Comparative transcriptomics reveal that adaptive evolution in immune genes drives the local adaptation and speciation of schizothoracine fish

Delin Qi[a]*, Rongrong Wu[a, b], Yan Chao[b], Mingzhe Xia[a], Qichang Chen [a], Zhiqin Zheng [a]

[a] State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University, Xining 810016, China

[b] Animal Science Department of Agriculture and Animal Husbandry College, Qinghai University, Xining 810016, China

* Corresponding Author: State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University, Xining 810016, China.

E-mail address: delinqi@126.com (D. Qi).

Supplemental material available at FigShare: https://doi.org/10.25387/.

ABSTRACT Transcriptomic information can increase our understanding of the molecular processes underlying speciation. The schizothoracine fish, the largest and most diverse taxon within the Qinghai-Tibetan Plateau (QTP) ichthyofauna, are widespread in drainages throughout the QTP. These fish thus serve as an ideal model group with which to investigate how molecular evolution drives local adaptation during speciation. Here, we performed an interspecific comparative analysis of the transcriptomes of 13 schizothoracine fish species, and identified the key positively selected genes (PSGs) associated with significantly enriched functions and metabolite pathway acting on the specific lineages (or species) in the schizothoracine fish. We generated 64,637,602–83,968,472 sequence reads per schizothoracine fish species using Illumina sequencing, yielding 95,251–145,805 unigenes per species. We identified 52 out of 2,064 orthologous genes as candidate genes, which have probably been subject to positive selection along the whole schizothoracine fish lineage. Nine of these candidate genes were significantly enriched in key GO functions and metabolite pathways, all of which were associated with the immune system. The lineage-specific evolution test showed species-specific differences among the nine candidate PSGs, probably due to ecological differences among drainages, as well as among micro-habitats in the same drainage (e.g., benthic and pelagic). Here, we provide evidence that the adaptive evolution of immune genes, along with the uplift of the QTP, allowed new schizothoracine species to colonize ecologically novel environments or to exploit vacant ecological niches during speciation.

KEYWORDS: Schizothoracine fish, the Qinghai-Tibetan Plateau, Transcriptome, Positive selection, Adaptation

The processes and mechanisms of speciation have long been a major question in evolutionary biology (Funk et al. 2002; Rocha et al. 2005; Baack and Stanton 2005; Tonnis et al. 2005; Nosil et al. 2009; Malinsky et al. 2015). Local adaptations, driven by small-scale ecological niches, often result in adaptive phenotypes and the genetic divergence of geographically isolated populations (Baack and Stanton 2005; Nosil et al. 2009; Collin and Fumagalli 2011). The accumulation of genetic differences may lead to the formation of new taxon (Zhang et al. 2013; Collin and Fumagalli 2011; Nosil et al. 2009; Kirkpatrick and Barton 2006). Genetic divergences associated with local adaptations allow species to colonize ecologically novel environments or to exploit vacant ecological niches, and thus play an important role in speciation (Zhou et al. 2012; Yang et al. 2014; Kang et al. 2017; Tong et al. 2017; Cai et al. 2013; Qu et al. 2013). Freshwater fish, especially fish lineages distributed in different ecological drainage systems within a given region, have provided key insights into the mechanisms whereby molecular evolution drives local adaptation during speciation (Yang et al. 2016; Kang et al. 2017; Tong et al. 2017; Tong et al. 2015; Ma et al. 2015; Guan et al. 2014; Xu et al. 2017).

Schizothoracine fish (Teleostei: Cyprinidae) are the largest and most diverse taxon within the Qinghai-Tibetan Plateau (QTP) ichthyofauna, including more than 70 recognized species (Wu and Wu 1992; Chen and Cao 2000). Schizothoracine fish are distributed throughout the majority of the drainage basins in the QTP, 1500–5500 m above sea level (Qi et al. 2006; Wu and Wu 1992; Chen and Cao 2000). The distribution range of these fishes include the Yellow River, Yangtze River, Indus River, Mekong River, Tsangpo River, Salween River, Brahmaputra River, Qiadam Basin, and isolated lakes (Wu and Wu 1992; Chen and Cao 2000). These water bodies represent rich ecological diversity and complexity (Wu and Wu 1992; Qi et al. 2012). From an evolutionary perspective, the schizothoracine fish can be divided into three groups: primitive, specialized, and highly specialized, corresponding with three particular stages of QTP geological evolution (Cao et al. 1981). It has been demonstrated that the uplifting of the QTP from 50 MYA had a profound effect on paleo-drainage and stimulated many paleo-environmental changes in the plateau, thus promoting

