Implications for Arctic Grayling Restoration. North American Journal of Fisheries Management, 38, 469–486. https://doi.org/10.1002/nafm.10049

Gomes-dos-Santos A, Froufe E, Gonçalves DV, Sousa R, Prié V, Ghamizi M, Benaissa H, Varandas S, Teixeira A, Lopes-Lima M (2019). Freshwater conservation assessments in (semi-) arid regions: Testing river intermittence and buffer strategies using freshwater mussels (Bivalvia, Unionida) in Morocco. Biological Conservation, 236, 420–434. https://doi.org/10.1016/j.biocon.2019.05.038

Good JM, Vanderpool D, Keeble S, Bi K (2015). Negligible nuclear introgression despite complete mitochondrial capture between two species of chipmunks. Evolution, 69, 1961–1972. https://doi.org/10.1111/evo.12712

Goodier SAM, Cotterill FPD, O'Ryan C, Skelton PH, de Wit MJ, Gemmell NJ (2011). Cryptic diversity of African tigerfish (genus *Hydrocynus*) reveals palaeogeographic signatures of linked Neogene geotec-tonic events. PLOS ONE, 6(12), e28775.

Goudet J (2005). HIERFSTAT, a package for R to compute and test hierarchical Fstatistics. Molecular Ecology Notes, 5, 184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x

Greenwood PH (1984). African cichlids and evolutionary theories. In: Echelle, A.A., Kornfield, I. (Eds.), Evolution of Fish Species Flocks. University of Maine at Orono Press, Orono, pp. 141–154.

Groot C (1996). Salmonid life histories. In: Pennell W, Bruce A (eds), Principles of Salmonid Culture. Elsevier, Amsterdam, pp. 97–230.

Grosswald MG (1998). Late-Weichselian ice sheets in Arctic and Pacific Siberia. Quaternary International, 45–46, 3–18. https://doi.org/10.1016/s1040-6182(97)00002-5

Grosswald MG (2009). Ice sheets in the Russian north and northeast during the last Great Chill. Data on Glaciological Studies, Vol. 106, Moscow, Nauka, pp. 152. [In Russian]

Gruchy CG, Vladykov VD (1968). Sexual Dimorphism in Anal Fin of Brown Trout, *Salmo trutta*, and Close Relatives. Journal of the Fisheries Research Board of Canada, 25, 813–815. https://doi.org/10.1139/f68-073

Gum B, Gross R, Geist J (2009). Conservation genetics and management implications for European grayling, *Thymallus thymallus*: Synthesis of phylogeography and population genetics. Fisheries Management and Ecology, 16(1), 37–51. https://doi.org/10.1111/j.1365-2400.2008.00641.x

Gundappa MK, To TH, Grønvold L, Martin SAM, Lien S, Geist J, Hazlerigg D, Sandve SR, Macqueen DJ (2022). Genome-Wide Reconstruction of Rediploidization Following Autopolyploidization across One Hundred Million Years of Salmonid Evolution. Molecular Biology and Evolution, 39(1), msab310. https://doi.org/10.1093/molbev/msab310

Guo B, Fang B, Shikano T, Momigliano P, Wang C, Kravchenko A, Merilä J (2019). A phylogenomic perspective on diversity, hybridization and evolutionary affinities in the stickleback genus *Pungitius*. Molecular Ecology, 28(17), 4046–4064. https://doi.org/10.1111/mec.15204

Hahn C, Bachmann L, Chevreux B (2013). Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nucleic Acids Research, 41(13), e129–e129. https://doi.org/10.1093/nar/gkt371

Hall TA (1999). BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. Nucleic Acids Symposium Series, 41, 95–98.

Hänfling B, Dümpelmann C, Bogutskaya NB, Brandl R, Brändle M (2009). Shallow phylogeographic structuring of *Vimba vimba* across Europe suggests two distinct refugia during the last glaciation. Journal of Fish Biology, 75(9), 2269–2286. https://doi.org/10.1111/j.1095-8649.2009.02415.x

Hantemirova EV, Heinze B, Knyazeva SG, Musaev AM, Lascoux M, Semerikov VL (2017). A new Eurasian phylogeographical paradigm? Limited contribution of southern populations to the recolonization of high latitude populations in *Juniperus communis* L. (Cupressaceae). Journal of Biogeography, 44(2), 271–282. https://doi.org/10.1111/jbi.12867

Hashemzadeh Segherloo I, Freyhof J, Berrebi P, Ferchaud A-L, Geiger M, Laroche J, Levin BA, Normandeau E, Bernatchez L (2021). A genomic perspective on an old question: *Salmo* trouts or *Salmo trutta* (Teleostei: Salmonidae)? Molecular Phylogenetics and Evolutio, 162, 107204.

Haugen TO, Vøllestad LA, (2001). Population differences in early life-history traits in grayling. Journal of Evolutionary Biology, 13(6), 897–905. https://doi.org/10.1046/j.1420-9101.2000.00242.x

He B, Su T, Wu Y, Xu J, Huang D (2018). Phylogenetic analysis of the mitochondrial genomes in bees (Hymenoptera: Apoidea: Anthophila). PLOS ONE, 13(8), e0202187. https://doi.org/10.1371/journal.pone.0202187

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences of the United States of America, 101(41), 14812–14817. https://doi.org/10.1073/pnas.0406166101

Hedrick AV, Temeles EJ (1989). The evolution of sexual dimorphism in animals: hypotheses and tests. Trends in Ecology & Evolution, 4, 136–138. https://doi.org/10.1016/0169-5347(89)90212-7

Heese T (1987). Morphology of two plastic forms of whitefish (*Coregonus lavaretus* (L.)) occurring in the Pomeranian Bay and Szczecin Lagoon. Acta Ichthyologica Et Piscatoria, 17, 3–28.

Hendry AP, Berg OK (1999). Secondary sexual characters, energy use, senescence, and the cost of reproduction in sockeye salmon. Canadian Journal of Zoology, 77, 1663–1675. https://doi.org/10.1139/z99-158

Hewitt GM (1999). Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society, 68(1-2), 87–112. https://doi.org/10.1111/j.1095-8312.1999.tb01160.x

Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society B: Biological Sciences, 359(1442), 183–195. https://doi.org/10.1098/rstb.2003.1388

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018). UFBoot2: Improving the ultrafast bootstrap approximation. Molecular Biology and Evolution, 35(2), 518–522. https://doi.org/10.1093/molbev/msx281

Horreo JL, Fitze PS (2018). Postglacial Colonization of Northern Europe by Reptiles. In: P. Pontarotti (Ed.), Origin and Evolution of Biodiversity (pp. 197–214). Springer, Cham. https://doi.org/10.1007/978-3-319-95954-2_12

Hubbs CL (1955). Hybridization between fish species in nature. Systematic Zoology, 4, 1–20. https://doi.org/10.2307/sysbio/4.1.1

Hughes NF (1992). Ranking of Feeding Positions by Drift-Feeding Arctic Grayling *Thymallus arcticus* in Dominance Hierarchies. Canadian Journal of Fisheries and Aquatic Sciences, 49, 1994–1998. https://doi.org/10.1139/f92-222

Hughes NF, Dill LM (1990). Position Choice by Drift-Feeding Salmonids: Model and Test for Arctic Grayling (*Thymallus arcticus*) in Subarctic Mountain Streams, Interior Alaska. Canadian Journal of Fisheries and Aquatic Sciences, 47, 2039–2048. https://doi.org/10.1139/f90-228

Hughes LC, Cardoso YP, Sommer JA, Cifuentes R, Cuello M, Somoza GM, González-Castro M, Malabarba LR, Cussac V, Habit EM, Betancur-R R, Ortí G (2020). Biogeography, habitat transitions and hybridization in a radiation of South American silverside fishes revealed by mitochondrial and genomic RAD data. Molecular Ecology, 29, 738–751. https://doi.org/10.1111/mec.15350

Huml VJ, Taylor MI, Harris, WE., Sen R, Ellis JS (2018). Neutral variation does not predict immunogenetic variation in the European grayling (*Thymallus thymallus*) - implications for management. Molecular Ecology, 27(21), 4157–4173. https://doi.org/10.1111/mec.14864

Illies J, (1961). Versuch einer allgemeinen biozönotischen Gliederung der Fließgewässer. Internationale Revue der gesamten Hydrobiologie, 46, 205–213. [in German] https://doi.org/10.1002/iroh.19610460205

International Commission on Zoological Nomenclature (ICZN) (1999). International Code of Zoological Nomenclature 4th Edition: adopted by the International Union of Biological Sciences. The International Trust for Zoological Nomenclature, London, pp. 106.

