



HAL
open science

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

Giulia Secci-Petretto, Gernot K Englmaier, Steven J Weiss, Alexander Antonov, Henri Persat, Gael P J Denys, Tamara Schenekar, Vladimir I Romanov, Eric B Taylor, Elsa Froufe

► To cite this version:

Giulia Secci-Petretto, Gernot K Englmaier, Steven J Weiss, Alexander Antonov, Henri Persat, et al.. Evaluating a species phylogeny using ddRAD SNPs: cytonuclear discordance and introgression in the salmonid genus *Thymallus* (Salmonidae). *Molecular Phylogenetics and Evolution*, 2023, 178, pp.107654. 10.1016/j.ympev.2022.107654 . hal-03847748

HAL Id: hal-03847748

<https://hal.science/hal-03847748>

Submitted on 22 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL
open science

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

Giulia Secci-Petretto, Gernot K Englmaier, Steven J Weiss, Alexander Antonov, Henri Persat, Gael P J Denys, Tamara Schenekar, Vladimir I Romanov, Eric B Taylor, Elsa Froufe

► To cite this version:

Giulia Secci-Petretto, Gernot K Englmaier, Steven J Weiss, Alexander Antonov, Henri Persat, et al.. Evaluating a species phylogeny using ddRAD SNPs: cytonuclear discordance and introgression in the salmonid genus *Thymallus* (Salmonidae). *Molecular Phylogenetics and Evolution*, 2023, 178, pp.107654. 10.1016/j.ympev.2022.107654 . hal-03847748

HAL Id: hal-03847748

<https://hal.science/hal-03847748>

Submitted on 22 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Giulia Secci-Petretto^{a, b}, Gernot K. Englmaier^c, Steven J. Weiss^{c, *}, Alexander Antonov^d, Henri Persat^e, Gael Denys^{f, g}, Tamara Schenekar^c, Vladimir I. Romanov^h, Eric B. Taylorⁱ, Elsa Froufe^{a, *}

^a 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

^b Department of Biology, Faculty of Sciences, U. Porto - University of Porto, Portugal

^c University of Graz, Institute of Biology, Universitätsplatz 2, 8010 Graz, Austria

^d Institute of Water and Ecological Problems, Far East Branch, Russian Academy of Sciences, ul. Kim Yu Chena 65, Khabarovsk, 680063 Russia

^e Société Française d'Ichthyologie, Muséum National d'Histoire Naturelle Paris, France, 57 rue Cuvier CP26, 75005 Paris, France

^f Unité Mixte de Service Patrimoine Naturel – Centre d'expertise et de données (UMS 2006 AFB – CNRS – MNHN), Muséum national d'Histoire naturelle, 36 rue Geoffroy Saint-Hilaire CP 41, 75005 Paris, France

^g Laboratoire de Biologie des organismes et écosystèmes aquatiques (BOREA), MNHN, CNRS, IRD, SU, UCN, UA, 57 rue Cuvier CP26, 75005 Paris, France

^h Tomsk State University, 36 Lenin Ave., Tomsk, 634050 Russia

ⁱ Department of Zoology, Biodiversity Research Centre and Beaty Biodiversity Museum, University of British Columbia, 6270 University Ave, Vancouver, BC V6T 1Z, Canada

ARTICLE INFO

Keywords:

Cyto-nuclear discordance
Introgression
Grayling
ddRADseq
Phylogenetics

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.

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

(Scribner and Avise, 1994; Avise, 1995; Asmussen et al., 2010; 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

* Corresponding authors.

E-mail addresses: steven.weiss@uni-graz.at (S.J. Weiss), elsafroufe@gmail.com (E. Froufe).

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 et al., 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, 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, 2020a, 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'pukhovskiy, 2005 (Knizhin et al., 2006); 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. tugariniae* Knizhin, Antonov, Safronov & Weiss, 2007 are found in sympatry (Froufe et al., 2003b).

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. Frag-

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
<i>T. aeliani</i>	Adige	Adige	46.616389, 11.184167	1
	Piave	Piave	46.394156, 12.357128	1
	Rienz	Isarco-Adige	46.759745, 12.051320	1
<i>T. arcticus</i>	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
<i>T. baicalensis</i>	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
	<i>T. baicalolenensis</i>	Bureya	Amur	51.916667, 134.883333
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
<i>T. brevicephalus</i>		Kaldzhir	Irtysh-Ob	48.416389, 85.183056
	Kara-Kaba	Irtysh-Ob	48.8, 86.516667	1
	Urunkhaika	Irtysh-Ob	48.766667, 86.116667	1

Table 1 (continued)

Species	Location	Drainage	Coordinates	N
<i>T. brevirostris</i>	Khukh	Zavkhan	47.516667, 98.45	1
<i>T. burejensis</i>	Bureya	Amur	51.715717, 134.309278	8
<i>T. flavomaculatus</i>	Urmi	Tunguska-Amur	48.713056, 134.266944	2
	Gobili	Anui-Amur	49.25, 138.316667	6
<i>T. grubii</i>	Koppi	Tatar Strait	48.615906, 139.590178	1
	Umalta Makit	Bureya-Amur	51.679767, 134.239539	2
<i>T. ligericus</i>	Bureya	Amur	51.715717, 134.309278	5
	Onon	Amur	48.663056, 110.425	3
<i>T. ligericus</i>	Alagnon	Loire	45.106389, 2.896667	1
	Sioule	Loire	46.025833, 2.904167	1
<i>T. nigrescens</i>	Chovsgul	Eg-Selenga-Baikal-Angara-Enisei	51.45, 100.65	2
<i>T. nikolskyi</i>	Biya	Ob	51.82, 87.154722	1
	Anuy	Ob	51.691417, 84.239306	1
<i>T. svetovidovi</i>	Belin	Enisei	52.19117, 98.6648	1
	Belin	Enisei	51.64529, 98.15946	1
<i>T. thymallus</i>	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
<i>T. thymallus</i>	Ain	Rhone	46.7451317, 5.809925	1
	Ure	Ouse	54.305556, -1.939722	1
	Padje	Kalixälven	67.703038, 18.600424	1
	Kaitumjaure	Mindyak	53.984444, 58.799444	1
<i>T. tugarinae</i>	Anui	Amur	49.283333, 137.916667	1
	Gobili	Anui-Amur	49.294994, 138.520103	3
<i>T. baicalolenensis</i> × <i>T. baicalensis</i>	Verkhnyaya Angara	Baikal-Angara-Enisei	55.864289, 108.896714	1
<i>T. baicalolenensis</i> × <i>T. baicalensis</i>	Barguzin	Baikal-Angara-Enisei	54.555367, 111.399297	2
<i>T. baicalolenensis</i> × <i>T. burejensis</i>	Bureya	Amur	51.715717, 134.309278	2
<i>B. lenok</i>	Urunkhaika	Irtysh-Ob	48.766667, 86.116667	1
<i>C. renka</i>	Achensee	Seeache-Isar-Danube	47.4755843, 11.7061217	1
<i>H. hucho</i>	Pielach	Danube	48.1993185, 15.4345248	1

