frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 25 August 2022
DOI 10.3389/fgene.2022.958076

Check for updates

OPEN ACCESS

EDITED BY
Genlou Sun,
Saint Mary's University, Canada

REVIEWED BY
Liandong Yang,
Institute of Hydrobiology (CAS), China
Jiuxuan Li,
Third Military Medical University, China

*CORRESPONDENCE
Baocheng Guo,
guobaocheng@ioz.ac.cn

SPECIALTY SECTION
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

RECEIVED 31 May 2022
ACCEPTED 15 July 2022
PUBLISHED 25 August 2022

CITATION
Jin L, Li Z, Wang C, Wang Y, Li X, Yang J, Zhao Y and Guo B (2022), Contrasting population differentiation in two sympatric Triplophysa loaches on the Qinghai–Tibet Plateau.
Front. Genet. 13:958076.
doi: 10.3389/fgene.2022.958076

COPYRIGHT
© 2022 Jin, Li, Wang, Wang, Li, Yang, Zhao and Guo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Contrasting population differentiation in two sympatric *Triplophysa* loaches on the Qinghai–Tibet Plateau

Ling Jin[1,2], Zitong Li[3], Chongnv Wang[1], Yingnan Wang[1], Xinxin Li[1], Jian Yang[4], Yahui Zhao[1] and Baocheng Guo[1,2,5,6]*

[1]Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]CSIRO Agriculture and Food, Canberra, ACT, Australia, [4]Assessment and Resource Conservation in Middle and Lower Reaches of the Yangtze River, Freshwater Fisheries Research Center, Chinese Academy of Fishery Sciences, Wuxi, China, [5]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China, [6]Academy of Plateau Science and Sustainability, Qinghai Normal University, Xining, China

Genetic differentiation in aquatic organisms is usually shaped by drainage connectivity. Sympatric aquatic species are thus expected to show similar population differentiation patterns and similar genetic responses to their habitats. Water bodies on the Qinghai–Tibet Plateau (QTP) have recently experienced dramatic physicochemical changes, threatening the biodiversity of aquatic organisms on the "roof of the world." To uncover ecological genetics in Tibetan loaches (*Triplophysa*)—the largest component of the QTP ichthyofauna—we characterized population differentiation patterns and adaptive mechanisms to salinity change in two sympatric and phylogenetically closely related Tibetan loaches, *T. stewarti* and *T. stenura*, by integrating population genomic, transcriptomic, and electron probe microanalysis approaches. Based on millions of genome-wide SNPs, the two Tibetan loach species show contrasting population differentiation patterns, with highly geographically structured and clear genetic differentiation among *T. stewarti* populations, whereas there is no such observation in *T. stenura*, which is also supported by otolith microchemistry mapping. While limited genetic signals of parallel adaption to salinity changes between the two species are found from either genetic or gene expression variation perspective, a catalog of genes involved in ion transport, energy metabolism, structural reorganization, immune response, detoxification, and signal transduction is identified to be related to adaptation to salinity change in *Triplophysa* loaches. Together, our findings broaden our understanding of the population characteristics and adaptive mechanisms in sympatric Tibetan loach species and would contribute to biodiversity conservation and management of aquatic organisms on the QTP.

KEYWORDS

population genomics, transcriptomics, otolith EPMA, linkage disequilibrium, salinity

Introduction

Aquatic organisms usually show population differentiation in accordance with the connectivity of water bodies. Therefore, for sympatric species, their genetic differentiation patterns and adaptive mechanisms to hydrological environments are expected to be similar unless affected by other factors (e.g., the degree of gene flow and effective population size difference). Freshwater organisms are vulnerable to drastic climate changes because they have limited abilities to disperse; and availability of water–their habitat–are highly climate-dependent (Sala et al., 2000; Dudgeon et al., 2006; Woodward et al., 2010; Poff et al., 2012; Tickner et al., 2020). The Qinghai–Tibet Plateau (QTP) is inevitably faced with the threat of climate change (Yao et al., 2012; Yao et al., 2019). For example, the mean air temperature has increased at a rate of 0.16–0.35 °C decade^{-1} from the early 1950s to 2014 on the QTP (Yao et al., 2012; Shen et al., 2015; Yao et al., 2019) in the context of global warming, which results in accelerated melting of glaciers and permafrost degeneration. To this end, the volume of many lakes on the QTP has increased rapidly in the past 40 years (Yang et al., 2017; Zhang et al., 2017; Zhang et al., 2020). Water storage of Lake Sêrling Co increased by 19–20 Gt from the 1970s to 2015 (Zhang et al., 2017). However, lakes in Southern Tibet (e.g., Lake Yumzhog Yumco) slightly decreased the water volume (Qiao et al., 2019). It suggests that water bodies on the QTP are susceptible to climate change. In addition, the salinity of lakes and rivers on the QTP varies widely, ranging from freshwater (e.g., Lake Mapam Yumco of 0.2 psu) to saltwater (e.g., Lake Lungmu Co of 138.6 psu) (Liu et al., 2021). The ongoing dramatic lake volume change may lead to rapid water salinity changes. For example, the salinity of Lake Sêrling Co has reduced from 18.5 g L^{-1} in 1979 to 12.4 g L^{-1} in 2017 (Zhu et al., 2019). Aquatic organisms on the QTP are thus facing challenges due to climate change.

Triplophysa loaches (Cypriniformes: Nemacheilidae), Schizothoracinae carps (Cypriniformes), and Sisoridae catfishes (Siluriformes) are the only native ichthyofauna on the QTP, with *Triplophysa* loaches being the largest component. *Triplophysa* loaches are mainly distributed on the QTP and its adjacent areas (Zhu, 1989). The genus *Triplophysa* originated ~23.5 million years ago (Mya) (Wang et al., 2016) and rapidly diversified into ~150 species accompanying the uplift of the QTP (Chen and Zhu, 1984; Chen et al., 1996; Fricke et al., 2019; GBIF, https://www.gbif.org/). *Triplophysa* loaches could be found in freshwater (e.g., River Lhasa He of 0.1 g L^{-1}; Zhang, 2017) to saltwater (e.g., Lake Ngangla Ringco of 12.8 psu; Liu et al., 2021). For example, *T. stewarti* could be found in Lake Bam Co and Lake Sêrling Co with salinity exceeding the salinity tolerance reported in *Triplophysa* species (4.17 g L^{-1} in *T. yarkandensis* (Yao et al., 2018) and 3.74 g L^{-1} in *T. dalaica* (Wu et al., 2017). Earlier studies in *Triplophysa* loaches mainly focused on taxonomy (Zhu, 1989; Wu and Wu, 1992; Prokofiev, 2010; Kottelat, 2012), systematics (Wang et al., 2016; Chen et al., 2019; Wang T. et al., 2020), and biogeography (Chen et al., 1992; Wang et al., 2012; Feng et al., 2019; Hu et al., 2020). Next-generation sequencing (NGS) data are now largely adopted to identify molecular signatures of adaptation to high altitude (Wang et al., 2015a; Wang et al., 2015b), cave (Zhao et al., 2020), and salt/alkali (Chen et al., 2020) in *Triplophysa* loaches with comparative genomic approach, and several *Triplophysa* genomes have been determined with NGS data (Yang L. et al., 2019; Yang X. et al., 2019; Yuan et al., 2020; Zhou et al., 2021). However, the population genetics and adaptive mechanisms to salinity changes using NGS data have been rarely investigated in *Triplophysa* loaches (but see Chen et al., 2020; Hu et al., 2020; Yuan et al., 2020; Huo et al., 2022). How salinity changes caused by climate changes on the QTP may influence the biodiversity of *Triplophysa* loaches is still underexplored, especially at an intraspecific level.

Population genomics has been widely applied to wild populations. With either outlier analysis or environmental association analysis (EAA), it is possible to identify genetic variants underlying environmental adaptation (Guo et al., 2015; Guo et al., 2016; Ahrens et al., 2018; Liberles et al., 2020). However, these classical methods can only identify loci discretely. Linkage disequilibrium- (LD-) based analysis can screen for contiguous adaptive loci as a cluster and also possibly uncover interacting genetic elements. The LD-based analysis thus has facilitated our understanding of adaptive mechanisms in non-model organisms (Kemppainen et al., 2015; Li et al., 2018; Fang et al., 2020; McKinney et al., 2020). By conducting LD-based network analysis among loci, evolutionary phenomena, including chromosomal rearrangements, local adaptation, and geographic structure, which leave LD signals on genomes, can be identified (Kemppainen et al., 2015; McKinney et al., 2020). However, these earlier LD-based network methods can only handle tens of thousands of loci and thus are usually applicable to restriction-site-associated DNA (RAD) sequencing data or low-coverage Illumina sequencing data due to computational constraints (Kemppainen et al., 2015; McKinney et al., 2020). To deal with more loci, they have to incorporate stepwise dimension reduction (Li et al., 2018; Fang et al., 2020). Herein, we develop a novel sparse graph construction and learning based LD approach to identify the interconnection between multiple loci and cluster correlated loci with millions of loci. Our LD graph learning approach applied sparse high dimensional regression (Meinshausen and Bühlmann, 2006) to estimate the covariance structure among loci by estimating the skeleton of an undirected graph following the terminology of graph representation learning. Then, a graph community detection algorithm (Pons and Latapy, 2006) was used to cluster loci into LD blocks. These inferences were conducted first on loci within each chromosome and then on the genome-wide level. The advantages of our method include utilizing small-scale computational resources, the possibility of utilizing parallel

