

RESEARCH ARTICLE

SPECIAL ISSUE
UNCOVERING DEVELOPMENTAL DIVERSITY

Phylogenomic analysis of the Lake Kronotskoe species flock of Dolly Varden charr reveals genetic and developmental signatures of sympatric radiation

Katherine C. Woronowicz¹, Evgeny V. Esin², Grigorii N. Markevich², Crisvelly Soto Martinez³, Sarah K. McMenamin³, Jacob M. Daane⁴, Matthew P. Harris^{1,*} and Fedor N. Shkil^{5,6,*}

ABSTRACT

Recent adaptive radiations provide experimental opportunities to parse the relationship between genomic variation and the origins of distinct phenotypes. Sympatric radiations of the charr complex (genus *Salvelinus*) present a trove for phylogenetic analyses as charrs have repeatedly diversified into multiple morphs with distinct feeding specializations. However, charr species flocks normally comprise only two to three lineages. Dolly Varden charr inhabiting Lake Kronotskoe represent the most extensive radiation described for the genus, containing at least seven lineages, each with defining morphological and ecological traits. Here, we perform the first genome-wide analysis of this species flock to parse the foundations of adaptive change. Our data support distinct, reproductively isolated lineages within the clade. We find that changes in genes associated with thyroid signaling and craniofacial development provided a foundational shift in evolution to the lake. The thyroid axis is further implicated in subsequent lineage partitioning events. These results delineate a genetic scenario for the diversification of specialized lineages and highlight a common axis of change biasing the generation of specific forms during adaptive radiation.

KEY WORDS: Craniofacial, Evolution, Thyroid, Variation

INTRODUCTION

The Salmonid fishes of the genus *Salvelinus* represent an exceptional example of parallel evolution and trophic adaptation to new environments (Klemetsen, 2010). Charr are remarkably variable with naturally occurring, phenotypically distinct populations that exhibit diverse life histories. These include complete and partial anadromy, as well as freshwater riverine and lacustrine residency (Taylor, 2016; Lecaudey et al., 2018; Osinov et al., 2021). Notably,

freshwater resident populations frequently establish species flocks showing stereotypical morphologies, each associated with an ecological niche (Nordeng, 1983; Walker et al., 1988; Sandlund et al., 1992; Maitland et al., 2007; Simonsen et al., 2017; Markevich et al., 2018; Esin et al., 2020; Jacobs et al., 2020). Even within sympatric populations, lineages with varying diets, depth preferences, and disparate spawning intervals are commonly observed (Jonsson and Jonsson, 2001; Klemetsen, 2010). Within independent lacustrine *Salvelinus* radiations, specific morphological adaptations have repeatedly evolved (Klemetsen, 2010). Such a propensity for repeated, independent radiations make the genus *Salvelinus* an attractive system to dissect the mechanisms that facilitate rapid generation of morphological variation.

Salvelinus charrs have repeatedly radiated into sympatric morphs with distinct evolved feeding specializations, akin to those of both African and South American cichlids (Barluenga et al., 2006; Malinsky et al., 2015, 2018) and African and Asian barbs (Myers, 1960; Nagelkerke and Sibbing, 2000; Levin et al., 2020). Variable traits include jaw size, mouth position, eye size, pigmentation, and others. The variability within and between charr radiations proffers important genetic case studies to uncover mechanisms of rapid morphological diversification. Such parallel events suggest that these populations experience similar selective pressures, and/or that there are genetic biases that limit the resulting morphologies. The contribution of standing variation from ancestral populations has been suggested as a mechanism for such phenotypic convergence in other radiations, including adaptation of stickleback populations to freshwater environments (Schluter and Conte, 2009). In contrast, the genetic mechanisms that enable charr to generate distinct, but repeated morphologies have not yet been well defined. It has been argued that shifts in developmental timing may underlie these common transitions to specialized morphologies among charrs (Esin et al., 2018). In this model, development biases radiations by constraining a common axis of change. Early and pleiotropic developmental shifts may give charr an exceptional capacity for adaptive radiation with only a small number of genetic modifications.

The Lake Kronotskoe radiation of Dolly Varden (*Salvelinus malma*) is unique among charrs, as it contains at least seven reproductively isolated phenotypes associating with different ecological niches (Markevich et al., 2018). This radiation is currently the broadest observed within resident lacustrine populations of this genus and, indeed, among salmonids more broadly (Markevich et al., 2018). Recent work has characterized genome-wide changes in species flocks of Alpine whitefishes (*Coregonus* spp.); however, those flocks comprise up to six species with varying morphologies (De-Kayne et al., 2022). Lake Kronotskoe is a volcanogenic lake of the Kamchatka peninsula (Fig. 1A) that formed approximately 12,000 years ago after a

¹Department of Orthopedics, Boston Children's Hospital and Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²Laboratory of Lower Vertebrate Ecology, Severtsov Institute, Moscow 119071, Russian Federation.

³Department of Biology, Boston College, Chestnut Hill, MA 02467, USA.

⁴Department of Biology and Biochemistry, University of Houston, Houston, TX 77204, USA. ⁵Laboratory of Evolutionary Morphology, Severtsov Institute, Moscow 119071, Russian Federation. ⁶Laboratory of Postembryonic Development, Koltzov Institute, Moscow 119071, Russian Federation.

*Authors for correspondence (harris@genetics.med.harvard.edu; fedorshkil@gmail.com)

© K.C.W., 0000-0002-5323-1617; E.V.E., 0000-0002-1200-6464; G.N.M., 0000-0002-6893-4286; C.S.M., 0009-0006-3148-2059; S.K.M., 0000-0002-1154-5810; J.M.D., 0000-0003-3631-4514; M.P.H., 0000-0002-7201-4693; F.N.S., 0000-0002-5637-8278

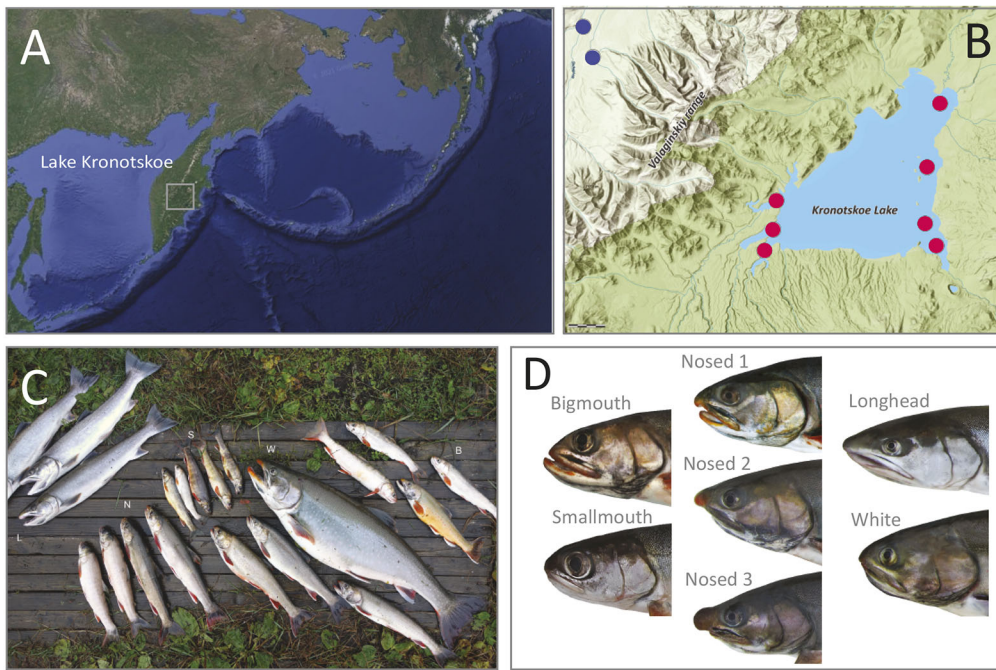


Fig. 1. Lake Kronotskoe radiation of Dolly Varden charr. (A) A map of the Kamchatka Peninsula. Lake Kronotskoe is centered within the highlighted box. (B) A map of Lake Kronotskoe geography and locations from which anadromous (blue dots) and resident lacustrine (red dots) Dolly Varden morphs were collected. (C) The Lake Kronotskoe morphs from left to right: L, longhead; N, nosed lineages; S, smallmouth; W, white; B, bigmouth. (D) Detailed images of representative adult heads for the seven sequenced lineages of the species flock.

lava flow dammed the ancestral river (Fig. 1B) (Braitseva et al., 1995). The resulting landlocked Dolly Varden population diversified within the new lacustrine environment, and encompassed prominent changes in craniofacial form supporting new feeding strategies (Fig. 1C,D) (Markevich et al., 2018; Esin et al., 2020). A major axis of change is proportional jaw length, seen in piscivorous longhead and deep-water benthivorous bigmouth morphs, as well as the modulation of frontonasal proportions found in the nosed morphs. The deep-water omnivorous smallmouth morph has an increase in eye size in relative proportion to the cranium as well as reduced jaw size (Fig. 1C,D). Enlarged eye size is widespread among deep-dwelling resident charr lineages in other lakes and may suggest specialization for foraging in low light conditions (Klemetsen, 2010). The varied Lake Kronotskoe charrs show strict natal homing with some morphs spawning in tributaries and other morphs spawning within the lake (Markevich et al., 2021). The morphs also exhibit differential annual timing windows for reproduction, which reinforces reproductive isolation and facilitates sympatric radiation (Esin et al., 2021a,b).

Here, we employ phylogenomics to dissect the genetic context of the Dolly Varden radiation within Lake Kronotskoe. We parse the shared genetic foundation for the morphological and physiological adaptations of these specialized lineages.

RESULTS

Genome-wide assessment of variability within and between Lake Kronotskoe Dolly Varden morphs

To investigate variation throughout the charr genome, we performed genome-wide targeted capture of coding and non-coding loci of Dolly Varden charr from Lake Kronotskoe using a custom, pan-Salmonidae capture array. Targeted capture of small population pools allows identification of lineage-defining variation via cross-clade sequence comparisons. Similar approaches have been recently used to assess variation in Belonifomes (Daane et al., 2021), rockfishes (Treaster et al., 2023) and notothenioids (Daane et al., 2019), allowing for clade-wide analysis of genomic variation and phylogenetic parsing of selective signals.

The pan-Salmonidae capture array consists of 558,882 bait sequences designed against conserved Atlantic salmon (*Salmo*

salar) and rainbow trout (*Oncorhynchus mykiss*) genomic sequences and targets 97 Mb of total sequence (Fig. S1). Coding regions constitute 83.8% of targeted regions with conserved non-coding, ultraconserved non-coding, and miRNA hairpins comprising the remaining 16.2% of targeted regions. With our capture methodology and analysis pipeline, bait sequences successfully hybridize with targeted elements harboring up to 15% deviation in sequence identity (Mason et al., 2011), thereby allowing recovery and subsequent analysis of sequence variability. Pairwise analyses between homologous loci were used to distinguish fixed and variable loci between morphs. Genes associating with variation were identified by making strategic comparisons within a phylogeny focused on character diversification (Daane et al., 2015). For outgroup analysis, we sampled anadromous Dolly Varden charr from the local Kamchatka river basin, which is adjacent to, but isolated from, that of Lake Kronotskoe (Fig. 1B). We also sequenced a single *Salvelinus leucomaenis* individual, which was collected from the same catchment as the Dolly Varden charr to serve as an outgroup.

For each group, we recovered approximately 90% of the targeted elements sequenced to a mean depth of 25–52 reads (Table S1). We detected substantial variation within each lineage as well as the putative ancestral population (avg π in Dolly Varden=0.003), supporting pairwise analysis to detect and compare shared and unique variation within each lineage. Loci under-represented in our capture were limited and represented as general gene classes (Table S2). These elements recovered with low coverage are similar to those observed in other broad capture approaches (Daane et al., 2019, 2021; Treaster et al., 2023).

Genetic differentiation and relationships among Dolly Varden morphs

As a first approach, we conducted principal components analysis (PCA) to visualize the relationships among our lacustrine and anadromous Dolly Varden samples and outgroup *S. leucomaenis*. PC1 separates *S. leucomaenis* from Dolly Varden and from the members of the Dolly Varden species flock (Fig. 2A), and PC2 broadly groups lineages according to ecological niche. Notably, all

three nosed morphs cluster tightly, suggesting that they are closely related.

We investigated relationships among Lake Kronotskoe charrs through reconstruction of the phylogeny given our substantial sequence data. We used IQ-TREE to derive a phylogeny from a dataset containing 622,831 variant sites, of which 22,701 variants were informative (Nguyen et al., 2015). Prior interpretations of this radiation argued for multiple-step diversification based on changes in resource utilization and accompanying feeding specializations leading to the extant morphs (Markevich et al., 2018). Our phylogenomic data support that each described morphology within the species flock represents a well-differentiated lineage, with each node having high bootstrap support (Fig. 2B). Our analysis clusters longhead and white morphs within a lineage that is an outgroup to the clade consisting of the bigmouth, smallmouth and nosed morphs. Nosed lineages differentiate as a distinct clade with nosed 2 as an outgroup to nosed lineages 1 and 3. As the nosed 2 population associates so closely with nosed 3 via PCA (Fig. 2A), we focused subsequent analyses on disentangling the differentiation between nosed morphs 1 and 3.

The anadromous Dolly Varden and lacustrine morphs are clearly genetically differentiated within ecologically specialized lineages, having pairwise mean fixation index (F_{st}) on par with distinct populations of alpine whitefishes (Fig. 2C) (Table S5) (Vonlanthen et al., 2012). The greatest pairwise differentiation was found between the anadromous Dolly Varden and the lacustrine deep-water bigmouth morph ($F_{st}=0.127$). Notably, each lacustrine lineage is more differentiated from the anadromous Dolly Varden population than from any other lacustrine lineage. Further, lacustrine lineages are similarly differentiated from the putative ancestral riverine Dolly Varden population. Within the species flock, the least differentiated pairing was between the lacustrine nosed 1 and nosed 3 morphs ($F_{st}=0.047$), consistent with their phylogenetic relationship. Although informative of broad patterns of divergence, these values are likely underestimates of genetic differentiation owing to the conserved element-based dataset biasing analysis to regions having an inherent constraint on variation.

Introgression analysis within the clade further supports the existence of distinct lineages, although some incomplete lineage sorting was detected. Significant introgression was identified in 80% of the 20 possible trios (DSuite; Holm-Bonferroni $FWER<0.01$) in patterns which deviated from the topology of the phylogeny (Table S3) (Malinsky et al., 2021). Despite observed significance, the introgression values were relatively small compared to the recent timeline of the Lake Kronotskoe radiation. The bigmouth and nosed 1 morphs showed the greatest excess of allele sharing ($D_{tree}=4.2\%$) (Table S3). We further calculated f_4 -admixture ratios and used f -branch statistics to disentangle the interrelatedness of admixture signals among morphs that share common internal phylogenetic branches (Fig. 2D). The greatest proportions of shared alleles were between the bigmouth morph and the nosed 1 morph (f -branch=9.0%), between the bigmouth morph and the nosed 3 morph (f -branch=8.3%), and between the white morph and the ancestor of the nosed 1 and nosed 3 morphs (f -branch=5.9%). The f -branch statistic further supported the interpretation that Lake Kronotskoe lineages differentiated while maintaining relative genetic isolation.

To determine the extent to which nosed 1 and nosed 3 share common introgressed loci with bigmouth, we calculated D in sliding windows (40 variant windows, 20 variants step-size) for smallmouth, bigmouth, nosed 1 (mean=0.046, s.d.=0.22) and smallmouth, bigmouth, nosed 3 (mean=0.041, s.d.=0.21). Anadromous Dolly Varden served as the outgroup for all analyses of introgression.