ments > 250 bp were selected from each individual using 1.8X AM-Pure 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 μ l 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 4.62×10^6 reads per sample, with a minimum of 2.12×10^6 and a maximum of 19.12×10^6 . 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 & Overcast, 2020) using a reference-based approach with the most recent available genome of *T. thymallus* (Linnaeus, 1758) (GCA_004348285.1; Sävilämmi 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 a 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 web-server 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 statistics 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 mCOLintF (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. brevisrostris* 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*.

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, Verkhnaya 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.

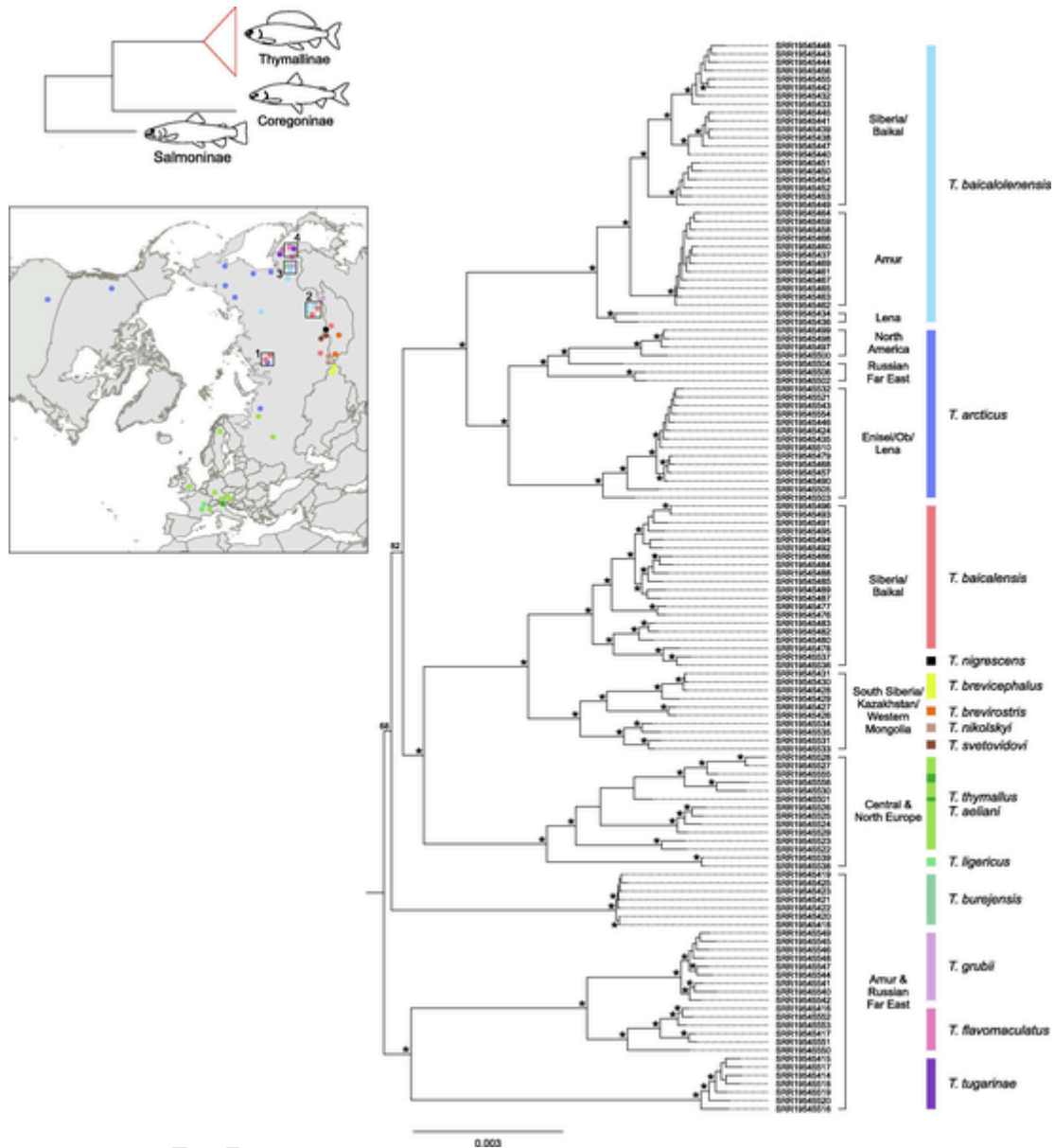


Fig. 1. Maximum Likelihood phylogeny inferred from genome-wide ddRADseq data for the genus *Thymallus*, including major drainages/geographic areas for each group of samples; asterisks indicate bootstrap values > 95 %. Outgroup taxa and phylogenetic placement in Salmonidae are depicted in the top left. 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*).

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 ($F_{ST} = 0.84$).

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. brevicephalus*, 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*.