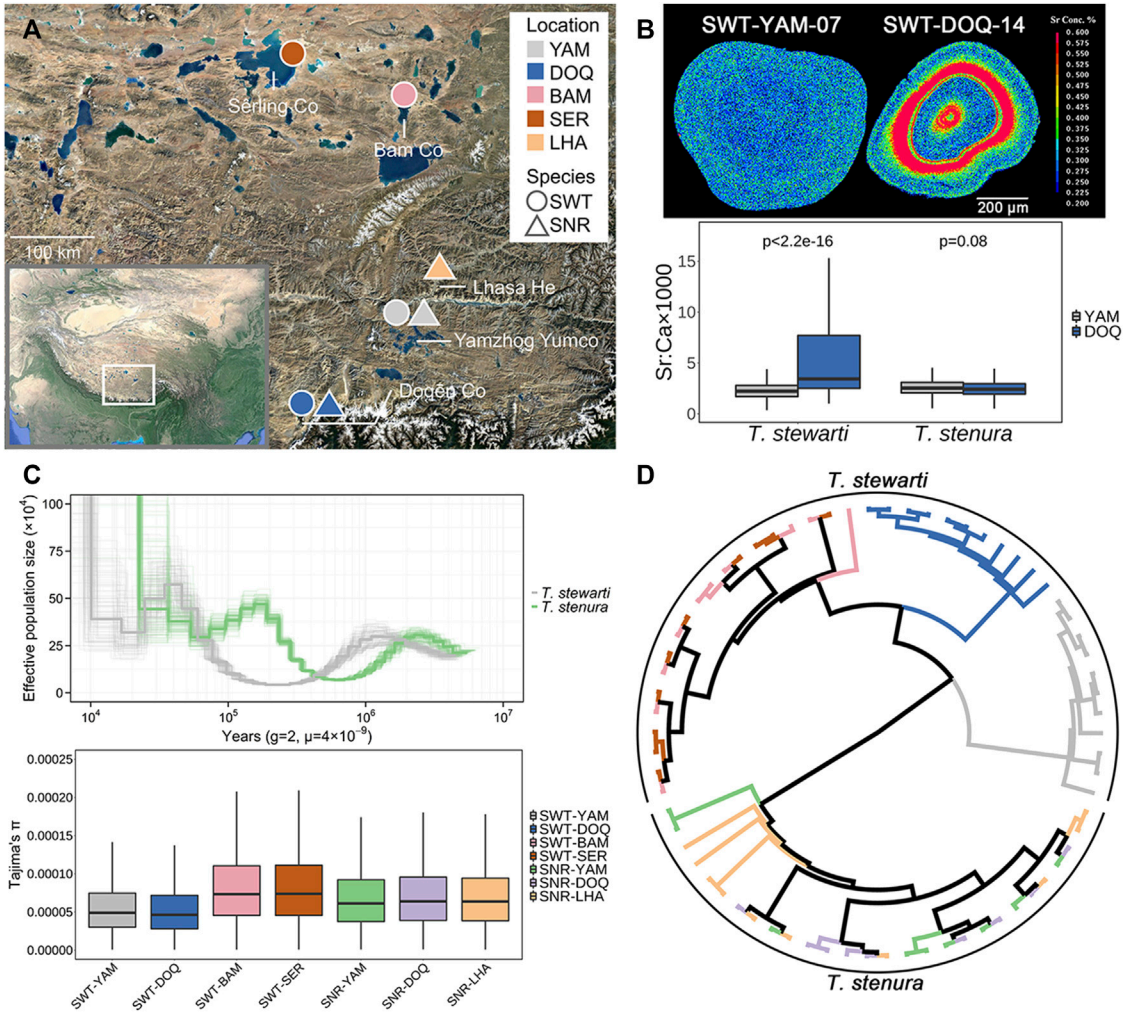

FIGURE 1
(A) Sampling sites in this study. SWT, *T. stewarti*; SNR, *T. stenura*; YAM, Lake Yamzhog Yumco; DOQ, Lake Doqên Co; BAM, Lake Bam Co; SER, Lake Sêrling Co; LHA, River Lhasa He. **(B)** Top panel: two-dimensional imaging of Sr concentrations using otolith EPMA mapping analyses (SWT-YAM-07 and SWT-DOQ-14 taken as examples). Sr concentrations are represented by 16 colors, from red (highest) to green to blue (lowest). Bottom panel: comparisons of otolith Sr:Ca ratios along line transects from the core to the edge between Lake Yamzhog Yumco and Lake Doqên Co populations of the two *Triplophysa* species. The significance test was complemented using the Wilcoxon Rank Sum Test with continuity correction. **(C)** Top panel: effective population size changes of *T. stewarti* and *T. stenura* estimated using PSMC. Thick lines represent population size inferences, and thin lines represent 100 bootstraps. Bottom panel: genetic diversity (Tajima's π) of each population. **(D)** IQ-TREE reconstructed using the "common" SNP dataset. The color of each branch corresponds to the population in C (bottom panel).

computation, avoiding choosing arbitrary LD-related parameters (e.g., without splitting SNPs into non-overlapping windows, or specifying the LD thresholds to cluster loci (Fang et al., 2020)), and the capability of simultaneously inferring the LD structure among multiple loci without the need to calculate pairwise LD values.

In this study, we aim to investigate the population genetics and adaptive mechanisms to salinity changes in two *Triplophysa* loaches, *T. stewarti* and *T. stenura*, and are particularly interested in whether the two *Triplophysa* species show similar population differentiation patterns and common adaptive mechanisms to salinity changes. Therefore, we incorporated population genomic, transcriptomic, and electron probe microanalysis (EPMA) approaches. First, we characterized spatial-temporal habitat changes of the two *Triplophysa* species from sympatric populations using the otolith EPMA technique. Then, we generated single nucleotide polymorphism datasets including 70 individuals of seven populations from five localities of the

TABLE 1 Sampling information of populations used in this study.

Species	Location	Population code	Coordinates	Number of individuals for DNA sequencing	Number of individuals for RNA sequencing	Number of individuals for otolith EPMA	Salinity ($mg \cdot L^{-1}$)
T. stewarti	Yamzhog Yumco	SWT-YAM	29.07°N, 90.39°E	11	7	8	1845.1[a]
T. stewarti	Doqēn Co	SWT-DOQ	28.22°N, 89.39°E	11	2	2	927[a]
T. stewarti	Sêrling Co	SWT-SER	31.67°N, 89.44°E	10	NA	NA	7662.84[a]
T. stewarti	Bam Co	SWT-BAM	31.35°N, 90.65°E	10	NA	NA	8233.89[a]
T. stenura	Yamzhog Yumco	SNR-YAM	29.07°N, 90.39°E	9	3	4	1845.1[a]
T. stenura	Doqēn Co	SNR-DOQ	28.22°N, 89.39°E	9	6	7	927[a]
T. stenura	Lhasa He	SNR-LHA	29.69°N, 90.89°E	10	NA	NA	100[b]

NA indicates that no individual is studied.
[a]Salinity of these locations was adopted from Yang (2018).
[b]Salinity of River Lhasa He was adopted from Zhang (2017).

two *Triplophysa* species to explore their population genetic features by utilizing whole-genome re-sequencing data and identified loci that bear similar signals of population genetic differentiation with a newly developed unsupervised LD graph learning method. Finally, we carried out RNA-seq in sympatric populations to investigate the transcriptomic responses to salinity changes in the two *Triplophysa* species.

Materials and methods

Sample collection

The animal procedures were approved by the Animal Experiment Board of the Institute of Zoology, Chinese Academy of Sciences. Fish specimens were collected in June–July 2018 and May–June 2019 from five locations, including seven populations of the two *Triplophysa* species, *T. stewarti* and *T. stenura* (Figure 1A; Table 1). The base map in Figure 1A was generated using Google Earth Pro 7.3.3.7786 (2020 Google LLC). The fishes were captured with hand seines and/or minnow traps, temporally raised in drinking mineral water, and delivered to a laboratory in Beijing. Fishes for re-sequencing were kept in 95% ethanol, and fin clips were dissected for DNA extraction and sequencing. Fish gills for RNA sequencing were dissected and instantly frozen in liquid nitrogen.

Electron probe microanalysis

The asteriscus otoliths were extracted from 21 individuals, including 18 individuals used in RNA-seq (Table 1). The experimental procedures followed the study of Liu et al. (2020). The extracted otoliths were embedded in epoxy resin (Epofix, Struers, Copenhagen, Denmark), mounted on a glass slide, and ground to expose the core using a grinding machine (70 μm/35 μm, Discoplan-TS, Struers, Copenhagen, Denmark). The otoliths were further polished on an automated polishing wheel (LaboPol-35, Struers, Copenhagen, Denmark). After polishing, the otoliths were cleaned in an ultrasonic bath and rinsed with Milli-Q water. Finally, the otoliths were dried and carbon-coated using a high vacuum evaporator (JEE-420, JEOL Ltd., Tokyo, Japan) before the EPMA examination. The line transect analyses, which measured Sr and Ca concentrations along the longest axis from the core to the edge of the otoliths, as well as the mapping analyses, which measured Sr concentrations evenly on the whole cross section, were complemented using a wavelength dispersive X-ray electron microprobe (JXA-8100, JEOL Ltd., Tokyo, Japan). Strontium titanate ($SrTiO_3$) and calcium carbonate ($CaCO_3$) were used as standards. For the line transect analyses, the accelerating voltage, beam current, and counting time were 15 kV, 2×10^{-8} A, and 15 s, respectively. The electron beam was 2 μm in diameter, spaced at 4 μm intervals. For the mapping analyses, the accelerating voltage, beam current, and counting time were 15 kV, 5×10^{-7} A, and 30 ms, with the electron beam of 2 μm in diameter. The pixel size for the mapping analyses was 3 × 3 μm.

DNA and RNA sequencing

DNA and RNA extraction, library construction, and sequencing were done by Novogene Technology Co., Ltd. Briefly, DNA was extracted from the fin clips and then

sheared into short fragments by ultrasonication. Short DNA fragments were end-repaired, phosphorylated, A-tailed, and ligated to index adapters. After purification and PCR amplification, the constructed libraries were sequenced on the Illumina HiSeq platform with a 150 bp paired-end strategy. In total, 408.84 G of raw data of 70 individuals from the seven populations were produced. Low-quality reads were removed according to the following criteria: 1) paired reads were removed when containing adapters; 2) paired reads were removed if the N content of either single-end read exceeds 10% of the total length; and 3) paired reads were removed if the number of low-quality bases (Q ≤ 5) of either single-end read exceeds 50% of the total length. Read quality control was double-checked using FastQC v0.11.9. Finally, 408.01 G of clean genomic data were obtained (Supplementary Table S1).

Total RNA was first extracted, and magnetic beads with oligo (dT) were used for enriching mRNA. Then, the mRNA was fragmented into short fragments. cDNA was synthesized using these mRNA fragments as templates and random hexamers as primers. The short cDNA fragments were purified, end-repaired, A-tailed, and then ligated to index adaptors. Fragments of 250–300 bp were selected for PCR amplification, purification, and cDNA library construction. These libraries were sequenced on Illumina NovaSeq 6000 platform with a 150 bp paired-end strategy. In total, 299.6 million raw reads of 18 transcriptomes from four populations were obtained (Supplementary Table S2). The number of clean reads for each transcriptome ranged from 11.1 and 19.8 million (Supplementary Table S2), and 286.0 million clean reads were obtained after quality filtering following the criteria of DNA read filtering.

SNP calling

The *T. tibetana* genome was retrieved from GenBank (accession number SOYY00000000; Yang X. et al., 2019) and used as the reference. The reference genome was first indexed using the "bwa index" in BWA Version 0.7.17-r1188 (Li, 2013), and then re-sequencing clean reads were aligned to the reference using "bwa mem." SAMtools (Version 1.8) (Li et al., 2009; Danecek et al., 2021) were used to convert SAM to BAM format, remove PCR duplicate reads, and calculate mapping rate and depth of coverage. The percentage of reads aligned to the reference genome ranged from 74.80% to 87.74%, and the depth of coverage was 4.53 × to 7.08 × (Supplementary Table S1).

Raw SNPs were obtained using BCFtools Version 1.8 (Li, 2011; Danecek et al., 2014; Danecek et al., 2021) commands "bcftools mpileup -C 50 -d 10,000 -q 20 -Q 20" and "bcftools call -m." SNPs were further filtered using "bcftools filter" in BCFtools and VCFtools Version 0.1.13 (Danecek et al., 2011), using the following criteria: 1) SNPs within 10 bp of an indel were removed; 2) only bialleles were retained; 3) only genotypes with a minimum depth of 2 and a maximum depth of 20 were retained; 4) SNPs with a proportion of missing data >50% across all populations and within each population were removed; 5) only SNPs with quality value ≥25 were retained; and 6) only SNPs with a minor allele frequency ≥0.05 were retained.

With the abovementioned criteria, an SNP dataset, hereinafter referred to as the "common" SNPs, was generated by including all 70 individuals from the two *Triplophysa* species, as well as "individual" SNP datasets for each of the two *Triplophysa* species. The "common" SNP dataset consists of 590,335 SNPs with 589,179 SNPs on 25 chromosomes and 1,156 SNPs on 267 scaffolds. The "individual" SNP dataset for *T. stewarti* contained 351,745 SNPs with 350,930 SNPs on 25 chromosomes and 815 SNPs on 267 scaffolds, and the SNP dataset for *T. stenura* contained 235,689 SNPs with 234,682 SNPs on 25 chromosomes and 1,007 SNPs on 267 scaffolds, respectively. The number of shared SNPs between the two "individual" SNP datasets was 18,428.