Isaac NJB, Mallet J, Mace GM (2004). Taxonomic inflation: its influence on macroecology and conservation. Trends in Ecology & Evolution, 19, 464–469. https://doi.org/10.1016/j.tree.2004.06.004

Iwasaki W, Fukunaga T, Isagozawa R, Yamada K, Maeda Y, Satoh T, Mabuchi K, Takeshima H, Miya M, Nishida M (2013). MitoFish and MitoAnnotator: A Mitochondrial Genome

Database of Fish with an Accurate and Automatic Annotation Pipeline. Molecular Biology and Evolution, 30(11), 2531–2540. https://doi.org/10.1093/molbev/mst141

Jacobs A, Carruthers M, Yurchenko A, Gordeeva NV, Alekseyev SS, Hooker O, Leong JS, Minkley DR, Rondeau EB, Koop BF, Adams CE, Elmer KR (2020). Parallelism in eco-morphology and gene expression despite variable evolutionary and genomic backgrounds in a Holarctic fish. PLOS Genetics, 16, e1008658. https://doi.org/10.1371/journal.pgen.1008658

Jacobson PC, Zimmer KD, Grow R, Eshenroder RL (2020). Morphological Variation of Cisco across Gradients of Lake Productivity. Transactions of the American Fisheries Society, 149, 462–473. https://doi.org/10.1002/tafs.10242

Janhunen M, Peuhkuri N, Piironen J (2009). Morphological variability among three geographically distinct Arctic charr (*Salvelinus alpinus* L.) populations reared in a common hatchery environment. Ecology of Freshwater Fish, 18, 106–116. https://doi.org/10.1111/j.1600-0633.2008.00329.x

Järvi T (1990). The Effects of Male Dominance, Secondary Sexual Characteristics and Female Mate Choice on the Mating Success of Male Atlantic Salmon *Salmo salar*. Ethology, 84, 123–132. https://doi.org/10.1111/j.1439-0310.1990.tb00789.x

Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, Li DZ (2020). GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biology, 21(1), 1–31. https://doi.org/10.1186/s13059-020-02154-5

Jombart T (2008). adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics, 24(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart T, Ahmed I (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics, 27(21), 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Johnson DE (1998). Applied Multivariate Methods for Data Analysts. Duxbury Press, University of Michigan.

Johnson SP, Carlson SM, Quinn TP (2006). Tooth size and skin thickness in mature sockeye salmon: evidence for habitat constraints and variable investment between the sexes. Ecology of Freshwater Fish, 15, 331–338. https://doi.org/10.1111/j.1600-0633.2006.00175.x

Jombart T (2008). adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics, 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jonsson B, Jonsson N (2015). Sexual size dimorphism in anadromous brown trout *Salmo trutta*. Journal of Fish Biology, 87, 187–193. https://doi.org/10.1111/jfb.12704

Joseph L (2021). Species limits in birds: Australian perspectives on interrelated challenges of allopatry, introgression of mitochondrial DNA, recent speciation, and selection. Ornithology, 138, 1–15. https://doi.org/10.1093/ORNITHOLOGY/UKAB012

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods, 14, 587–589. https://doi.org/10.1038/NMETH.4285

Kapli P, Lutteropp S, Zhang J, Kobert K, Pavlidis P, Stamatakis A, Flouri T (2017). Multi-rate Poisson Tree Processes for single-locus species delimitation under Maximum Likelihood and Markov Chain Monte Carlo. Bioinformatics, 33, 1630–1638, https://doi.org/10.1093/bioinformatics/btx025

Karjalainen J, Marjomäki TJ (2018). Communal pair spawning behaviour of vendace (*Coregonus albula*) in the dark. Ecology of Freshwater Fish, 27, 542–548. https://doi.org/10.1111/eff.12368

Kasapoglu N, Çankırılıgil EC, Çakmak E, Özel OT (2020). Meristic and morphometric characteristics of the Black Sea salmon, *Salmo labrax* Pallas, 1814 culture line: an endemic species for Eastern Black Sea. Journal of Fisheries, 8, 935–939. http://journal.bdfish.org/index.php/fisheries/article/view/JFish20250

Katoh K, Standley DM (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution, 30(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kaus A, Michalski S, Hänfling B, Karthe D, Borchardt D, Durka W (2019). Fish conservation in the land of steppe and sky: Evolutionarily significant units of threatened salmonid species in Mongolia mirror major river basins. Ecology and Evolution, 9(6), 3416–3433. https://doi.org/10.1002/ece3.4974

Kautt AF, Kratochwil CF, Nater A, Machado-Schiaffino G, Olave M, Henning F, Torres-Dowdall J, Härer A, Hulsey CD, Franchini P, Pippel M, Myers EW, Meyer A (2020). Contrasting signatures of genomic divergence during sympatric speciation. Nature, 588 (7836), 106–111.

Kazyak DC, Hilderbrand RH, Holloway AE (2013). Rapid Visual Assessment to Determine Sex in Brook Trout. North American Journal of Fisheries Management, 33, 665–668. https://doi.org/10.1080/02755947.2013.785998

Kerr KCR, Stoeckle MY, Dove CJ, Weigt LA, Francis CM, Hebert PDN (2007). Comprehensive DNA barcode coverage of North American birds. Molecular Ecology Notes, 7(4), 535–543. https://doi.org/10.1111/j.1471-8286.2007.01670.x

Kirillov AF, Knizhin IB (2014). Ichthyofauna of the Lena River (Laptev Sea Basin): Modern composition and historical formation. Journal of Ichthyology, 54(7), 433–445. https://doi.org/10.1134/S0032945214040031

Kishino H, Hasegawa M (1989). Evaluation of the maximum Likelihood estimate of the evolutionary tree topologies from DNA sequence. Journal of Molecular Evolution, 29, 170–179. https://doi.org/10.1007/BF02100115

Kishino H, Miyata T, Hasegawa M (1990). Maximum Likelihood inference of protein phylogeny and the origin of chloroplasts. Journal of Molecular Evolution, 31, 151–160. https://doi.org/10.1007/BF02109483

Knebelsbeger T, Dunz AR, Neumann D, Geiger MF (2015). Molecular diversity of Germany's freshwater fishes and lampreys assessed by DNA barcoding. Molecular Ecology Resources, 15(3), 562–572. https://doi.org/10.1111/1755-0998.12322

Knizhin IB (2009). Graylings (*Thymallus* Cuvier, 1829) of the Holarctic (Systematics, Phylogeography, and Ecological features) [= Хариусы (*Thymallus* Cuvier, 1829) Голарктики (систематика, филогеография, особенности экологии)]. Dissertation, Severtsov Institute of Ecology and Evolution, Russian Academy of Science, pp. 268. [in Russian]

Knizhin IB, Weiss SJ, Antonov A., Froufe E (2004). Morphological and Genetic Diversity of Amur Graylings (*Thymallus*, Thymallidae). Journal of Ichthyology, 44(1), 52–69.

Knizhin IB, Antonov AL, Weiss SJ (2006a). A new subspecies of the Amur grayling *Thymallus grubii flavomaculatus* ssp. nova (Thymallidae). Journal of Ichthyology, 46(8), 555–562. https://doi.org/10.1134/S0032945206080017

Knizhin IB, Weiss SJ, Sušnik S (2006b). Graylings of Baikal Lake basin (*Thymallus*, Thymallidae): Diversity of forms and their taxonomic status. Journal of Ichthyology, 46(6), 418–435. https://doi.org/10.1134/S0032945206060026

Knizhin IB, Weiss SJ, Bogdanov BE, Samarina SS, Froufe E (2006c). Finding a new form of the grayling *Thymallus arcticus* (Thymallidae) in the basin of Lake Baikal. Journal of Ichthyology, 46, 34–43. https://doi.org/10.1134/S003294520601005X

Knizhin IB, Kirillov AF, Weiss SJ (2006d). On the diversity and taxonomic status of graylings (*Thymallus* Thymallidae) from the Lena River. Journal of Ichthyology, 46, 234–246. https://doi.org/10.1134/S0032945206030039

Knizhin IB, Bogdanov BE, Vasil'eva EA (2006e). Biological and Morphological Characteristic of the Arctic Grayling *Thymallus arcticus* (Thymallidae) from Alpine Lakes of the Basin of the Upper Reaches of the Angara River. Journal of Ichthyology, 46, 709–721. https://doi.org/10.1134/S0032945206090037

Knizhin IB, Antonov AL, Safronov SN, Weiss SJ (2007). New species of grayling *Thymallus tugarinae* sp. nova (Thymallidae) from the Amur River Basin. Journal of Ichthyology, 47(2), 123–139. https://doi.org/10.1134/S0032945207020014

Knizhin IB, Weiss SJ, Bogdanov BE, Kopun T, Muzalevskaya OV (2008a). Graylings (Thymallidae) of water bodies in western Mongolia: Morphological and genetic diversity. Journal of Ichthyology, 48(9), 714–735. https://doi.org/10.1134/S0032945208090038

Knizhin IB, Weiss SJ, Bogdanov BE, Kopun T (2008b). New data on the distribution of the Upper Lena form of grayling (Thymallidae) in the basin of Lake Baikal and its taxonomic status. Journal of Ichthyology, 48(3), 217–223. https://doi.org/10.1134/S003294520803003X

Knizhin IB, Weiss SJ (2009). A new species of grayling *Thymallus svetovidovi* sp. nova (Thymallidae) from the Yenisei basin and its position in the genus *Thymallus*. Journal of Ichthyology, 49(1), 1–9. https://doi.org/10.1134/S0032945209010019

Koblmüller S, Egger B, Sturmbauer C, Sefc KM (2010). Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic Lake Tanganyika cichlid tribe Tropheini. Molecular Phylogenetics and Evolution, 55, 318–334. https://doi.org/10.1016/j.ympev.2009.09.032

Komatsu G, Baker VR, Arzhannikov SG, Gallagher R, Arzhannikova A, Oguchi T (2016). Catastrophic flooding, palaeolakes, and late Quaternary drainage reorganization in northern Eurasia. International Geology Review, 58(14), 1693–1722. https://doi.org/10.1080/00206814.2015.1048314

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I, (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources, 15, 1179–1191. https://doi.org/10.1111/1755-0998.12387

Koseki Y, Maekawa K (2000). Sexual selection on mature male parr of masu salmon (*Oncorhynchus masou*): does sneaking behavior favor small body size and less-developed sexual characters? Behavioral Ecology and Sociobiology, 48, 211–217. https://doi.org/10.1007/s002650000231

Koskinen MT, Ranta E, Piironen J, Veselov A, Titov S, Haugen TO, Nilsson J, Carlstein M, Primmer CR (2000). Genetic lineages and postglacial colonization of grayling (*Thymallus thymallus*, Salmonidae) in Europe, as revealed by mitochondrial DNA analyses. Molecular Ecology, 9, 1609–1624. https://doi.org/10.1046/j.1365-294x.2000.01065.x

Koskinen MT, Knizhin I, Primmer CR, Schlötterer C, Weiss S (2002). Mitochondrial and nuclear DNA phylogeography of *Thymallus* spp. (grayling) provides evidence of ice-age

mediated environmental perturbations in the world's oldest body of fresh water, Lake Baikal. Molecular Ecology, 11(12), 2599–2611. https://doi.org/10.1046/j.1365-294X.2002.01642.x

Kottelat M (2006). Fishes of Mongolia – A check-list of the fishes known to occur in Mongolia with comments on systematics and nomenclature. Environment and Social Development Sector, East Asia and Pacific Region, The World Bank, Washington, pp. 103.