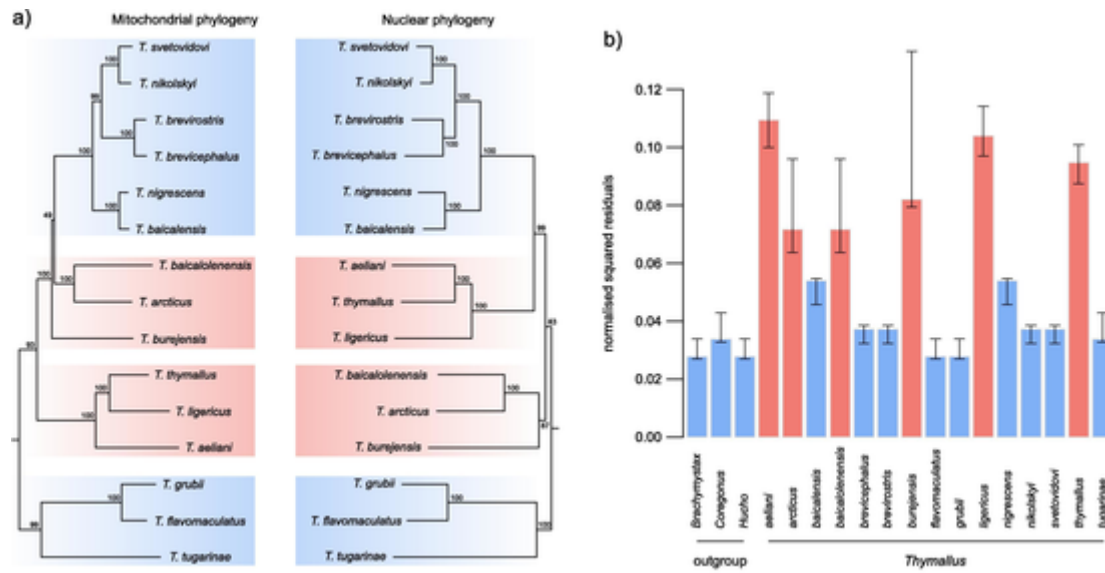


Fig. 2. Comparison between mitochondrial and nuclear phylogenies for *Thymallus* spp. a) ML phylogenies using a single individual per species. b) values of normalized squared residuals inferred with PACo for each species. Outliers with normalized square residuals above the program's standard threshold (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.)

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 analysis 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 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., 2020; 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).

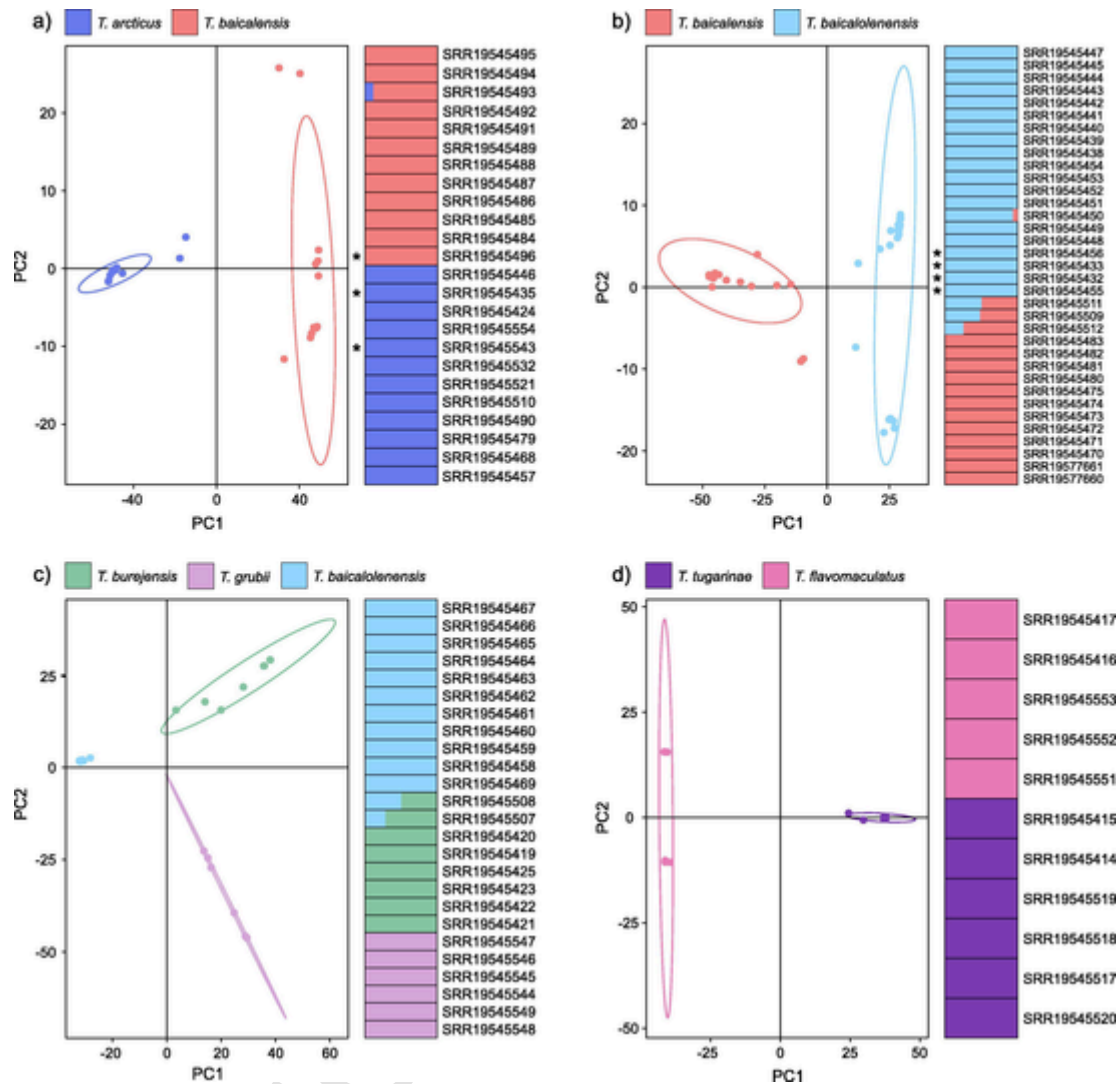


Fig. 3. PCA and fastSTRUCTURE results for a) *T. arcticus* and *T. baicalensis* from the Kantaiskoye Lake system (based on 27,457 SNPs); b) *T. baicalensis* and *T. baicalolenensis* from the Lake Baikal basin (based on 16,093 SNPs); c) *T. baicalolenensis*, *T. burejensis* and *T. grubii* from the upper 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.