Among sliding windows with $D\geq 0.8$ (32 of 10,617 windows for smallmouth, bigmouth, nosed 1; 28 of 10,790 windows for smallmouth, bigmouth, nosed 3), we observed shared, isolated regions of the genome with signatures of admixture (Table S6). Among these, we detected a large interval with evidence for shared introgression between bigmouth and nosed morphs 1 and 3 (mean D for the interval=0.53 for smallmouth, bigmouth, nosed 1; 0.54 for smallmouth, bigmouth, nosed 3) with more than one-third of windows with $D>3$ s.d. from the mean (Fig. S2). However, most introgressed intervals were more spatially restricted. One locus potentially associating with craniofacial differentiation spanned an interval centered on *zyg11b*, which has been implicated in craniofacial microsomia (Fig. 2E) (Tingaud-Sequeira et al., 2020). Another locus included multiple genes of interest, including *zswim5*, which is expressed in cranial neural crest (Wong et al., 2016), *rnf152*, which is involved in regulating neural crest formation (Yoon et al., 2022), and *mc4r*, which is a crucial regulator of appetite and metabolism via leptin and thyroid signaling modulation (Fig. 2F) (Decherf et al., 2010). We also investigated coding changes within these introgressed regions and identified variation of interest in *mc4r*. The particular residue is broadly conserved as proline among fishes, reptiles, birds, and mammals, yet the riverine Dolly Varden variably encode arginine at this position (Fig. 2F'). This residue is differentially present among lake lineages, being found variable in bigmouth but fixed in white and nosed. This amino acid substitution is predicted to affect protein function by the SIFT algorithm. This variant residue is fixed in smallmouth charr, which are generally reduced in size compared with their congeners. This suggests that this *mc4r* variant was inherited from standing variation in the founding Dolly Varden population and is consistent with an altered function of metabolic regulation in this lineage controlled by *mc4r*. This suite of craniofacial and metabolic genes may have functioned like a superlocus that was spread via shared introgression between the bigmouth and nosed lineages and subsequently reinforced the differentiation of distinct lineages.

Patterns of differentiation between river and lake populations of Dolly Varden charr

To understand shared variation that differentiates the Lake Kronotskoe residents from the riverine anadromous population of Dolly Varden, pairwise F_{st} was calculated per coding (25,373 genes) and conserved non-coding elements (CNEs) (22,575 CNEs) from our targeted capture (Fig. 3A,B). F_{st} values were computed by grouping all variation contained within the lacustrine species flock and comparing against the variation contained in the riverine anadromous Dolly Varden charr population (lake versus river). There were 327 genes (Table S7) and 80 CNEs (Table S8) with $F_{st}>0.5$. Each CNE was assigned to flanking target genes by GREAT (McLean et al., 2010). Some intervals harbored multiple highly differentiating CNEs. For example, the close proximity of two CNEs led to two hits for *cholinesterase-like* and *dync2h1* (Fig. 3B). Interestingly, craniofacial (orange) and thyroid-related (blue) genes (five craniofacial genes of top 20, two thyroid genes of top 20) and CNEs (12 craniofacial CNEs of top 20, three thyroid CNEs of top 20) were highly represented across the most differentiating elements (Fig. 3A,B).

Craniofacial genes assigned to highly differentiating CNEs by GREAT analyses included both those with predicted roles in neural crest (*tfap2a*) (Rothstein and Simoes-Costa, 2020), as well as frontonasal (*meis2*, *pitx2*) (Evans and Gage, 2005; Fabik et al., 2020) and splanchnocranium (*faf1*) development (Ma et al., 2017). Of note, *pitx2* is also involved in release of thyroid stimulating hormone from the pituitary (Castinetti et al., 2011). *dync2h1* and



implicated in craniofacial development (*ofcc1* and *tfap2a*; *meis2* and *spred1*; *fam172a* and *nr2f1a*). Genes connected to ion transport (*myh6*, *khnyn*, *or11a1*, *endod1*, *cacnb2* and *hrh1*) and vesicular transport (*man2a1*, *shld2*, *gdpd2*, *onecut2* and *clint1a*) according to KEGG and Reactome reconstructions were also well represented

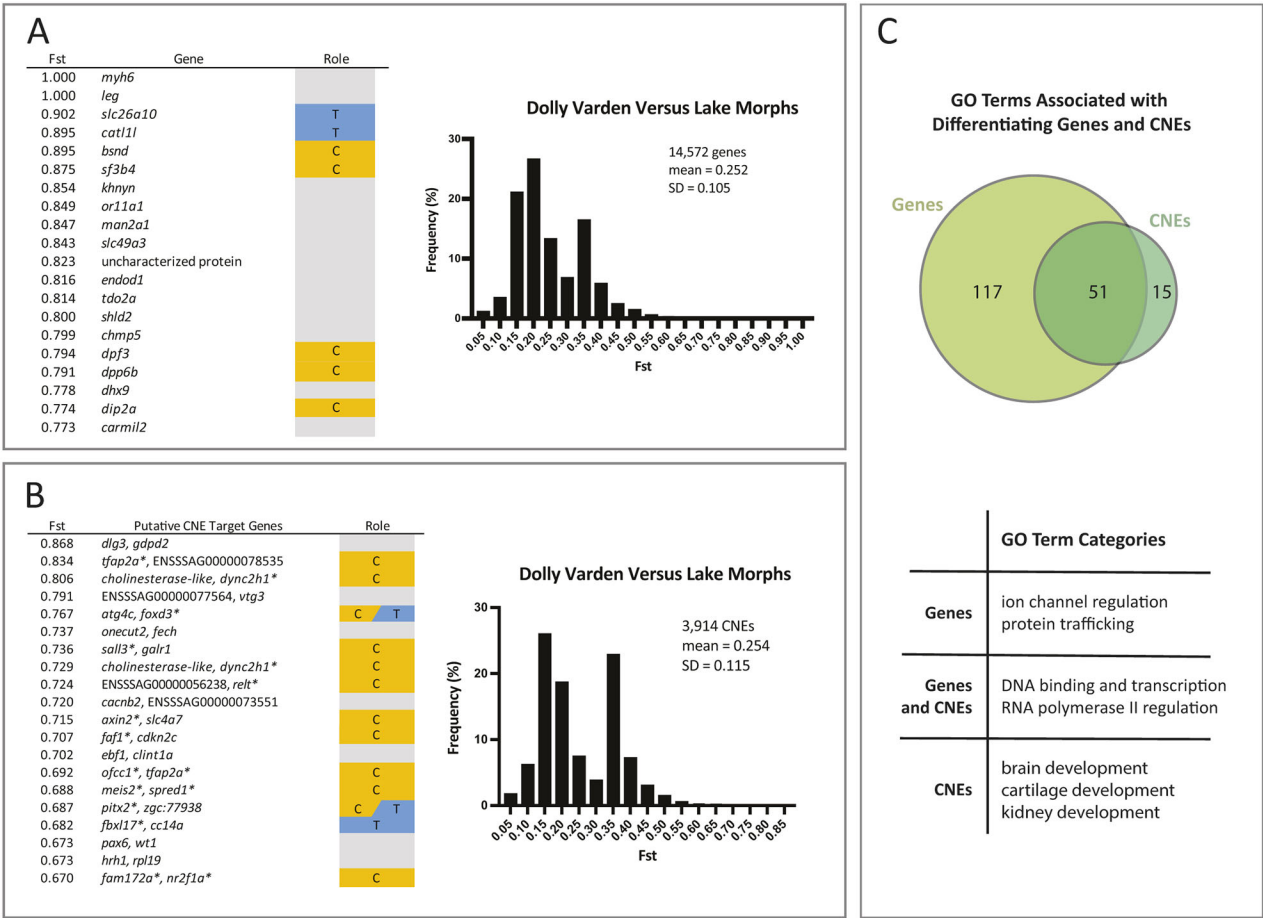


Fig. 3. Coding and non-coding conserved elements differentiating anadromous Dolly Varden charr from the Lake Kronotskoe species flock. (A) Table showing the top 20 genes differentiating riverine Dolly Varden from Lake Kronotskoe inhabitants along with distribution of Fst values of genes. (B) Table of the top 20 differentiating CNEs and the distribution of Fst values. (C) Overview of shared and unique GO terms associated with genes and CNEs with Fst>0.5. Venn diagram values represent GO terms that occur six or more times. Table summarizes broad categories of most frequent GO term associations. The distributions and the numbers of elements represent all non-zero Fst values; C (orange), genes with known roles in modulating craniofacial morphology; T (blue), genes with known roles in thyroid function. Asterisk denotes gene classified as C, T, or both.

among the top differentiating genes and as targets of the top differentiating CNEs. The majority of nucleotide changes identified were single single nucleotide polymorphisms (SNPs) with unknown functional significance. However, high Fst is reflective of potential selection or bottleneck at the locus.

Gene Ontology (GO) terms were assigned to all genes (2825 GO terms) and CNEs (1419 GO terms) that differentiate anadromous Dolly Varden from Lake Kronotskoe residents (Fst>0.5). Among GO terms appearing six or more times, specific themes emerged (Fig. 3C). Highly differentiating genes were associated with roles in ion channel regulation and protein trafficking. Highly differentiating CNEs were associated with brain, kidney and cartilage development. Other GO terms shared among highly differentiating coding and non-coding elements included functions such as DNA binding, transcription factors, and regulation of RNA polymerase II (Table S9).

Within the candidate genes, there were signals related to regulation of thyroid signaling in development. For example, the gene *slc26a10* encodes a sulfate transporter that functions in thyroid hormone synthesis and also acts downstream of Thyroid hormone receptor alpha (THRα) (Richard et al., 2020). We found that *slc26a10* has segregating nonsynonymous changes in conserved residues that are highly differentiated between riverine Dolly Varden and lacustrine morphs (Fig. 4). However, the function of this

gene is not well understood, and alignment-based analyses predict this amino acid substitution to have little effect on function (neutral by PROVEAN, tolerated by SIFT) (Ng and Henikoff, 2003; Choi and Chan, 2015). By Fst score, the next candidate gene, *cat11l*, encodes Cathepsin L1-like and is orthologous to Cathepsin L1. This gene may be involved in the processing of thyroid prohormone (Friedrichs et al., 2003). Next, Splicing factor 3b subunit 4 (*sf3b4*) is a key developmental regulator of frontonasal patterning and growth and is associated with craniofacial Nager syndrome in humans (Bernier et al., 2012; Petit et al., 2014). Within this locus, we found synonymous variants at high Fst in all lacustrine morphs, and fixed within benthic groups (Fig. 5). These elevated footprints of clustered, highly differentiating variants may indicate the presence of further modifications within poorly conserved non-coding regions in linkage to these sites, which were not targeted by our bait sequences.

Thyroid hormone signaling activity is associated with shifts in craniofacial form

As modulation of thyroid signaling has been implicated in phenotypic specialization of charrs in different environments (Esin et al., 2021a,b), the pattern of fixation in thyroid-associated genes and non-coding elements in lacustrine morphs compared with riverine Dolly

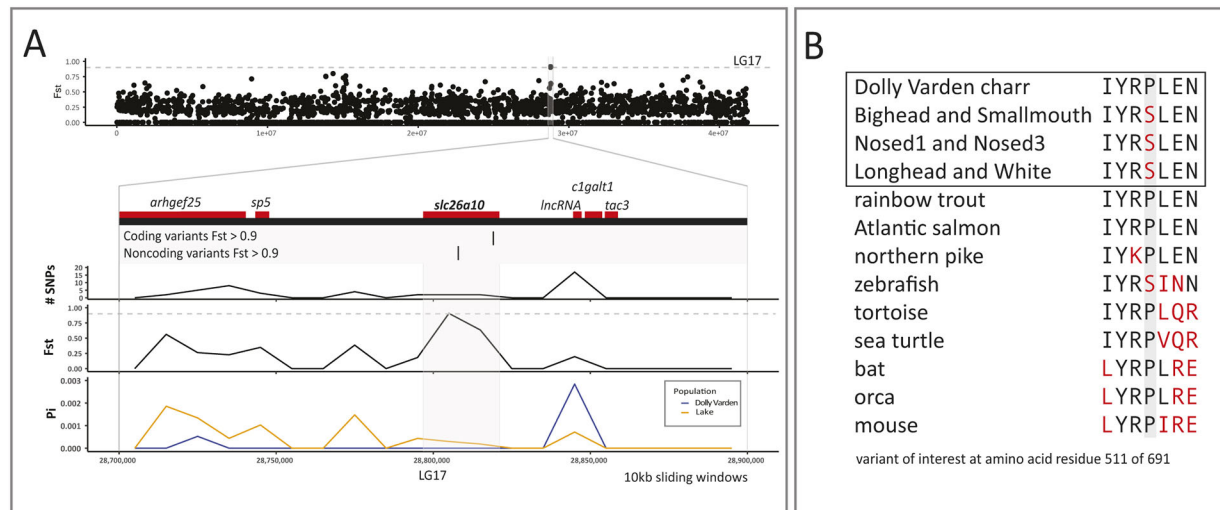


Fig. 4. Differentiation of *slc26a10* in pairwise comparisons between anadromous Dolly Varden charr and the Lake Kronotskoe species flock.

(A) The *slc26a10* locus shows high differentiation (F_{st}). The gene locus contains one highly differentiating non-coding and one coding variant. Dolly Varden *slc26a10* is homologous to human pendrin (SLC26A4), a known thyroid regulator. The broader locus has low nucleotide diversity as illustrated by sliding window plots of Tajima's P_i . Horizontal dashed gray line represents $F_{st}=0.9$. Plots are included for 10 kb sliding windows along LG17, the coding elements within the broader locus, the number of variants per sliding window, F_{st} , and Tajima's P_i per sliding window in pairwise comparisons between riverine Dolly Varden charr and the Lake Kronotskoe species flock. (B) *Slc26a10* contains a fixed amino acid substitution in a conserved proline that differentiates Dolly Varden charr from each of the major clades of the lacustrine morphs.

Varden warranted further investigation. To address whether changes in thyroid metabolism are associated with different morphs in Lake Kronotskoe, we assessed levels of circulating thyroid hormone in adult riverine and lacustrine charr individuals. Intriguingly, we found a significant decrease in triiodothyronine (T_3 ; the most genomically active form of thyroid hormone) hormone in specific lacustrine populations (Fig. 6A). The pattern of reduced T_3 abundance across the species flock correlates with a clear change in craniofacial proportions: nosed and smallmouth morphs, with sub-terminally positioned mouths, have significantly decreased T_3 levels. This is in stark contrast to riverine Dolly Varden and the lacustrine White

morph, which have comparatively 'wild-type' craniofacial form, and the piscivorous longhead and deep-water benthivorous bigmouth morphs, which are highly specialized (Fig. 1D).