Kovács J, Némethc K, Szabóa P, Kocsisd L, Gábor Kereszturic G, Újvárie G, Vennemann T, (2020). Volcanism and paleoenvironment of the pula maar complex: A pliocene terrestrial fossil site in Central Europe (Hungary). Palaeogeography, Palaeoclimatology, Palaeoecology, 537, 109398. https://doi.org/10.1016/j.palaeo.2019.109398

Kovalchuk OM (2015). The First Appearance of *Hucho* (Salmonidae) in the Fossil Record of Eastern Europe. Vestnik Zoologii, 49(5), 413–420.

Kratt LF, Smith RJF (1978). Breeding Tubercles Occur on Male and Female Arctic Grayling (*Thymallus arcticus*). Copeia, 1:185–188. https://doi.org/10.2307/1443852

Kratt LF, Smith RJF (1979) Agonistic behaviour of age 0, age 1 and non-breeding adult Arctic grayling *Thymallus arcticus* (Pallas). Journal of Fish Biology, 15, 389–404. https://doi.org/10.1111/j.1095-8649.1979.tb03622.x

Kratt LF, Smith RJF (1980). An analysis of the spawning behaviour of the Arctic grayling *Thymallus arcticus* (Pallas) with observations on mating success. Journal of Fish Biology, 17, 661–666. https://doi.org/10.1111/j.1095-8649.1980.tb02799.x

Krézsek C, Olariu C (2021). Filling of sedimentary basins and the birth of large rivers: The lower Danube network in the Dacian Basin, Romania. Global and Planetary Change, 197, 103391. https://doi.org/10.1016/j.gloplacha.2020.103391

Krijgsman W, Stoica M, Vasiliev I, Popov VV (2010). Rise and fall of the Paratethys Sea during the Messinian Salinity Crisis. Earth and Planetary Science Letters, 290, 183–191. https://doi.org/10.1016/j.epsl.2009.12.020

Kucheruk AI, Mruk AI, Kovalenko VO (2015). Morphometric characteristics of European grayling (*Thymallus thymallus* L.) in Transcarpathian rivers. Scientific reports of NULES of Ukraine.

Kumar S, Stecher G, Tamura K (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution, 33(7), 1870–1874. https://doi.org/10.1093/molbev/msw054

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular Biology and Evolution, 35(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Laikre L, Schwartz MK, Waples RS, Ryman N, GeM Working Group (2010). Compromising genetic diversity in the wild: Unmonitored large-scale release of plants and animals. Trends in Ecology & Evolution, 25, 520–529. https://doi.org/10.1016/j.tree.2010.06.013

Lanfear R, Calcott B, Ho SYW, Guindon S (2012). PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. Molecular Biology and Evolution, 29(6), 1695–1701. https://doi.org/10.1093/molbev/mss020

Lazarev S, de Leeuw A, Stoica M, Mandic O, van Baak CGC, Vasiliev I, Krijgsman W (2020). From Khersonian drying to Pontian "flooding": late Miocene stratigraphy and palaeoenvironmental evolution of the Dacian Basin (Eastern Paratethys). Global and Planetary Change, 192, 103224. https://doi.org/10.1016/j.gloplacha.2020.103224

Lecaudey LA, Schliewen UK, Osinov AG, Taylor EB, Bernatchez L, Weiss SJ (2018). Molecular Phylogenetics and Evolution Inferring phylogenetic structure, hybridization and divergence times within Salmoninae (Teleostei: Salmonidae) using RAD-sequencing. Molecular Phylogenetics and Evolution, 124, 82–99. https://doi.org/10.1016/j.ympev.2018.02.022

Lemey P, Rambaut A, Drummond AJ, Suchard MA (2009). Bayesian phylogeography finds its roots. PLOS Computational Biology, 5(9). https://doi.org/10.1371/journal.pcbi.1000520

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. Frontiers in Zoology, 10, 1–14. https://doi.org/10.1186/1742-9994-10-34

Lerceteau-Köhler E, Schliewen U, Kopun T, Weiss S (2013). Genetic variation in brown trout Salmo trutta across the Danube, Rhine, and Elbe headwaters: a failure of the phylogeographic paradigm? BMC Ecology and Evolution, 13, 176. http://www.biomedcentral.com/1471-2148/13/176

Li H (2012). seqtk Toolkit for processing sequences in FASTA/Q formats. Available online at: https://github.com/lh3/seqtk

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics, 25 (16), 2078–2079.

Li YL, Liu JX (2018). StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. Molecular Ecology Resources, 18(1), 176–177. https://doi.org/10.1111/1755-0998.12719

Li JY, Li WX, Wang AT, Zhang Y (2021). MitoFlex: an efficient, high-performance toolkit for animal mitogenome assembly, annotation and visualization. Bioinformatics, 37(18), 3001–3003. https://doi.org/10.1093/bioinformatics/btab111

Liedigk R, Kolleck J, Böker KO, Meijaard E, Md-Zain BM, Abu, M, Abdul-Latiff B, Ampeng A, Lakim M, Abdul-Patah P, Tosi AJ, Brameier M, Zinner D, Roos C (2015). Mitogenomic phylogeny of the common long-tailed macaque (*Macaca fascicularis fascicularis*). BMC Genomics, 16(1), 222. https://doi.org/10.1186/s12864-015-1437-0

Lien S, Koop BF, Sandve SR, Miller JR, Kent MP, Nome T, Hvidsten TR, Leong JS, Minkley DR, Zimin A, Grammes F, Grove H, Gjuvsland A, Walenz B, Hermansen RA, von Schalburg K, Rondeau EB, Di Genova A, Samy JKA, Vik JO, Vigeland MD, Caler L, Grimholt U, Jentoft S, Våge ID, De Jong P, Moen T, Baranski M, Palti Y, Smith DR, Yorke JA, Nederbragt AJ, Tooming-Klunderud A, Jakobsen KS, Jiang X, Fan D, Hu Y, Liberles DA, Vidal R, Iturra P, Jones SJM, Jonassen I, Maass A, Omholt SW, Davidson WS (2016). The Atlantic salmon genome provides insights into rediploidization. Nature, 533, 200–205. https://doi.org/10.1038/nature17164

Linke S, Lehner B, Ouellet Dallaire C, Ariwi J, Grill G, Anand M, Beames P, Burchard-Levine V, Maxwell S, Moidu H, Tan F, Thieme M (2019). Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. Scientific Data, 6(1), 283. https://doi.org/10.1038/s41597-019-0300-6

Liu Y, Li Y, Meng W, Liu L (2016). The complete mitochondrial DNA sequence of Xinjiang arctic grayling *Thymallus arcticus grubei*. Mitochondrial DNA Part B, 1:1, 724–725. https://doi.org/10.1080/23802359.2016.1229589

Lopes-Lima M, Burlakova L, Karatayev A, Gomes-dos-Santos A, Zieritz A, Froufe E, Bogan AE (2019). Revisiting the North American freshwater mussel genus *Quadrula sensu lato* (Bivalvia Unionidae): Phylogeny, taxonomy and species delineation. Zoologica Scripta, 48(3), 313–336. https://doi.org/10.1111/zsc.12344

Ma B, Jiang H, Sun P, Chen J (2015). Complete mitochondrial genome of *Thymallus grubii* (Salmonidae: Thymallinae). Mitochondrial DNA, 26(6), 815–816. https://doi.org/10.3109/19401736.2013.855907

Ma B, Jiang H, Sun P, Chen J, Li L, Zhang X, Yuan L (2016). Phylogeny and dating of divergences within the genus *Thymallus* (Salmonidae: Thymallinae) using complete mitochondrial genomes. Mitochondrial DNA Part A, 27(5), 3602–3611. https://doi.org/10.3109/19401736.2015.1079824

MacGuigan DJ, Near TJ (2019). Phylogenomic signatures of ancient introgression in the rogue lineage of darters (Teleostei: Percidae). Systematic Biology, 68 (2), 329–346. https://doi.org/10.1093/sysbio/syy074

Magreiter H (1951). Der Geschlechtsdimorphismus der Äschen. Österreichs Fischerei 2:28–30. [in German]

Mäkinen H, Papakostas S, Vøllestad LA, Leder EH (2015). Plastic and Evolutionary Gene Expression Responses Are Correlated in European Grayling (*Thymallus thymallus*) Subpopulations Adapted to Different Thermal Environments. Plastic and Evolutionary Gene Expression Responses Are Correlated in European Grayling. Journal of Heredity, 107, 81–89. https://doi.org/10.5061/dryad.7bp03