Effective population size, genetic diversity, and population structure analysis

In order to estimate effective population size changes in the two species, two individuals of SWT-YAM-10 and SNR-YAM-03 were further DNA-sequenced to ~100 × depth of coverage. Read mapping, sorting, and PCR duplicates removing were complemented, as shown in the above section. Effective population size changes were estimated using Pairwise Sequentially Markovian Coalescent (PSMC) Version 0.6.5-r67 (Li and Durbin, 2011). Consensus calling was done using "samtools mpileup -C50 -Q20 -q20 -d10000," "bcftools view -c," and "vcfutils.pl vcf2fq -d20 -D120 -Q20." Fastq format was converted to a psmc input file format using fq2psmcfa. PSMC was run with options of -N25 -t15 -r5 -p "4+25*2+4+6". Bootstrap was conducted by randomly sampling with replacement from 500 kb sequence segments generated by the splitfa tool, and 100 rounds were performed. The effective population size changes were calculated with the generation time of 2 years and 4×10^{-9} substitutions per synonymous site per year following the study on *T. bleekeri* (Yuan et al., 2020) and visualized using psmc_plot.pl and R package "ggplot2" v3.3.2 (Wickham, 2016).

Non-overlapping 100 kb sliding window Tajima's π was calculated based on "individual" SNP datasets using VCFtools to estimate genetic diversity in each population. To investigate genome-wide differentiation in each species, Weir and Cockerham's F_{ST} in non-overlapping 100 kb sliding windows between populations was calculated by VCFtools using "individual" SNP datasets. Windows with 15–205 SNPs, comprising ~90% of genome-wide windows, were included in genetic diversity and F_{ST} calculations. Principal component analysis (PCA) was performed using PLINK Version

1.90b6.18 (Chang et al., 2015; Purcell and Chang, 2020) with the "common" dataset and two "individual" datasets, respectively. Population stratification and genetic admixture were inferred based on "individual" SNP datasets of the two species using ADMIXTURE v1.3.0 (Alexander et al., 2009). Genetic clusters (K) ranging from 1 to $n + 1$ (n is the number of populations in the dataset) were applied. The optimal K was determined through a cross-validation procedure (Alexander and Lange, 2011). Maximum-likelihood trees were reconstructed using IQ-TREE multicore Version 2.0.6 (Nguyen et al., 2014; Minh et al., 2020), with 1,000 ultrafast bootstrap replicates to determine branch confidence values (Hoang et al., 2017). The best-fitting model was estimated using ModelFinder and selected based on the corrected Akaike information criterion (cAIC) (Kalyaanamoorthy et al., 2017).

Identifying genetic differentiation with LD-based analyses

An unsupervised linkage disequilibrium graph learning method (hereinafter referred to as LD graph learning) was developed to identify loci showing similar genetic differentiation signals. The object of LD graph learning is first to build an undirected correlation network among SNPs on the basis of their pairwise LDs and then use a graph-based clustering approach to classify SNPs into LD blocks. The proposed LD graph learning comprises two steps: chromosome-wise LD network construction and genome-wide LD network construction.

First, the local LD network was constructed on the basis of SNPs separately in each chromosome using an adapted version of the sparse high dimensional graph estimation approach introduced in Meinshausen and Bühlmann (2006). In turn, each SNP j (j = 1, .., p_c; p_c is the total number of the SNPs in the chromosome c) is considered a response variable, and all the SNPs other than SNP j are considered explanatory variables in a multinomial LASSO regression model (Friedman et al., 2010), defined as

$$\max_{\beta_k}\left[\frac{1}{n}\sum_{i=1}^{n} y_{ik}\log p\left(x_{ij}=k\middle|x_{il}\right)-\lambda\sum_{l\neq j}\left|\beta_{kl}\right|\right] \tag{1}$$

with $p(x_{ij}=k|x_{il}) = \frac{\exp(\sum_{l\neq j} x_{il}\beta_{kl})}{\sum_{s\in\{0,1,2\}}\exp(\sum_{l\neq j} x_{il}\beta_{sl})}$,where x_{ij} is the genotype value of individual i and SNP j, coded as x_{ij} = 0, 1, 2 for genotypes AA, AB, and BB, respectively. y_{ik} is an indicator variable: if $x_{ij} = k$ (k = 0, 1, or 2), y_{ik} = 1; otherwise, y_{ik} = 0. β_{kl} is the regression parameter of SNP x_{il}, $\lambda\sum_{l\neq j}|\beta_{kl}|$ is a penalty term that can shrink many of the regression parameters to zero, when there is no correlation between these SNPs and the SNP j, and λ is a shrinkage factor that decides the number of SNPs to have zero regression coefficients. In practice, Equation 1 was solved by the coordinate descent algorithm, implemented in the R package "glmnet" (Friedman et al., 2010). The optimal value of the shrinkage factor λ was determined by maximizing the Akaike information criterion (AIC). From Equation 1, the non-zero regression coefficients $\sqrt{\sum_{k\in(0,1,2)}\beta_{kl}^2} > 0$ indicate that the SNP j and l are in LD with each other. If SNP j and l are presented as two nodes in an undirected graph, there is an edge between j and l. Consequently, on the basis of the output of Equation 1, we can generate an undirected and unweighted graph, which can be represented by a $p_c \times p_c$ adjacency matrix A, with element A_{lj} = 1 indicating SNP l and j are in LD and A_{lj} = 0 indicating l and j are not in LD. The graph is expected to be sparse, meaning that most of the elements of A equal zero. We then applied a random-walk-based clustering approach (Pons and Latapy, 2006) to define LD blocks, using the function "cluster_walktrap" in the R package "igraph" (Csardi and Nepusz, 2006).

Next was to construct the genome-wide network based on the chromosome-wise clusters. To reduce the dimension of the data, in each chromosome-wise cluster, we calculated the principal components of the SNPs within that cluster and then took the first PCs that accumulatively explained 80% of the genetic variance to be used as covariates in the genome-wide LD network analysis. Considering each chromosome-wise cluster C_j (j = 1, . . . , M) as the basic unit (like "SNP"), a group LASSO regression (Yuan and Lin, 2006) was then applied to detect the correlation between each cluster j and all other clusters as

$$\min_{\alpha_j}\frac{1}{2n}\left(z_{ij1}-\sum_{l\neq j}\sum_{k=1}^{m_j} z_{ilk}\alpha_{lk}\right)^2+\lambda\sum_{l\neq j}\left\|\sum_{k=1}^{m_j}\alpha_{lk}^2\right\|_2 \tag{2}$$

where z_{ij1} (j = 1, . . . , M; M is the total number of chromosome-wise clusters) represents the first PC within the cluster j, z_{ilk} (l = 1, . . . , m_l; m_l is the total number of PCs in cluster l) are the PCs defined in clusters other than j, α_{lk} are the regression coefficients that define the level of correlation between the first PC in cluster j and all the PCs in cluster l, and $\lambda\sum_{l\neq j}\|\sum_{k=1}^{m_j}\alpha_{lk}^2\|_2 = \lambda\sum_{l\neq j}\sqrt{\sum_{k=1}^{m_j}\alpha_{lk}^2}$ is the group LASSO penalty, and it can shrink a group of regression parameters defined within a cluster to zero if that cluster was not correlated with the cluster j. The optimal shrinkage was again decided by AIC. Here, algorithm (2) generates a sparse undirected and unweighted graph of M nodes representing M chromosome-wise clusters, and the edge between a pair of clusters tells whether the two are correlated with each other. The genome-wide SNP clusters were then obtained using the same random-walk-based clustering approach as in chromosome-wise LD network construction.

To further validate the robustness of our LD graph learning approach in detecting clusters of highly differentiated SNPs, an earlier developed LD-based network analysis, three-step LDna (Kemppainen et al., 2015; Fang et al., 2020), was used. Three-step LDna was performed following the procedures described in Fang et al. (2020). In brief, pairwise LD (measured by r^2

statistic) between SNPs was first calculated using ngsLD v1.1.1 (Fox et al., 2019) within the maximum distance of 100 SNPs between any two SNPs. SNPs were then divided into non-overlapping windows within each chromosome using the "LDnClusteringEL" function in LDna package v0.65 (Kemppainen et al., 2015; Li et al., 2018). The desired number of SNPs per window ("nSNPs" parameter) was set as 800 for *T. stewarti* populations and 1,000 for *T. stenura* populations, with other parameters of w1 = 10, LD_threshold1 = 0.5, and LD_threshold2 = 0.8. Maximally connected loci (MCL) that exhibited the highest median LD with all other SNPs in the same window were chosen. Next, MCL were used to perform LDna within each chromosome. The parameters were set to $\lambda_{lim} = 8$ and $|E|_{min} = 50$ to extract single outlier clusters (SOCs). Finally, only MCL in each SOC was selected to perform genome-wide LDna. SOCs were grouped into genome-wide LD clusters, and original SNPs were extracted. In *T. stenura*, only 4.513 windows were obtained in LDnClusteringEL, so genome-wide LDna was directly performed using the MCL of these 4,513 windows. Pairwise F_{ST} per site of these LD clusters was calculated using VCFtools. Absolute differentiation D_{xy} was calculated using the script popgenWindows.py (https://github.com/simonhmartin/genomics_general/blob/master/popgenWindows.py).

In order to identify genetic variation associated with salinity variation in each of the two *Triplophysa* species, both the LD graph learning and three-step LDna were performed with "individual" SNP datasets by considering different scenarios of parallel adaptation according to salinity in population locations (Figure 1A; Table 1). Seven populations from the five water bodies of the two species were thus classified as a low-salinity group of River Lhasa He population, mid-salinity group of Lake Yamzhog Yumco and Lake Doqēn Co populations, and high-salinity group of Lake Bam Co and Lake Sêrling Co populations, which was used in the LD graph learning and three-step LDna for adaptation signal detection.

Environmental association analysis

BAYENV2 (Coop et al., 2010; Günther and Coop, 2013) was used to analyze the association between genetic variation and salinity variation with "individual" SNP datasets. The two "individual" datasets were first filtered to estimate the null model, allowing no missing data and limited LD ($r^2 < 0.2$). SNPs showing the maximum/minimum 5% of average F_{ST}, which might indicate a directional or balancing selection, were also removed. The remaining neutral SNPs (9,448 SNPs in *T. stewarti* and 38,047 SNPs in *T. stenura*) were used to estimate covariance matrices from 100,000 iterations. Covariance matrix estimation was independently run three times. By using the mean matrix as the null model and original "individual" SNP datasets, correlations between genetic variation and salinity variation were detected. BAYENV2 was independently run three times using different random seeds for 50,000 iterations. Bayes factors were averaged across the three independent runs. SNPs associated with salinity adaptation using the above unsupervised and supervised methods were annotated using ANNOVAR (version date 7/6/2020; Yang and Wang, 2015).

Differential gene expression analysis

RNA-seq reads were aligned to the reference genome using HISAT2 version 2.2.1 (Kim et al., 2019). 31.53%–67.76% of reads were aligned to the reference genome (Supplementary Table S2). Next, transcripts were assembled and quantified for each transcriptome using StringTie v2.1.4 (Pertea et al., 2015; Pertea et al., 2016; Kovaka et al., 2019) with default parameters and with the reference annotation guiding the assembly processes. The assembled transcripts for each sample were then merged, and the merged transcripts were used to estimate transcript abundances for each sample. The "prepDE.py" script in the StringTie package was used to create a gene count matrix as input for differential expression analysis.