Our findings identify a disproportionate number of differentiated loci in nosed lineages that are known to regulate thyroid signaling (Fig. 6B). We assessed specific mutations identified as fixed or nearly fixed ($F_{st}>0.9$) within nosed morphs that differentiate them from longhead or white morphs. A candidate locus that differentiates the low T_3 nosed 1 and nosed 3 morphs from the longhead and white outgroups is a previously unknown *leptin* homolog (Fig. 6C). All salmonids have two *leptin* A ohnologs

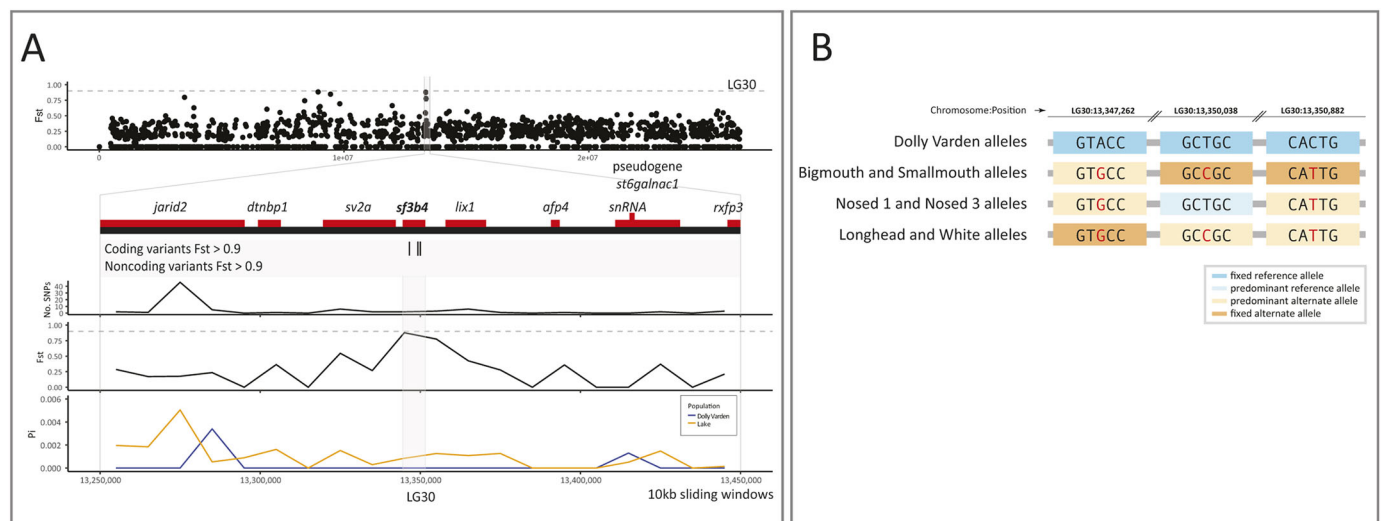


Fig. 5. Fixation of variation in *sf3b4* between anadromous Dolly Varden charr and the Lake Kronotskoe species flock. (A) Three highly differentiating synonymous variants lie within *sf3b4*. F_{st} is plotted in non-overlapping 10 kb sliding windows along the length of the chromosome LG30. *sf3b4* is contained within a 200 kb interval of low nucleotide diversity as illustrated by sliding window plots of Tajima's P_i . Horizontal dashed gray line represents $F_{st}=0.9$. The broad locus encompassing *sf3b4* is shown in detail including plots of the number of variants, F_{st} , and Tajima's P_i in 10 kb non-overlapping sliding windows. (B) Three synonymous variants in *sf3b4* are fixed in lacustrine Dolly Varden charr; dark orange, non-reference allele fixed; light orange, non-reference allele predominant (alternate allele frequency >50%) in sequence pool; light blue, a lineage for which the reference allele is predominant (reference allele frequency >50%) at the locus.

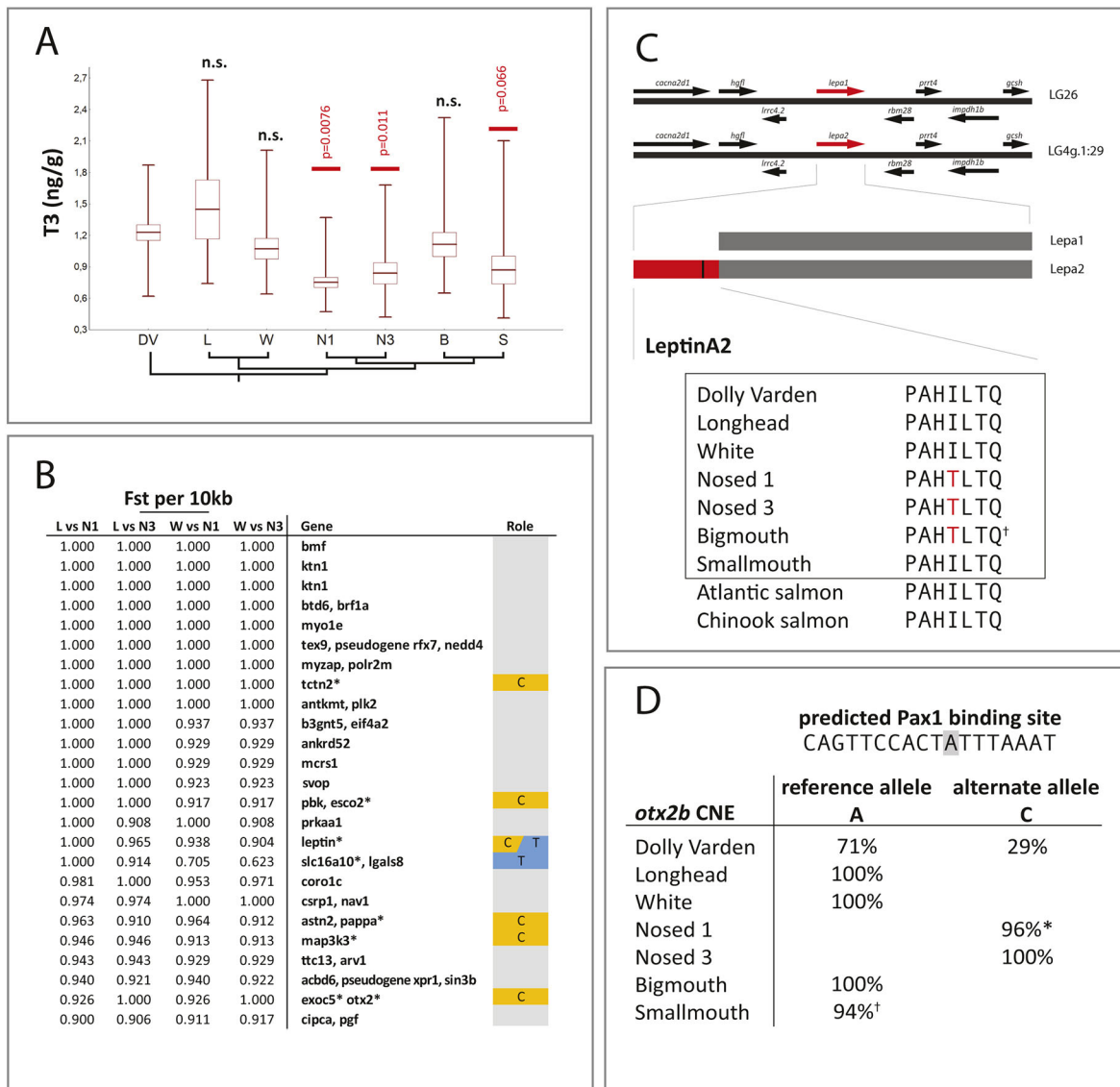


Fig. 6. Thyroid hormone (T_3) levels and associating highly differentiating candidates in lacustrine Dolly Varden morphs related to thyroid function and craniofacial morphology. (A) Serum T_3 levels were significantly lower in nosed 1, nosed 3 and smallmouth lineages relative to anadromous Dolly Varden; mean $\pm$ 1 s.d., with boundary values indicated (min/max). $n=16$ bigmouth, 17 Dolly Varden, 7 longhead, 24 nosed 1, 14 nosed 3, 17 smallmouth, and 16 white morphs. n.s., not significant. P -values determined by one-way ANOVA with post-hoc Tukey honestly significant difference (HSD) test. (B) Top 20 candidate regions differentiating nosed lineages from longhead and white lineages. Sites uniquely differentiating longhead morphs from nosed 1 and nosed 3 morphs, and white morphs from nosed 1 and nosed 3 morphs identify loci associated with thyroid function (blue, T) or craniofacial developmental (orange, C) modulating elements. L, longhead; N1, nosed 1; N3, nosed 3; W, white. (C) Schematic of Leptin A paralogs in salmonids detailing a previously unknown gene, *leptina2*, encoding a unique 66 amino acid N-terminus extension compared to its paralog. Nosed lineages contain a fixed, non-synonymous SNP in this conserved N-terminal sequence. (D) An *otx2b* CNE contains fixed or highly differentiating variants within a predicted Pax1-binding domain that associates with lake morphs exhibiting significantly different thyroid signaling activities; percentage of reference and alternate allele reads indicated per morph. Low-level detection of variant or reference alleles noted [one read for reference allele (asterisk), two reads for non-reference allele (dagger)].

derived from the shared salmonid whole-genome duplication. Leptin A2 encodes an N-terminal sequence extended by 66 amino acids compared to the Leptin A1 or B orthologs within salmonids. The nonsynonymous change differentiating nosed lineages lies within a conserved residue of the unique sequence of this leptin paralog. As leptin reciprocally modulates thyroid hormone activity, both endocrine signaling pathways affect global metabolic activity. Notably, *mc4r*, which shows evidence of allele sharing between the bigmouth and nosed lineages, and harbors coding variation within the radiation, serves as a relay through which leptin stimulates the thyroid axis, suggesting that these hormonal axes may be modulated at multiple levels in differentiated lineages (Decherf et al., 2010).

Exogenous leptin and thyroid hormone treatments are associated with enlarged craniofacial dimensions, marking an intriguing relationship between these pathways and nosed lineage diversification (Yagasaki et al., 2003; Zimmermann-Belsing et al., 2003; Copeland et al., 2011; Shkil et al., 2012; Keer et al., 2019).

We also identified a highly differentiated variant within a conserved non-coding enhancer of *otx2b* (Fig. 6D). Nosed 1 and nosed 3 morphs are fixed or nearly fixed for a variant allele that lies in a predicted Pax1 transcription factor binding site. *otx2b* is involved in development of the skull and the anterior pituitary gland, which regulates hormonal signaling including the thyroid axis (Diaczok et al., 2008; Bando et al., 2020). Due to the significant

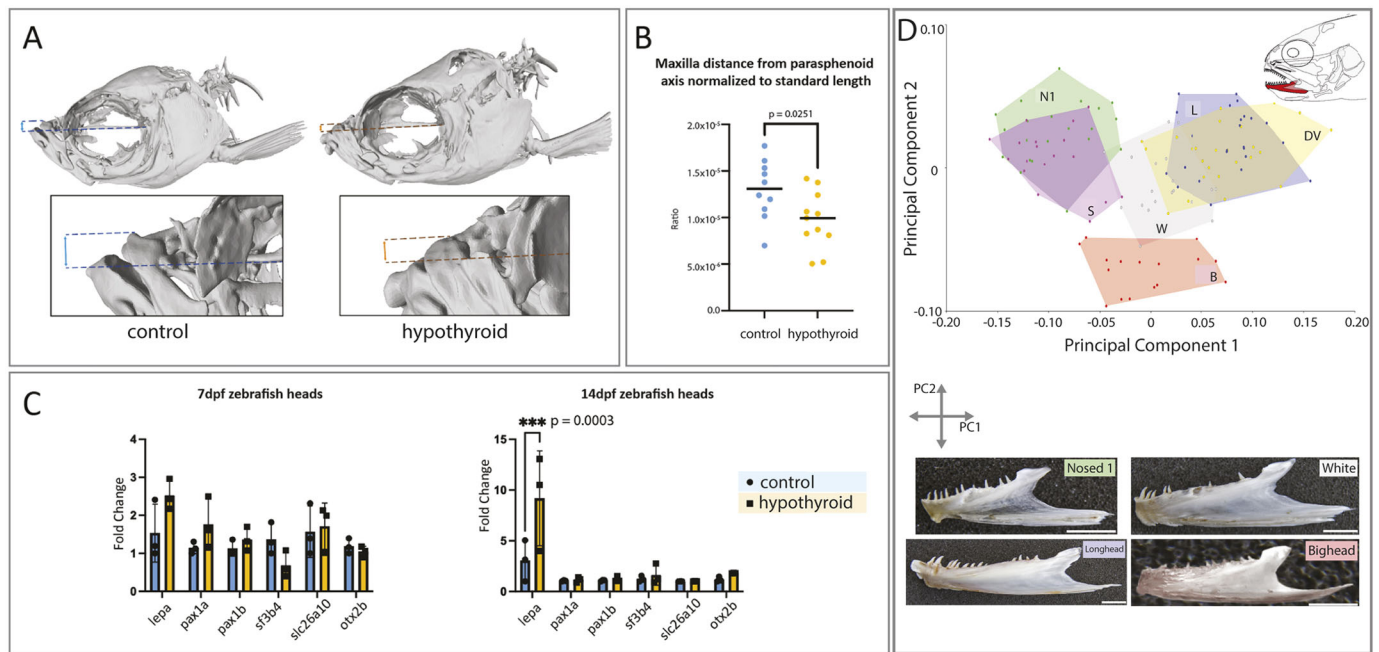


Fig. 7. Effect of thyroid follicle ablation upon morphology and gene expression. (A) In thyroid-ablated adult zebrafish (hypothyroid), there is a significant shift in the position of the maxilla towards a more terminal position. Dashed lines indicate the axis of the parasphenoid and the parallel line intersecting the dorsalmost point on the maxilla. (B) In hypothyroid individuals, there is a significant reduction in the distance from the dorsal-most position of the maxilla to the long axis of the body. All distances are relative to the standard length. $n=10$ control and 11 hypothyroid individuals. P -value determined by a two-tailed, unpaired t -test. (C) In hypothyroid individuals, there is a significant increase of *lepa* expression in the head at 14 dpf. $n=3$ pools of 20 heads for each condition. Differences in mean assessed by a two-way ANOVA followed by Šidák's multiple comparison test. (D) Plotting principal components 1 and 2 reveals significant differences in dentary shape (Procrustes ANOVA $F_{40;944}=40.84$ $P<0.0001$) between morphs. Images show representative examples of dentary bones. Scale bars: 10 mm.

overlapping domains of *pax1* and *otx2* expression in the pharyngeal arches and the oral endoderm from which the pituitary gland arises, the differentiating variant identified provides a plausible regulatory shift associated with evolution of craniofacial morphology (Liu et al., 2013, 2020).

We asked whether the subterminal mouth positions exhibited by the low- T_3 lineages (smallmouth, nosed 1 and nosed 3) could be caused by the reduced plasma thyroid hormone levels in these morphs. We used transgenic thyroid ablation to determine whether experimental hypothyroidism (McMenamin et al., 2014) caused any parallel shifts in subterminal mouth position in the zebrafish system. Although zebrafish skulls are anatomically and kinematically different than charr, chiefly in their ability for jaw protrusion, hypothyroid zebrafish showed a significant shift in the maxilla position, moving ventrally from a supra-terminal position (Fig. 7A,B). Further, to test whether the top candidate genes were altered in a hypothyroid context, we extracted mRNA from the heads of control and hypothyroid larval zebrafish at 7 and 14 days postfertilization (dpf) and quantified expression levels by RT-qPCR (Fig. 7C). While there is strong genetic differentiation between the lake morphs and the riverine Dolly Varden for *slc26a10* and *sf3b4*, we did not detect any significant difference in gene expression levels. We also quantified gene expression for *otx2b* and its putative regulators, *pax1a* and *pax1b*. While these genes differentiated the nosed morphs, we did not detect significant differences in gene expression levels in the head. However, *lepa* was significantly upregulated in the heads of 14 dpf hypothyroid larvae, indicating that under normal developmental conditions, thyroid hormone suppresses *lepa* expression in the head.

Work in zebrafish previously identified skeletal elements that are phenotypically and functionally sensitive to thyroid hormone titer

(Keer et al., 2019, 2022; Conith et al., 2022); we used geometric morphometrics to evaluate these skeletal elements in charrs. Among charr lineages, we compared the shapes of the dentary, anguloarticular, hyomandibula and parasphenoid (Fig. S3). Charrs significantly differed in the shape of TH-sensitive bones (Procrustes ANOVA for dentary and anguloarticular: $F_{40;944}=40.84$ $P<0.0001$; $F_{80;1760}=18.88$ $P<0.0001$, respectively) and display subtle differences in the shape of TH-insensitive bone (Procrustes ANOVA for hyomandibula: $F_{60;1212}=6.55$ $P<0.001$) (Fig. 7D, Figs S4, S5). The parasphenoid, which forms the neurocranial base, was not variable (Fig. S6). These data were supported by pairwise calculation of Procrustes distances between morphs, which demonstrated significant differences ($P<0.001$) in the shape of jaw bones between all morphs, excluding pairs (longhead/Dolly Varden and nosed 1/smallmouth), and absence of differences in the shape of parasphenoid between most of morphs, excluding pairs formed by bigmouth, smallmouth and piscivorous morphs (longhead and white). Performing PCA on the dentary and anguloarticular revealed that the distribution of morphs along PC1 (a component explaining 66.9% and 57.7% of the variance, respectively) (Fig. 7B) corresponded to their distribution along the T_3 value axis (Fig. 6A). Overall, these data mirror patterns observed in TH-modulation during zebrafish development, and detail specific jaw patterning variation among the different charr lineages within these parameters (Nguyen et al., 2023 preprint).