the speciation of the schizothoracine fish endemic to this region and shaping current species distributions (Qi et al. 2012; Li et al. 2013). Due to the extensive ecological diversity of the distribution drainages, and because speciation events occurred during the short period of QTP uplift, the schizothoracine fish is an ideal model group with which to investigate how molecular evolution drives local adaptation during speciation. Previous studies, using transcriptomic comparisons between fish species endemic to the QTP and other fish species endemic to lowland areas, have provided evidence of the genetic adaptation of schizothoracine fish to high-altitude environments (Yang et al. 2014; Tong et al. 2015; Tong et al. 2017). However, previous studies have only focused on the adaptive genetic differences between QTP schizothoracine fish and low-altitude fish species; the potentially adaptive genetic differences among the schizothoracine fish remain unclear. In addition, previous studies have only included a single schizothoracine species, which may have been insufficient to fully characterize the genetic signals of local adaptation during speciation in this group.

In this study, we sequenced, assembled, and merged the transcriptomes of four tissues (head kidney, muscle, brain, and liver) from 13 schizothoracine species. We then conducted a comparative transcriptomic analysis to identify key positively selected genes (PSGs) associated with significantly enriched functions and metabolite pathways acting on specific lineages (or species) of schizothoracine fish. We aimed to provide evidence of the genetic adaptation of schizothoracine fish to their local aquatic environment, leading to speciation.

MATERIALS AND METHODS

Experimental Animals

Fish were collected from across their ranges using gill nets or cast nets (Figure 1). During field sampling, a blow to the head was used to stun fish. The liver, muscle, brain, and head kidney were then removed and frozen in liquid nitrogen until use. The specimens used in this study included 13 species, representing the primitive, specialized, and highly specialized groups of schizothoracine fishes. All of the procedures involving animals followed the guidelines of, and

were conducted under the approval of, the Animal Care and Use Committee, Qinghai University, China.