PCR duplicate reads in BAM files were removed using SAMtools. SNP calling and filtering were complemented using BCFtools and VCFtools with parameters the same as in re-sequencing data analysis. The filtering criteria were the same as described in the "SNP filtering" section, except that genotypes with depth >20 were also retained and SNPs with a proportion of missing data >20% across all populations were removed. PCA was performed using PLINK.

Differential expression analysis based on the negative binomial distribution was done by R package "DESeq2" v1.26.0 (Love et al., 2014). The DESeq2 model was corrected internally for library size differences. In order to avoid the distance measure being dominated by a few highly variable genes, rlog-transformed data were used to calculate Euclidean distance between samples. The significance level was adjusted with the Benjamini and Hochberg correction (Benjamini and Hochberg, 1995) to control for the false discovery rate in multiple pairwise comparisons. Only genes with raw count >0 for each sample of each species were kept for extracting differential expressed genes to avoid the impacts of a relatively low depth of coverage. Expression was deemed significantly different for genes with a corrected *p*-value below 0.05.

Results

Ecological and population genetic characteristics

Microchemical mapping with EPMA showed that *T. stewarti* from Lake Doqēn Co has a significantly higher Sr:Ca ratio in

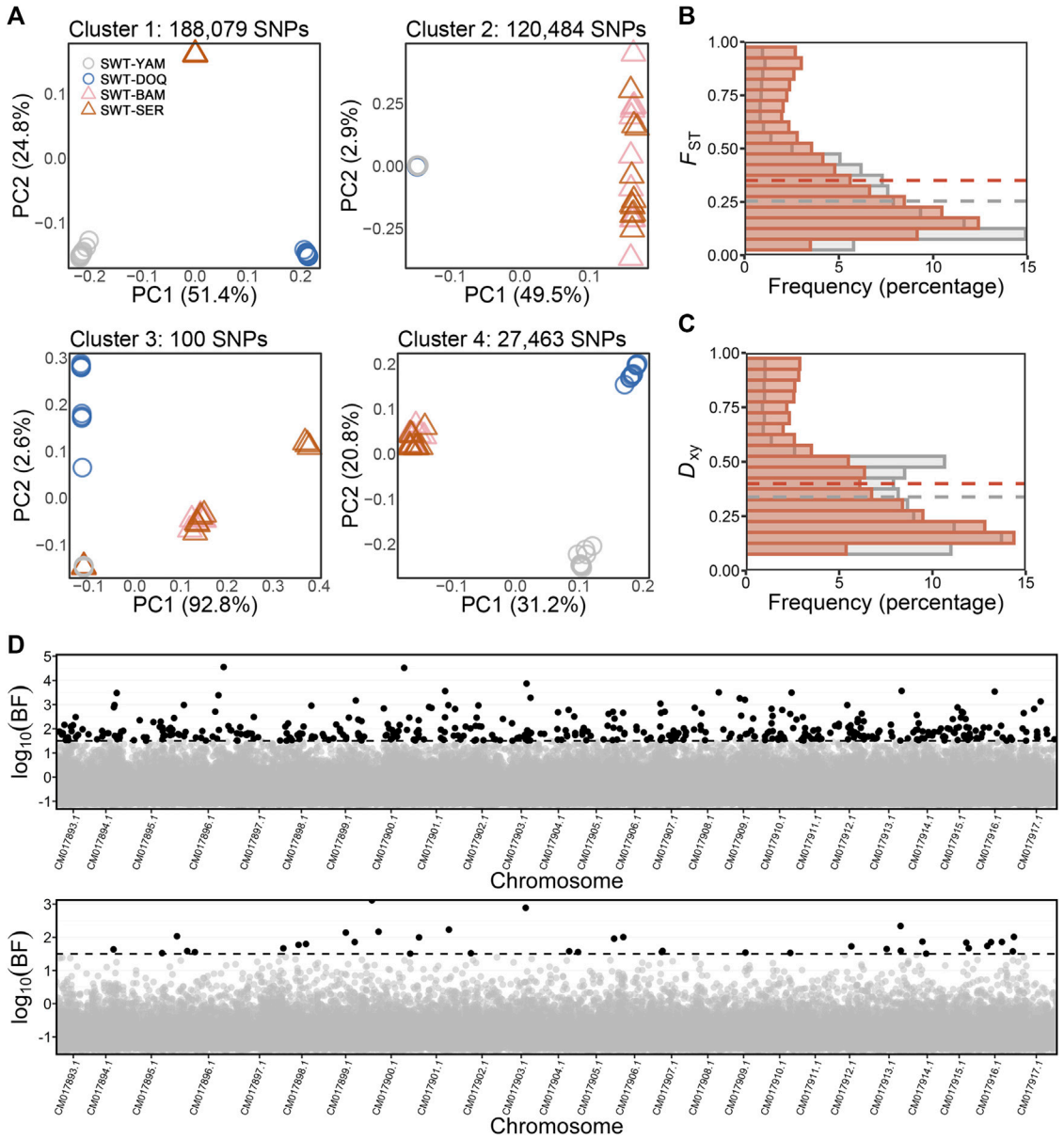

FIGURE 2
(A) PCA of genome-wide SNP clusters identified using the LD graph learning method in all four populations of *T. stewarti*, with cluster 2 showing the pattern of parallel adaptation to salinity between high- (triangles) and mid-salinity (circles) groups. **(B,C)** Distributions of fixation index F_{ST} **(B)** and absolute population differentiation D_{xy} **(C)** of SNPs in genome-wide SNP cluster 2 (red bars) compared with all SNPs (gray bars) between high- and mid-salinity groups. Red and gray dash lines represent the average F_{ST}/D_{xy} of SNPs in genome-wide SNP cluster 2 and F_{ST}/D_{xy} of all SNPs, respectively. **(D)** Environmental association analysis of *T. stewarti* (upper panel) and *T. stenura* (lower panel). SNPs with a Bayesian factor above the threshold ($\log_{10}$(BF) = 1.5, dash line) are colored in black.

otolith than *T. stewarti* from Lake Yamzhog Yumco (the Wilcoxon Rank Sum Test, $p < 2.2e\text{-}16$), whereas the Sr:Ca ratio is not significantly different between populations of *T. stenura* (the Wilcoxon Rank Sum Test, $p = 0.08$). The Sr:Ca ratio in otolith is significantly different between *T. stewarti* and *T. stenura* populations either from different lakes or from the same lakes (the Wilcoxon Rank Sum Tests, $p < 0.001$) (Figure 1B). PSMC showed that both species in Lake Yamzhog Yumco experienced population contraction from about 1–2 Mya, though *T. stenura* population expanded to 5×10^5 at 200 Kya, whereas *T. stewarti* maintained a relatively low effective population size at the same period (Figure 1C). Population

genetic diversity was lower in populations SWT-DOQ and SWT-YAM compared to SNR-DOQ and SNR-YAM, respectively (paired *t*-tests, $p < 2.2e\text{-}16$; Figure 1C). Population genetic relationships were inferred with the "common" and "individual" SNP datasets obtained. PCA and phylogenetic inference with the "common" dataset showed that the two species were explicitly separated into two clusters (Supplementary Figure S1 and Figure 1D). The four *T. stewarti* populations were grouped into three clusters, with two populations (SWT-BAM and SWT-SER) in one cluster. The three populations of *T. stenura* were mixed together (Figure 1D and Supplementary Figure S1). Bayesian clustering inference in ADMIXTURE and pairwise F_{ST} of "individual" SNP datasets also supported the above pattern of population differentiation, with *T. stewarti* populations deeply diverged but *T. stenura* populations unseparated (Supplementary Figure S1, Supplementary Figures S2, S3).

Genetic differentiation associated with salinity variation

Using the newly developed unsupervised LD graph learning approach, four genome-wide SNP clusters were identified with an "individual" SNP dataset of the 350,930 SNPs for the four *T. stewarti* populations, of which the genome-wide SNP cluster 2 containing 120,484 SNPs showed clear genetic differentiation between high- and mid-salinity groups (Figure 2A). SNPs in cluster 2 showed significantly higher differentiation than all SNPs between the two salinity groups according to F_{ST} and D_{xy} values (Figures 2B, C). A total of 45 out of 71 chromosome-wide SNP clusters within this genome-wide SNP cluster 2, composed of 63,344 SNPs, showed higher weighted F_{ST} than all SNPs between high- and mid-salinity groups. The newly developed unsupervised LD graph learning approach was also performed in pairwise high- and mid-salinity *T. stewarti* populations. A total of 350 genome-wide SNP clusters were identified in comparison between SWT-SER and SWT-YAM, 311 in comparison between SWT-SER and SWT-DOQ, 275 in SWT-BAM and SWT-YAM, and 258 in SWT-BAM and SWT-DOQ, of which 15, 10, 14, and 11 genome-wide SNP clusters showed higher F_{ST} than weighted F_{ST} calculated using all SNPs between the two populations, respectively. 1,640 SNPs were found in these highly differentiated genome-wide SNP clusters in all the four *T. stewarti* population pairs (Supplementary Figure S3). Of these 1,640 SNPs, 1,615 SNPs (98.5%) were included in genome-wide SNP cluster 2 when LD graph learning was done with all of the four *T. stewarti* populations, and 1,583 SNPs (96.5%) were in 45 highly differentiated chromosome-wide clusters of genome-wide cluster 2. The unsupervised three-step LDna with the "individual" SNP dataset of the 350,930 SNPs in the four *T. stewarti* populations identified 11 LD clusters, and LD cluster 2, containing 20,311 SNPs, showed similar genetic signatures between high- and mid-salinity populations as the genome-wide SNP cluster 2 identified by LD graph learning (Supplementary Figures S3, S4). 18,545 out of 20,311 SNPs (91.3%) in this LD cluster 2 identified by three-step LDna were also included in the genome-wide SNP cluster 2 identified by LD graph learning. Using these abovementioned unsupervised methods, we identified 1,443 common SNPs showing high genetic differentiation between salinity groups. Furthermore, the supervised EAA method with BAYENV2 identified 410 SNPs, which were strongly associated with salinity variation, and 20 of the 410 SNPs were highly differentiated loci and associated with salinity variation with the abovementioned unsupervised methods (Figure 2D and Supplementary Figure S3). These 20 SNPs were annotated within or close to genes that function in response to osmotic stress or salinity acclimation, for example, hepatocyte nuclear factor 4-alpha (*HNF4A*), facilitated glucose transporter member 1 (*GLUT1*), gamma-glutamyl hydrolase (*GGH*), and dipeptidyl aminopeptidase-like protein 6 (*DPP6*) (Supplementary Table S4).