DISCUSSION

Lake Kronotskoe harbors a unique radiation of Dolly Varden charr that provides a powerful new case to study the genetic and developmental foundations supporting vertebrate radiations. Previous models

centered on morphology, ecology and feeding behavior suggest two lacustrine clades: a deep-water clade consisting of smallmouth and bigmouth morphs, and a shallow-water (pelagic and littoral) clade composed of longhead, white and nosed morphs (Markevich et al., 2018). Our data re-define the evolutionary relationships among the lake morphs and support differentiated true-breeding lineages.

Lake Kronotskoe arose from a volcanogenic event and presently drains via a waterfall considered impassable by charr (Viktorovskii, 1978). Pairwise *F_{st}* analyses found that each lake lineage is more differentiated from riverine Dolly Varden than from any other lineage. Furthermore, each lake lineage is differentiated from riverine Dolly Varden charr at roughly equivalent levels. We determined that spatiotemporal differences in spawning of the Lake Kronotskoe lineages (Markevich et al., 2021) ensure their reproductive isolation and low hybridization. By contrast, the much older Lake Malawi cichlid radiation, has *f*-branch values commonly exceeding 5% and up to 14.2% (Malinsky et al., 2018) and *Coregonus* salmonids also have extensive introgression (De-Kayne et al., 2022). These new data suggest that the species flock in Lake Kronotskoe was established by a single founding population having shared genetic signatures that quickly established and maintained reproductively isolated populations.

We found specific genetic differentiation between lacustrine resident lineages and riverine Dolly Varden populations with selective signatures in genes regulating craniofacial development and thyroid function. PROVEAN and SIFT predicted that the serine substitution shared among the lake morphs would have little consequence upon *slc26a10* function. However, such alignment-based prediction methods tend to underestimate potential effects from substitutions if that variation is shared in other lineages (Ng and Henikoff, 2003; Choi and Chan, 2015). The fact that zebrafish also encode for a serine at that residue indicates that this amino acid substitution is likely not deleterious – reflected in low predicted impact scores – but it does not necessarily mean that function is not modulated.

While the relationship to morphological radiation is less clear, the abundance of ion transport and protein trafficking genes among the highly differentiating genes suggests that these processes may also be important drivers of charr evolution. Crucially, ion homeostasis is central to effective osmoregulation during freshwater adaptation (McCormick et al., 2019). Proper ion channel expression is also a factor in chondrocyte maturation and homeostasis (Dicks et al., 2023; Brylka et al., 2024). Whether and how protein trafficking contributes to adaptations in Lake Kronotskoe is less obvious.

The data show that CNEs are particularly influential in the evolution of kidney function and cartilage morphology within the Lake Kronotskoe radiation. The kidneys are indispensable for osmoregulation and obligate freshwater populations of salmonids place different demands on their kidneys than their anadromous counterparts (Tipsmark et al., 2010). Cartilage templates lay foundations for many craniofacial structures. In concert, variation of these three traits may contribute to differential behavior, physiology and morphology characterized within this adaptive radiation.

Charr lineages differ metabolically, especially in energy storage and expenditure that broadly track with our genetic findings (Esin et al., 2024a). Longhead and white, with relatively normal *T₃* levels and without the *lepa2* substitution, are metabolically advanced and accumulate lipids; the smallmouth lineage, which was found to have significantly reduced *T₃* levels, grows slower and accumulates glycogen; the nosed charr lineages, which have significantly reduced *T₃* levels as well as a fixed amino acid substitution in *lepa2* operate at the lowest metabolic rate and do not accumulate

sparse substances; *T₃* levels remain unchanged in the bigmouth lineage, but the population does share the *lepa2* variant, which is fixed in the nosed morphs having some of the more differentiated structural changes (Esin et al., 2024a). Regulators of energy metabolism, such as thyroid hormones and leptin, are prime pleiotropic mechanisms for generating increased morphological and physiological variance (Esin et al., 2023; Zwahlen et al., 2024). In this initial selection, the shift of these metabolic regulators may have expanded the ecological opportunity of an ancestral population of lacustrine charr.

As many specialized morphologies are hypothesized to arise from heterochronic shifts in development (Simonsen et al., 2017), the thyroid axis may prove to be a common mechanism underlying the adaptive potential and may contribute to the remarkable similarity of morphologies exhibited by lacustrine charr species flocks (Esin et al., 2021a,b). Indeed, the pairs of morphs in each of the Lake Kronotskoe clades exhibit alternative heterochronic tendencies. For example, the smallmouth morph with proportionally large eyes and blunt, rounded rostra shows hallmarks of paedomorphosis, while the sister bigmouth morph possesses peramorphic traits, such as overdeveloped lower jaw. The enlarged jaws and frontonasal protrusion of the longhead morph, as well as the drastically modulated frontonasal proportion of the nosed 3 morph, are peramorphic features in comparison to their sister lineages. In other charr radiations, there are similar examples of coupled shifts in craniofacial form alongside variable TH levels. In Lake El'gytgyn, there resides an extremely low TH-content small-mouth charr (*S. elgyticus*) with big eyes and a blunt, rounded rostrum, while the closely related boganida charr (*S. boganidae*) has a high TH content and elongated jaws (Esin et al., 2021a,b, 2024b). In another example from charr, the piscivorous stone charr lineage, dwelling in sympatry with typical Dolly Varden in the Kamchatka river, is characterized by a high TH level and proportionally large jaws. These and other examples demonstrate that developmental and morphological traits vary in independent charr groups and in Dolly Varden populations residing in volcanically polluted rivers (Melnik et al., 2020; Esin et al., 2023).

Indeed, thyroid hormone-induced adaptive morphologies are found in phylogenetically distant fishes, the large African barbs (g. *Labeobarbus*; Cypriniformes; Teleostei), inhabiting Lake Tana (Nagelkerke and Sibbing, 2000). The age of the Lake Tana species flock of barbs is comparable to the Lake Kronotskoe species flock, yet genetic differences between Lake Tana morphs are comparatively subtle (de Graaf et al., 2010; Nagelkerke et al., 2015), and ecomorphological differentiation is a result of heterochronic shifts presumably induced by thyroid axis alterations (Shkil et al., 2015). Such similarities suggest that genetic modification of thyroid signaling may be a widespread mechanism facilitating rapid freshwater teleost adaptive radiations; thyroid modifications could provide a pleiotropic foundation from which more specialized morphologies may be further elaborated. Our experimentally induced hypothyroidism lend support to this possibility: ablating thyroid follicles in the zebrafish creates a shift towards a subterminal mouth position, mirroring the morphology of the low-*T₃* charr lineages. Further, the craniofacial elements variable among the Lake Kronotskoe charrs are the same bones that are known to be sensitive to thyroid hormone alterations in a zebrafish context (Keer et al., 2019, 2022; Borisov and Shkil, 2024 preprint). Furthermore, the significant increase in *lepa* expression in hypothyroid zebrafish suggests that endocrine signaling axes may synergize, further expanding the array of potential adult morphologies attainable along a shared axis of change.

In anadromous salmonids, smoltification in preparing for migration out to sea requires orchestration of hormonal and physiological switches. A hypothesis stemming from lacustrine populations is that selective pressure on the ancestral riverine charr population was relaxed upon colonization of the lake, as the newly resident population of charr became obligate freshwater residents. Such an initial shift in developmental programs may constitute a common node among lacustrine-locked charr, biasing the direction of adaptation to generate similar forms among independent lineages, while simultaneously laying the foundation upon which more trophically specialized morphologies may arise. In this context, exclusion of the smoltification stage from the charr life cycle may permit lacustrine adaptive diversification. In support of this model, thyroid hormone signaling is selectively modified during freshwater colonization and subsequent adaptive radiations of the threespine stickleback *Gasterosteus aculeatus* (Kitano et al., 2010).

The lineage-specific pattern of highly differentiating loci identified in nosed morphs, suggests that an initial developmental state of extensive modifications to thyroid signaling and craniofacial development was further refined in these lineages. The fixation of variation in *leptina2* and a predicted *otx2b* regulatory region found in nosed morphs over other lake groups suggests further modulation of the thyroid signaling in these lineages. The data are supported by findings of altered T_3 levels within these lineages as adults and the presence of *mc4r* within an interval of excess allele sharing with the bigmouth morph. Thus, beginning with an initial suite of shared genetic variants, lineage-specific, secondary elaborations may have accumulated and further catalyzed the exceptional species flock diversification in Lake Kronotskoe.

Such repeated and parallel derivations of morphotypes across the *Salvelinus* complex suggest that there is an underlying genetic framework biasing the radiations of resident lacustrine populations. Our data suggest that shared modulation of the thyroid signaling axis in tandem with craniofacial regulators may enforce such biases. The patterns of variation identified in the Lake Kronotskoe radiation point to a fundamental genetic groundwork for craniofacial evolution and a common axis for morphological change.

MATERIALS AND METHODS

Field material collection

Adult charrs that passed the spawning season without spawning changes in color and head and reproductive states were sampled in Lake Kronotskoe. Adult riverine Dolly Varden charrs were collected in the nearest watercourse draining the opposing slope of Valaginskiy range. Blood, pectoral fin tissue and photographs were collected. Blood for thyroid hormone test was carefully collected from the caudal vessel with a VACUETTE serum tube. The distal part of the right pectoral fin (finclip, 0.2–0.3 cm²) was taken with scissors and fixed in pure alcohol for DNA analysis. Fish were photographed, treated with antibacterial and antifungal solution (Melafix and Pimafix, API) for 30 min, and released if the fish did not display any signs of injury and/or infection in 48 h. All catches were carried out in accordance with the Russian Federal Act on Specially Protected Natural Areas (OOPT) N33-Φ₃ 14/03/1995, article 10, and Plan of the research activities of Kronotsky Nature Reserve. The procedures with fish were approved by the Bioethics Commission of the AN Severtsov Institute of Ecology and Evolution, Russian Academy of Science.

PhyloChip targeted sequence enrichment design

We aimed to create a pan-Salmoniformes targeted sequence capture design that can enrich sequencing libraries for conserved genetic regions across a broad diversity of available salmon genomes. This design targets protein-coding exons as well as a set of conserved non-protein coding elements (CNEs), miRNA hairpins, and ultraconservative non-coding elements

(UCNEs). The majority of capture baits were derived from the Atlantic salmon genome (*Salmo salar*, ICSASG_v2) (Davidson et al., 2010), with inclusion of regions from the genome of rainbow trout (*Oncorhynchus mykiss*, AUL_PRJEB4421_v1) (Berthelot et al., 2014) that were either not represented in the Atlantic salmon genome or were <85% identity to a capture target within the rainbow trout genome. As these fish bracket both sides of the salmon phylogeny (Fig. S1A), the ‘PhyloChip’ design strategy enables DNA from the majority of salmonids target regions to be efficiently enriched using this one capture design.

As the Atlantic salmon genome was not annotated at the time of capture design, annotated coding sequences were isolated from the rainbow trout and northern pike (*Esox lucius*, GCF_000721915.2_ASM72191v2) (Rondeau et al., 2014) genomes. These were then identified within the Atlantic salmon genome via BLASTN (ncbi-blast-2.2.30+; parameters ‘-max_target_seqs 1 -outfmt 6’), and these hits used in the capture design. Genes from rainbow trout that were not identified in Atlantic salmon or that had <85% identity to the best BLAST hit within the Atlantic salmon genome were also retained in the capture design. CNEs were defined from the constrained regions in the Ensembl compara 11-way teleost alignment (Ensembl release-84) (Herrero et al., 2016). To reserve space in the capture design, only CNEs ≥75 bp in length were included in the capture baits. These CNEs were extracted from the Japanese medaka (*Oryzias latipes*, MEDAKA1), three-spined stickleback (*Gasterosteus aculeatus*, BRAOD S1), and zebrafish (*Danio rerio*, GRCz10.84) genomes using Bedtools (v2.23.0) intersectBed (Quinlan and Hall, 2010). miRNA hairpins were extracted from miRbase and UCNEs from UCNEbase (Kozomara and Griffiths-Jones, 2010; Dimitrieva and Bucher, 2013). As with protein-coding exons, these elements were identified within each reference genome using BLASTN (ncbi-blast-2.2.30+; parameters ‘-max_target_seqs 1 -outfmt 6’). miRNA hairpins were padded to be at least 100 bp to improve capture specificity.

From these targeted regions, the specific SeqCap EZ Developer (06471684001) oligonucleotide capture baits were made in collaboration with the NimbleGen design team. Capture baits are strategically designed to standardize oligo annealing temperature, remove low-complexity DNA regions and to reduce the oligo sequence redundancy. The capture design targeted sequence from 558,882 genomic regions (97,049,118 total bp) across the two salmonid genomes. This included including 460,210 protein coding exons, 93,973 CNEs, 1082 miRNAs and 3617 UCNEs (Fig. S1).

DNA extraction and preparation of sequencing libraries

Tissue from finclips was digested and genomic DNA was column purified using QIAGEN DNeasy Blood & Tissue Kit (QIAGEN, 69506). Genomic DNA was extracted from finclips of one *S. leucomaenis*, three riverine Dolly Varden charr, eight bigmouth morphs, ten longhead morphs, five nosed 1 morphs, seven nosed 2 morphs, five nosed 3 morphs, six smallmouth morphs and six white morphs. Pools of genomic DNA were produced for each lineage such that genomic DNA from every individual in a lineage pool was equally represented. The pooled samples were sheared to a target size of 200 bp in Tris-HCl EDTA shearing buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). Mechanical shearing was performed using a Covaris E220 ultrasonicator (duty cycle, 10%; intensity, 5; cycles/burst, 200; time, 400 s; temperature, 8°C) and Covaris microTUBE Snap-Cap tubes (Covaris, 520045). Sequencing libraries were produced using the KAPA HyperPrep Kit (Roche, 07137923001) using 500 ng of starting material for each library. Library preparation was conducted by following the SeqCap EZ HyperCap Workflow Version 1.2. The sequencing library for the nosed 2 samples utilized enzymatic shearing using the KAPA HyperPlus Kit (Roche, 07962401001) and 100 ng of starting material (SeqCap EZ HyperCap Workflow Version 3.0). Fragment size and DNA concentration were quantified using Agilent 2100 BioAnalyzer and High Sensitivity DNA Chips (Agilent, 5067-4626). Paired-end, 150bp Illumina HiSeq sequencing was performed on a pool consisting of multiple barcoded libraries.

Trimming adapters and read mapping

Illumina adapter sequences were removed from reads using Trimmomatic v.0.36 (Bolger et al., 2014). Trimmed and masked reads were aligned to the *Salvelinus alpinus* reference genome (RefSeq Assembly accession: GCF_002910315.2) using NextGenMap v0.5.5 (Sedlazeck et al., 2013).

The flag - -strata 1 was so only the highest scoring read mappings were recorded in the alignment file.

Variant calling and filtering

The variants used to reconstruct the phylogeny and to conduct the PCA were derived from a sample of *S. leucomaenis*, anadromous Dolly Varden, and all seven members of the species flock. samtools v.1.15.1 was used to fix mates, mark duplicates, and filter reads below the minimum mapping quality set to -q 30 (Danecek et al., 2021). bcftools v1.13 was used to call and filter variants (Danecek et al., 2021). Only the variants with quality scores ≥ 20 , depth of coverage on a per-sample basis between ten and 500 reads, fraction of missing genotypes $F_{\text{MISSING}} \leq 0.72$. SNPs within 2 bp of indels and other variant types were excluded, and minor allele frequency > 0.05 . The quality-filtered VCF file contained 623,619 variants. The set of variants used for introgression analyses, quantification of F_{st} , π and GO term enrichment analyses were called and filtered from alignments of the anadromous Dolly Varden, bigmouth, longhead, nosed 1, nosed 3, smallmouth and white lineages (*S. leucomaenis*, and nosed 2 were excluded). The calling and filtering criteria were identical to the conditions described above except for the depth thresholds. Those were filtered on a per site basis for coverage between 70 and 3500 reads. This VCF file contained 526,811 variants.