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-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,

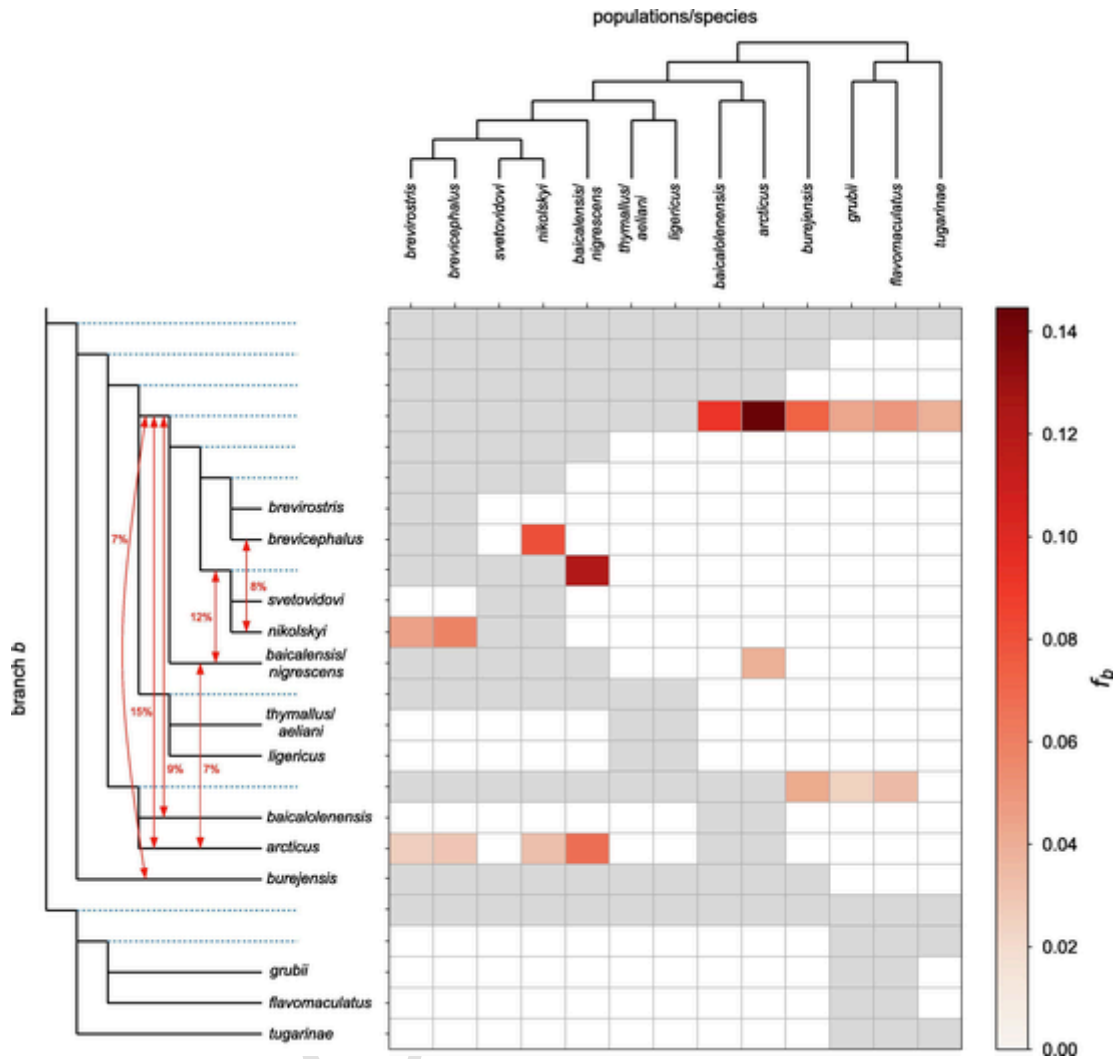


Fig. 4. Results of Dsuite f -branch statistics for *Thymallus* spp. The phylogenetic tree used is shown along the x- and y-axes (the y-axis showing an expanded form of the tree where both terminal- and internal branches point to a corresponding row). Grey squares represent trio combinations that cannot be tested for due to topological constraints.

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., 2006, 2008). Knizhin et al. (2008) 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. (2008) 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-Mcnelly 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 seems 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, 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. tu-*

garinae) 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., 2006; 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 rearrangements (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.

Uncited references

Eaton et al. (2020), Lopes-Lima et al. (2019), Shimodaira and Goldman (2002), Takahashi et al. (2020).

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 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.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

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.

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.