The newly developed unsupervised LD graph learning, three-step LDna, and supervised EAA with BAYENV2 were also carried out to identify genetic differentiation associated with salinity variation in *T. stenura*. Because PCA of genome-wide SNP clusters did not separate the low-salinity group (SNR-LHA) from the mid-salinity group (SNR-YAM and SNR-DOQ), F_{ST} instead of PCA was used as a benchmark in LD graph learning. 142 out of 368 genome-wide SNP clusters containing 77,730 SNPs showed higher weighted F_{ST} than those of all SNPs. Totally, 61,858 SNPs were commonly found to be in highly differentiated genome-wide clusters in LD graph learning in the two pairwise *T. stenura* populations, SNR-YAM and SNR-LHA, as well as SNR-DOQ and SNR-LHA. Moreover, 22,213 SNPs were commonly found to be in highly differentiated genome-wide SNP clusters when LD graph learning was performed with all populations and between population pairs in *T. stenura*. Using three-step LDna, 16 SOCs containing 2,265 SNPs were identified (Supplementary Figure S5). A total of 1,122 SNPs in six SOCs out of these 2,265 SNPs showed higher genetic differentiation, of which 433 SNPs overlapped with the 22,213 SNPs identified using LD graph learning. BAYENV2 detected 38 SNPs strongly associated with salinity in *T. stenura*, with only one SNP overlapping with SNPs identified using the above-unsupervised methods of LD graph learning and three-step LDna (Figure 2D). No common SNP showing genetic differentiation associated with salinity variation was identified between the two species *T. stewarti* and *T. stenura*.

Transcriptome responses to salinity change

The two species of *T. stewarti* and *T. stenura* were separated by the first PC based on 8,025 SNPs from RNA-seq data (Figure 3A).

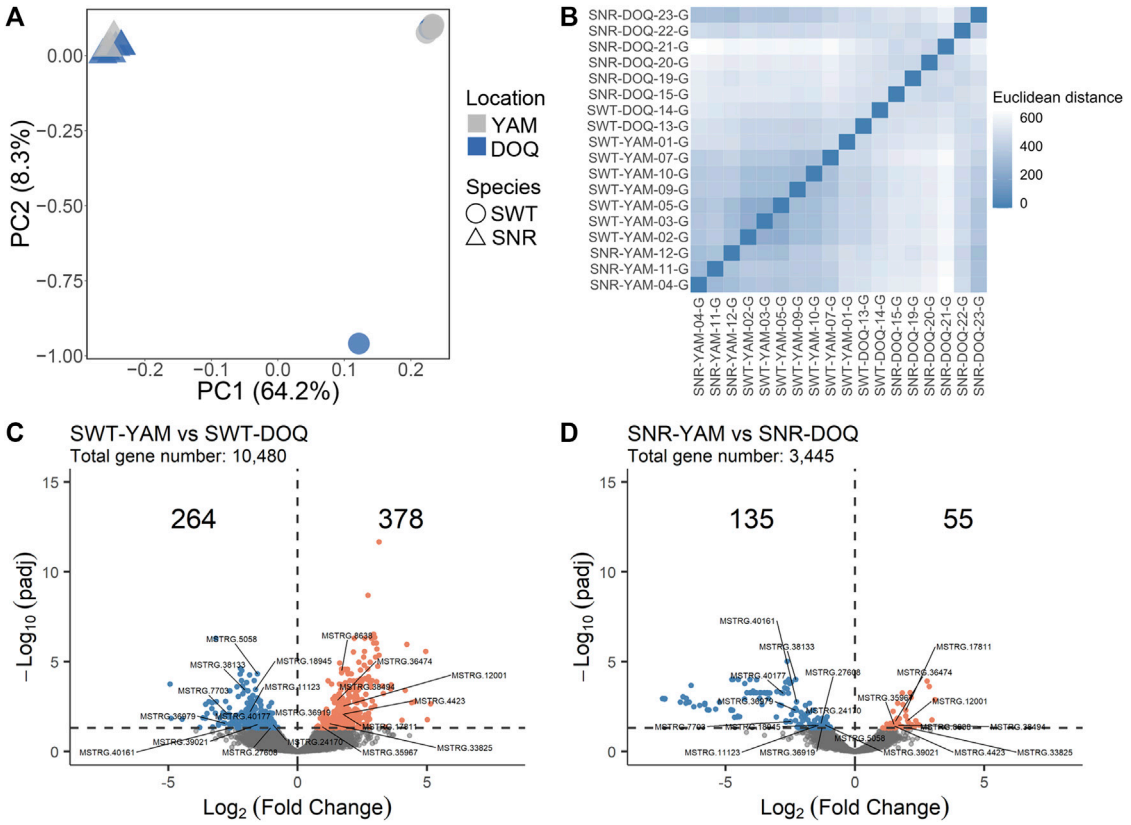

FIGURE 3
(A) PCA of four populations using SNPs called from RNA-seq data. **(B)** Euclidean distances among RNA-seq samples calculated using the gene count matrix after DESeq2 correction for library differences and rlog-transformation. **(C,D)** Volcano plots showing DEGs between Lake Yamzhog Yumco and Lake Doqēn Co populations of *T. stewarti* **(C)** and *T. stenura* **(D)**. The number of genes upregulated (orange dots) and downregulated (blue dots) in Lake Yamzhog Yumco populations are marked. The 20 gene IDs labeled are common DEGs between the two species.

The expression profiles of samples from the same location showed higher similarity (Figure 3B); that is, Euclidean distances among Lake Yamzhog Yumco samples were closer compared to those between samples from Lake Yumzhog Yumco and Lake Doqēn Co., and the same for Lake Doqēn Co samples. In *T. stewarti*, 642 out of 10,480 genes were identified as differentially expressed genes (DEGs) between population SWT-YAM and SWT-DOQ, with 264 downregulated and 378 upregulated in the Lake Yamzhog Yumco population (Figure 3C). In *T. stenura*, 3,445 genes remained after filtering, and 190 genes were identified as DEGs between population SNR-YAM and SNR-DOQ, with 135 downregulated and 55 upregulated genes in the Lake Yamzhog Yumco population (Figure 3D). Twenty genes were identified as common DEGs in both species (Figures 3C,D). Of these 20 genes, eight were novel genes assembled by StringTie and 12 genes were genes with annotations, including peroxiredoxin-1 (*PRDX1*), C member 5 multi-specific organic anion transporter (*ABCC5*), and adenomatous polyposis coli protein (*APC*).

Discussion

The most salient finding in this study is that the two sympatric and phylogenetically closely related Tibetan loach species show contrasting population differentiation patterns, which broaden our understanding of the Tibetan loach diversification on the QTP. In addition, a catalog of genes involved in ion transport, energy metabolism, structural reorganization, immune response, detoxification, and signal transduction is found to be related to adaptation to salinity change in *Triplophysa* loaches, whereas the two phylogenetically closely related Tibetan loach species show limited genetic signals of parallel adaption to salinity changes from either genetic or gene expression variation perspective. We will discuss these findings in detail below. Finally, we highlight the utilization of our newly developed unsupervised LD graph learning approach with large SNP datasets in the molecular ecological study of non-model organisms.

Causes of contrasting population differentiation

Phylogeography provides the key to understanding the evolutionary history of organisms and responses of populations and species to geological events (Schluter, 2000). The phylogeography of *Triplophysa* loaches on the QTP is thought to result from the occurrence of geological events associated with the QTP uplift (Wang et al., 2016). Based on millions of genome-wide SNPs, *T. stewarti* populations are deeply diverged into three clusters of the BAM/SER populations, the YAM population, and the DOQ population (Figure 1D), which may be attributed to geographical barriers. The Gandisê Mountains and the Nyainqêntanglha Mountains divided these lakes and rivers into the northern part (including Lake Sêrling Co and Lake Bam Co) and the southern part (including Lake Yamzhog Yumco, Lake Doqēn Co, and River Lhasa He). These two parts were once connected when the Qiangtang Palaeolake, merging Sêrling Co and Bam Co together, outflew into the external drainage system of southern Tibet during 40–35 ka BP, but then disconnected due to lake level decrease (Zhao et al., 2002a; Zhao et al., 2002b). As a result, the geographical barriers facilitate the genetic divergence between the northern (BAM/SER) and southern (YAM and DOQ) populations in *T. stewarti*. As lake volume in the northern part is rapidly increasing in the past 40 years, the genetic mixture between BAM and SER populations in *T. stewarti* suggests that the drainage is still connected between Sêrling Co and Bam Co nowadays. Interestingly, the three *T. stenura* populations are genetically mixed, which is different from sympatrically distributed *T. stewarti* populations in Yamzhog Yumco and Doqēn Co (Figure 1D). The difference between sympatric populations in the two species of Tibetan loaches is also seen in trajectories of changes in effective population size. It showed that *T. stewarti* had experienced a recent population contraction; meanwhile, *T. stenura* experienced population expansion in Yamzhog Yumco, and as such, a lower genetic diversity is found in *T. stewarti* than that in *T. stenura* in Yamzhog Yumco (Figure 1C). In addition, otolith Sr:Ca ratio EPMA showed that sympatric populations in the two species possess different ecological characteristics. Otolith Sr:Ca ratio EPMA reflects the spatial-temporal habitat changes in fish (Campana, 1999; Secor and Rooker, 2000; Yang et al., 2006). Thus, genetic divergence patterns in sympatric populations in the two species are in accordance with their ecological divergence patterns. Furthermore, according to Wu and Wu (1992), *T. stewarti* and *T. stenura* possess different morphs of the posterior chamber of the swim bladder. The posterior chamber of the swim bladder in *T. stewarti* is well developed but degenerated in *T. stenura*. With a well-developed swim bladder, *T. stewarti* might prefer still water or argodromile, whereas *T. stenura* prefers flowing water or riffle streams. *T. stewarti* thus may possess restricted dispersal capability compared with *T. stenura*. Therefore, the morphological difference in the posterior chamber of the swim bladder between the two *Triplophysa* loaches might explain their explicit ecological and genetic differentiation between sympatric populations. Herein, we further confirmed that the deeply diverged *T. stewarti* populations belong to one taxonomic species rather than different ones. According to Wu and Wu (1992) and Zhu (1989), *Triplophysa* species (*T. brevicauda*, *T. leptosoma*, *T. microps*, *T. stoliczkai*, *T. stenura*, and *T. orientalis*), showing sympatric distributions with our sampling sites, are morphologically distinct to *T. stewarti*, with caudal peduncle laterally compressed and/or swim bladder reduced, whereas *T. stewarti* showed more slender caudal peduncle with roughly round-shaped cross-section and developed posterior chamber of the swim bladder. Moreover, the genetic distance among *T. stewarti* populations measured by Kimura 2-parameter (K2P) distance of mitochondrial *COI* gene sequences ranged from 0.0013 to 0.0170, within the range of intraspecific genetic distance in *Triplophysa* (Wang T. et al., 2020). In fact, many species in *Triplophysa* loaches are distributed in sympatry as the two species *T. stewarti* and *T. stenura* in our study, with the posterior chamber of the swim bladder well developed in one species but reduced in another (Zhu, 1989; Wu and Wu, 1992). It seems that divergence in functional traits, such as swim bladder, might have played an important role in *Triplophysa* diversification along with geological events.