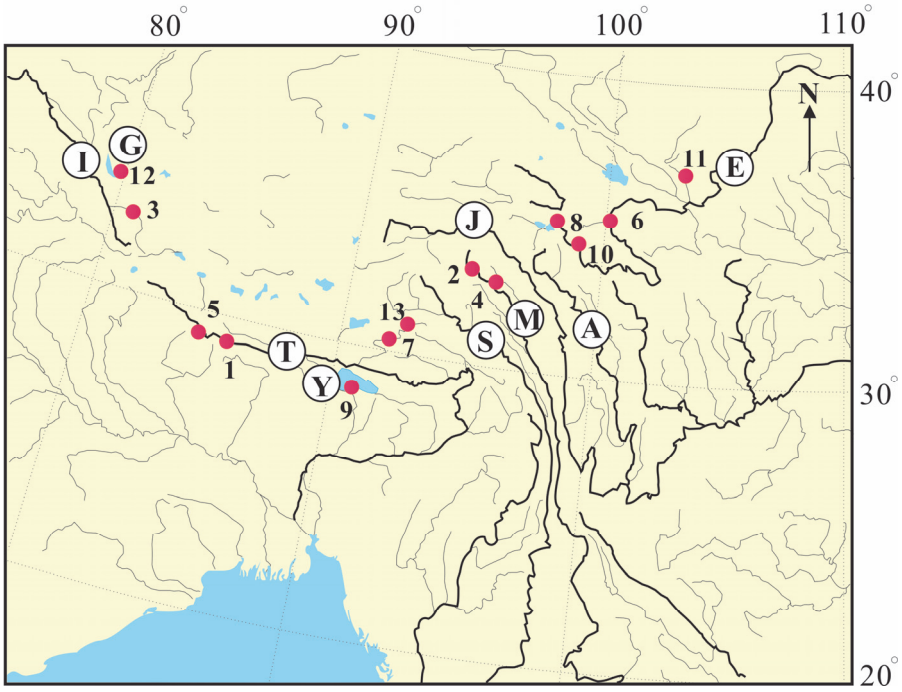

Figure 1 Location of fish sampling sites. Sites: A, Yalung River; E, Yellow River; G, Bangongco Lake; I, Indus River; J, Jinsha River; M, Mekong River; S, Salween River; T, Tsangpo River; Y, Yomzho Lake. Species sampled: 1, *Schizothorax macropogon*; 2, *Schizothorax lissolabiatus*; 3, *Schizothorax labiatus*; 4, *Schizothorax lantsangensis*; 5, *Ptychobarbus dipogon*; 6, *Gymnodiptychus pachycheilus*; 7, *Oxygymnocypris stewartii*; 8, *Gymnocypris eckloni*; 9, *Gymnocypris waddelli*; 10, *Platypharodon extremus*; 11, *Schizopygopsis pylzovi*; 12, *Schizopygopsis stoliczkai*; 13, *Schizopygopsis younghusbandi*.

RNA extraction, Illumina library preparation, and sequencing

Total RNA was isolated from each sample using the Ambion Magmax-96 total RNA isolation kit (Life Sciences, USA), following the manufacturer's instructions. RNA degradation and contamination was monitored using 1% agarose gels. RNA purity was checked using a Nano Photometer spectrophotometer (IMPLEN, CA, USA). RNA concentration was measured using Qubit RNA Assay Kits in a Qubit 2.0 Flurometer (Life Technologies, CA, USA). A single pooled RNA sample, containing equal volumes of RNA from each organ (liver, muscle, brain and head kidney), was prepared for each species. We used 3.0 μg of each pooled RNA sample used to prepare one Illumina sequencing library per species.

Sequencing libraries were generated using the NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, USA), following manufacturer's instructions. Index codes were added to attribute sequences to each sample. PCR products were purified using an AMPure XP system (Beckman Coulter, USA) and library quality was assessed with a Bioanalyzer 2100 (Agilent). The index-coded samples were clustered with a cBot Cluster Generation System using a TruSeq PE Cluster Kit v3-cBot-HS (Illumina), following the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina Hiseq 2500 platform, generating 125 bp paired-end reads. All sequence reads were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive database under bioproject number PRJNA494936.

Quality control and de novo assembly

Raw data (raw reads) were sorted by individual species and processed using self-written Perl scripts. In this step, data (reads) were cleaned by removing reads containing adapter sequences, reads containing ploy-N sequences, and low-quality reads. All of the downstream analyses were based on high-quality clean data. De novo transcriptome assembly was performed with the short read assembly program Trinity (Grabherr et al. 2011), with min_kmer_cov set to 2 by default and all of the other parameters set to default (Grabherr et al. 2011). In brief, reads of a certain length with overlapping areas were first joined to form longer fragments (contigs) without gaps. Then, paired-end reads were mapped back to contigs. Finally, the contigs were connected until sequences could no longer be extended. These sequences were termed unigenes.

Transcriptome annotation

All of the unigenes from the 13 schizothoracine fish were annotated against the NCBI non-redundant (Nr) protein database (ftp://ftp.ncbi.nih.gov/blast/db/) with BLAST (Altschul et al. 1997), setting an E-value cut-off of 1E-5. We also used BLAST to align and annotate unigene sequences against the nucleotide (Nt) database (ftp://ftp.ncbi.nih.gov/blast/db/), the Clusters of Orthologous Groups of proteins (KOG) database (ftp://ftp.ncbi.nih.gov/pub/COG/ KOG/), Swiss-Prot (http://www.uniprot.org/), the Kyoto Encyclopedia of Genes and Genomes (KEGG)

database (http://www.genome.jp/kegg/), and the Gene Ontology (GO) database (http://www.geneontology.org/). The results of the BLAST search against the GO database were imported into Blast2GO 3.2 (Gotz et al. 2008) for GO term mapping. The clusters of orthologous groups (COG) database were then used to identify putative functions for the unigenes, based on known orthologous gene products (Tatusov et al. 2003). The KEGG pathways were analyzed using the online KEGG Automatic Annotation Sever (http://www.genome.jp/kegg/kaas/), with the bi-directional best-hit method and an E-value cut-off of 1E-10.