Coverage of targeted elements

Coverage statistics were derived using the BEDTools v.2.21.1 coverage function (Quinlan and Hall, 2010). Alignment files were intersected with a bed file containing the positions of each targeted element. From this intersection, the average depth of coverage was quantified per base.

Deriving phylogeny

The phylogeny was derived using IQ-TREE v.1.6.12 (Nguyen et al., 2015). The input consisted of 622,831 nucleotide sites, including 22,701 parsimony informative variants. The ModelFinder function (Kalyaanamoorthy et al., 2017) determined the base substitution model of best fit to be a transversion model with empirical base frequencies and a proportion of invariable sites (TVM+F+I). 1000 ultrafast bootstrap replicates (Hoang et al., 2017) quantified support for the phylogeny.

PCA of sequence variation

PLINK v.1.90b7 was utilized to conduct PCA (Purcell et al., 2007). Linkage pruning was conducted using 50 kb windows, 10 bp step size, and $R^2 > 0.1$. The linkage pruned dataset consisted of 65,488 variant sites. The PLINK eigenvector and eigenvalue outputs were plotted in R.

Introgression analysis

Introgression was quantified for all trios in the phylogeny using Dsuite v.0.4 (Malinsky et al., 2021) Dtrios. Riverine Dolly Varden was specified as the outgroup. The f -branch statistic was depicted as a matrix by taking the output from Dsuite Fbranch and running the Dsuite dtools.py script to generate a plot. Dinvestigate was used to generate sliding windows of 40 variants per window and 50% overlap.

Calculating pairwise F_{st} and Tajima's P_i

To quantify genetic differentiation, pairwise F_{st} was calculated using PoPoolation2 v.1201 (Kofler et al., 2011a). The software package allows F_{st} to be quantified in sliding windows or in a genewise manner. Sequencing alignment data were converted into the mpileup format using SAMtools v.1.13 (Li et al., 2009). The PoPoolation2 program mpileup2sync (-min-qual 20) generated the sync file used as input to calculate F_{st} in sliding windows. Popoolation2 calculated F_{st} based on allele frequency (Hartl et al., 1997). The PoPoolation2 program fst-sliding (-min-coverage 20 -min-count 3 -max-coverage 200) was used to calculate F_{st} in non-overlapping sliding windows. To assess which genes were most differentiating between populations, the function 'create-genewise-sync' was utilized to intersect the sync file with a gtf containing all targeted regions in the *S. alpinus* genome, and filtered according to the same depth and minor allele count criteria as sliding windows analyses. Prior to filtering for depth, F_{st} was calculated for 59,478 genes and 22,590 CNEs. Non-coding elements were associated with putative regulatory targets by following the GREAT workflow to establish basal regulatory

windows. A BED file of CNE loci was intersected with intervals spanning 5 kb upstream of and 1 kb downstream from transcriptional start sites, with up to a 1 Mb extension (McLean et al., 2010). To quantify nucleotide diversity, Tajima's P_i was calculated using PoPoolation v1.2.2 (Kofler et al., 2011b). The software also enables quantification of Tajima's P_i in sliding windows. The same depth criteria that were used for F_{st} sliding window quantifications were used to calculate Tajima's P_i in sliding windows for individual lineages.

Enzyme-linked immunosorbent assay for thyroid hormone in blood

Serum samples were transferred to 2 ml specimen tubes and centrifuged at 12,000 g for 10 min with Eppendorf MiniSpin. The number of samples collected was 16 bigmouth, 17 Dolly Varden, seven longhead, 24 nosed 1, 14 nosed 3, 17 smallmouth, and 16 white morphs. Serum was then collected into Eppendorf 1.5 ml tubes and placed in a freezer at -24 to -26°C . The total T_3 (bioactive form of thyroid hormone) concentration in plasma was evaluated by enzyme-linked immunosorbent assay Monobind Total Triiodothyronine (tT_3) test system (Monobind Inc.) and the hormone was measured in accordance with the manufacture protocol using StatFax 303 Plus strip reader (Awareness Technology Inc.). The differences between means were analyzed by one-way ANOVA with a post-hoc Tukey honestly significant difference (HSD) test.

Zebrafish thyroid follicle ablations

Danio rerio were all of the line *Tg(tg:nVenus-2a-nfnB)wp.rt8* (McMenamin et al., 2014). Briefly, clutches of transgenic embryos were sorted for nVenus expression at 4 dpf then treated overnight with either 1% DMSO (for control euthyroid fish) or with 1% DMSO and 10 mM metronidazole, which induces conditional thyroid ablation in the *nfnB*-expressing thyroid follicle cells. Thyroid ablation was visually confirmed at 5 dpf. This animal work was overseen by the Boston College IACUC board and this work was approved under protocol 2007-006-01.

Quantification of maxilla position

AMIRA (v.6.0.0) was used to visualize μCT scans of adult zebrafish skulls (Blythe et al., 2022; Nguyen et al., 2022). A line was drawn intersecting the parasphenoid at its proximal- and distal-most points to approximate the long-axis of the body. A perpendicular line was drawn from the dorsal-most position of the maxilla to intersect the parasphenoid axis. For each individual, this distance was normalized to the standard length. To reduce any quantitative bias stemming from the size of the denominator during normalization to standard length, we selected roughly size matched individuals. Wild-type samples ranged in size from 20.0 mm to 20.5 mm and hypothyroid individuals ranged in size from 19.0 mm to 20.5 mm. $n=10$ DMSO control individuals and $n=11$ metronidazole-treated individuals.

Gene expression quantification

Euthyroid controls and hypothyroid siblings were decapitated at 7 or 14 dpf posterior to the operculae, and three sets of 20 heads for each condition were stored in RNAlater™ Stabilization Solution (Thermo Fisher Scientific) at -20°C . RNA was extracted using a Quick-RNA™ Microprep Kit (Thermo Fisher Scientific) and cDNA libraries synthesized using SuperScript™ IV Reverse Transcriptase (Thermo Fisher Scientific). Using primer sequences (Table S4) for *actinB1*, *lepa*, *otx2b*, *pax1b*, *pax1a*, *sf3b4* and *slc26a10*, qPCR was performed with PowerUp™ SYBR™ Green Master Mix (Thermo Fisher Scientific) on a QuantStudio™ 3 Real-Time PCR System (Thermo Fisher Scientific) with three technical and biological replicates. Results were analyzed using DataConnect Software. Relative gene expression was calculated using the $\Delta\Delta\text{CT}$ method with *actinB1* serving as the housekeeping gene (Livak and Schmittgen, 2001).

Geometric morphometrics

Images of dry osteological samples were used for landmarking (Saltykova et al., 2015). Using TPSdig v2.0 (Rohlf, 2015), we digitized landmarks (LMs), most of which have been used for the homologous bones of zebrafish (Keer et al., 2022) and charrs (Jónsdóttir et al., 2024): six LMs for dentary ($n=16$ bigmouth, 25 Dolly Varden, 24 longhead, 23 nosed 1, 18 smallmouth, 21 white); ten LMs for anguloarticular ($n=22$ bigmouth, 21

Dolly Varden, 19 longhead, 15 nosed 1, 19 smallmouth, 20 white); eleven LM for hyomandibula ($n=24$ bigmouth, 27 Dolly Varden, 13 longhead, 20 nosed 1, 18 smallmouth, 30 white); and eight LM for parasphenoid ($n=20$ bigmouth, 14 Dolly Varden, 16 longhead, 20 nosed 1, 17 smallmouth, 21 white) (Fig. S6). Shape analysis was performed in MorphoJ v.1.06d (Klingenberg, 2011). We implemented Generalized Procrustes superimposition and assessed variation in the shape with PCA. For better visualization of shape variability along PC1/PC2, we created a wireframe mesh connecting landmarks. To estimate the shape differences between the morphs, we implemented Procrustes ANOVA and Canonical Variate (CV) analysis with a calculation of pairwise Procrustes distances (10,000 permutation rounds).

Acknowledgements

We owe thanks to Dr Stacy Nguyen for providing μ CT data and to Dr E. A. Saltykova for providing images of dry osteological samples. The study was conducted within the framework of the research programs of the IEE RAS (0089-2021-0003) and IDB RAS (0088-2021-0019).

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: K.C.W., C.S.M., S.K.M., J.M.D., M.P.H., F.N.S.; Methodology: K.C.W., J.M.D., F.N.S.; Validation: K.C.W., C.S.M.; Formal analysis: K.C.W., C.S.M., F.N.S.; Investigation: K.C.W.; Resources: E.V.E., G.N.M., C.S.M., J.M.D., F.N.S.; Data curation: K.C.W.; Writing - original draft: K.C.W.; Writing - review & editing: E.V.E., G.N.M., C.S.M., S.K.M., J.M.D., M.P.H., F.N.S.; Visualization: K.C.W., F.N.S.; Supervision: S.K.M., M.P.H., F.N.S.; Funding acquisition: K.C.W., S.K.M.

Funding

This work was supported by the National Institutes of Health [R35GM146467 to S.K.M., 5F32DE029362 to K.C.W.] and the National Science Foundation [NSF 1845513 to S.K.M.]. Deposited in PMC for release after 12 months.

Data availability

All targeted sequencing reads are available at NCBI Sequence Read Archive under BioProject PRJNA1155113.

Peer review history

The peer review history is available online at <https://journals.biologists.com/dev/lookup/doi/10.1242/dev.203002.reviewer-comments.pdf>

Special Issue

This article is part of the Special Issue 'Uncovering developmental diversity', edited by Cassandra Extavour, Liam Dolan and Karen Sears. See related articles at <https://journals.biologists.com/dev/issue/151/20>.

References

- Bando, H., Gergics, P., Bohnsack, B. L., Toolan, K. P., Richter, C. E., Shavit, J. A. and Camper, S. A. (2020). Otx2b mutant zebrafish have pituitary, eye and mandible defects that model mammalian disease. *Hum. Mol. Genet.* **29**, 1648-1657. doi:10.1093/hmg/ddaa064
- Barluenga, M., Stölting, K. N., Salzburger, W., Muschick, M. and Meyer, A. (2006). Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* **439**, 719-723. doi:10.1038/nature04325
- Bernier, F. P., Caluseriu, O., Ng, S., Schwartzentruber, J., Buckingham, K. J., Innes, A. M., Jabs, E. W., Innis, J. W., Schuette, J. L., Gorski, J. L. et al. (2012). Haploinsufficiency of SF3B4, a component of the pre-mRNA spliceosomal complex, causes Nager syndrome. *Am. J. Hum. Genet.* **90**, 925-933. doi:10.1016/j.ajhg.2012.04.004
- Berthelot, C., Brunet, F., Chalopin, D., Juanchich, A., Bernard, M., Noël, B., Bento, P., Da Silva, C., Labadie, K., Alberti, A. et al. (2014). The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat. Commun.* **5**, 3657. doi:10.1038/ncomms4657
- Blythe, J., Nguyen, S. and Mcmenamin, S. (2022). Thyroid hormone mediates juvenile and adult craniofacial shape change in zebrafish. *FaceBase Consortium*. doi:10.25550/1Q-X114
- Bolger, A. M., Lohse, M. and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120. doi:10.1093/bioinformatics/btu170
- Borisov, V. and Shkil, F. (2024). Effects and phenotypic consequences of transient thyrotoxicosis and hypothyroidism at different stages of zebrafish *Danio rerio* (Teleostei: Cyprinidae) skeleton development. *bioRxiv*. 2024.2006.2008.598073. doi:10.1101/2024.06.08.598073

- Braitseva, O. A., Melekestsev, I. V., Ponomareva, V. V. and Sulerzhitsky, L. D. (1995). Ages of calderas, large explosive craters and active volcanoes in the Kuril-Kamchatka region, Russia. *Bull. Volcanol.* **57**, 383-402. doi:10.1007/s004450050100
- Brylka, L. J., Alimiy, A.-R., Tschaffon-Müller, M. E. A., Jiang, S., Ballhause, T. M., Baranowsky, A., Von Kroge, S., Delsmann, J., Pawlus, E., Eghbalian, K. et al. (2024). Piezo1 expression in chondrocytes controls endochondral ossification and osteoarthritis development. *Bone Res.* **12**, 12. doi:10.1038/s41413-024-00315-x
- Castinetti, F., Brinkmeier, M. L., Gordon, D. F., Vella, K. R., Kerr, J. M., Mortensen, A. H., Hollenberg, A., Brue, T., Ridgway, E. C. and Camper, S. A. (2011). PITX2 and PITX1 regulate thyrotroph function and response to hypothyroidism. *Mol. Endocrinol.* **25**, 1950-1960. doi:10.1210/me.2010-0388
- Choi, Y. and Chan, A. P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **31**, 2745-2747. doi:10.1093/bioinformatics/btv195
- Conith, M. R., Ringo, D., Conith, A. J., Deleon, A., Wagner, M., Mcmenamin, S., Cason, C. and Cooper, W. J. (2022). The evolution of feeding mechanics in the danioninae, or why giant danios don't suck like zebrafish. *Integr. Organismal Biol.* **4**, obac049. doi:10.1093/iob/obac049
- Copeland, D. L., Duff, R. J., Liu, Q., Prokop, J. and Londraville, R. L. (2011). Leptin in teleost fishes: an argument for comparative study. *Front. Physiol.* **2**, 26. doi:10.3389/fphys.2011.00026
- Daane, J. M., Rohner, N., Konstantinidis, P., Djuranovic, S. and Harris, M. P. (2015). Parallelism and epistasis in skeletal evolution identified through use of phylogenomic mapping strategies. *Mol. Biol. Evol.* **33**, 162-173. doi:10.1093/molbev/msv208
- Daane, J. M., Dornburg, A., Smits, P., Macguigan, D. J., Brent Hawkins, M., Near, T. J., William Detrich III, H. and Harris, M. P. (2019). Historical contingency shapes adaptive radiation in Antarctic fishes. *Nat. Ecol. Evol.* **3**, 1102-1109. doi:10.1038/s41559-019-0914-2
- Daane, J. M., Blum, N., Lanni, J., Boldt, H., Iovine, M. K., Higdon, C. W., Johnson, S. L., Lovejoy, N. R. and Harris, M. P. (2021). Modulation of bioelectric cues in the evolution of flying fishes. *Curr. Biol.* **31**, 5052-5061.e5058. doi:10.1016/j.cub.2021.08.054
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M. et al. (2021). Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008. doi:10.1093/gigascience/giab008
- Davidson, W. S., Koop, B. F., Jones, S. J. M., Iturra, P., Vidal, R., Maass, A., Jonassen, I., Lien, S. and Omholt, S. W. (2010). Sequencing the genome of the Atlantic salmon (*Salmo salar*). *Genome Biol.* **11**, 403. doi:10.1186/gb-2010-11-9-403
- De Graaf, M., Megens, H.-J., Samallo, J. and Sibbing, F. (2010). Preliminary insight into the age and origin of the Labeobarbus fish species flock from Lake Tana (Ethiopia) using the mtDNA cytochrome b gene. *Mol. Phylogenet. Evol.* **54**, 336-343. doi:10.1016/j.ympev.2009.10.029
- De-Kayne, R., Selz, O. M., Marques, D. A., Frei, D., Seehausen, O. and Feulner, P. G. D. (2022). Genomic architecture of adaptive radiation and hybridization in Alpine whitefish. *Nat. Commun.* **13**, 4479. doi:10.1038/s41467-022-32181-8
- Decherf, S., Seugnet, I., Kouidhi, S., Lopez-Juarez, A., Clerget-Froidevaux, M.-S. and Demeneix, B. A. (2010). Thyroid hormone exerts negative feedback on hypothalamic type 4 melanocortin receptor expression. *Proc. Natl. Acad. Sci. USA* **107**, 4471-4476. doi:10.1073/pnas.0905190107
- Diazok, D., Romero, C., Zunich, J., Marshall, I. and Radovick, S. (2008). A novel dominant negative mutation of OTX2 associated with combined pituitary hormone deficiency. *J. Clin. Endocrinol. Metab.* **93**, 4351-4359. doi:10.1210/jc.2008-1189
- Dicks, A. R., Maksaev, G. I., Harissa, Z., Savadipour, A., Tang, R., Steward, N., Liedtke, W., Nichols, C. G., Wu, C.-L. and Guilak, F. (2023). Skeletal dysplasia-causing TRPV4 mutations suppress the hypertrophic differentiation of human iPSC-derived chondrocytes. *eLife* **12**, e71154. doi:10.7554/eLife.71154
- Dimitrova, S. and Bucher, P. (2013). UCNEbase—a database of ultraconserved non-coding elements and genomic regulatory blocks. *Nucleic Acids Res.* **41**, D101-D109. doi:10.1093/nar/gks1092
- Esin, E. V., Markevich, G. N. and Pichugin, M. Y. (2018). Juvenile divergence in adaptive traits among seven sympatric fish ecomorphs arises before moving to different lacustrine habitats. *J. Evol. Biol.* **31**, 1018-1034. doi:10.1111/jeb.13283
- Esin, E. V., Bocharova, E. S., Borisova, E. A. and Markevich, G. N. (2020). Interaction among morphological, trophic and genetic groups in the rapidly radiating Salvelinus fishes from Lake Kronotskoe. *Evol. Ecol.* **34**, 611-632. doi:10.1007/s10682-020-10048-y
- Esin, E. V., Markevich, G. N., Zlenko, D. V. and Shkil, F. N. (2021a). Thyroid-mediated metabolic differences underlie ecological specialization of extremophile salmonids in the arctic lake El'gygytyn. *Front. Ecol. Evol.* **9**.
- Esin, E. V., Markevich, G. N., Melnik, N. O., Kapitanova, D. V. and Shkil, F. N. (2021b). Natural toxic impact and thyroid signalling interplay orchestrates riverine adaptive divergence of salmonid fish. *J. Anim. Ecol.* **90**, 1004-1019. doi:10.1111/1365-2656.13429
- Esin, E. V., Shulgina, E. V. and Shkil, F. N. (2023). Rapid hyperthyroidism-induced adaptation of salmonid fish in response to environmental pollution. *J. Evol. Biol.* **36**, 1471-1483. doi:10.1111/jeb.14220