Genetic mechanisms underlying adaptation to salinity change

Lakes and rivers on the QTP exhibit large salinity heterogeneity (Liu et al., 2021). Worldwide climate change has recently led to drastic lake volume changes on the QTP (Yang et al., 2017; Zhang et al., 2017; Zhang et al., 2020), which may result in continuous salinity changes in these lakes. Thus, adaptation to salinity changes is key to the biodiversity evolving in aquatic organisms on the QTP. In this study, we investigated the genetic mechanisms of salinity adaptation in the two *Triplophysa* species: *T. stewarti* and *T. stenura*. First, the two *Triplophysa* species showed limited genetic signals of parallel adaption to salinity changes from either genetic or gene expression variation. A total of 39 genes were identified to harbor highly differentiated loci associated with adaptation to salinity changes in both *T. stewarti* and *T. stenura* with millions of genome-wide SNPs using unsupervised genetic differentiation methods. With RNA-seq data, 20 common DEGs in gill were identified between *T. stewarti* and *T. stenura*, of which few seem to be associated with salinity adaptation (Figures 3C,D). In fact, non-parallel genetic changes more frequently occur at the interspecific

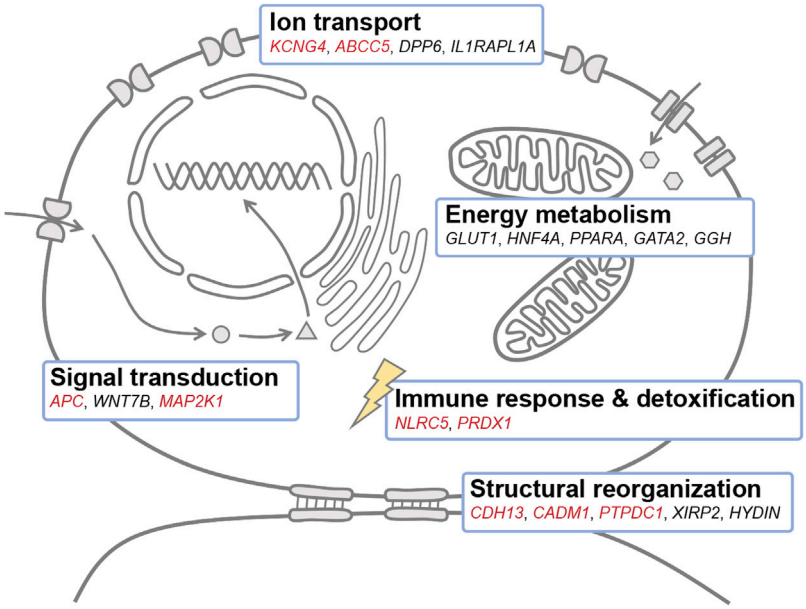

FIGURE 4
Candidate genes associated with adaptation to salinity change in *Triplophysa* loaches. Genes marked in red were associated with adaptation to salinity changes in both *T. stewarti* and *T. stenura*. Genes marked in black were identified in *T. stewarti*. The schematic model in which genes are assigned to different functional categories was adapted from Nguyen et al. (2016). *KCNG4*: potassium voltage-gated channel subfamily G member 4, *ABCC5*: C member 5 multi-specific organic anion transporter 5, *DPP6*: dipeptidyl aminopeptidase-like protein 6, *IL1RAPL1A*: interleukin-1 receptor accessory protein-like 1-A, *GLUT1*: facilitated glucose transporter member 1, *HNF4A*: hepatocyte nuclear factor 4-alpha, *PPARA*: peroxisome proliferator-activated receptor alpha, *GATA2*: GATA-binding factor 2, *GGH*: gamma-glutamyl hydrolase, *NLRC5*: NOD-like receptor C5, *PRDX1*: peroxiredoxin-1, *CDH13*: cadherin-13, *CADM1*: cell adhesion molecule 1, *PTPDC1*: protein tyrosine phosphatase domain-containing protein 1, *XIRP2*: Xin actin-binding repeat-containing protein 2, *HYDIN*: hydrocephalus-inducing protein-like protein, *APC*: adenomatous polyposis coli protein, *MAP2K1*: dual specificity mitogen-activated protein kinase 1.

level, especially with adaptation to given selection pressure involving polygenes, such as adaptation to salinity change in fishes (Wang and Guo, 2019), although parallelism often underlies adaptation to a similar environment (Stern, 2013). For example, genetic changes associated with repeated marine-freshwater divergence in two geographically coexisting and ecologically similar stickleback species, nine- and three-spined sticklebacks, are largely non-parallel (Wang Y. et al., 2020; Kemppainen et al., 2021). Second, we noticed that prevalent genetic changes associated with salinity change were identified in *T. stewarti* but few in *T. stenura* (Figure 2D), which might be due to *T. stewarti* having populations adapted to Lake Bam Co and Lake Sêrling Co with ~8 mg L^{-1} of salinity and high genetic connectivity among *T. stenura* populations (Figure 1D). It suggests that *T. stewarti* is a better model for understanding genetic mechanisms of adaptation to salinity change on the QTP.

A catalog of genes involved in ion transport, energy metabolism, structural reorganization, immune response, detoxification, and signal transduction may be responsible for adaptation to salinity change in *Triplophysa* loaches (Figure 4). A total of 26 genes were commonly identified with both unsupervised and supervised approaches to be associated with adaptation to salinity change in *T. stewarti*, and they are involved in energy metabolism to maintain osmotic pressure, ion transport, structural reorganization, and signal transduction to transduce signals in osmosensing (Supplementary Table S4). Of genes involved in energy metabolism, *HNF4A* is involved in the transcriptional regulation of long-chain polyunsaturated fatty acid biosynthesis (Dong et al., 2016), and it is assumed to play an important regulatory role in osmotic acclimation in rabbitfish (Wang et al., 2018), killifish (Whitehead et al., 2012), and striped catfish (Schmitz et al., 2017). *GLUT1* is responsible for glucose uptake and energy supplementation under salinity stress (Mueckler, 1994), and earlier studies showed that the expression of *GLUT1* is upregulated in response to salinity increase in sea bream (Balmaceda-Aguilera et al., 2012) and striped catfish (Nguyen et al., 2016). Of genes related to ion transport, *DPP6* promotes cell surface expression and modulates the activity and gating characteristics of the potassium channel KCND2 (Soh and Goldstein, 2008; Seikel and Trimmer, 2009). Of 39 genes identified to be associated with adaptation to salinity changes in both *T. stewarti* and *T. stenura* using unsupervised genetic differentiation methods, *KCNG4* is presumed to be involved in ion transport under osmotic stress (Mederos et al.,

2009), and genes (e.g., *CDH13*, *CADM1*, and *PTPDC1*) are involved in structural reorganization during cell volume change under salinity stress (Nguyen et al., 2016). Genes, *CBFB* and *NLRC5*, are identified to be associated with salinity change with BAYENV2 in both *Triplophysa* species. *NLRC5* was involved in the immune response during the smoltification process in Atlantic salmon (Pontigo et al., 2016), marine-freshwater divergence in three-spined stickleback (Terekhanova et al., 2014), and biological invasions in round goby (Adrian-Kalchhauser et al., 2020). Of 20 common DEGs identified in both *Triplophysa* species, the *APC* gene may function in Wnt signaling by promoting the rapid degradation of CTNNB1 (Aoki and Taketo, 2007), and Wnt signaling has been demonstrated to involve a marine-freshwater divergence in three-spined sticklebacks (Yu et al., 2009; Jones et al., 2012). Interestingly, *WNT7B* involved in Wnt signaling is associated with adaptation to salinity change in *T. stewarti* (Supplementary Table S4).

Methodological consideration

Graph learning has rarely been used in population genomics data analysis (but see McKinney et al., 2020). However, the graph learning-based method should be applied to identify groups of loci that showed similar evolutionary signals. Our newly developed LD graph learning method produced consistent results with the previously developed three-step LDna method (Fang et al., 2020). The large proportion of overlapping highly differentiated SNPs between our LD graph learning and three-step LDna (Supplementary Figure S3) and disproportionally more SNPs identified as salinity-associated using EAA also identified as highly differentiated in LD graph learning and three-step LDna validate the effectiveness of our newly developed LD graph learning method. Instead of calculating LD values measured by r^2 statistic between every two pairwise SNPs or between SNPs within a limited distance (e.g., a distance of 100 SNPs in three-step LDna), we modeled the correlation between one SNP and all the other SNPs simultaneously and used these correlation coefficient values to represent the correlation between SNPs. In addition, our LD graph learning method incorporates only two steps and uses fewer parameters (e.g., no need to set λ_{lim} and $|E|_{min}$ in LDna and the derived three-step LDna). According to the LDna method (Kemppainen et al., 2015), the values of λ_{lim} and $|E|_{min}$ will affect the clustering of SNPs, though LD clusters are robust to these parameter settings. Compared with early developed LDna (Kemppainen et al., 2015) and network analysis of linkage disequilibrium (McKinney et al., 2020), our LD graph learning method can handle up to millions of loci, which is memory-saving and more applicable to re-sequencing data. Our script also supports parallel computing.

In conclusion, by integrating population genomic, transcriptomic, and electron probe microanalysis approaches, we unraveled the population differentiation patterns and possible adaptive mechanisms to salinity change in two *Triplophysa* species: *T. stewarti* and *T. stenura*. Our results showed contrasting population differentiation patterns between the two species, which may be attributed to their different ecological characteristics and population histories. Using both unsupervised and supervised genetic differentiation methods, including our brand-new LD graph learning method, we found that the two *Triplophysa* species showed limited genetic signals of parallel adaption to salinity changes from either genetic or gene expression variation perspective. However, a catalog of genes involved in iron transport, energy metabolism, structural reorganization, immune response, and signal transduction may function in adaptation to salinity change in *Triplophysa*, of which several genes are also identified in other fishes. These findings help predict how aquatic organisms will respond to similar selective pressures on the QTP when facing challenges from climate changes.

Data availability statement

SNPs and quantitative expression data based on RNA-seq underlying this study have been deposited in Dryad https://doi.org/10.5061/dryad.x3ffbg7mr. Code for the newly developed unsupervised linkage disequilibrium graph learning method has been deposited in Zenodo https://doi.org/10.5281/zenodo.6422021.

Ethics statement

The animal study was reviewed and approved by the Animal Experiment Board of the Institute of Zoology, Chinese Academy of Sciences.

Author contributions

BG conceived the project. LJ, ZL, CW, YW, XL, JY, and YZ analyzed the data. LJ and BG wrote the paper with input from all other authors. All authors read and approved the final manuscript.

Funding

The study was supported by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP, Grant no. 2019QZKK0501), the China Three Gorges Corporation Research Program (Grant no. 0704177), the National Natural Science Foundation of China (Grant no. 32022009), and the CAS Pioneer Hundred Talents.

Acknowledgments

We thank Haibo Liu, Yongli Ding, Jie Bai, and Peiyao Jia for help in sampling, Bohao Fang for help in data analyses, Tao Jiang and Zhongya Xuan for help in otolith EPMA, and Xiong Xiong for sharing environmental information.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.958076/full#supplementary-material

References

Adrian-Kalchhauser, I., Blomberg, A., Larsson, T., Musilova, Z., Peart, C. R., Pippel, M., et al. (2020). The round goby genome provides insights into mechanisms that may facilitate biological invasions. *BMC Biol.* 18 (1), 11. doi:10.1186/s12915-019-0731-8

Ahrens, C. W., Rymer, P. D., Stow, A., Bragg, J., Dillon, S., Umbers, K. D. L., et al. (2018). The search for loci under selection: Trends, biases and progress. *Mol. Ecol.* 27 (6), 1342–1356. doi:10.1111/mec.14549

Alexander, D. H., and Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinforma.* 12 (1), 246. doi:10.1186/1471-2105-12-246

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19 (9), 1655–1664. doi:10.1101/gr.094052.109

Aoki, K., and Taketo, M. M. (2007). Adenomatous polyposis coli (*APC*): A multi-functional tumor suppressor gene. *J. Cell Sci.* 120 (19), 3327–3335. doi:10.1242/jcs.03485

Balmaceda-Aguilera, C., Martos-Sitcha, J. A., Mancera, J. M., and Martinez-Rodriguez, G. (2012). Cloning and expression pattern of facilitative glucose transporter 1 (GLUT1) in gilthead sea bream *Sparus aurata* in response to salinity acclimation. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 163 (1), 38–46. doi:10.1016/j.cbpa.2012.04.026

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57 (1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Campana, S. E. (1999). Chemistry and composition of fish otoliths: Pathways, mechanisms and applications. *Mar. Ecol. Prog. Ser.* 188, 263–297. doi:10.3354/meps188263

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., Lee, J. J., et al. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* 4 (1), 7. doi:10.1186/s13742-015-0047-8

Chen, J., and Zhu, S. (1984). Phylogenetic relationships of the subfamilies in the loach family Cobitidae (Pisces). *Zool. Syst.* 9 (2), 201–208.