Ortholog identification and sequence alignment

Translated amino acid sequences from the 13 schizothoracine fish were used to construct a database, along with sequences from another two fish species: zebrafish (*Danio rerio*) and medaka (*Oryzias latipes*), obtained from the Ensembl database (release 89). Next, we used performed a self-to-self BLASTP against all of the amino acid sequences, with a E-value cutoff of $1e^{-5}$; hits with identity < 30% and coverage < 30% were removed. One-to-one orthologs between 13 schizothoracine fish were determined using OrthoMCL v2.0.9 software (Li et al. 2003) with default settings. Finally, putative single copy orthologs across the 13 schizothoracine fish were obtained. For genes with multiple transcripts, the longest transcript was chosen. Each orthologous gene set was aligned using MUSCLE v. 3.8.31 (Edgar 2004), and trimmed using Gblocks (Castresana 2000) with the parameter "–t = c". We deleted all of the gaps and ambiguous bases ("N") from the alignments to reduce the effects of these elements on positive selection inference. After this deletion process, trimmed alignments shorter than 150 bp (50 codons) were discarded.

Phylogeny construction and PSGs identification across the schizothoracine fish lineage

We constructed a phylogenetic tree of the 13 schizothoracine fish using the concatenated orthologous sequences using a maximum-likelihood (ML) analysis in MEGA 6.0 (Tamura et al. 2013). The ML tree was then used in a PAML analysis (Yang 2007). We examined the schizothoracine fish phylogeny for evidence of natural selection by comparing non-synonymous/synonymous substitution ratios (ω = Ka/Ks), where ω = 1, <1, and>1 indicated

neutral evolution, purifying selection, and positive selection, respectively (Yang and Nielsen 2002; Yang et al. 2005). To determine how many orthologous genes had undergone positive selection across the whole lineage of schizothoracine fishes, we applied site-specific ML models using the codeml in PAML (version 4.7) (Yang 2007). Site-specific models that allowed ω to vary among sites were used to detect site-dependent evolution across the schizothoracine fish lineage. Thus, all the putative single-copy genes identified in the 13 schizothoracine fishes were separately tested for positive selection using the neutral model (M7) and the selection model (M8).

PSG GO enrichment and KEGG pathways analysis

GO functional enrichment and KEGG pathway analyses were performed for genes shown to be under positive selection across the lineage of the schizothoracine fishes. GO enrichment analysis was performed using GOseq (Young et al. 2010), which is based on a Wallenius non-central hyper-geometric distribution. KEGG pathways were analyzed using KOBAS v2.0.12 (Xie et al. 2011), with FDR set to BH.

Lineage-specific evolution of the key PSGs associated significantly enriched functions and metabolite pathways

Positive selection pressure on genes is one of the fundamental processes underlying adaptive changes in genes and genomes, resulting in evolutionary innovations and species differences (Marra et al. 2017). During speciation and dispersal, the different lineages (or species) of the schizothoracine fish have been subject to a variety of different selection pressures, including low temperatures, hypoxic conditions, intense ultraviolet radiation, and unique pathogens (Wu and Wu 1992; Qi et al. 2012). To investigate which of the key PSGs were associated with significantly enriched functions and metabolite pathways acting on specific lineages (or species), which may have driven the adaption of specific taxa to local environmental conditions, we used branch-specific ML models (Zhang 2005). Branch-specific models allow the ω ratio to vary among branches in a phylogeny, and are useful for detecting positive selection pressures acting on particular lineages. Thus, for the positively selected genes significantly enriched in GO terms and KEGG pathways, we evaluated the ω ratio for every lineage of the

schizothoracine fish, using branch-specific ML models with branch labels.