- Esin, E. V., Markevich, G. N., Shulgina, E. V., Baskakova, Y. A., Artemov, R. V. and Shkil, F. N. (2024a). Differences in energy storage in sympatric salmonid morphs with contrasting lifestyles. *Evol. Biol.* doi:10.1007/s11692-024-09641-8
- Esin, E. V., Shkil, F. N., Zlenko, D. V., Medvedev, D. A., Korostelev, N. B. and Markevich, G. N. (2024b). Quaternary history and radiation of Salvelinus fish in the ancient Arctic lake El'gygytyn. *Hydrobiologia* 851, 2389-2404. doi:10.1007/s10750-023-05464-4
- Evans, A. L. and Gage, P. J. (2005). Expression of the homeobox gene Pitx2 in neural crest is required for optic stalk and ocular anterior segment development. *Hum. Mol. Genet.* 14, 3347-3359. doi:10.1093/hmg/ddi365
- Fabik, J., Kovacova, K., Kozmik, Z. and Machon, O. (2020). Neural crest cells require Meis2 for patterning the mandibular arch via the Sonic hedgehog pathway. *Biol. Open* 9, bio052043. doi:10.1242/bio.052043
- Friedrichs, B., Tepel, C., Reinheckel, T., Deussing, J., Von Figura, K., Herzog, V., Peters, C., Saffig, P. and Brix, K. (2003). Thyroid functions of mouse cathepsins B, K, and L. *J. Clin. Invest.* 111, 1733-1745. doi:10.1172/JCI15990
- Hartl, D. L., Clark, A. G. and Clark, A. G. (1997). *Principles of Population genetics*. Sunderland, MA: Sinauer associates.
- Herrero, J., Muffato, M., Beal, K., Fitzgerald, S., Gordon, L., Pignatelli, M., Vilella, A. J., Searle, S. M. J., Amode, R., Brent, S. et al. (2016). Ensembl comparative genomics resources. *Database* 2016, bav096. doi:10.1093/database/bav096
- Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q. and Vinh, L. S. (2017). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* 35, 518-522. doi:10.1093/molbev/msx281
- 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. et al. (2020). Parallelism in eco-morphology and gene expression despite variable evolutionary and genomic backgrounds in a Holarctic fish. *PLoS Genet.* 16, e1008658. doi:10.1371/journal.pgen.1008658
- Jónsdóttir, G. Ó., von Elm, L. M., Ingimarsson, F., Tersigni, S., Snorrason, S. S., Pálsson, A. and Steele, S. E. (2024). Diversity in the internal functional feeding elements of sympatric morphs of Arctic charr (*Salvelinus alpinus*). *PLoS One* 19, e0300359. doi:10.1371/journal.pone.0300359
- Jonsson, B. and Jonsson, N. (2001). Polymorphism and speciation in Arctic charr. *J. Fish Biol.* 58, 605-638. doi:10.1111/j.1095-8649.2001.tb00518.x
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A. and Jermin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587-589. doi:10.1038/nmeth.4285
- Keer, S., Cohen, K., May, C., Hu, Y., Mcmenamin, S. and Hernandez, L. P. (2019). Anatomical assessment of the adult skeleton of zebrafish reared under different thyroid hormone profiles. *Anat. Rec.* 302, 1754-1769. doi:10.1002/ar.24139
- Keer, S., Storch, J. D., Nguyen, S., Prado, M., Singh, R., Hernandez, L. P. and McMenamin, S. K. (2022). Thyroid hormone shapes craniofacial bones during postembryonic zebrafish development. *Evol. Dev.* 24, 61-76. doi:10.1111/ede.12399
- Kitano, J., Lema, S. C., Luckenbach, J. A., Mori, S., Kawagishi, Y., Kusakabe, M., Swanson, P. and Peichel, C. L. (2010). Adaptive divergence in the thyroid hormone signaling pathway in the stickleback radiation. *Curr. Biol.* 20, 2124-2130. doi:10.1016/j.cub.2010.10.050
- Klemetsen, A. (2010). The Charr problem revisited: exceptional phenotypic plasticity promotes ecological speciation in postglacial lakes. *Freshwater Rev.* 3, 49-74. doi:10.1608/FRJ-3.1.3
- Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Mol. Ecol. Resour.* 11, 353-357. doi:10.1111/j.1755-0998.2010.02924.x
- Kofler, R., Pandey, R. V. and Schlötterer, C. (2011a). PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* 27, 3435-3436. doi:10.1093/bioinformatics/btr589
- Kofler, R., Orozco-Terwengel, P., De Maio, N., Pandey, R. V., Nolte, V., Futschik, A., Kosiol, C. and Schlötterer, C. (2011b). PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. *PLoS ONE* 6, e15925. doi:10.1371/journal.pone.0015925
- Kozomara, A. and Griffiths-Jones, S. (2010). miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* 39 suppl_1, D152-D157. doi:10.1093/nar/gkq1027
- Lecaudey, L. A., Schlieven, U. K., Osinov, A. G., Taylor, E. B., Bernatchez, L. and Weiss, S. J. (2018). Inferring phylogenetic structure, hybridization and divergence times within Salmoninae (Teleostei: Salmonidae) using RAD-sequencing. *Mol. Phylogenet. Evol.* 124, 82-99. doi:10.1016/j.ympev.2018.02.022
- Levin, B. A., Simonov, E., Dgebuadze, Y. Y., Levina, M. and Golubtsov, A. S. (2020). In the rivers: multiple adaptive radiations of cyprinid fishes (Labeobarbus) in Ethiopian Highlands. *Sci. Rep.* 10, 7192. doi:10.1038/s41598-020-64350-4
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Subgroup, G. P. D. P. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079. doi:10.1093/bioinformatics/btp352
- Liu, X., Wang, H., Li, G., Huang, H.-Z. and Wang, Y.-Q. (2013). The function of *DrPax1b* gene in the embryonic development of zebrafish. *Genes Genet. Syst.* 88, 261-269. doi:10.1266/ggs.88.261
- Liu, Y.-H., Lin, T.-C. and Hwang, S.-P. L. (2020). Zebrafish Pax1a and Pax1b are required for pharyngeal pouch morphogenesis and ceratobranchial cartilage development. *Mech. Dev.* 161, 103598. doi:10.1016/j.mod.2020.103598
- Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCT} method. *Methods* 25, 402-408. doi:10.1006/meth.2001.1262
- Ma, Z., Zhu, P., Pang, M., Guo, L., Chang, N., Zheng, J., Zhu, X., Gao, C., Huang, H., Cui, Z. et al. (2017). A novel inducible mutagenesis screen enables to isolate and clone both embryonic and adult zebrafish mutants. *Sci. Rep.* 7, 10381. doi:10.1038/s41598-017-10968-w
- Maitland, P. S., Winfield, I. J., McCarthy, I. D. and Igoe, F. (2007). The status of Arctic charr *Salvelinus alpinus* in Britain and Ireland. *Ecol. Freshw. Fish.* 16, 6-19. doi:10.1111/j.1600-0633.2006.00167.x
- Malinsky, M., Challis, R. J., Tyers, A. M., Schiffels, S., Terai, Y., Ngatunga, B. P., Miska, E. A., Durbin, R., Genner, M. J. and Turner, G. F. (2015). Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science* 350, 1493-1498. doi:10.1126/science.aac9927
- Malinsky, M., Svardal, H., Tyers, A. M., Miska, E. A., Genner, M. J., Turner, G. F. and Durbin, R. (2018). Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nat. Ecol. Evol.* 2, 1940-1955. doi:10.1038/s41559-018-0717-x
- Malinsky, M., Matschiner, M. and Svardal, H. (2021). Dsuite - Fast D-statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour.* 21, 584-595. doi:10.1111/1755-0998.13265
- Markevich, G., Esin, E. and Anisimova, L. (2018). Basic description and some notes on the evolution of seven sympatric morphs of Dolly Varden *Salvelinus malma* from the Lake Kronotskoe Basin. *Ecol. Evol.* 8, 2554-2567. doi:10.1002/ece3.3806
- Markevich, G. N., Zlenko, D. V., Shkil, F. N., Schlieven, U. K., Anisimova, L. A., Sharapkova, A. A. and Esin, E. V. (2021). Natural barriers and internal sources for the reproductive isolation in sympatric salmonids from the lake-river system. *Evol. Biol.* 48, 407-421. doi:10.1007/s11692-021-09546-w
- Mason, V. C., Li, G., Helgen, K. M. and Murphy, W. J. (2011). Efficient cross-species capture hybridization and next-generation sequencing of mitochondrial genomes from noninvasively sampled museum specimens. *Genome Res.* 21, 1695-1704. doi:10.1101/gr.120196.111
- McCormick, S. D., Regish, A. M., Arden, W. R., Björnsson, B. T. and Bernier, N. J. (2019). The evolutionary consequences for seawater performance and its hormonal control when anadromous Atlantic salmon become landlocked. *Sci. Rep.* 9, 968. doi:10.1038/s41598-018-37608-1
- McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M. and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* 28, 495-501. doi:10.1038/nbt.1630
- McMenamin, S. K., Bain, E. J., McCann, A. E., Patterson, L. B., Eom, D. S., Waller, Z. P., Hamill, J. C., Kuhlman, J. A., Eisen, J. S. and Parichy, D. M. (2014). Thyroid hormone-dependent adult pigment cell lineage and pattern in zebrafish. *Science* 345, 1358-1361. doi:10.1126/science.1256251
- Melnik, N. O., Markevich, G. N., Taylor, E. B., Loktyushkin, A. V. and Esin, E. V. (2020). Evidence for divergence between sympatric stone charr and Dolly Varden along unique environmental gradients in Kamchatka. *J. Zool. Syst. Evol. Res.* 58, 1135-1150. doi:10.1111/jzs.12367
- Myers, G. S. (1960). The endemic fish fauna of Lake Lanao, and the evolution of higher taxonomic categories. *Evolution* 14, 323-333. doi:10.2307/2405975
- Nagelkerke, L. A. J. and Sibbing, F. A. (2000). The large barbs (*Barbus* spp., cyprinidae, teleostei) of Lake Tana (Ethiopia), with a description of a new species, *Barbus osseensis*. *Neth. J. Zool.* 50, 179-214. doi:10.1163/156854200505946
- Nagelkerke, L. A. J., Leon-Kloosterziel, K. M., Megens, H. J., De Graaf, M., Diekmann, O. E. and Sibbing, F. A. (2015). Shallow genetic divergence and species delineations in the endemic *Labeobarbus* species flock of Lake Tana, Ethiopia. *J. Fish Biol.* 87, 1191-1208. doi:10.1111/jfb.12779
- Ng, P. C. and Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 31, 3812-3814. doi:10.1093/nar/gkg509
- Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A. and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268-274. doi:10.1093/molbev/msu300
- Nguyen, S. V., Lanni, D., Xu, Y., Michaelson, J. S. and McMenamin, S. K. (2022). Dynamics of the zebrafish skeleton in three dimensions during juvenile and adult development. *Front. Physiol.* 13, 875866. doi:10.3389/fphys.2022.875866
- Nguyen, S., Lee, R. S., Mohlmann, E., Petruccio, G., Blythe, J., Ranieri, I. and Mcmenamin, S. (2023). Craniofacial diversity across Danionins and the effects of TH status on craniofacial morphology of two *Danio* species. *bioRxiv*, 2023.2008.2009.552728.
- Nordeng, H. (1983). Solution to the 'Char Problem' based on Arctic Char (*Salvelinus alpinus*) in Norway. *Can. J. Fish. Aquat. Sci.* 40, 1372-1387. doi:10.1139/f83-159
- Osinov, A. G., Volkov, A. A. and Mugue, N. S. (2021). Charrs of the genus *Salvelinus* (Salmonidae): hybridization, phylogeny and evolution. *Hydrobiologia* 848, 705-726. doi:10.1007/s10750-020-04478-6
- Petit, F., Escande, F., Jourdain, A. S., Porchet, N., Amiel, J., Doray, B., Delrue, M. A., Flori, E., Kim, C. A., Marlin, S. et al. (2014). Nager syndrome: confirmation