Malinsky M, Svardal H, Tyers AM, Miska EA, Genner MJ, Turner GF, Durbin R (2018). Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. Nature Ecology & Evolution, 2, 1940–1955. https://doi.org/10.1038/s41559-018-0717-x

Malinsky M, Matschiner M, Svardal H (2021). Dsuite - Fast D-statistics and related admixture evidence from VCF files. Molecular Ecology Resources, 21, 584–595. https://doi.org/10.1111/1755-0998.13265

Mamcarz A, Nowak M (1986). Rearing of coregonid fishes (Coregonidae) in illuminated lake cages: VI. Characteristics of the spawners of whitefish (*Coregonus lavaretus* L.) and peled (*Coregonus peled* Gmel.) from cage culture. Aquaculture, 55, 51–58. https://doi.org/10.1016/0044-8486(86)90055-4

Margold M, Jansen JD, Codilean AT, Preusser F, Gurinov AL, Fujioka T, Fink D (2018). Repeated megafloods from glacial Lake Vitim, Siberia, to the Arctic Ocean over the past 60,000 years. Quaternary Science Reviews, 187, 41–61. https://doi.org/10.1016/J.QUASCIREV.2018.03.005

Marić S, Razpet A, Nikolić V, Simonović P (2011). Genetic differentiation of European grayling (*Thymallus thymallus*) populations in Serbia, based on mitochondrial and nuclear DNA analyses. Genetics Selection Evolution, 43, 2. https://doi.org/10.1186/1297-9686-43-2

Marić S, Kalamujić B, Snoj A, Razpet A, Lukić-Bilela L, Pojskić N, Sušnik Bajec S (2012). Genetic variation of European grayling (*Thymallus thymallus*) populations in the Western Balkans. Hydrobiologia, 691, 225–237. https://doi.org/10.1007/s10750-012-1076-2

Marić S, Askeyev IV, Askeyev OV, Monakhov SP, Bravničar J, Snoj A (2014). Phylogenetic and population genetic analysis of *Thymallus thymallus* (Actinopterygii, Salmonidae) from

the middle Volga and upper Ural drainages. Hydrobiologia, 740, 167–176. https://doi.org/10.1007/s10750-014-1951-0

Mason AJ, Grazziotin FG, Zaher H, Lemmon AR, Lemmon ER, Parkinson CL (2019). Reticulate evolution in nuclear Middle America causes discordance in the phylogeny of palm-pitvipters (Viperidae: *Bothriechis*). Journal of. Biogeography, 46 (5), 833–844. https://doi.org/10.1111/jbi.13542

Matzke NJ (2013). Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Frontiers of Biogeography, 5(4), 242–248. https://doi.org/10.21425/F5FBG19694

May-Mcnally SL, Quinn TP, Taylor EB, (2015). Low levels of hybridization between sympatric Arctic char (*Salvelinus alpinus*) and Dolly Varden char (*Salvelinus malma*) highlights their genetic distinctiveness and ecological segregation. Ecology and Evolution, 5, 3031–3045. https://doi.org/10.1002/ece3.1583

McAllister DE, Harington CR (1969). Pleistocene grayling, *Thymallus* , from Yukon, Canada. Canadian Journal of Earth Sciences, 6(5), 1185–1190. https://doi.org/10.1139/e69-119

McCart P (1965). Growth and Morphometry of Four British Columbia Populations of Pygmy Whitefish (*Prosopium coulteri*ournal of the Fisheries Research Board of Canada, 22, 1229–1259. https://doi.org/10.1139/f65-107

McPhail JD, Murray CB (1979). The early life history and ecology of Dolly Varden (*Salvelinus malma*) in the upper Arrow Lakes. Univ. British Columbia Rep. for British Columbia Hydro, Vancouver, pp. 113.

McPhail JD, Baxter JS (1996). A Review of Bull Trout (*Salvelinus confluentus*) Life-history and Habitat Use in Relation to Compensation and Improvement Opportunities. Fisheries Management Report No. 104, Univ. Boulevard, Vancouver, pp. 39.

Mendes SL, Machado MP, Coelho MM, Sousa VC (2021). Genomic data and multispecies demographic modelling uncover past hybridization between currently allopatric freshwater species. Heredity 127, 401–412. https://doi.org/10.1038/s41437-021-00466-1

Meraner A, Gandolfi A (2012). Phylogeography of European grayling, *Thymallus thymallus* (Actinopterygii, Salmonidae), within the Northern Adriatic basin: evidence for native and exotic mitochondrial DNA lineages. Hydrobiologia, 693, 205–221. https://doi.org/10.1007/s10750-012-1109-x

Meraner A, Unfer G, Gandolfi A (2013). Good news for conservation: mitochondrial and microsatellite DNA data detect limited genetic signatures of inter-basin fish transfer in

*Thymallus thymallus* (Salmonidae) from the Upper Drava River. Knowledge and Management of Aquatic Ecosystems, 409, 01. https://doi.org/10.1051/kmae/2013046

Meraner A, Cornetti L, Gandolfi A (2014). Defining conservation units in a stocking-induced genetic melting pot: unraveling native and multiple exotic genetic imprints of recent and historical secondary contact in Adriatic grayling. Ecology and Evolution, 4(8), 1313–1327. https://doi.org/10.1002/ece3.931

Merz JE, Merz WR (2004). Morphological features used to identify Chinook salmon sex during fish passage. The Southwestern Naturalist, 49, 197–202. https://www.jstor.org/stable/3672685

Meulenkamp JE, Sissingh W (2003). Tertiary palaeogeography and tectonostratigraphic evolution of the Northern and Southern Peri-Tethys platforms and the intermediate domains of the African–Eurasian convergent plate boundary zone. Palaeogeography, Palaeoclimatology, Palaeoecology, 196(1–2), 209–228. https://doi.org/10.1016/S0031-0182(03)00319-5

Meyer CP (2003). Molecular systematics of cowries (Gastropoda: Cypraeidae) and diversification patterns in the tropics. Biological Journal of the Linnean Society, 79, 401–459. https://doi.org/10.1046/j.1095-8312.2003.00197.x

Mikheev PB (2009). Sexual Dimorphism of Lower Amur Grayling *Thymallus tugarinae* (Thymallidae). Amurian zoological journal, 1, 386–391.

Miller MA, Pfeiffer W, Schwartz T (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In 2010 Gateway Computing Environments Workshop (GCE) (pp. 1–8). IEEE. https://doi.org/10.1109/GCE.2010.5676129

Miller JM, Malenfant RM, Moore SS, Coltman DW (2012). Short Reads, Circular Genome: Skimming SOLiD Sequence to Construct the Bighorn Sheep Mitochondrial Genome. Journal of Heredity, 103(1), 140–146. https://doi.org/10.1093/jhered/esr104

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Molecular Biology and Evolution, 37(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Miya M, Nishida M (2014). The mitogenomic contributions to molecular phylogenetics and evolution of fishes: a 15-year retrospect. Ichthyological Research, 62(1), 29–71. https://doi.org/10.1007/s10228-014-0440-9

Mohammed AR, Verbyla KL, Al-Mamun HA, McWilliam S, Evans B, King H, Kube P, Kijas JW (2019). Polygenic and sex specific architecture for two maturation traits in farmed Atlantic salmon. BMC Genomics, 20, 139. https://doi.org/10.1186/s12864-019-5525-4

Monet G, Uyanik A, Champigneulle A (2006). Geometric morphometrics reveals sexual and genotypic dimorphisms in the brown trout. Aquatic Living Resources, 19, 47–57. https://doi.org/10.1051/alr:2006004

Morbey YE (2018). Female-biased dimorphism in size and age at maturity is reduced at higher latitudes in lake whitefish *Coregonus clupeaformis*. Journal of Fish Biology, 93, 40–46. https://doi.org/10.1111/jfb.13675

Mueller M, Pander J, Geist J (2018). Comprehensive analysis of >30 years of data on stream fish population trends and conservation status in Bavaria, Germany. Biological Conservation, 226, 311–320. https://doi.org/10.1016/j.biocon.2018.08.006

Muir AM, Blackie CT, Marsden JE, Krueger CC (2012). Lake charr *Salvelinus namaycush* spawning behaviour: new field observations and a review of current knowledge. Reviews in Fish Biology and Fisheries, 22, 575–593. https://doi.org/10.1007/s11160-012-9258-6

Myoung J-G, Hong K-P, Kim Y-U (1993). Morphological study of *Oncorhynchus* spp. (Pisces: Salmonidae) in Korea- III. Sexual dimorphism of chum salmon, *Oncorhynchus keta*. Korea Journal of Ichthyology, 5, 85–95.

Næsje TF, Hansen LP, Järvi T (1988). Sexual dimorphism in the adipose fin of Atlantic salmon, *Salmo salar* L. Journal of Evolutionary Biology, 33, 955–956. https://doi.org/10.1111/j.1095-8649.1988.tb05547.x

Nelson JS, (2006). Fishes of the World, Fourth Edition. John Wiley & Sons, Inc., Hoboken, New Jersey, pp. 601.

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ, (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution, 32(1), 268–274. https://doi.org/10.1093/molbev/msu300

Nikulina YS, Polyaeva KV (2020). Morphology, biology and parasite fauna of the least cisco (*Coregonus sardinella*) of the Yenisei River. Biosystems Diversity, 28, 230–237. https://doi.org/10.15421/012030

Nitychoruk JM, Gutowsky LFG, Harrison PM, Hossie TJ, Power M, Cooke SJ (2013). Sexual and seasonal dimorphism in adult adfluvial bull trout (*Salvelinus confluentus*). Canadian Journal of Zoology, 91, 480–488. https://doi.org/10.1139/cjz-2012-0294

Normandeau DA (1963). The Life history of the round whitefish *Prosopium cylindraceum* (Pallas) of Newfound Lake, New Hampshire. Dissertation, University of New Hampshire, pp. 115.