RESULTS

Illumina sequencing and de novo transcriptome assembly

We generated 64,637,602–83,968,472 sequence reads for each of the 13 schizothoracine fish with Illumina sequencing. After trimming and filtering low-quality sequences, we obtained 62,487,860–81,238,722 high-quality clean sequence reads for each of the 13 species. Due to the absence of genome information for the schizothoracine fishes, we de novo assembled the sequence reads using Trinity, obtaining 129,982–207,887 for each of the 13 species. Each of these transcripts included 95,251–145,805 unigenes (Table 1). Detailed unigene information is summarized in Supplementary Material 1.

Table 1 Summary of transcriptome data for 13 schizothoracine fish

Species	Raw reads	Clean reads	Clean bases (Gb)	Transcripts	Unigenes
S. lantsangensis	83968472	81238722	12.19	149827	107266
S. macropogon	66249882	62964616	9.44	150512	109939
S. lissolabiatus	64637602	62875610	9.43	129982	95251
S. labiatus	74677296	70975858	10.65	149487	111736
G. pachycheilus	71034408	68403344	10.26	171564	125690
P. dipogon	81626182	78155270	11.72	190151	125289
O. stewartii	68519742	65848700	9.88	131558	99448
P. extremus	80057188	75758954	11.36	166210	117964
S. pylzovi	66998630	62487860	9.37	134550	96467
S. stoliczkai	71929644	69194100	10.38	169946	121641
S. younghusbandi	76338196	73209828	10.98	164357	116520
G. waddelli	80912802	78367978	11.76	207887	145805
G. eckloni	72359070	70142108	10.52	162235	110231

Unigene annotation

All of the unigene sets across the 13 schizothoracine fish were annotated based on similarity to sequences in six public databases. We found the largest number of unigene matches in the Nt

database, ranging from 115,232 (79.03%: *Gymnocypris waddelli*) to 85,026 (89.26%: *Schizothorax lissolabiatus*). We found that 81.43%–90.94% of the unigenes from each species had a match in at least one database. We used the unigenes BLAST hits from the Nr database for subsequent analyses, as the Nr database had the maximum number of protein reference sequences per species (Table 2).

Table 2 Functional annotation of unigenes for 13 schizothoracine fish.

Species	Nr	Nt	Swiss-Prot	KEGG	GO	COG	Total
S. lantsangensis	36506 (30.04%)	95438 (88.97%)	29097 (27.12%)	18209 (16.97%)	28577 (26.64%)	17084 (15.92%)	96925 (90.35%)
S. macropogon	38093 (34.64%)	97898 (89.04%)	30783 (28.0%)	19207 (17.47%)	29542 (26.87%)	17799 (16.18%)	99454 (90.46%)
S. lissolabiatus	36571 (38.39%)	85026 (89.26%)	30006 (31.5%)	19035 (19.98%)	28284 (29.69%)	17193 (18.05%)	86623 (90.94%)
S. labiatus	43895 (39.28%)	99055 (88.65%)	35911 (32.13%)	22916 (20.5%)	32614 (29.18%)	19639 (17.57%)	101007 (90.39%)
G. pachycheilus	36678 (29.18%)	102017 (81.16%)	30350 (24.14%)	19180 (15.52%)	30542 (24.29%)	17831 (14.18%)	104791 (83.37%)
P. dipogon	43240 (34.51%)	101998 (81.41%)	33871 (27.03%)	20852 (16.64%)	32562 (25.98%)	19336 (15.43%)	104812 (83.65%)
O. stewartii	37748 (37.95%)	82070 (82.52%)	30815 (31.32%)	19800 (19.9%)	28338 (28.49%)	17364 (17.46%)	84213 (84.68%)
P. extremus	42824 (36.3%)	94790 (80.35%)	34494 (29.24%)	22361 (18.95%)	31591 (26.78%)	18632 (15.79%)	97520 (82.66%)
S. pylzovi	42038 (43.57%)	81162 (84.13%)	34629 (35.89%)	22262 (23.07%)	30444 (31.55%)	19076 (19.77%)	83356 (86.4%)
S. stoliczkai	37381 (30.73%)	98143 (80.68%)	30118 (24.75%)	18849 (15.49%)	30213 (24.83%)	17797 (14.63%)	100900 (82.94%)
S. younghusbandi	43254 (37.12%)	94808 (81.36%)	35206 (30.21%)	22452 (19.26%)	33117 (28.42%)	19961 (17.13%)	97992 (84.09%)
G. waddelli	41076 (28.17%)	115232 (79.03%)	32775 (22.47%)	20568 (14.1%)	33045 (22.66%)	18884 (16.95%)	118732 (81.43%)
G. eckloni	34532 (31.32%)	89171 (80.89%)	27505 (24.95%)	16578 (15.0%)	28141 (25.52%)	16869 (15.3%)	91554 (83.05)