- of SF3B4 haploinsufficiency as the major cause. *Clin. Genet.* **86**, 246-251. doi:10.1111/cge.12259
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., Daly, M. J. et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559-575. doi:10.1086/519795
- Quinlan, A. R. and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842. doi:10.1093/bioinformatics/btq033
- Richard, S., Guyot, R., Rey-Millet, M., Prieux, M., Markossian, S., Aubert, D. and Flamant, F. (2020). A pivotal genetic program controlled by thyroid hormone during the maturation of GABAergic neurons. *iScience* **23**, 100899-100899. doi:10.1016/j.isci.2020.100899
- Rohlf, F. J. (2015). The tps series of software. *Hystrix* **26**, 9-12.
- Rondeau, E. B., Minkley, D. R., Leong, J. S., Messmer, A. M., Jantzen, J. R., Von Schalburg, K. R., Lemon, C., Bird, N. H. and Koop, B. F. (2014). The genome and linkage map of the Northern Pike (*Esox lucius*): conserved synteny revealed between the Salmonid sister group and the Neoteleostei. *PLoS ONE* **9**, e102089. doi:10.1371/journal.pone.0102089
- Rothstein, M. and Simoes-Costa, M. (2020). Heterodimerization of TFAP2 pioneer factors drives epigenomic remodeling during neural crest specification. *Genome Res.* **30**, 35-48. doi:10.1101/gr.249680.119
- Saltykova, E., Markevich, G. and Kuzishchin, K. (2015). Divergent skull morphology between trophic separated lacustrine forms of Dolly Varden charr from Lake Kronotskoe, Kamchatka, Russia. *Environ. Biol. Fishes* **98**, 559-570. doi:10.1007/s10641-014-0287-y
- Sandlund, O. T., Gunnarsson, K., Jónasson, P. M., Jonsson, B., Lindem, T., Magnússon, K. P., Malmquist, H. J., Sigurjónsdóttir, H., Skúlason, S. and Snorrason, S. S. et al. (1992). The Arctic Charr *Salvelinus alpinus* in Thingvallavatn. *Oikos* **64**, 305-351. doi:10.2307/3545056
- Schluter, D. and Conte, G. L. (2009). Genetics and ecological speciation. *Proc. Natl Acad. Sci. USA* **106** suppl._1, 9955-9962. doi:10.1073/pnas.0901264106
- Sedlazeck, F. J., Rescheneder, P. and Von Haeseler, A. (2013). NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics* **29**, 2790-2791. doi:10.1093/bioinformatics/btt468
- Shkil, F. N., Kapitanova, D. V., Borisov, V. B., Abdissa, B. and Smirnov, S. V. (2012). Thyroid hormone in skeletal development of cyprinids: effects and morphological consequences. *J. Appl. Ichthyol.* **28**, 398-405. doi:10.1111/j.1439-0426.2012.01992.x
- Shkil, F. N., Lazebnyi, O. E., Kapitanova, D. V., Abdissa, B., Borisov, V. B. and Smirnov, S. V. (2015). Ontogenetic mechanisms of explosive morphological divergence in the Lake Tana (Ethiopia) species flock of large African barbs (Labeobarbus; Cyprinidae; Teleostei). *Russian J. Dev. Biol.* **46**, 294-306. doi:10.1134/S1062360415050069
- Simonsen, M. K., Siwertsson, A., Adams, C. E., Amundsen, P.-A., Præbel, K. and Knudsen, R. (2017). Allometric trajectories of body and head morphology in three sympatric Arctic charr (*Salvelinus alpinus* (L.)) morphs. *Ecol. Evol.* **7**, 7277-7289. doi:10.1002/ece3.3224
- Taylor, E. B. (2016). The Arctic char (*Salvelinus alpinus*) 'complex' in North America revisited. *Hydrobiologia* **783**, 283-293. doi:10.1007/s10750-015-2613-6
- Tingaud-Sequeira, A., Trimouille, A., Marlin, S., Lopez, E., Berenguer, M., Gherbi, S., Arveiler, B., Lacombe, D. and Rooryck, C. (2020). Functional and genetic analyses of ZYG11B provide evidences for its involvement in OAVS. *Mol. Genet. Genomic Med.* **8**, e1375. doi:10.1002/mgg3.1375
- Tipsmark, C. K., Sørensen, K. J. and Madsen, S. S. (2010). Aquaporin expression dynamics in osmoregulatory tissues of Atlantic salmon during smoltification and seawater acclimation. *J. Exp. Biol.* **213**, 368-379. doi:10.1242/jeb.034785
- Treaster, S., Deelen, J., Daane, J. M., Murabito, J., Karasik, D. and Harris, M. P. (2023). Convergent genomics of longevity in rockfishes highlights the genetics of human life span variation. *Sci. Adv.* **9**, eadd2743. doi:10.1126/sciadv.add2743
- Viktorovskii, R. M. (1978). *Mekhanizmy vidoobrazovaniya u gol'tsov Kronotskogo Oзера*. Изд-во Наука.
- Vonlanthen, P., Bittner, D., Hudson, A. G., Young, K. A., Müller, R., Lundsgaard-Hansen, B., Roy, D., Di Piazza, S., Largiadere, C. R. and Seehausen, O. (2012). Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature* **482**, 357-362. doi:10.1038/nature10824
- Walker, A. F., Greer, R. B. and Gardner, A. S. (1988). Two ecologically distinct forms of arctic charr *Salvelinus alpinus* (L.) in Loch Rannoch, Scotland. *Biol. Conserv.* **43**, 43-61. doi:10.1016/0006-3207(88)90077-8
- Wong, T. C. B., Rebbert, M., Wang, C., Chen, X., Heffer, A., Zarelli, V. E., Dawid, I. B. and Zhao, H. (2016). Genes regulated by potassium channel tetramerization domain containing 15 (Kctd15) in the developing neural crest. *Int. J. Dev. Biol.* **60**, 159-166. doi:10.1387/ijdb.160058id
- Yagasaki, Y., Yamaguchi, T., Watahiki, J., Konishi, M., Katoh, H. and Maki, K. (2003). The role of craniofacial growth in leptin deficient (ob/ob) mice. *Orthod. Craniofac. Res.* **6**, 233-241. doi:10.1034/j.1600-0544.2003.00260.x
- Yoon, G.-H., Kim, K., Park, D.-S. and Choi, S.-C. (2022). RNF152 negatively regulates Wnt/ β -catenin signaling in *Xenopus* embryos. *BMB Rep.* **55**, 232-237. doi:10.5483/BMBRep.2022.55.5.187
- Zimmermann-Belsing, T., Brabant, G., Holst, J. J. and Feldt-Rasmussen, U. (2003). Circulating leptin and thyroid dysfunction. *Eur. J. Endocrinol.* **149**, 257-271. doi:10.1530/eje.0.1490257
- Zwahlen, J., Gairin, E., Vianello, S., Mercader, M., Roux, N. and Laudet, V. (2024). The ecological function of thyroid hormones. *Philos. Trans. R. Soc. B Biol. Sci.* **379**, 20220511. doi:10.1098/rstb.2022.0511

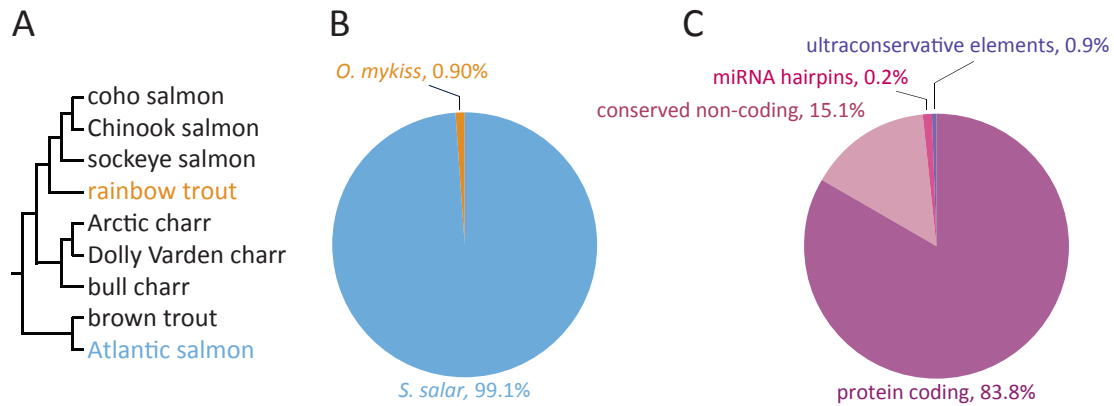


Fig. S1. Design of pan-Salmoniformes targeted capture array. (A) Conserved bait sequences were derived from Atlantic salmon (*Salmo salar*) and rainbow trout (*Oncorhynchus mykiss*) reference genomes. **(B)** Relative contributions of *S. salar* and *O. mykiss* derived bait sequences to the capture array. **(C)** Classifications and relative abundances of conserved elements targeted for capture.

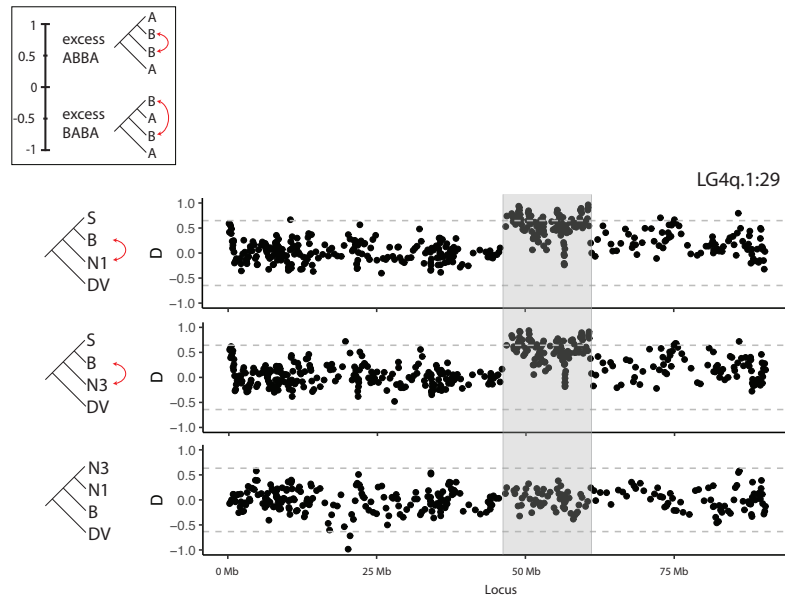


Fig. S2. Sliding window plots of D showing a large interval of excess allele sharing. Sliding window plots for trios consisting of Smallmouth (S), Bigmouth (B), and Nosed 1 (N1), of S, B, and Nosed 3 (N3), and of N3, N1, and B. Positive D values indicate an excess of the ABBA pattern (red arrows), while negative values indicate an excess of the BABA pattern. The three plots show a common pattern of excess of allele sharing overlapping with between B and N1 and B and N3, while there is no excess of allele sharing between B and N1 over B and N3. Horizontal lines signify 3SDs from the mean.

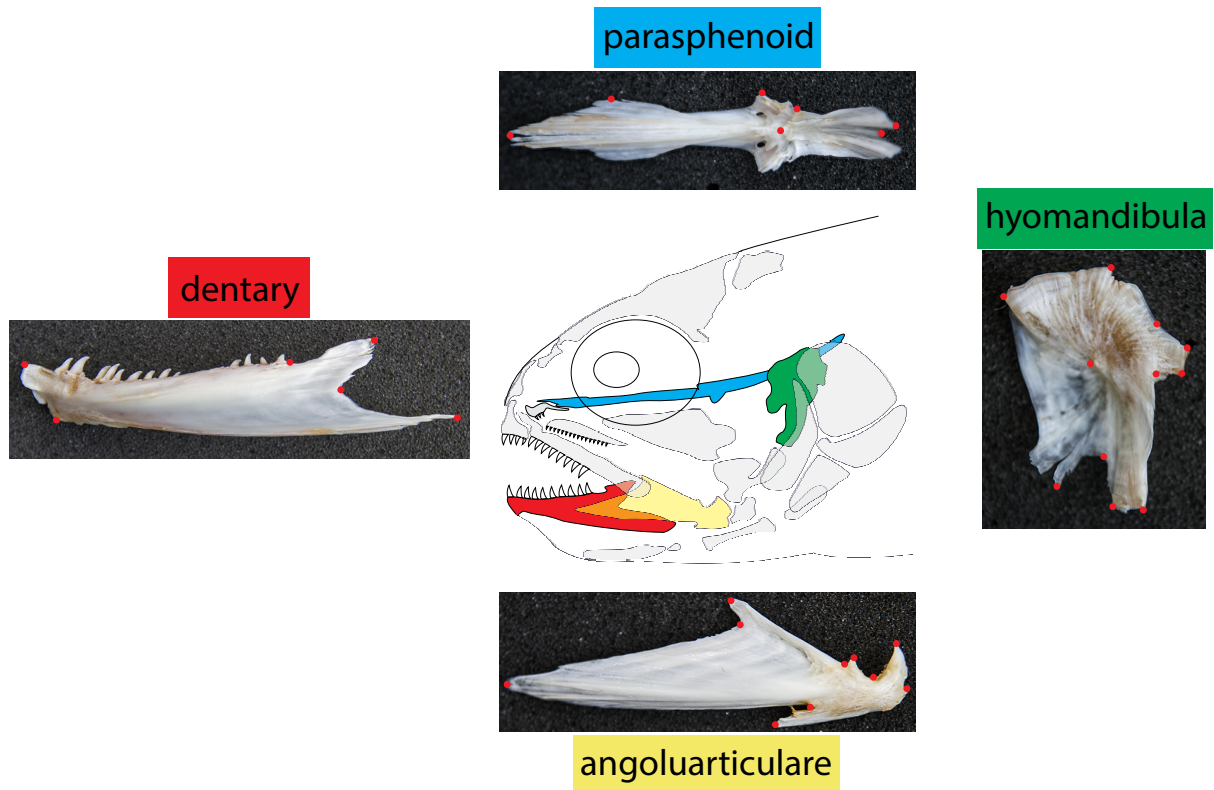


Fig. S3. Summary of landmarks used for morphometrics of charr lineages.

Comparisons dentary, parasphenoid, hyomandibula, and anguloarticular were made among Dolly Varden, Bigmouth, Longhead, Nosed1, Smallmouth, and White lineages. Landmarks are represented by red dots.

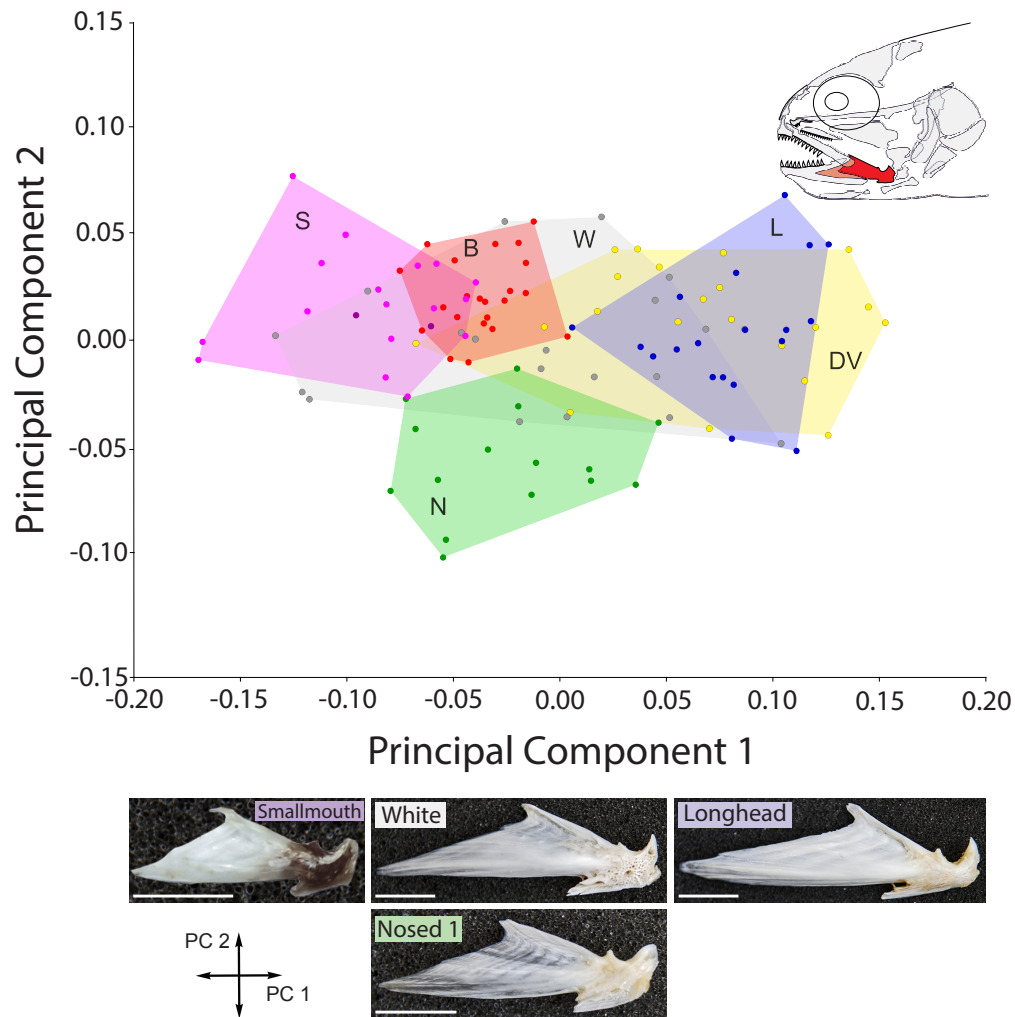


Fig. S4. Principal components analysis of anguloarticular shape.

Anguloarticular is highlighted in red in *S. malma* skull schematic in the upper right. Principal components 1 and 2 are depicted in the graph. Representative examples of dry osteological samples are shown below.