Appendix A. Supplementary material

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

References

- Abell, R., Thieme, M.L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., Coad, B., Mandrak, N., Balderas, S.C., Bussing, W., Stiassny, M.L.J., Skelton, P., Allen, G.R., Unmack, P., Naseka, A., Ng, R., Sindorf, N., Robertson, J., Armijo, E., Higgins, J.V., Heibel, T.J., Wikramanayake, E., Olson, D., López, H.L., Reis, R.E., Lundberg, J.G., Sabaj Pérez, M.H., Petry, P., 2008. Freshwater Ecoregions of the World: A New Map of Biogeographic Units for Freshwater Biodiversity Conservation. *Bioscience* 58, 403–414. <https://doi.org/10.1641/B580507>.
- Alekseyev, S.S., Samusenok, V.P., Matveev, A.N., Andreev, R.S., Samusenok, I. V., Vokin, A.I., Yur'ev, A.L., 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. *J. Ichthyol.* 61, 396–406. <https://doi.org/10.1134/S0032945221030024>.
- Andrews, S., 2010. FastQC: a quality control tool for high throughput sequence data. [<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>].
- Antonov, A.L., 2004. A new species of grayling *Thymallus burejensis* sp. nov. (Thymallidae) from the Amur Basin. *J. Ichthyol.* 44 (6), 401–411.
- Antonov, A.L., Mikheev, I.E., 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. *J. Ichthyol.* 56 (2), 208–216.
- Asmussen, M.A., Arnold, J., Avise, J.C., 2010. Definition and properties of disequilibrium statistics for associations between nuclear and cytoplasmic genotypes. *Mol. Ecol. Evol. Orig. Side Sel. Writings from Avis. Lab.* 419–432. https://doi.org/10.1142/9789814317764_0007.
- Avise, J.C., 1995. Mitochondrial DNA Polymorphism and a Connection Between Genetics and Demography of Relevance to Conservation. *Conserv. Biol.* 9, 686–690. <https://doi.org/10.1046/j.1523-1739.1995.09030686.x>.
- Balbuena, J.A., 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, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. *Mol. Ecol.* 13, 729–744. <https://doi.org/10.1046/j.1365-294X.2003.02063.x>.
- Bangs, M.R., Douglas, M.R., Brunner, P.C., Douglas, M.C., 2020. Reticulate evolution as a management challenge: Patterns of admixture with phylogenetic distance in endemic fishes of western North America. *Evol. Appl.* 13 (6), 1400–1419.
- Bochkarev, N.A., Zuykova, E.I., Pestryakova, L.A., Zakharov, E.S., Romanov, V.I., Sokolov, V.V., Politov, D.V., 2018. Siberian Whitefish (*Coregonus lavaretus pidschian*, Coregonidae) from the Anabar River: Morphogenetic Structure of the Population. *Russ J Genet* 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. *Mol. Phylogenet. Evol.* 40 (3), 856–865. <https://doi.org/10.1016/j.ympcv.2006.04.020>.
- Bushnell, B., 2014. BBMap: A Fast, Accurate, Splice-Aware Aligner. United States. <https://www.osti.gov/biblio/1241166>.
- Cambell, C.S.M., Dutoit, L., King, T.M., Craw, D., Burridge, C.P., Wallis, G.P., Waters, J.M., 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 Distributions* 0,000-000. <https://doi.org/10.1111/ddi.13629>.
- Costedoat, C., Gilles, A., 2009. Quaternary Pattern of Freshwater Fishes in Europe: Comparative Phylogeography and Conservation Perspective. *Open Conserv. Biol. J.* 3, 36–48. <https://doi.org/10.2174/1874839200903010036>.
- De Ré, F.C., Robe, L.J., Wallau, G.L., Loreto, E.L.S., 2017. Inferring the phylogenetic position of the *Drosophila flavopilosa* group: Incongruence within and between mitochondrial and nuclear multilocus datasets. *J. Zool. Syst. Evol. Res.* 55, 208–221. <https://doi.org/10.1111/jzs.12170>.
- Denys, G.P., Dettai, A., Persat, H., Daszkiewicz, P., Hautecoeur, M., Keith, P., 2020. Revision of *Phoxinus* in France with the description of two new species (Teleostei, Leuciscidae). *Cybiurn* 44, 205–237. <https://doi.org/10.26028/cybiurn/2020-443-003i>.
- Dias, R.M., Tófoli, R.M., Barbosa da Silva, J.C., Gomes, L.C., Agostinho, A.A., 2022. Effects of habitat complexity on trophic interactions of three congeneric fish species. *Aquat Ecol* 56, 877–889. <https://doi.org/10.1007/s10452-022-09954-w>.
- Durand, E.Y., Patterson, N., Reich, D., Slatkin, M., 2011. Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* 28, 2239–2252. <https://doi.org/10.1093/molbev/msr048>.
- Eaton, D.A.R., Overcast, I., 2020. ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36, 2592–2594. <https://doi.org/10.1093/bioinformatics/btz966>.
- Englmaier, G.K., Antonov, A., Weiss, S.J., 2022. General patterns of sexual dimorphism in graylings (*Thymallus*), with a comparison to other salmonid species. *Rev. Fish Biol. Fish.* 32, 645–667. <https://doi.org/10.1007/s11160-021-09694-4>.
- Feller, A.F., Haesler, M.P., Peichel, C.L., Seehausen, O., 2020. Genetic architecture of a key reproductive isolation trait differs between sympatric and non-sympatric sister species of Lake Victoria cichlids. *Proc. R. Soc. B Biol. Sci.* 287 (1924), 20200270.
- 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. *Divers. Distrib.* 9 (4), 269–282. <https://doi.org/10.1046/j.1472-4642.2003.00024.x>.
- Froufe, E., Knizhin, I., Koskinen, M.T., Primmer, C.R., Weiss, S., 2003b. Identification of reproductively isolated lineages of Amur grayling (*Thymallus grubii* Dybowski 1869): Concordance between phenotypic and genetic variation. *Mol. Ecol.* 12 (9), 2345–2355.