Functional annotation and classification

Based on sequence homology against the Nr database, the 28,141 (*Gymnocypris eckloni*) to 33,117 (*Schizopygopsis younghusbandi*) unigenes from the 13 schizothoracine fish were annotated to 56 GO categories (Table S2). Of these, there were 24 biological process (BP) sub-categories, 18 cellular component (CC) sub-categories, and 14 molecular function (MF) sub-categories. Binding (GO:0005488) was the most represented MF category; cellular process (GO:0009987) was the most represented BP category, and cell (GO:0005623) and cell part (GO:0044464) were the most represented CC categories (Supplementary Material 2).

We functionally classified 16,578 (*G. eckloni*) to 22,916 (*Schizothorax labiatus*) unigenes from the 13 schizothoracine fish into five KEGG categories and 32 functional sub-categories. Among the 32 subcategories, signal transduction were the most highly represented, followed by endocrine system, cell communication, nervous system, and immune system (Supplementary Material 3).

Putative orthologs, phylogeny, and positively selected genes

We divided 477,274 proteins from the 13 schizothoracine fish and two other fish species (*D. rerio* and *O. latipes*) into 52,508 orthologous groups (gene families) using OrthoMCL (Li et al. 2003), following the self-self-comparison with BLASTP. After alignment and trimming for quality control, we compared the orthologous groups among the 13 fish species, and identified 2,064 putative single-copy genes in each fish species. These single-copy genes were used in the subsequent phylogenetic and evolutionary analyses. The ML phylogenetic tree, based on 2,064 orthologous genes, was well supported with 100% bootstrap values at all of the nodes (Figure 2). This ML tree was congruent with previous trees based on morphological characters (Wu and Wu 1992) and on mitochondrial genes (Saitoh et al. 2006; Qi et al. 2012).

Positive selection analysis pinpointed the genes that were associated with a functional or environmental shift. For the 2,064 orthologous genes that harbored both synonymous and non-synonymous substitutions, we identified 52 genes with Ka/Ks ratios >1, and 187 genes with Ka/Ks ratios between 0.5 and 1 (Table 3 and Figure 3). The 52 genes (with Ka/Ks ratios > 1) were considered candidate genes, that probably underwent positive selection along the

schizothoracine fish lineage, and which may have influenced the adaptation of these fish to the aquatic environment of the QTP (Table 3 and Figure 3).