NOW Community (2022). New and Old Worlds Database of Fossil Mammals (NOW). Licensed under CC BY 4.0. Retrieved from https://nowdatabase.org/now/database/ [downloaded July 2022].

Ocock J, Baasanjav G, Baillie JEM, Erbenebat M, Kottelat M, Mendsaikhan B, Smith K (2006). Mongolian Red List of Fishes. Zoological Society of London, London, UK.

Olivier J-M, Dole-Olivier M-J, Amoros C, Carrel G, Malard F, Lamouroux N, Bravard J-P (2009). Chapter 7 - The Rhône River Basin. In K Tockner, U Uehlinger, CT Robinson (Eds.), Rivers of Europe (pp. 247–295). Academic Press. https://doi.org/10.1016/B978-0-12-369449-2.00007-2

Olson KW, Krabbenhoft TJ, Hrabik TR, Mendsaikhan B, Jensen OP (2019). Pelagic–littoral resource polymorphism in Hovsgol grayling *Thymallus nigrescens* from Lake Hovsgol, Mongolia. Ecology of Freshwater Fish, 28(3). 411–423. https://doi.org/10.1111/eff.12464

Oke KB, Motivans E, Quinn TP, Hendry AP (2019). Sexual dimorphism modifies habitat-associated divergence: Evidence from beach and creek breeding sockeye salmon. Journal of Evolutionary Biology, 32, 227–242. https://doi.org/10.1111/jeb.13407

Osinov AG, Lebedev VS (2000). Genetc divergence and phylogeny of the Salmoninae based on allozyme data. Journal of Fish Biology, 57(2), 354–381. https://doi.org/10.1111/j.1095-8649.2000.tb02177.x

Paetzold C, Wood KR, Eaton DAR, Wagner WL, Appelhans MS (2019). Phylogeny of Hawaiian Melicope (Rutaceae): RAD-seq Resolves Species Relationships and Reveals Ancient Introgression. Frontiers in Plant Science, 10, 1074. https://doi.org/10.3389/FPLS.2019.01074/BIBTEX

Pagel M (1999). Inferring the historical patterns of biological evolution. Nature, 401, 877–884. https://doi.org/10.1038/44766

Palandačić A, Naseka A, Ramler D, Ahnelt H (2017). Contrasting morphology with molecular data: an approach to revision of species complexes based on the example of European *Phoxinus* (Cyprinidae). BMC Ecology and Evolution, 17, 184. https://doi.org/10.1186/s12862-017-1032-x

Palandačić A, Witman K, Spikmans F (2022). Molecular analysis reveals multiple native and alien *Phoxinus* species (Leusciscidae) in the Netherlands and Belgium. Biological Invasions, 24, 2273–2283. https://doi.org/10.1007/s10530-022-02784-9

Palcu DV, Vasiliev I, Stoica M, Krijgsman W (2019). The end of the Great Khersonian Drying of Eurasia: Magnetostratigraphic dating of the Maeotian transgression in the Eastern Paratethys. Basin Research, 31(1), 33–58. https://doi.org/10.1111/bre.12307

Pan T, Sun Z, Lai X, Orozcoterwengel P, Yan P, Wu G, Wang H, Zhu W, Wu X, Zhang B (2019). Hidden species diversity in *Pachyhynobius*: A multiple approaches species delimitation with mitogenomes. Molecular Phylogenetics and Evolution, 137, 138–145. https://doi.org/10.1016/J.YMPEV.2019.05.005

Papakostas S, Vøllestad LA, Bruneaux M, Aykanat T, Vanoverbeke J, Ning M, Primmer CR, Leder EH (2014). Gene pleiotropy constrains gene expression changes in fish adapted to different thermal conditions. Nature Communications, 5(1), 4071. https://doi.org/10.1038/ncomms5071

Patil TS, Jamdade RA, Patil SM, Govindwar SP, Muley DV (2018). DNA barcode based delineation of freshwater fishes from northern Western Ghats of India, one of the world's biodiversity hotspots. Biodiversity and Conservation, 27(13), 3349–3371. https://doi.org/10.1007/s10531-018-1604-0

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D (2012). Ancient admixture in human history. Genetics, 192, 1065–1093. https://doi.org/10.1534/genetics.112.145037

Perdices A, Bohlen J, Šlechtová V, Doadrio I (2016). Molecular evidence for multiple origins of the European spined loaches (Teleostei, Cobitidae). PLOS ONE, 11(1), 1–16. https://doi.org/10.1371/journal.pone.0144628

Pérez-Escobar OA, Balbuena JA, Gottschling M (2016). Rumbling Orchids: How To Assess Divergent Evolution Between Chloroplast Endosymbionts and the Nuclear Host. Systematic Biology, 65, 51–65. https://doi.org/10.1093/SYSBIO/SYV070

Persat H (1977). Ecologie de l'ombre commun. Bulletin français de la pêche et de la pisciculture, 266, 11–20. https://doi.org/10.1051/kmae:1977004

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

Persat H, Mattersdorfer K, Charlat S, Schenekar T, Weiss S (2016). Genetic integrity of the European grayling (*Thymallus thymallus*) populations within the Vienne River drainage basin after five decades of stockings. Cybium, 40(1), 7–20. https://doi.org/10.26028/cybium/2016-401-001

Persat H, Weiss S, Froufe E, Secci-Petretto G, Denys GP (2019). A third European species of grayling (Actinopterygii, Salmonidae), endemic to the Loire River basin (France), *Thymallus ligericus* n. sp. Cybium, 43(3), 233–238. https://doi.org/10.26028/cybium/2019-433-004

Persat H, Keith P, Denys G (2020). Biogéographie et historique de la mise en place des peuplements ichtyologiques de France métropolitaine. Chapitre 2. In P. Keith, N. Poulet, G. Denys, T. Changeux, E. Feunteun & H. Persat (Eds.), Les Poissons d'eau douce de France - 2ème edition (pp. 43–97). Biotope / Publications scientifiques du MNHN. [in French]

Perry WB, Solberg MF, Besnier F, Dyrhovden L, Matre IH, Fjelldal PG, Ayllon F, Creer S, Llewellyn M, Taylor MI, Carvalho G, Glover KA (2019). Evolutionary drivers of kype size in Atlantic salmon (*Salmo salar*): domestication, age and genetics. Royal Society Open Science, 6, 190021. http://dx.doi.org/10.1098/rsos.190021

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. PLOS ONE, 7(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Pfennig KS (2016). Reinforcement as an initiator of population divergence and speciation. Current Zoology, 62, 145–154. https://doi.org/10.1093/cz/zow033

Pincheira-Donoso D, Harvey LP, Grattarola F, Jara M, Cotter SC, Tregenza T, Hodgson DJ (2021). The multiple origins of sexual size dimorphism in global amphibians. Global Ecology and Biogeography, 30, 443–458. https://doi.org/10.1111/geb.13230

Popov SV, Rögl F, Rozanov AY, Steininger FF, Shcherba IG, Kovac M (2004). Lithological-Paleogeographic maps of Paratethys. Courier Forschungsinstitut Senckenberg, 250, 1–46.

Pravdin IF (1966). Guide on Fish Study [= Rukovodstvo po izucheniyu ryb]. Pishchevaya Promyshlennost, Moscow. [in Russian]

Proulx R, Magnan P (2004). Contribution of phenotypic plasticity and heredity to the trophic polymorphism of lacustrine brook charr (*Salvelinus fontinalis* M.). Evolutionary Ecology Research, 6, 503–522.

Puillandre N, Lambert A, Brouillet S, Achaz G (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular Ecology, 21(8), 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

Punzalan D, Hosken DJ (2010). Sexual Dimorphism: Why the Sexes Are (and Are Not) Different. Current Biology, 20, R972–R973. https://doi.org/10.1016/j.cub.2010.09.067

Quinn TP, Foote CJ (1994). The effects of body size and sexual dimorphism on the reproductive behaviour of sockeye salmon, *Oncorhynchus nerka*. Animal Behaviour, 48, 751–761. https://doi.org/10.1006/anbe.1994.1300

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: https://www.R-project.org/

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: https://www.R-project.org/

Raj A, Stephens M, Pritchard JK (2014). FastSTRUCTURE: Variational inference of population structure in large SNP data sets. Genetics, 197, 573–589. https://doi.org/10.1534/genetics.114.164350

Rambaut A (2012). FigTree v1.4. Molecular evolution, phylogenetics and epidemiology. Edinburgh, UK: Retrieved from http://tree.bio.ed.ac.uk/software/figtree.

Rambaut A (2018). FigTree v. 1.4.4. https://github.com/rambaut/figtree/releases [downloaded June 2022]

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard, MA (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Systematic Biology, 67(5), 901–904. https://doi.org/10.1093/sysbio/syy032

Rand PS, Berejikian BA, Bidlack A, Bottom D, Gardner J, Kaeriyama M, Lincoln R, Nagata M, Pearsons TN, Schmidt M, Smoker WW, Weitkamp LA, Zhivotovsky LA (2012). Ecological interactions between wild and hatchery salmonids and key recommendations for research and management actions in selected regions of the North Pacific. Environmental Biology of Fishes, 94, 343–358. https://doi.org/10.1007/s10641-012-9988-2

Ratnasingham S, Hebert PD (2007). BARCODING: bold: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes, 7(3), 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Rawson DS (1950). The grayling (*Thymallus signifer*) in northern Saskatchewan. The Canadian Fish Culturist, 6, 3–10.

van Rensburg AJ, Robin M, Phillips B, Van Buskirk J (2021). European common frog (*Rana temporaria*) recolonized Switzerland from multiple glacial refugia in northern Italy via trans- and circum-Alpine routes. Ecology and Evolution, 11, 15984–15994.