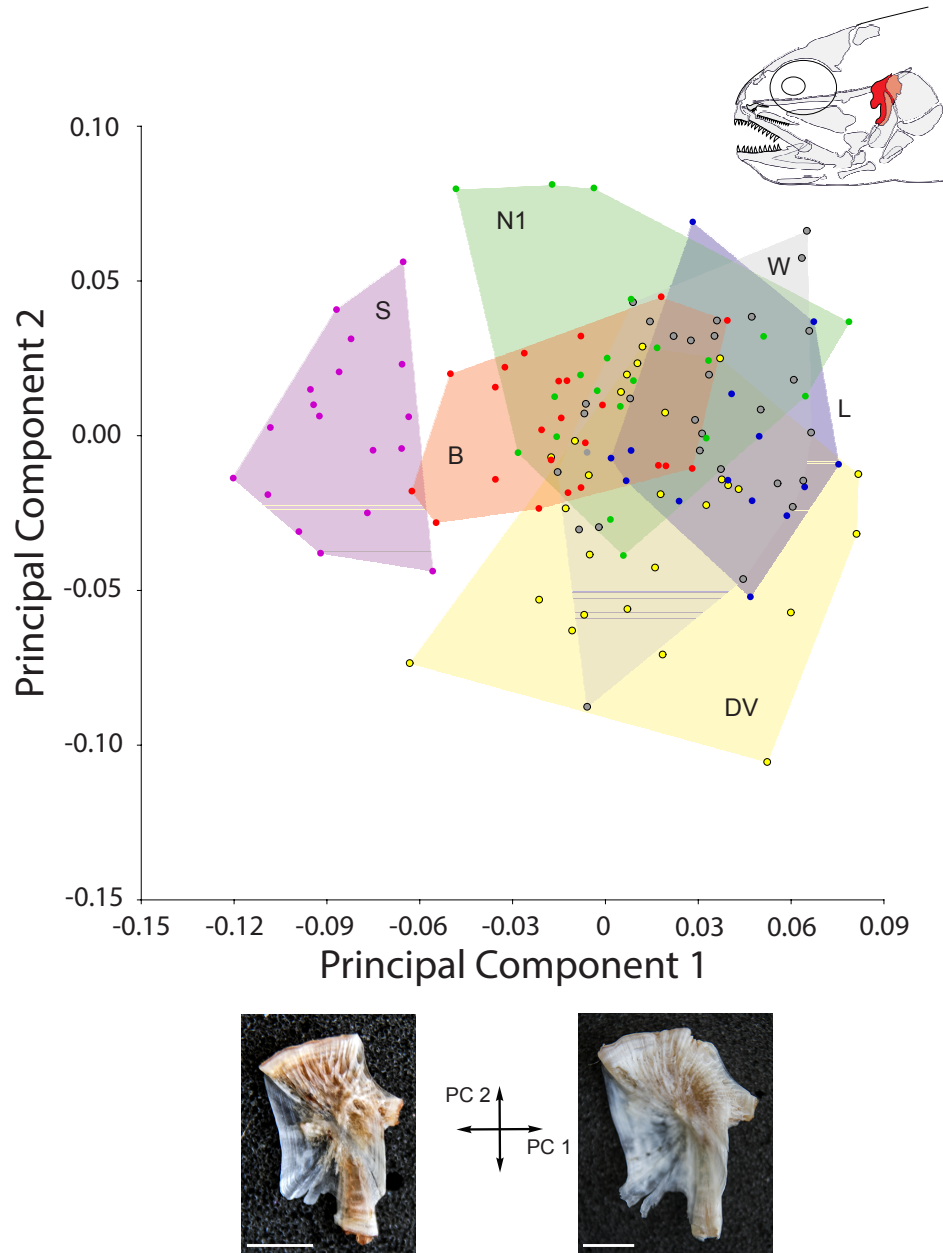


Fig. S5. Principal components analysis of hyomandibula shape. Hyomandibula is highlighted in red in *S. malma* skull schematic in the upper right. Principal components 1 and 2 are depicted in the graph. Representative examples of dry osteological samples are shown below.

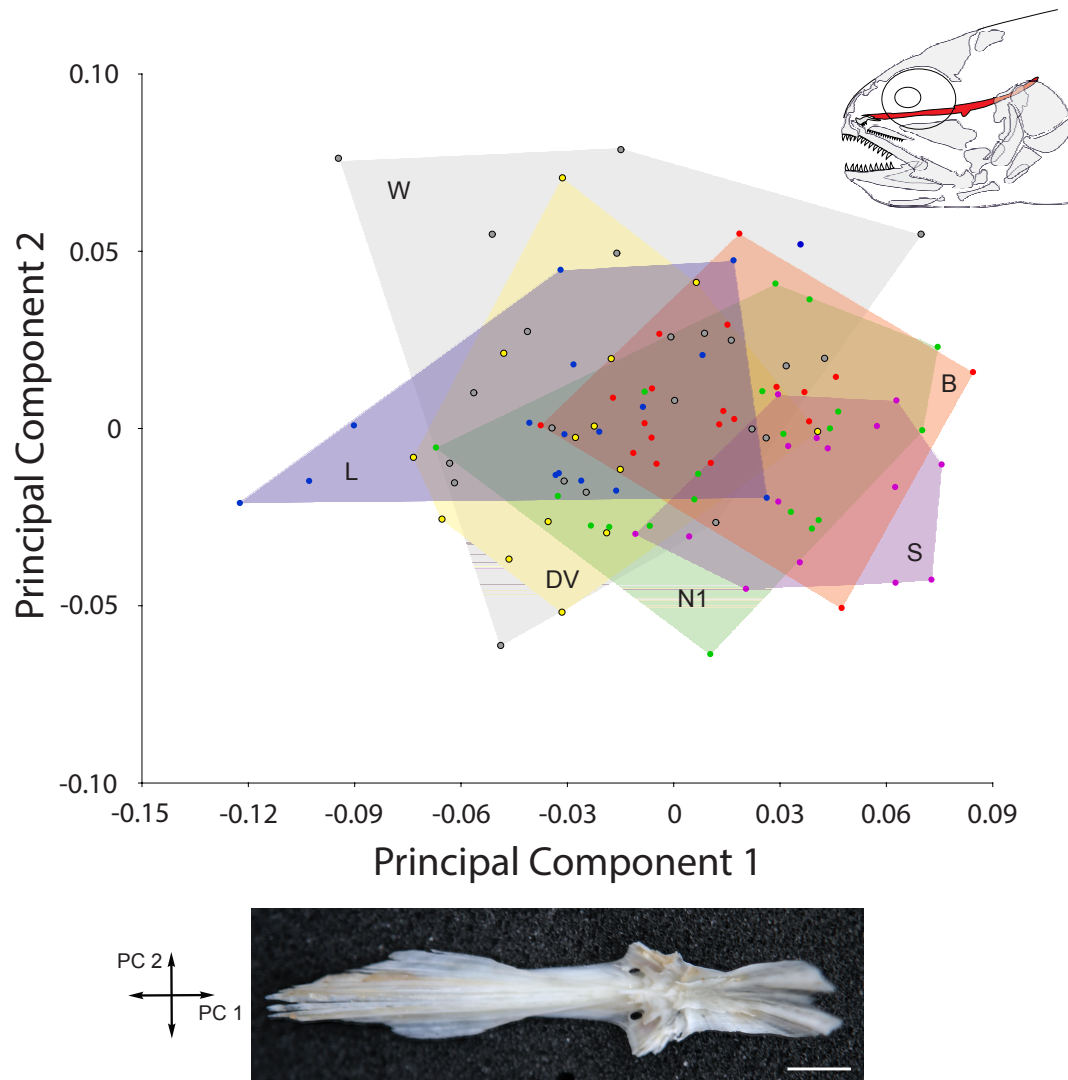


Fig. S6. Principal components analysis of parasphenoid shape. Parasphenoid is highlighted in red in *S. malma* skull schematic in the upper right. Principal components 1 and 2 are depicted in the graph. Representative example of parasphenoid is shown below.

Table S1. Summary of reads aligned and targeted element coverage.

Lineage	No. Individuals Pooled	No. Paired Raw Reads (million)	No. Reads Mapped (million)	Percent of reads mapped	Median Depth	Mean Depth	"Conservome" Coverage 2x (percent)	"Conservome" Coverage 10x (percent)
<i>S. leucomaenis</i>	1	55.9	25.5	0.46	50 reads	67 reads	90.50%	85.90%
Bigmouth	8	72.3	35.3	0.49	66 reads	88 reads	91.60%	88.50%
Dolly Varden	3	64.1	32.4	0.51	61 reads	82 reads	91.60%	88.30%
Longhead	10	46.1	22.8	0.49	41 reads	56 reads	91.10%	85.70%
Nosed 1	5	68.8	34.9	0.51	65 reads	87 reads	91.70%	88.50%
Nosed 2	7	49.5	17.5	0.35	16 reads	20 reads	88.70%	65.50%
Nosed 3	5	68.4	34.1	0.5	63 reads	84 reads	91.60%	88.50%
Smallmouth	6	43.1	21.1	0.49	38 reads	52 reads	90.80%	84.90%
White	6	36.9	18.5	0.5	34 reads	46 reads	90.60%	83.80%

Table S2. Table of significantly underrepresented GO terms. This table contains the set of significantly underrepresented GO terms among all sample libraries (Benjamini-Hochberg FDR < 0.05).

GO Name	GO ID	Bin Size
activation of MAPK activity	GO:0000187	15
anaphase-promoting complex	GO:0005680	51
anaphase-promoting complex-dependent catabolic process	GO:0031145	41
ATP-dependent chromatin remodeling	GO:0043044	133
ATPase activity	GO:0016887	262
calcium ion binding	GO:0005509	1185
calcium, potassium:sodium antiporter activity	GO:0008273	52
carbohydrate binding	GO:0030246	111
carbohydrate metabolic process	GO:0005975	204
cell adhesion	GO:0007155	541
chemokine activity	GO:0008009	72
cullin-RING ubiquitin ligase complex	GO:0031461	23
DNA binding	GO:0003677	2972
DNA integration	GO:0015074	1571
DNA-templated transcription, initiation	GO:0006352	36
fatty acid amide hydrolase activity	GO:0017064	6
ferric iron binding	GO:0008199	24
fructose-bisphosphate aldolase activity	GO:0004332	27
hexose metabolic process	GO:0019318	34
homophilic cell adhesion via plasma membrane adhesion molecules	GO:0007156	362
iron ion transport	GO:0006826	23
isomerase activity	GO:0016853	45
MAP kinase activity	GO:0004707	52
nucleic acid binding	GO:0003676	2731
oligopeptide transport	GO:0006857	20
phosphatidylinositol metabolic process	GO:0046488	49
phosphatidylinositol phosphate kinase activity	GO:0016307	47
plasma membrane	GO:0005886	454
receptor tyrosine kinase binding	GO:0030971	28
regulation of mitotic metaphase/anaphase transition	GO:0030071	44
sensory perception of sound	GO:0007605	69
skeletal muscle fiber development	GO:0048741	44
SWI/SNF complex	GO:0016514	147
transcription coactivator activity	GO:0003713	164
transcription factor TFIID complex	GO:0005669	38
transmembrane transport	GO:0055085	1148
transmembrane transporter activity	GO:0022857	306
transporter activity	GO:0005215	218
transposition, DNA-mediated	GO:0006313	1551
ubiquitin protein ligase binding	GO:0031625	28
vesicle-mediated transport	GO:0016192	250

Table S3. Table of D_{tree} scores, f_4 -admixture ratios, and Z-scores for each of the 20 trios contained within the Lake Kronotskoe species flock. 16 trios were found to have a significant, though minimal, contribution of introgressed alleles (asterisks) (Holm-Bonferoni, FWER < 0.01). B, Bigmouth; L, Longhead; N1, Nosed 1; N3, Nosed 3; S, Smallmouth; W, White.

Pop1	Pop2	Pop3	Dtree(%)	Z-score	p-value	f4-ratio
Smallmouth	Nosed3	White	3.9%	13.0	0	8.6%
Smallmouth	Nosed1	White	3.7%	11.4	0	8.2%
Smallmouth	Nosed3	Longhead	2.8%	11.4	0	6.3%
Smallmouth	Nosed1	Longhead	2.2%	8.3	1.18E-16	4.9%
Smallmouth	Bigmouth	Nosed1	4.2%	8.0	1.56E-15	9.0%
Smallmouth	Bigmouth	White	2.2%	7.7	1.26E-14	4.7%
Bigmouth	Nosed3	White	1.8%	7.6	2.44E-14	4.1%
Longhead	White	Nosed1	2.2%	7.3	2.73E-13	4.5%
Longhead	White	Bigmouth	1.7%	7.3	2.76E-13	3.9%
Bigmouth	Nosed1	White	1.7%	7.0	2.11E-12	3.7%
Smallmouth	Bigmouth	Nosed3	3.6%	6.7	2.31E-11	8.3%
Longhead	White	Nosed3	1.8%	5.9	4.65E-09	3.8%
Bigmouth	Nosed3	Longhead	1.8%	5.0	6.52E-07	4.0%
Smallmouth	Bigmouth	Longhead	1.1%	3.9	0.00010063	2.4%
Bigmouth	Nosed1	Longhead	1.2%	3.6	0.00035939	2.6%
Nosed1	Nosed3	Longhead	0.7%	3.3	0.00112553	1.5%
Longhead	White	Smallmouth	0.5%	3.0	0.00286287	1.4%
Nosed1	Nosed3	Smallmouth	0.4%	1.9	0.0572252	0.9%
Nosed1	Nosed3	White	0.2%	1.1	0.278755	0.5%
Nosed3	Nosed1	Bigmouth	0.3%	1.1	0.285829	0.7%

Table S4. Sequences for primers used in qPCR

Target	Forward Primer	Reverse Primer
<i>lepa</i>	TGACGGGCAAAATTTACTTCCA	AGTGTGGATAGATCTCGGCG
<i>otx2b</i>	CAAGCAACCACCTTACACGG	GAGGAGTCGCTGGGTATCC
<i>pax1b</i>	AGTACACCCAGGCTTCATCA	TGTCCACCGTAAACACCGTA
<i>pax1a</i>	TTGGGGTGTCAATAGAGCGA	GTCGACGAAGGCTGAGGG
<i>sf3b4</i>	ACAGGACAACACCAGGGTTAT	GGGCTTGCCGTAAAGTTTGA
<i>slc26a10</i>	CTGCTTCACAAGAGACTGCC	AAAGCAAACGCCATCCCTTG
<i>actinB1</i>	CGACCAGAAGCGTACAGAGA	AATCCCAAAGCCAACAGAGA

Table S5. This file contains all pairwise Fst values calculated in 10kb sliding windows.

Available for download at

<https://journals.biologists.com/dev/article-lookup/doi/10.1242/dev.203002#supplementary-data>

Table S6. This file contains all sliding windows where $D \geq 0.8$ for the indicated trios. Sliding windows analysis was conducted over intervals containing 40 variants and with 50% overlap between intervals.

Available for download at

<https://journals.biologists.com/dev/article-lookup/doi/10.1242/dev.203002#supplementary-data>

Table S7. This file contains all genewise Fst values for Dolly Varden vs. lake lineages where $F_{st} \geq 0.5$.

Available for download at

<https://journals.biologists.com/dev/article-lookup/doi/10.1242/dev.203002#supplementary-data>

Table S8. This file contains Fst values per CNE for Dolly Varden vs. lake lineages where $F_{st} \geq 0.5$. Genes were assigned to each CNE using GREAT.

Available for download at

<https://journals.biologists.com/dev/article-lookup/doi/10.1242/dev.203002#supplementary-data>

Table S9. This file contains all GO terms assigned to the genes (Table S7) and CNEs (Table S8) with $F_{st} \geq 0.5$. This file further classifies the GO terms that contain a minimum of 6 genes as being unique or shared between the genes and CNEs.

Available for download at

<https://journals.biologists.com/dev/article-lookup/doi/10.1242/dev.203002#supplementary-data>