- Froufe, E., 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. *Mol. Phylogenet. Evol.* 34, 106–117. <https://doi.org/10.1016/j.ympcv.2004.09.009>.
- Froufe, E., Prié, V., Faria, J., Ghamizi, M., Gonçalves, D.V., Gürlek, M.E., Karauzas, I., Kebapçı, 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*). *Mol. Phylogenet. Evol.* 100, 322–332.
- Good, J.M., 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, S.A.M., Cotterill, F.P.D., O'Ryan, C., Skelton, P.H., de Wit, M.J., Gemmill, N.J., 2011. Cryptic diversity of African tigerfish (genus *Hydrocynus*) reveals palaeogeographic signatures of linked Neogene geotectonic events. *PLoS ONE* 6 (12), e28775.
- Goudet, J., 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Mol. Ecol. Notes* 5, 184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>.
- Greenwood, P.H., 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.
- Gundappa, M.K., To, T.H., Grønvald, L., Martin, S.A.M., Lien, S., Geist, J., Hazlerigg, D., Sandve, S.R., Macqueen, D.J., 2022. Genome-Wide Reconstruction of Rediploidization Following Autopolyploidization across One Hundred Million Years of Salmonid Evolution. *Mol. Biol. Evol.* 39. <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*. *Mol. Ecol.* 28, 4046–4064. <https://doi.org/10.1111/mec.15204>.
- Hashemzadeh Segherloo, I., Freyhof, J., Berrebi, P., Ferchaud, A.-L., Geiger, M., Laroche, J., Levin, B.A., Normandeau, E., Bernatchez, L., 2021. A genomic perspective on an old question: *Salmo trouts* or *Salmo trutta* (Teleostei: Salmonidae)? *Mol. Phylogenet. Evol.* 162, 107204.
- Hoang, D., Chernomor, O., von Haeseler, A., Quang Minh, B., Sy Vinh, L., Rosenberg, M.S., 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* 35, 518–522. <https://doi.org/10.105281/zenodo.854445>.
- Hubbs, C.L., 1955. Hybridization between fish species in nature. *Syst. Zool.* 4, 1–20. <https://doi.org/10.2307/sysbio.4.1.1>.
- Hughes, L.C., Cardoso, Y.P., Sommer, J.A., Cifuentes, R., Cuello, M., Somoza, G.M., González-Castro, M., Malabarba, L.R., Cussac, V., Habit, E.M., 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. *Mol. Ecol.* 29, 738–751. <https://doi.org/10.1111/mec.15350>.
- Jacobs, A., Carruthers, M., Yurchenko, A., Gordeeva, N.V., Alekseyev, S.S., Hooker, O., Leong, J.S., Minkley, D.R., Rondeau, E.B., Koop, B.F., Adams, C.E., Elmer, K.R., Buerkle, A., 2020. Parallelism in eco-morphology and gene expression despite variable evolutionary and genomic backgrounds in a Holarctic fish. *PLOS Genet.* 16 (4), e1008658.
- 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>.
- 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, B.Q., Wong, T.K.F., von Haeseler, A., Jermini, S., 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. <https://doi.org/10.1038/nmeth.4285>.
- Kautt, A.F., Kratochwil, C.F., Nater, A., Machado-Schiaffino, G., Olave, M., Henning, F., Torres-Dowdall, J., Härer, A., Hulseley, C.D., Franchini, P., Pippel, M., Myers, E.W., Meyer, A., 2020. Contrasting signatures of genomic divergence during sympatric speciation. *Nature* 588 (7836), 106–111.
- Kerr, K.C.R., Stoeckle, M.Y., Dove, C.J., Weigt, L.A., Francis, C.M., Hebert, P.D.N., 2007. Comprehensive DNA barcode coverage of North American birds. *Mol. Ecol. Notes* 7, 535–543. <https://doi.org/10.1111/j.1471-8286.2006.01670.x>.
- Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum Likelihood estimate of the evolutionary tree topologies from DNA sequence. *J. Mol. Evol.* 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. *J. Mol. Evol.* 31, 151–160. <https://doi.org/10.1007/BF02109483>.
- Knizhin, I.B., Weiss, S.J., Antonov, A.L., Froufe, E., 2004. Morphological and Genetic Diversity of Amur Graylings (*Thymallus*, Thymallidae). *J. Ichthyol.* 44, 52–69. <https://doi.org/10.1134/S0032945207020014>.
- Knizhin, I.B., Weiss, S.J., Sušnik, S., 2006. Graylings of Baikal Lake basin (*Thymallus*, Thymallidae): Diversity of forms and their taxonomic status. *J. Ichthyol.* 46, 418–435. <https://doi.org/10.1134/S0032945206060026>.
- Knizhin, I.B., Weiss, S.J., Bogdanov, B.E., Kopun, T., 2008. New data on the distribution of the Upper Lena form of grayling (Thymallidae) in the basin of Lake Baikal and its taxonomic status. *J. Ichthyol.* 48, 217–223. <https://doi.org/10.1007/s11489-008-3003-y>.
- Kobl Müller, S., Egger, B., Sturm Bauer, C., Sefc, K.M., 2010. Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic Lake Tanganyika cichlid tribe Tropheini. *Mol. Phylogenet. Evol.* 55, 318–334. <https://doi.org/10.1016/j.ympev.2009.09.032>.
- Komatsu, G., Baker, V.R., Arzhannikov, S.G., Gallagher, R., Arzhannikova, A.V., Murana, A., Oguchi, T., 2016. Catastrophic flooding, palaeolakes, and late Quaternary drainage reorganization in northern Eurasia. *Int. Geol. Rev.* 58, 1693–1722. <https://doi.org/10.1080/00206814.2015.1048314>.
- Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., Mayrose, I., 2015. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* 15, 1179–1191. <https://doi.org/10.1111/1755-0998.12387>.
- Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Mol. Biol. Evol.* 29, 1695–1701. <https://doi.org/10.1093/molbev/mss020>.