Revell LJ (2012). Phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution, 3, 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x

Reyes-Gavilán FG, Ojanguren AF, Braña F (1997). The ontogenetic development of body segments and sexual dimorphism in brown trout (*Salmo trutta* L.). Canadian Journal of Zoology, 75, 651–655. https://doi.org/10.1139/z97-083

Ricker WE (1938). “Residual” and kokanee salmon in Cultus Lake. Journal of the Biological Board of Canada, 4, 192–218.

Ridder WP (1989). Age, Length, Sex, and Abundance of Arctic Grayling in the Goodpaster River, 1956 through 1988. Fishery data series No. 94. Alaska Department of Fish and Game Division of Sport Fish Juneau, Alaska.

Ritchie AM, Lo N, Ho SYW (2016). Examining the sensitivity of molecular species delimitations to the choice of mitochondrial marker. Organisms Diversity and Evolution, 16(3), 467–480. https://doi.org/10.1007/s13127-016-0275-5

Robertson FM, Gundappa MK, Grammes F, Hvidsten TR, Redmond AK, Lien S, Martin SAM, Holland PWH, Sandve SR, Macqueen DJ (2017). Lineage-specific rediploidization is a mechanism to explain time-lags between genome duplication and evolutionary diversification. Genome Biology, 18, 1–14. https://doi.org/10.1186/s13059-017-1241-z

Roman I, Bourgeois Y, Reyes-Velasco J, Jensen OP, Waldman J, Boissinot S (2018). Contrasted patterns of divergence and gene flow among five fish species in a Mongolian rift lake following glaciation. Biological Journal of the Linnean Society, 125, 115–125. https://doi.org/10.1093/biolinnean/bly083/5051965

Romanov VI (2016). Morphological Differences of Arctic Grayling *Thymallus arcticus* (Pallas, 1776) from Lakes in Khatanga River Basin Associated with the Adaptations to the Mountain and Plain Habitats [= Морфологическая изменчивость сибирского хариуса (*Thymallus arcticus* (Pallas, 1776)) из разнотипных озёр бассейна реки Хатанги (полуостров Таймыр) в связи с адаптациями к горным и равнинным условиям]. The Bulletin of Irkutsk State University, Series Biology, Ecology, 18, 45–57. [in Russian]

Romanov VI (2020). Fauna of grayling fish of Khantai lake. Bull. Lampetra, 9, 161–177.

Ronquist F, Huelsenbeck JP (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 19(12), 1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna, S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology, 61(3), 539–542. https://doi.org/10.1093/sysbio/sys029

Rose JP, Cassio AP, Lemmon EP, Lemmon AR, Sytsma KJ (2021). Out of sight, Out of mind: Widespread Nuclear and Plastid-nuclear discordance in the flowering plant genus *Polemonium* (Polemoniaceae) suggests widespread historical gene flow despite limited nuclear signal. Systematic Biology, 70 (1), 162–180. https://doi.org/10.1093/sysbio/syaa049

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Molecular Biology and Evolution, 34(12), 3299–3302, https://doi.org/10.1093/molbev/msx248

Rückert-Ülkümen N, Kaya O (1993). Neue Beiträge zur Tertiär-Stratigraphie und Otolithenfauna der Umgebung von Istanbul (Küçükçekmece-Büyükçekmece See), Türkei. Mitteilungen der Bayerischen Staatssammlung für Paläontologie und Historischen Geologie, 33, 51-89. [in German].

Rückert-Ülkümen N, Kowalke T, Matzke-Karasz R, Witt W, Yiğitbaş E (2006). Biostratigraphy of the Paratethyan Neogene at Yalova (İzmit-Province, NW-Turkey). Newsletters on Stratigraphy, 42(1), 43–68. http://dx.doi.org/10.1127/0078-0421/2006/0042-0043

Rundle HD, Schluter D (1998). Reinforcement of stickleback mate preferences: Sympatry breeds contempt. Evolution, 52(1), 200–208. https://doi.org/10.1111/j.1558-5646.1998.tb05153.x

Ruxton GD (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. Behavioral Ecology, 17, 688–690. https://doi.org/10.1093/beheco/ark016

Rylková K, Kalous L, Bohlen J, Lamatsch DK, Petrtýl M (2013). Phylogeny and biogeographic history of the cyprinid fish genus *Carassius* (Teleostei: Cyprinidae) with focus on natural and anthropogenic arrivals in Europe. Aquaculture, 380–383, 13–20. https://doi.org/10.1016/j.aquaculture.2012.11.027

Sakun OF, Butskaya NA (1968). Determination of Stage of Maturation and Study of Sexual Cycles of Fishes [= Opredelenie stadii zrelosti i izuchenie polovykh tsiklov ryb]. Murmansk: PINRO. [in Russian]

Sambrook J, Fritsch EF, Maniatis T (1989). Extraction and purification of plasmid DNA. Molecular Cloning: A Laboratory Manual, 1(1), 21–152.

dos Santos AM, Cabezas AP, Tavares AI, Xavier R, Branco M (2016). tcsBU: a tool to extend TCS network layout and visualization. Bioinformatics, 32(4), 627–628. https://doi.org/10.1093/bioinformatics/btv636

Sävilammi T, Primmer CR, Varadharajan S, Guyomard R, Guiguen Y, Sandve SR, Asbjørn Vøllestad L, Papakostas S, Lien S (2019). The Chromosome-Level Genome Assembly of European Grayling Reveals Aspects of a Unique Genome Evolution Process Within Salmonids. G3: Genes, Genomes, Genetics, 9(5), 1283–1294. https://doi.org/10.1534/g3.118.200919

Schedel FDB (2020). Phylogeny, divergence time estimates and systematics of African cichlids (Cichlidae: Pseudocrenilabrinae), with a focus on the rheophilic cichlids of East and Central Africa (Doctoral thesis). Ludwig-Maximilian University of Munich, Germany, pp. 240.

Schroeter JC, Maloy AP, Rees CB, Bartron ML (2019). Fish mitochondrial genome sequencing: expanding genetic resources to support species detection and biodiversity monitoring using environmental DNA. Conservation Genetics Resources, 1–14. https://doi.org/10.1007/s12686-019-01111-0

Scribner KT, Avise JC (1994). Cytonuclear genetics of experimental fish hybrid zones inside biosphere 2. Proceedings of the National Academy of Sciences, 91(11), 5066–5069. https://doi.org/10.1073/pnas.91.11.5066

Scribner KT, Page KS, Bartron ML (2000). Hybridization in freshwater fishes: A review of case studies and cytonuclear methods of biological inference. Reviews in Fish Biology and Fisheries, 10, 293–323. https://doi.org/10.1023/A:1016642723238.

Secci-Petretto G, Englmaier GK, Weiss SJ, Antonov A, Persat H, Denys G, Schenekar T, Romanov VI, Taylor EB, Froufe E (2023). Evaluating a species phylogeny using ddRAD SNPs: cytonuclear discordance and introgression in the salmonid genus *Thymallus* (Salmonidae). Molecular Phylogenetics and Evolution, 178, 107654. https://doi.org/10.1016/j.ympev.2022.107654

Seifertová M, Bryja J, Vyskočilová M, Martínková N, Šimková A (2012). Multiple Pleistocene refugia and postglacial colonization in the European chub (*Squalius cephalus*) revealed by combined use of nuclear and mitochondrial markers. Journal of Biogeography, 39(6), 1024–1040. https://doi.org/10.1111/j.1365-2699.2011.02661.x

Selz OM, Dönz CJ, Vonlanthen P, Seehausen O (2020). A taxonomic revision of the whitefish of lakes Brienz and Thun, Switzerland, with descriptions of four new species (Teleostei, Coregonidae). ZooKeys, 989, 79–162. https://doi.org/10.3897/zookeys.989.32822

Semenchenko AA (2005). Biological Peculiarities of Arctic Grayling *Thymallus arcticus grubii* Dybowski of the Samarga River [= Особенности биологии амурского хариуса *Thymallus arcticus grubii* Dybowski реки Самарга]. Readings in Memory of Vladimir Yakovlevich Levanidov (Vladimir Ya. Levanidov's Biennial Memorial Meetings) 1, 217–228. [in Russian]

Shapovalov L, Taft AC (1954). The life histories of the steelhead rainbow trout (*Salmo gairdneri gairdner*) and silver salmon (*Oncorhynchus kisutch*) with special reference to Wadell Creek, California, and recommendations regarding their management. Fish Bulletin, 98.

Sharma E, Künstner A, Fraser BA, Zipprich G, Kottler VA, Henz SR, Weigel D, Dreyer C (2014). Transcriptome assemblies for studying sex-biased gene expression in the guppy, *Poecilia reticulata*. BMC Genomics, 15, 400. https://doi.org/10.1186/1471-2164-15-400

Shimodaira H (2002). An approximately unbiased test of phylogenetic tree selection. Systematic Biology, 51(3), 492–508. https://doi.org/10.1080/10635150290069913

Shimodaira H, Hasegawa M (1999). Multiple comparisons of log-Likelihoods with applications to phylogenetic inference. Molecular Biology and Evolution, 16, 1114–1116. https://doi.org/10.1093/oxfordjournals.molbev.a026201.

Shubin PN, Zakharov AB (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.