Chen, S.-A., Hou, J., Yao, N., Xie, C., and Li, D. (2020). Comparative transcriptome analysis of *Triplophysa yarkandensis* in response to salinity and alkalinity stress. *Comp. Biochem. Physiol. Part D. Genomics Proteomics* 33, 100629. doi:10.1016/j.cbd.2019.100629

Chen, W., Yang, J., Li, Y., and Li, X. (2019). Exploring taxonomic diversity and biogeography of the family Nemacheilinae (Cypriniformes). *Ecol. Evol.* 9 (18), 10343–10353. doi:10.1002/ece3.5553

Chen, Y., Chen, Y., and Liu, H. (1996). Studies of the position of the Qinghai-Xizang Plateau region in zoogeographic divisions and its eastern demarcation line. *Acta Hydrobiol. Sin.* 20 (2), 97–103.

Chen, Y., Yang, J., and Xu, G. (1992). A new blind loach of *Triplophysa* from Yunnan Stone Forest with comments on its phylogenetic relationship. *Zool. Res.* 13 (1), 17–23.

Coop, G., Witonsky, D., Di Rienzo, A., and Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics* 185 (4), 1411–1423. doi:10.1534/genetics.110.114819

Csardi, G., and Nepusz, T. (2006). *The igraph software package for complex network research*. InterJournal, Complex Systems, 1695. Available at: https://igraph.org.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27 (15), 2156–2158. doi:10.1093/bioinformatics/btr330

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *GigaScience* 10 (2), giab008. doi:10.1093/gigascience/giab008

Danecek, P., Schiffels, S., and Durbin, R. (2014). *Multiallelic calling model in bcftools (-m)*. Available at: http://samtools.github.io/bcftools/call-m.pdf.

Dong, Y., Wang, S., Chen, J., Zhang, Q., Liu, Y., You, C., et al. (2016). Hepatocyte nuclear factor 4α (HNF4α) is a transcription factor of vertebrate fatty acyl desaturase gene as identified in marine teleost *Siganus canaliculatus*. *PLoS ONE* 11 (7), e0160361. doi:10.1371/journal.pone.0160361

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., et al. (2006). Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biol. Rev. Camb. Philos. Soc.* 81 (2), 163–182. doi:10.1017/S1464793105006950

Fang, B., Kemppainen, P., Momigliano, P., Feng, X., and Merilä, J. (2020). On the causes of geographically heterogeneous parallel evolution in sticklebacks. *Nat. Ecol. Evol.* 4 (8), 1105–1115. doi:10.1038/s41559-020-1222-6

Feng, C., Zhou, W., Tang, Y., Gao, Y., Chen, J., Tong, C., et al. (2019). Molecular systematics of the *Triplophysa robusta* (cobitoidea) complex: Extensive gene flow in a depauperate lineage. *Mol. Phylogenet. Evol.* 132, 275–283. doi:10.1016/j.ympev.2018.12.009

Fox, E. A., Wright, A. E., Fumagalli, M., and Vieira, F. G. (2019). ngsLD: evaluating linkage disequilibrium using genotype likelihoods. *Bioinformatics* 35 (19), 3855–3856. doi:10.1093/bioinformatics/btz200

Fricke, R., Eschmeyer, W. N., and Van der Laan, R. (2019). *Eschmeyer's catalog of fishes: Genera, species, references, version (12/2019)*. Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33 (1), 1–22. doi:10.18637/jss.v033.i01

Günther, T., and Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics* 195 (1), 205–220. doi:10.1534/genetics.113.152462

Guo, B., DeFaveri, J., Sotelo, G., Nair, A., and Merilä, J. (2015). Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. *BMC Biol.* 13 (1), 19. doi:10.1186/s12915-015-0130-8

Guo, B., Li, Z., and Merilä, J. (2016). Population genomic evidence for adaptive differentiation in the Baltic Sea herring. *Mol. Ecol.* 25 (12), 2833–2852. doi:10.1111/mec.13657

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 (2), 518–522. doi:10.1093/molbev/msx281

Hu, J., Liu, M., and He, D. (2020). Phylogeography of *Triplophysa stenura* (Nemacheilidae): Responded to the mid-pleistocene climate transition in the Qinghai-Tibetan plateau. *Zool. Stud.* 59, e67. doi:10.6620/ZS.2020.59-67

Huo, B., Liu, X., Chen, S. A., Liu, J. Y., Zhou, Q., Shen, J. Z., et al. (2022). Population structure, genetic diversity and differentiation of *Triplophysa tenuis* in Xinjiang Tarim River. *Front. Genet.* 13, 860678. doi:10.3389/fgene.2022.860678

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., et al. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484 (7392), 55–61. doi:10.1038/nature10944

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14 (6), 587–589. doi:10.1038/nmeth.4285

Kemppainen, P., Knight, C. G., Sarma, D. K., Hlaing, T., Prakash, A., Maung Maung, Y. N., et al. (2015). Linkage disequilibrium network analysis (LDna) gives a global view of chromosomal inversions, local adaptation and geographic structure. *Mol. Ecol. Resour.* 15 (5), 1031–1045. doi:10.1111/1755-0998.12369

Kemppainen, P., Li, Z., Rastas, P., Löytynoja, A., Fang, B., Yang, J., et al. (2021). Genetic population structure constrains local adaptation in sticklebacks. *Mol. Ecol.* 30 (9), 1946–1961. doi:10.1111/mec.15808

Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37 (8), 907–915. doi:10.1038/s41587-019-0201-4

Kottelat, M. (2012). Conspectus cobitidum: An inventory of the loaches of the world (teleostei: Cypriniformes: Cobitoidei). *Raffles Bull. Zool.* 1, 1–175.

Kovaka, S., Zimin, A. V., Pertea, G. M., Razaghi, R., Salzberg, S. L., Pertea, M., et al. (2019). Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* 20 (1), 278. doi:10.1186/s13059-019-1910-1

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27 (21), 2987–2993. doi:10.1093/bioinformatics/btr509

Li, H. (2013). *Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM*. arXiv.

Li, H., and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature* 475 (7357), 493–496. doi:10.1038/nature10231

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25 (16), 2078–2079. doi:10.1093/bioinformatics/btp352

Li, Z., Kemppainen, P., Rastas, P., and Merilä, J. (2018). Linkage disequilibrium clustering-based approach for association mapping with tightly linked genomewide data. *Mol. Ecol. Resour.* 18 (4), 809–824. doi:10.1111/1755-0998.12893

Liberles, D. A., Chang, B., Geiler-Samerotte, K., Goldman, A., Hey, J., Kaçar, B., et al. (2020). Emerging frontiers in the study of molecular evolution. *J. Mol. Evol.* 88 (3), 211–226. doi:10.1007/s00239-020-09932-6

Liu, C., Zhu, L., Wang, J., Ju, J., Ma, Q., Qiao, B., et al. (2021). *In-situ* water quality investigation of the lakes on the Tibetan Plateau. *Sci. Bull. (Beijing).* 66 (17), 1727–1730. doi:10.1016/j.scib.2021.04.024

Liu, H., Jiang, T., Chen, X., and Yang, J. (2020). Otolith microchemistry of two *Triplophysa* species in Tongtianhe River. *Southwest China J. Agric. Sci.* 33 (09), 2132–2137. doi:10.16213/j.cnki.scjas.2020.9.037

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (12), 550. doi:10.1186/s13059-014-0550-8

McKinney, G., McPhee, M. V., Pascal, C., Seeb, J. E., and Seeb, L. W. (2020). Network analysis of linkage disequilibrium reveals genome architecture in chum salmon. *G3* 10 (5), 1553–1561. doi:10.1534/g3.119.400972

Mederos, y. S. M., Rinné, S., Skrobek, L., Renigunta, V., Schlichthörl, G., Derst, C., et al. (2009). Mutation of histidine 105 in the T1 domain of the potassium channel Kv2.1 disrupts heteromerization with Kv6.3 and Kv6.4. *J. Biol. Chem.* 284 (7), 4695–4704. doi:10.1074/jbc.M808786200

Meinshausen, N., and Bühlmann, P. (2006). High-dimensional graphs and variable selection with the Lasso. *Ann. Stat.* 34 (3), 1436–1462. doi:10.1214/009053606000000281

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37 (5), 1530–1534. doi:10.1093/molbev/msaa015

Mueckler, M. (1994). Facilitative glucose transporters. *Eur. J. Biochem.* 219 (3), 713–725. doi:10.1111/j.1432-1033.1994.tb18550.x

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2014). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32 (1), 268–274. doi:10.1093/molbev/msu300

Nguyen, T. V., Jung, H., Nguyen, T. M., Hurwood, D., and Mather, P. (2016). Evaluation of potential candidate genes involved in salinity tolerance in striped catfish (*Pangasianodon hypophthalmus*) using an RNA-Seq approach. *Mar. Genomics* 25, 75–88. doi:10.1016/j.margen.2015.11.010

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11 (9), 1650–1667. doi:10.1038/nprot.2016.095

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., Salzberg, S. L., et al. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33 (3), 290–295. doi:10.1038/nbt.3122

Poff, N. L., Olden, J. D., and Strayer, D. L. (2012). "Climate change and freshwater fauna extinction risk," in *Saving a million species: Extinction risk from climate change*. Editor L. Hannah (Washington, DC: Island Press/Center for Resource Economics), 309–336.

Pons, P., and Latapy, M. (2006). Computing communities in large networks using random walks. *J. Graph Algorithms Appl.* 10 (2), 191–218. doi:10.7155/jgaa.00124

Pontigo, J. P., Aguero, M. J., Sanchez, P., Oyarzun, R., Vargas-Lagos, C., Mancilla, J., et al. (2016). Identification and expressional analysis of NLRC5 inflammasome gene in smolting Atlantic salmon (*Salmo salar*). *Fish. Shellfish Immunol.* 58, 259–265. doi:10.1016/j.fsi.2016.09.031

Prokofiev, A. M. (2010). Morphological classification of loaches (Nemacheilinae). *J. Ichthyol.* 50 (10), 827–913. doi:10.1134/S0032945210100012

Purcell, S., and Chang, C. (2020). *PLINK [v1.90b6.18 64-bit]*. Available at: https://www.cog-genomics.org/plink/1.9/.

Qiao, B., Zhu, L., and Yang, R. (2019). Temporal-spatial differences in lake water storage changes and their links to climate change throughout the Tibetan Plateau. *Remote Sens. Environ.* 222, 232–243. doi:10.1016/j.rse.2018.12.037

Sala, O. E., Stuart Chapin, F., Iii, Armesto, J. J., Berlow, E., Bloomfield, J., DiRzo, R., et al. (2000). Global biodiversity scenarios for the year 2100. *Science* 287 (5459), 1770–1774. doi:10.1126/science.287.5459.1770

Schluter, D. (2000). *The ecology of adaptive radiation*. London: Oxford University Press.