- Lecaudey, L.A., Schlieven, U.K., Osinov, A.G., Taylor, E.B., Bernatchez, L., Weiss, S.J., 2018. Inferring phylogenetic structure, hybridization and divergence times within Salmoninae (Teleostei: Salmonidae) using RAD-sequencing. *Mol. Phylogenet. Evol.* 124, 82–99. <https://doi.org/10.1016/j.ympev.2018.02.022>.
- Leray, M., Yang, J.Y., Meyer, C.P., Mills, S.C., Agudelo, N., Ranwez, V., Boehm, J.T., Machida, R.J., 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. *Front. Zool.* 10, 1–14. <https://doi.org/10.1186/1742-9994-10-34>.
- 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, Y.L., Liu, J.X., 2018. StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol. Ecol. Resour.* 18, 176–177. <https://doi.org/10.1111/1755-0998.12719>.
- Lien, S., Koop, B.F., Sandve, S.R., Miller, J.R., Kent, M.P., Nome, T., Hvidsten, T.R., Leong, J.S., Minkley, D.R., Zimin, A., Grammes, F., Grove, H., Gjuvsland, A., Walenz, B., Hermansen, R.A., Von Schalburg, K., Rondeau, E.B., Di Genova, A., Samy, J.K.A., Olav Vik, J., Vigeland, M.D., Caler, L., Grimholt, U., Jentoft, S., Inge Våge, D., De Jong, P., Moen, T., Baranski, M., Palti, Y., Smith, D.R., Yorke, J.A., Nederbragt, A.J., Tooming-Klunderud, A., Jakobsen, K.S., Jiang, X., Fan, D., Hu, Y., Liberles, D.A., Vidal, R., Iturrar, P., Jones, S.J.M., Jonassen, I., Maass, A., Omholt, S.W., Davidson, W.S., 2016. The Atlantic salmon genome provides insights into rediploidization. *Nature* 533, 200–205. <https://doi.org/10.1038/nature17164>.
- Lopes-Lima, M., Burlakova, L., Karatayev, A., Gomes-dos-Santos, A., Zieritz, A., Froufe, E., Bogan, A.E., 2019. Revisiting the North American freshwater mussel genus *Quadrula* sensu lato (Bivalvia Unionidae): Phylogeny, taxonomy and species delineation. *Zool. Scr.* 48, 313–336. <https://doi.org/10.1111/zsc.12344>.
- MacGuigan, D.J. Near T. J., 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>.
- Malinsky, M., Svardal, H., Tyers, A.M., Miska, E.A., Genner, M.J., Turner, G.F., Durbin, R., 2018. Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nat. Ecol. Evol.* 2, 1940–1955. <https://doi.org/10.1038/s41559-018-0717-x>.
- Malinsky, M., Matschner, M., Svardal, H., 2021. Dsuite - Fast D-statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour.* 21, 584–595. <https://doi.org/10.1111/1755-0998.13265>.
- Margold, M., Jansen, J.D., Codilean, A.T., Preusser, F., Gurinov, A.L., Fujioka, T., Fink, D., 2018. Repeated megafloods from glacial Lake Vitim, Siberia, to the Arctic Ocean over the past 60,000 years. *Quat. Sci. Rev.* 187, 41–61. <https://doi.org/10.1016/j.quascirev.2018.03.005>.
- Marić, S., Askeyev, I.V., Askeyev, O.V., Monakhov, S.P., 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, A.J., Graziotin, F.G., Zaher, H., Lemmon, A.R., Lemmon, E.R., Parkinson, C.L., 2019. Reticulate evolution in nuclear Middle America causes discordance in the phylogeny of palm-pitvipers (Viperidae: Bothriechis). *J. Biogeogr.* 46 (5), 833–844. <https://doi.org/10.1111/jbi.13542>.
- May-McNally, S.L., Quinn, T.P., Taylor, E.B., 2015. Low levels of hybridization between sympatric Arctic char (*Salvelinus alpinus*) and Dolly Varden char (*Salvelinus malma*) highlights their genetic distinctiveness and ecological segregation. *Ecol. Evol.* 5, 3031–3045. <https://doi.org/10.1002/ece3.1583>.
- Mendes, S.L., Machado, M.P., Coelho, M.M., Sousa, V.C., 2021. Genomic data and multi-species 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., 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. *Ecol. Evol.* 4, 1313–1327. <https://doi.org/10.1002/ece3.931>.
- Meyer, C.P., 2003. Molecular systematics of cowries (Gastropoda: Cypraeidae) and diversification patterns in the tropics. *Biol. J. Linn. Soc.* 79, 401–459. <https://doi.org/10.1046/j.1095-8312.2003.00197.x>.
- Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* 32, 268–274. <https://doi.org/10.1093/molbev/msu300>.
- Olson, K.W., Krabbenhoft, T.J., Hrabik, T.R., Mendsaikhan, B., Jensen, O.P., 2019. Pelagic-littoral resource polymorphism in Hovsgol grayling *Thymallus nigrescens* from Lake Hovsgol, Mongolia. *Ecol. Freshw. Fish* 28, 411–423. <https://doi.org/10.1111/eff.12464>.
- Paetold, C., Wood, K.R., Eaton, D.A.R., Wagner, W.L., Appelhans, M.S., 2019. Phylogeny of Hawaiian Melicope (Rutaceae): RAD-seq Resolves Species Relationships and Reveals Ancient Introgression. *Front. Plant Sci.* 10, 1074. <https://doi.org/10.3389/fpls.2019.01074/BIBTEX>.
- Palandačić, A., Witman, K., Spikmans, F., 2022. Molecular analysis reveals multiple native and alien *Phoxinus* species (Leuciscidae) in the Netherlands and Belgium. *Biol. Invasions* 24 (8), 2273–2283.
- 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>.
- Pérez-Escobar, O.A., Balbuena, J.A., Gottschling, M., 2016. Rumbling Orchids: How To Assess Divergent Evolution Between Chloroplast Endosymbionts and the Nuclear Host. *Syst. Biol.* 65, 51–65. <https://doi.org/10.1093/sysbio/syv070>.
- Persat, H., Weiss, S., Froufe, E., Secci-Petretto, G., Denys, G.P., 2019. A third European species of grayling (Actinopterygii, Salmonidae), endemic to the Loire river basin (France). *Thymallus ligericus* n. sp. *Cybum* 43, 233–238. <https://doi.org/10.26028/cybum.2019-433-004>.
- Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., Orlando, L., 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.
- Pfennig, K.S., 2016. Reinforcement as an initiator of population divergence and speciation. *Curr. Zool.* 62, 145–154. <https://doi.org/10.1093/cz/zow033>.