Sidlauskas BL, Mol JH, Vari RP (2011). Dealing with allometry in linear and geometric morphometrics: a taxonomic case study in the *Leporinus cylindriformis* group (Characiformes: Anostomidae) with description of a new species from Suriname. Zoological Journal of the Linnean Society, 162, 103–130. https://doi.org/10.1111/j.1096-3642.2010.00677.x

Simonsen MK, Siwertsson A, Adams CE, Amundsen PA, Praebel K, Knudsen R (2017). Allometric trajectories of body and head morphology in three sympatric Arctic charr (*Salvelinus alpinus* (L.)) morphs. Ecology and Evolution, 7(18), 7277–7289. https://doi.org/10.1002/ece3.3224

Siwertsson A, Knudsen R, Adams CE, Præbel K, Amundsen P-A (2013). Parallel and non-parallel morphological divergence among foraging specialists in European whitefish (*Coregonus lavaretus*). Ecology and Evolution, 3, 1590–1602. https://doi.org/10.1002/ece3.562

Skog A, Vøllestad LA, Stenseth NC, Kasumyan A, Jakobsen KS (2014). Circumpolar phylogeography of the northern pike (*Esox lucius*) and its relationship to the Amur pike (*E. reichertii*). Frontiers in Zoology, 11(1), 1–8. https://doi.org/10.1186/s12983-014-0067-8

Skurikhina LA, Mednikov BM, Tugarina PY (1985). Geneticheskaya divergentziya khariusov (*Thymallus* Cuvier) Evrazii i seti vidov. Zoologicheskii Zhurnal, 64(2), 245. [In Russian]

Šlechtova V, Bohlen J, Freyhof J, Persat H, Delmastro GB (2004). The Alps as barrier to dispersal in cold-adapted freshwater fishes? Phylogeographic history and taxonomic status of the bullhead in the Adriatic freshwater drainage. Molecular Phylogenetics and Evolution, 33(1), 225–239. https://doi.org/10.1016/j.ympev.2004.05.005

Šlechtová V, Bohlen J, Perdices A (2008). Molecular phylogeny of the freshwater fish family Cobitidae (Cypriniformes: Teleostei): Delimitation of genera, mitochondrial introgression and evolution of sexual dimorphism. Molecular Phylogenetics and Evolution, 47(2), 812–831. https://doi.org/10.1016/j.ympev.2007.12.018

Smith GR (1992). Introgression in Fishes: Significance for Paleontology, Cladistics, and Evolutionary Rates. Systematic Biology, 41(1), 41–57. https://doi.org/10.1093/sysbio/41.1.41

Sommer RS, Zachos FE (2009). Fossil evidence and phylogeography of temperate species: 'glacial refugia' and post-glacial recolonization. Journal of Biogeography, 36, 2013–2020. https://doi.org/10.1111/j.1365-2699.2009.02187.x

Srivathsan A, Meier R (2012). On the inappropriate use of Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. Cladistics, 28(2), 190–194. https://doi.org/10.1111/j.1096-0031.2011.00370.x

Stamatakis A (2016). The RAxML v8.2.X Manual. Retrieved from http://sco.h-its.org/exelixis/web/software/raxml/index.html

Stamford MD, Taylor EB. (2004). Phylogeographical lineages of Arctic grayling (*Thymallus arcticus*) in North America: Divergence, origins and affinities with Eurasian *Thymallus*. Molecular Ecology, 13(6), 1533–1549. https://doi.org/10.1111/j.1365-294X.2004.02174.x

Stange M, Sanchez-Villagra MR, Salzburger W, Matschiner M (2018). Bayesian Divergence-Time Estimation with Genome-Wide Single-Nucleotide Polymorphism Data of Sea Catfishes (Ariidae) Supports Miocene Closure of the Panamanian Isthmus. Systematic Biology, 67, 681–699. https ://doi.org/10.1093/sysbi o/syy006

Stefaniak K, Kovalchuk O, Kotusz J, Stachowicz-Rybka R, Mirosław-Grabowska J, Winter H, Niska M, Sobczyk A, Barkaszi Z., Kotowski A, Malkiewicz M, Alexandrowicz WP, Raczyński P, Badura J, Przybylski B, Ciszek D, Urbański K (2020). Pleistocene freshwater environments of Poland: a comprehensive study of fish assemblages based on multi-proxy approach. Boreas, 50(2), 457–476. https://doi.org/10.1111/bor.12489

Van Steenberge MW, Vanhove MPM, Chocha Manda A, Larmuseau MHD, Swart BL, Khang'Mate F, Arndt A, Hellemans B, Van Houdt J, Micha J-C, Koblmüller S, Roodt-Wilding R, Volckaert, FAM (2020). Unravelling the evolution of Africa's drainage basins through a widespread freshwater fish, the African sharptooth catfish *Clarias gariepinus*. Journal of Biogeography, 47(8), 1739–1754. https://doi.org/10.1111/jbi.13858

Straffin E. Blum M, Colls A, Stokes S (1999). Alluvial stratigraphy of the Loire and Arroux rivers (Burgundy, France). Quaternaire, 10(4), 271–282. https://doi.org/10.3406/quate.1999.1648

Strimmer K, Rambaut A, (2002). Inferring confidence sets of possibly misspecified gene trees. Proceedings of the Royal Society B: Biological Sciences, 269, 137–142. https://doi.org/10.1098/rspb.2001.1862

Suchard MA, Rambaut A (2009). Many-core algorithms for statistical phylogenetics. Bioinformatics, 25(11), 1370–1376. https://doi.org/10.1093/bioinformatics/btp244

Sukumaran J, Knowles LL (2017). Multispecies coalescent delimits structure, not species. Proceedings of the National Academy of Sciences of the United States of America, 114(7), 1607–1612. https://doi.org/10.1073/pnas.1607921114

Sušnik S, Berrebi P, Dovč P, Hansen MM, Snoj A (2004). Genetic introgression between wild and stocked salmonids and the prospects for using molecular markers in population rehabilitation: The case of the Adriatic grayling (*Thymallus thymallus* L. 1785). Heredity, 93(3), 273–282. https://doi.org/10.1038/sj.hdy.6800500

Sušnik S, Weiss S, Odak T, Dellilng B, Treer T, Snoj A (2006). Reticulate evolution: ancient introgression of the Adriatic brown trout mtDNA in softmouth trout *Salmo obtusirostris* (Teleosei: Salmonidae). Biological Journal of the Linnean Society, 90, 139–152. https://doi.org/10.1111/j.1095-8312.2007.00717.x

Svetovidov AN (1936). Euroasian Graylings (Genus *Thymallus* Cuvier). Trudy Zoologicheskogo Instituta. Akademiya Nauk SSSR 3:183–301.

Swanson EM, McElhinny TL, Dworkin I, Weldele ML, Glickman SE, Holekamp KE (2013). Ontogeny of sexual size dimorphism in the spotted hyena (*Crocuta crocuta*). Journal of Mammalogy, 94, 1298–1310. https://doi.org/10.1644/12-MAMM-A-277.1

Tack SL (1973). Distribution, Abundance, and Natural History of the Arctic Grayling in the Tanana River Drainage. Alaska Department of Fish and Game, Annual Progress Report, 14(R-I).

Takács P, Bánó B, Czeglédi I, Erős T, Ferincz A, Gál B, Bánó-Kern B, Kovács B, Nagy AA, Nyeste K, Lente V, Preiszner B, Sipos S, Staszny A, Vitál Z, Weiperth A, Csoma E (2022). The mixed phylogenetic origin of northern pike (*Esox lucius* Linnaeus 1758) populations in the Middle Danubian drainage. BMC Zoology, 7, 28. https://doi.org/10.1186/s40850-022-00129-6

Takahashi K, Terai Y, Nishida M, Okada N (2001). Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons. Molecular Biology and Evolution, 18 (11), 2057–2066. https://doi.org/10.1093/oxfordjournals.molbev.a003747

Tamate T, Maekawa K (2004). Female-biased mortality rate and sexual size dimorphism of migratory masu salmon, *Oncorhynchus masou*. Ecology of Freshwater Fish,13, 96–103. https://doi.org/10.1111/j.1600-0633.2004.00043.x

Tchernavin V (1944). The Breeding Characters of Salmon in Relation to their Size. Proceedings of the Zoological Society of London, 113: 206–132. https://doi.org/10.1111/j.1096-3642.1944.tb00845.x

Thioulouse J, Chessel D, Dolédec S, Olivier JM (1997). ADE-4: a multivariate analysis and graphical display software. Statistics and Computing, 7, 75–83. https://doi.org/10.1023/A:1018513530268

Thorn MW, Morbey YE (2016). Evidence for the secondary sexual development of the anal fin in female kokanee salmon *Oncorhynchus nerka*. Journal of Fish Biology, 88, 448–458. https://doi.org/10.1111/jfb.12800

Toews DPL, Brelsford A (2012). The biogeography of mitochondrial and nuclear discordance in animals. Molecular Ecology, 21, 3907–3930. https://doi.org/10.1111/J.1365-294X.2012.05664.X

Tsogtsaikhan P, Mendsaikhan B, Jargalmaa G, Ganzorig B, Weidel BC (2017). Age and growth comparisons of Hovsgol grayling (*Thymallus nigrescens* Dorogostaisky, 1923), Baikal grayling (*T. baicalensis* Dybowski, 1874), and lenok (*Brachymystax lenok* Pallas, 1773) in lentic and lotic habitats of Northern Mongolia. Journal of Applied Ichthyology, 33, 108–115.