Schmitz, M., Ziv, T., Admon, A., Baekelandt, S., Mandiki, S. N. M., L'Hoir, M., et al. (2017). Salinity stress, enhancing basal and induced immune responses in striped catfish *Pangasianodon hypophthalmus* (Sauvage). *J. Proteomics* 167, 12–24. doi:10.1016/j.jprot.2017.08.005

Secor, D. H., and Rooker, J. R. (2000). Is otolith strontium a useful scalar of life cycles in estuarine fishes? *Fish. Res.* 46 (1), 359–371. doi:10.1016/S0165-7836(00)00159-4

Seikel, E., and Trimmer, J. S. (2009). Convergent modulation of Kv4.2 channel α subunits by structurally distinct DPPX and KChIP auxiliary subunits. *Biochemistry* 48 (24), 5721–5730. doi:10.1021/bi802316m

Shen, S. S. P., Yao, R., Ngo, J., Basist, A. M., Thomas, N., Yao, T., et al. (2015). Characteristics of the Tibetan Plateau snow cover variations based on daily data during 1997–2011. *Theor. Appl. Climatol.* 120 (3), 445–453. doi:10.1007/s00704-014-1185-0

Soh, H., and Goldstein, S. A. N. (2008). ISA channel complexes include four subunits each of DPP6 and Kv4.2. *J. Biol. Chem.* 283 (22), 15072–15077. doi:10.1074/jbc.M706964200

Stern, D. L. (2013). The genetic causes of convergent evolution. *Nat. Rev. Genet.* 14 (11), 751–764. doi:10.1038/nrg3483

Terekhanova, N. V., Logacheva, M. D., Penin, A. A., Neretina, T. V., Barmintseva, A. E., Bazykin, G. A., et al. (2014). Fast evolution from precast bricks: Genomics of young freshwater populations of threespine stickleback *Gasterosteus aculeatus*. *PLoS Genet.* 10 (10), e1004696. doi:10.1371/journal.pgen.1004696

Tickner, D., Opperman, J. J., Abell, R., Acreman, M., Arthington, A. H., Bunn, S. E., et al. (2020). Bending the curve of global freshwater biodiversity loss: An emergency recovery plan. *Bioscience* 70 (4), 330–342. doi:10.1093/biosci/biaa002

Wang, J., Tang, Q., Wang, Z., Zhang, Y., Wu, Q., Peng, Z., et al. (2012). The complete mitogenome sequence of a cave loach *Triplophysa rosa* (Teleostei, Balitoridae, Nemacheilinae). *Mitochondrial DNA* 23 (5), 366–368. doi:10.3109/19401736.2012.696628

Wang, S., Chen, J., Jiang, D., Zhang, Q., You, C., Tocher, D. R., et al. (2018). Hnf4α is involved in the regulation of vertebrate LC-PUFA biosynthesis: Insights into the regulatory role of Hnf4α on expression of liver fatty acyl desaturases in the marine teleost *Siganus canaliculatus*. *Fish. Physiol. Biochem.* 44 (3), 805–815. doi:10.1007/s10695-018-0470-8

Wang, T., Zhang, Y.-p., Yang, Z.-y., Liu, Z., and Du, Y.-y. (2020a). DNA barcoding reveals cryptic diversity in the underestimated genus *Triplophysa* (Cypriniformes: Cobitidae, Nemacheilinae) from the northeastern Qinghai-Tibet Plateau. *BMC Evol. Biol.* 20 (1), 151. doi:10.1186/s12862-020-01718-0

Wang, Y., and Guo, B. (2019). Adaption to extreme environments: A perspective from fish genomics. *Rev. Fish. Biol. Fish.* 29, 735–747. doi:10.1007/s11160-019-09577-9

Wang, Y., Shen, Y., Feng, C., Zhao, K., Song, Z., Zhang, Y., et al. (2016). Mitogenomic perspectives on the origin of Tibetan loaches and their adaptation to high altitude. *Sci. Rep.* 6 (1), 29690. doi:10.1038/srep29690

Wang, Y., Yang, L., Wu, B., Song, Z., and He, S. (2015a). Transcriptome analysis of the plateau fish (*Triplophysa dalaica*): Implications for adaptation to hypoxia in fishes. *Gene* 565 (2), 211–220. doi:10.1016/j.gene.2015.04.023

Wang, Y., Yang, L., Zhou, K., Zhang, Y., Song, Z., He, S., et al. (2015b). Evidence for adaptation to the Tibetan Plateau inferred from Tibetan loach transcriptomes. *Genome Biol. Evol.* 7 (11), 2970–2982. doi:10.1093/gbe/evv192

Wang, Y., Zhao, Y., Wang, Y., Li, Z., Guo, B., Merilä, J., et al. (2020b). Population transcriptomics reveals weak parallel genetic basis in repeated marine and freshwater divergence in nine-spined sticklebacks. *Mol. Ecol.* 29 (9), 1642–1656. doi:10.1111/mec.15435

Whitehead, A., Roach, J. L., Zhang, S. J., and Galvez, F. (2012). Salinity- and population-dependent genome regulatory response during osmotic acclimation in the killifish (*Fundulus heteroclitus*) gill. *J. Exp. Biol.* 215 (8), 1293–1305. doi:10.1242/jeb.062075

Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. New York: Springer-Verlag.

Woodward, G., Perkins, D. M., and Brown, L. E. (2010). Climate change and freshwater ecosystems: Impacts across multiple levels of organization. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365 (1549), 2093–2106. doi:10.1098/rstb.2010.0055

Wu, P., Geng, L., Jiang, H., Tong, G., Li, C., Xu, W., et al. (2017). Tolerance of three Cobitidae fish species to high salinity and alkalinity. *J. Fish. Sci. China* 24 (2), 248. doi:10.3724/sp.j.1118.2017.16151

Wu, Y., and Wu, C. (1992). *The fishes of the qinghai-xizang plateau*. Chengdu, China: Sichuan Publishing House of Science and Technology.

Yang, H., and Wang, K. (2015). Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat. Protoc.* 10 (10), 1556–1566. doi:10.1038/nprot.2015.105

Yang, J., Arai, T., Liu, H., Miyazaki, N., and Tsukamoto, K. (2006). Reconstructing habitat use of *Coilia mystus* and *Coilia ectenes* of the Yangtze River estuary, and of *Coilia ectenes* of Taihu Lake, based on otolith strontium and calcium. *J. Fish. Biol.* 69 (4), 1120–1135. doi:10.1111/j.1095-8649.2006.01186.x

Yang, L., Wang, Y., Wang, T., Duan, S., Dong, Y., Zhang, Y., et al. (2019a). A chromosome-scale reference assembly of a Tibetan loach, *Triplophysa siluroides*. *Front. Genet.* 10, 991. doi:10.3389/fgene.2019.00991

Yang, R., Zhu, L., Wang, J., Ju, J., Ma, Q., Turner, F., et al. (2017). Spatiotemporal variations in volume of closed lakes on the Tibetan Plateau and their climatic responses from 1976 to 2013. *Clim. Change* 140 (3), 621–633. doi:10.1007/s10584-016-1877-9

Yang, X. (2018). *Database of lake diatom – salinity of lake water conversion relationship on the Tibetan Plateau (1723–2001)*. National Tibetan Plateau Data Center.

Yang, X., Liu, H., Ma, Z., Zou, Y., Zou, M., Mao, Y., et al. (2019b). Chromosome-level genome assembly of *Triplophysa tibetana*, a fish adapted to the harsh high-altitude environment of the Tibetan Plateau. *Mol. Ecol. Resour.* 19 (4), 1027–1036. doi:10.1111/1755-0998.13021

Yao, N., Song, Y., Wang, S., Chen, S. A., and Liu, J. Y., (2018). Toxicity of salinity and alkalinity to *Triplophysa* (*hedinichthys) yarkandensis* (day) of tarim river. *Southwest China J. Agric. Sci.* 31 (2), 423–428.

Yao, T., Thompson, L., Yang, W., Yu, W., Gao, Y., Guo, X., et al. (2012). Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. *Nat. Clim. Chang.* 2 (9), 663–667. doi:10.1038/nclimate1580

Yao, T., Xue, Y., Chen, D., Chen, F., Thompson, L., Cui, P., et al. (2019). Recent third pole's rapid warming accompanies cryospheric melt and water cycle intensification and interactions between monsoon and environment: Multidisciplinary approach with observations, modeling, and analysis. *Bull. Am. Meteorol. Soc.* 100 (3), 423–444. doi:10.1175/bams-d-17-0057.1

Yu, J., Carroll, T. J., Rajagopal, J., Kobayashi, A., Ren, Q., McMahon, A. P., et al. (2009). A Wnt7b-dependent pathway regulates the orientation of epithelial cell division and establishes the cortico-medullary axis of the mammalian kidney. *Development* 136 (1), 161–171. doi:10.1242/dev.022087

Yuan, D., Chen, X., Gu, H., Zou, M., Zou, Y., Fang, J., et al. (2020). Chromosomal genome of *Triplophysa bleekeri* provides insights into its evolution and environmental adaptation. *GigaScience* 9 (11), giaa132. doi:10.1093/gigascience/giaa132

Yuan, M., and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. B* 68, 49–67. doi:10.1111/j.1467-9868.2005.00532.x

Zhang, F. (2017). *Hydrochemical characteristics of Lhasa River Basin and source apportionment of heavy metals in water*. Tianjin: Tianjin University. Master thesis.

Zhang, G., Yao, T., Piao, S., Bolch, T., Xie, H., Chen, D., et al. (2017). Extensive and drastically different alpine lake changes on Asia's high plateaus during the past four decades. *Geophys. Res. Lett.* 44 (1), 252–260. doi:10.1002/2016GL072033

Zhang, G., Yao, T., Xie, H., Yang, K., Zhu, L., Shum, C., et al. (2020). Response of Tibetan plateau lakes to climate change: Trends, patterns, and mechanisms. *Earth. Sci. Rev.* 208, 103269. doi:10.1016/j.earscirev.2020.103269

Zhao, Q., Zhang, R., Xiao, Y., Niu, Y., Shao, F., Li, Y., et al. (2020). Comparative transcriptome profiling of the loaches *Triplophysa bleekeri* and *Triplophysa rosa* reveals potential mechanisms of eye degeneration. *Front. Genet.* 10, 1334. doi:10.3389/fgene.2019.01334

Zhao, X., Wu, Z., Zhu, D., and Hu, D. (2002a). Quaternary glaciations in the west nyaiqentanglha Mountains. *Quat. Sci.* 22 (05), 424–433.

Zhao, X., Zhu, D., Wu, Z., and Ma, Z. (2002b). The development of nam Co lake in Tibet since late pleistocene. *Acta Geosci. Sin.* 23 (04), 329–334.

Zhou, C., Hu, B., Tang, Y., Yang, C., Ma, W., Wang, X., et al. (2021). The chromosome-level genome of *Triplophysa dalaica* (Cypriniformes: Cobitidae) provides insights into its survival in extremely alkaline environment. *Genome Biol. Evol.* 13 (8), evab153. doi:10.1093/gbe/evab153

Zhu, L., Wang, J., Ju, J., Ma, N., Zhang, Y., Liu, C., et al. (2019). Climatic and lake environmental changes in the Serling Co region of Tibet over a variety of timescales. *Sci. Bull. (Beijing).* 64 (7), 422–424. doi:10.1016/j.scib.2019.02.016

Zhu, S. (1989). *The loaches of the subfamily nemacheilinae in China (cypriniformes: Cobitidae)*. Nanjing, China: Jiangsu Science and Technology Publishing House.