- Raj, A., Stephens, M., Pritchard, J.K., 2014. FastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics* 197, 573–589. <https://doi.org/10.1534/genetics.114.164350>.
- Robertson, F.M., Gundappa, M.K., Grammes, F., Hvidsten, T.R., Redmond, A.K., Lien, S., Martin, S.A.M., Holland, P.W.H., Sandve, S.R., Macqueen, D.J., 2017. Lineage-specific rediploidization is a mechanism to explain time-lags between genome duplication and evolutionary diversification. *Genome Biol.* 18, 1–14. <https://doi.org/10.1186/s13059-017-1241-z>.
- Roman, I., Bourgeois, Y., Reyes-Velasco, J., Jensen, O.P., Waldman, J., Boissinot, S., 2018. Contrasted patterns of divergence and gene flow among five fish species in a Mongolian rift lake following glaciation. *Biol. J. Linn. Soc.* 125, 115–125. <https://doi.org/10.1093/biolinnean/bly083>.
- Romanov, V.I., 2020. Fauna of grayling fish of Khantai lake. *Bull. Lampetra* 9, 161–177.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574. <https://doi.org/10.1093/bioinformatics/btg180>.
- Rose, J.P., Cassio, A.P., Lemmon, E.P., Lemmon, A.R., Sytsma, K.J., 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. *Syst. Biol.* 70 (1), 162–180. <https://doi.org/10.1093/sysbio/syaa049>.
- Rundle, H.D., Schluter, D., 1998. Reinforcement of stickleback mate preferences: Sympatry breeds contempt. *Evolution* N. Y. 52, 200–208. <https://doi.org/10.1111/j.1558-5646.1998.tb05153.x>.
- Sävilämmi, T., Primmer, C.R., Varadharajan, S., Guyomard, R., Guiguen, Y., Sandve, S.R., 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. Genet.* 9, 1283–1294. <https://doi.org/10.1534/g3.118.200919>.
- Schedel, F.D.B., 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 dissertation). University of Monaco, Germany.
- Scribner, K.T., Avise, J.C., 1994. Cytonuclear genetics of experimental fish hybrid zones inside biosphere 2. *Proc. Natl. Acad. Sci. U. S. A.* 91, 5066–5069. <https://doi.org/10.1073/pnas.91.11.5066>.
- Scribner, K.T., Page, K.S., Bartron, M.L., 2000. Hybridization in freshwater fishes: A review of case studies and cytonuclear methods of biological inference. *Rev. Fish Biol. Fish.* 10, 293–323. <https://doi.org/10.1023/A:1016642723238>.
- Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* 51, 492–508. <https://doi.org/10.1080/10635150290069913>.
- Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-Likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1116. <https://doi.org/10.1093/oxfordjournals.molbev.a026201>.
- Simonsen, M.K., Siwertsson, A., Adams, C.E., Amundsen, P.A., Praelik, 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>.
- Strimmer, K., Rambaut, A., 2002. Inferring confidence sets of possibly misspecified gene trees. *Proc. R. Soc. B Biol. Sci.* 269, 137–142. <https://doi.org/10.1098/rspb.2001.1862>.
- Sušnik, S., Berrebi, P., Dovč, P., Hansen, M.M., 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* (Edinb). 93, 273–282. <https://doi.org/10.1038/sj.hdy.6800500>.
- Sušnik, S., Weiss, S., Odak, T., Delliling, B., Treer, T., Snoj, A., 2006. Reticulate evolution: ancient introgression of the Adriatic brown trout mtDNA in softmouth trout *Salmo obtusirostris* (Teleostei: Salmonidae). *Biol. J. Linn. Soc.* 90, 139–152. <https://doi.org/10.1111/j.1095-8312.2007.00717.x>.
- Takahashi, T., Nagano, A.J., Kawaguchi, L., Onikura, N., Nakajima, J., Miyake, T., Suzuki, N., Kanoh, Y., Tsuruta, T., Tanimoto, T., Yasui, Y., Oshima, N., Kawamura, K., 2020. A ddRAD-based population genetics and phylogenetics of an endangered freshwater fish from Japan. *Conserv. Genet.* 21 (4), 641–652.
- Team, R.C., 2021. R: A language and environment for statistical computing.
- Toews, D.P.L., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* 21, 3907–3930. <https://doi.org/10.1111/J.1365-294X.2012.05664.X>.
- Wallis, G.P., Cameron-Christie, S.R., Kennedy, H.L., Palmer, G., Sanders, T.R., Winter, D.J., 2017. Interspecific hybridization causes long-term phylogenetic discordance between nuclear and mitochondrial genomes in freshwater fishes. *Mol. Ecol.* 26 (12), 3116–3127.
- 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. *Mol. Ecol.* 11, 1393–1407. <https://doi.org/10.1046/j.1365-294X.2002.01544.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. *J. Fish Biol.* 71, 371–386. <https://doi.org/10.1111/j.1095-8649.2007.01662.x>.
- Weiss, S., Grimm, J., Gonçalves, D.V., Secci-Petretto, G., Englmaier, G.K., 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.J., 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. *J. Fish Biol.* 82, 505–521. <https://doi.org/10.1111/jfb.12007>.
- Weiss, S.J., Gonçalves, D.V., Secci-Petretto, G., Englmaier, G.K., Gomes-Dos-Santos, A., Denys, G.P.J., Persat, H., Antonov, A., Hahn, C., Taylor, E.B., Froufe, E., 2021. Global systematic diversity, range distributions, conservation and taxonomic assessments of graylings (Teleostei: Salmonidae; *Thymallus* spp.). *Org. Divers. Evol.* 21 (1), 25–42.
- 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. *Zool. Scr.* 49, 117–128. <https://doi.org/10.1111/zsc.12393>.
- Winstanley, C., Rapley, R., 2003. Extraction and Purification of Plasmid DNA. *Nucleic Acid Protoc. Handbook* 1 (1), 327–331. <https://doi.org/10.1385/1-59259-038-1:327>.