Turan D, Kottelat M, Bektaş Y (2011). *Salmo tigridis*, a new species of trout from the Tigris River, Turkey (Teleostei: Salmonidae). Zootaxa, 2993, 23–33. https://doi.org/10.11646/zootaxa.2993.1.2

Turan D, Kottelat M, Engin S (2012). The trouts of the Mediterranean drainages of southern Anatolia, Turkey, with description of three new species (Teleostei: Salmonidae). Ichthyological Exploration of Freshwaters, 23, 219–236.

Turan D, Doğan E, Kaya C, Kanyılmaz M (2014). *Salmo kottelati*, a new species of trout from Alakır Stream, draining to the Mediterranean in southern Anatolia, Turkey (Teleostei, Salmonidae). ZooKeys, 462, 135–151. https://doi.org/10.3897/zookeys.462.8177

Turan D, Kalaycı G, Bektaş Y, Kaya C, Bayçelebi E (2020). A new species of trout from the northern drainages of Euphrates River, Turkey (Salmoniformes: Salmonidae). Journal of Fish Biology, 96, 1454–1462. https://doi.org/10.1111/jfb.14321

Van Houdt JKJ, De Cleyn L, Perretti A, Volckaert FAM (2005). A mitogenic view on the evolutionary history of the Holarctic freshwater gadoid, burbot (*Lota lota*). Molecular Ecology, 14(8), 2445–2457. https://doi.org/10.1111/j.1365-294X.2005.02590.x

Varadharajan S, Sandve SR, Gillard GB, Tørresen OK, Mulugeta TD, Hvidsten TR, Lien S, Vøllestad LA, Jentoft S, Nederbragt AJ, Jakobsen KS (2018). The grayling genome reveals selection on gene expression regulation after whole genome duplication. Genome Biology and Evolution, 10(10), 2785–2800. https://doi.org/10.1093/gbe/evy201

Wallis GP, Cameron-Christie SR, Kennedy HL., Palmer G, Sanders TR, Winter DJ (2017). Interspecific hybridization causes long-term phylogenetic discordance between nuclear and mitochondrial genomes in freshwater fishes. Molecular Ecology, 26, 3116–3127. https://doi.org/10.1111/mec.14096

Ward JC (1951). The biology of the Arctic grayling in the southern Athabaska drainage. M.S. Thesis, University of Alberta, Edmonton, pp. 71.

Ward RD (2009). DNA barcode divergence among species and genera of birds and fishes. Molecular Ecology Resources, 9(4), 1077–1085. https://doi.org/10.1111/j.1755-0998.2009.02541.x

Warton DI, Wright IJ, Falster DS, Westoby M (2006). Bivariate line-fitting methods for allometry. Biological Reviews, 81, 259–291. https://doi.org/10.1017/S1464793106007007

Wedekind C, Evanno G, Urbach D, Jacob A, Müller R (2008). 'Good-genes' and 'compatible-genes' effects in an Alpine whitefish and the information content of breeding tubercles over the course of the spawning season. Genetica, 132, 199–208. https://doi.org/10.1007/s10709-007-9164-3

Weiss S, Ferrand N (2002). Phylogeography of Southern European Refugia: Evolutionary perspectives on the origins and conservation of European biodiversity. Springer, Dordrecht, The Netherlands, pp. 377.

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 mitochondrial DNA control region. Molecular Ecology, 11(8), 1393–1407. https://doi.org/10.1046/j.1365-294X.2002.01544.x

Weiss S, Knizhin I, Kirillov A, Froufe E (2006). Phenotypic and genetic differentiation of two major phylogeographical lineages of arctic grayling *Thymallus arcticus* in the Lena River, and

surrounding Arctic drainages. Biological Journal of the Linnean Society, 88(4), 511–525. https://doi.org/10.1111/j.1095-8312.2006.00621.x

Weiss S, Knizhin I, Romanov V, Kopun T (2007). Secondary contact between two divergent lineages of grayling *Thymallus* in the lower Enisey basin and its taxonomic implications. Journal of Fish Biology, 71(C), 371–386. https://doi.org/10.1111/j.1095-8649.2007.01662.x

Weiss SJ, Kopun T, Sušnik Bajec S (2013). Assessing natural and disturbed population structure in European grayling *Thymallus thymallus*: Melding phylogeographic, population genetic and jurisdictional perspectives for conservation planning. Journal of Fish Biology, 82(2), 505–521. https://doi.org/10.1111/jfb.12007

Weiss S, Grimm J, Gonçalves DV, Secci-Petretto G, Englmaier GK, Baimukanov M, Froufe E (2020a). Comparative genetic analysis of grayling (*Thymallus* spp. Salmonidae) across the paleohydrologically dynamic river drainages of the Altai-Sayan mountain region. Hydrobiologia, 847, 2823–2844. https://doi.org/10.1007/s10750-020-04273-3

Weiss S, Secci-Petretto G, Antonov A, Froufe E (2020b). Multiple species of grayling (*Thymallus* sp.) found in sympatry in a remote tributary of the Amur River. Zoologica Scripta, 49(1), 117–128. https://doi.org/10.1111/zsc.12393

Weiss SJ, Gonçalves DV, Secci-Petretto G, Englmaier GK, Gomes-Dos-Santos A, Denys GPJ, Persat H, Antonov A, Hahn C, Taylor EB, Elsa F (2021). Global systematic diversity, range distributions, conservation and taxonomic assessments of graylings (Teleostei: Salmonidae; *Thymallus* spp.). Organisms Diversity & Evolution, 21, 25–42. https://doi.org/10.1007/s13127-020-00468-7

Wilhelm G, Handschuh S, Plantl J, Nemeschkal HL (2011). Sexual dimorphism in head structures of the weevil *Rhopalapion longirostre* (Olivier 1807) (Coleoptera: Curculionoidea): a response to ecological demands of egg deposition. Biological Journal of the Linnean Society, 104, 642–660. https://doi.org/10.1111/j.1095-8312.2011.01751.x

Willson MF (1997). Variation in salmonid life histories: patterns and perspectives. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, pp. 50. https://doi.org/10.2737/PNW-RP-498

Winans GA, Pollard S, Kuligowski DR (2003). Two reproductive life history types of kokanee, *Onchorynchus nerka*, exhibit multivariate morphometric and protein genetic differentiation. Environmental Biology of Fishes, 77, 87–100. https://doi.org/10.1023/A:1024401102141

Winstanley C, Rapley R (2003). Extraction and Purification of Plasmid DNA. Nucleic Acid Protoc. Handbook 1, 327–331.

Xu G-F, Zhen-Bo M, Qi-Rui H, Han Y (2017). Complete mitochondrial genome of *Coregonus muksum*. Mitochondrial DNA Part A, 28(1), 9–10. https://doi.org/10.3109/19401736.2015.1106518

Yamamoto S, Kitano S, Maekawa K, Koizumi I, Morita K (2006). Introgressive hybridization between Dolly Varden *Salvelinus malma* and white-spotted charr *Salvelinus leucomaenis* on Hokkaido Island, Japan. Journal of Fish Biology, 68(A), 68–85. https://doi.org/10.1111/j.0022-1112.2006.00994.x

Yamamoto T, Hirohashi N, Fujiwara R, Suzuki T, Maruta H, Omiya H, Kitanishi S (2017). Relationships between body size and secondary sexual characters, and sperm characters in male Dolly Varden char (*Salvelinus malma*). Ecology of Freshwater Fish, 26, 397–402. https://doi.org/10.1111/eff.12283

Yang L, Meng F, Wang R, Shi G (2017). Complete mitochondrial genome of the *Salvelinus malma* sp. (Salmoniformes, Salmonidae) with phylogenetic consideration. Mitochondrial DNA Part B, 2(2), 889–890. https://doi.org/10.1080/23802359.2017.1403865

Yasuike M, Jantzen S, Cooper GA, Leder E, Davidson WS, Koop, BF (2010). Grayling (Thymallinae) phylogeny within salmonids: Complete mitochondrial DNA sequences of *Thymallus arcticus* and *Thymallus thymallus*. Journal of Fish Biology, 76(2), 395–400. https://doi.org/10.1111/j.1095-8649.2009.02494.x

Zhang J, Kapli P, Pavlidis P, Stamatakis A (2013). A general species delimitation method with applications to phylogenetic placements. Bioinformatics, 29(22), 2869–2876. https://doi.org/10.1093/bioinformatics/btt499

Zhang S, Wei Q, Du H, Li L (2016). The complete mitochondrial genome of the Endangered *Hucho bleekeri* (Salmonidae: Huchen). Mitochondrial DNA, 27(1), 124–125. https://doi.org/10.3109/19401736.2013.878906

Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT (2020). PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Molecular Ecology Resources, 20, 348–355. https://doi.org/10.1093/molbev/msu300

Zhao J-W (2015). The complete mitochondrial genome of the *Thymallus grubii* (Amur grayling). Mitochondrial DNA, 26(5), 799–800. https://doi.org/10.3109/19401736.2013.855899

Zhivotovsky LA, Kim HYu (2015). Morphological Markers of Sex in Pink Salmon *Oncorhynchus gorbuscha* (Salmonidae). Journal of Ichthyology, 55, 131–133.

Zinovjev EA (2012). Ecology of grayling in the Perm Kama region: a monograph [= Экология хариусов Пермского Прикамья: монография] Perm State National Research University, Perm, pp. 445. [in Russian]

Zinovjev EA, Bogdanov VD (2012). Morphobiological Features of the Siberian Grayling (*Thymallus arcticus*, Thymallidae) of the Polar Urals [= Морфобиологические особенности сибирского хариуса (*Thymallus arcticus*, Thymallidae) Полярного Урала]. Veterinaria Kubani 4. [in Russian]