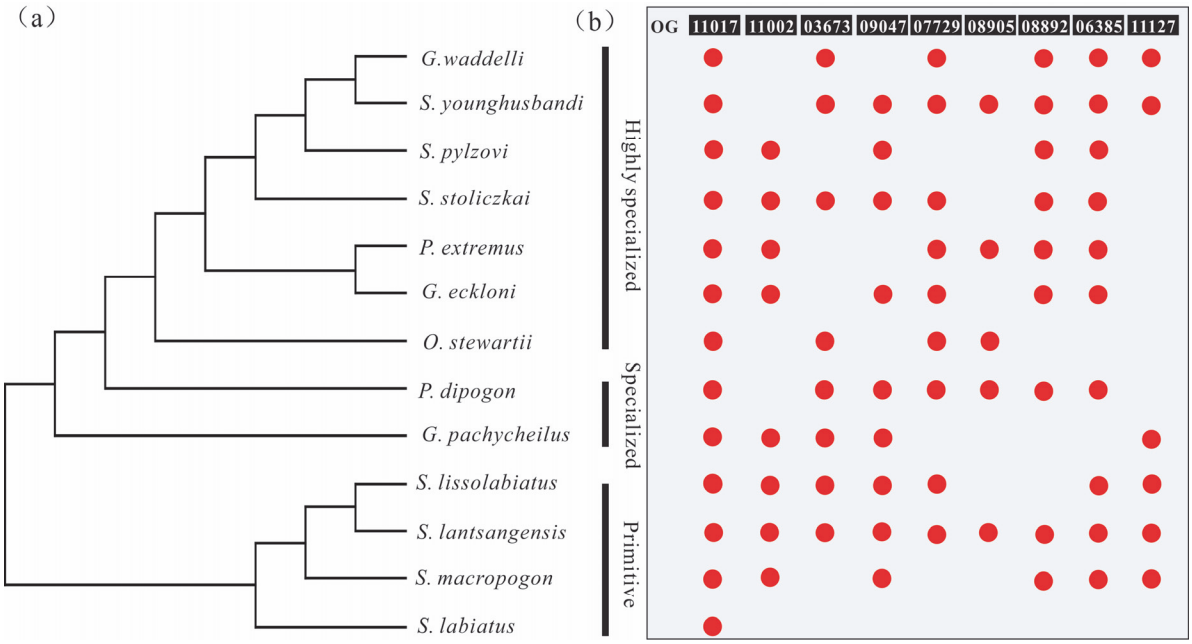

Figure 2 (a) Maximum likelihood (ML) phylogenetic tree, inferred based on 2064 orthologous genes. (b) Lineage-specific evolution test of the nine candidate positively selected genes (PSGs). Red dots represent genes with a Ka/Ks ratio >1 in a specific lineage.

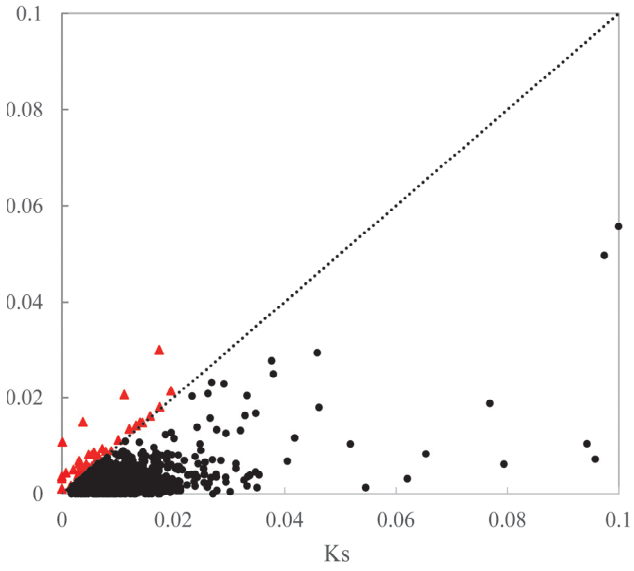

Figure 3 Plot of nonsynonymous (Ka) vs. synonymous (Ks) substitutions. The dotted line suggests neutrality, values above the line (red triangles) are subject to positive selection, and values below the line (black dots) are subject to purifying selection.

Table 3 Identification of candidate genes under positive selection (Ka/Ks>1).

Orthologous unigenes	Gene name	ka/ks
OG10207	Immunoglobulin-like domain containing receptor 1b	999
OG09786	Uncharacterized protein LOC751751 isoform X1	999
OG11255	Beta-catenin-interacting protein 1 (ICAT)	999
OG08139	Extracellular matrix protein 2	999
OG12459	Photosystem II 10 kda polypeptide	36.36
OG09166	ceramide galactosyltransferase (CGT)	6.23
OG11017	C-C motif chemokine 25 (CCL25)	4.0
OG11439	Centrosomal protein of 164 kda	2.52
OG12633	Vasodilator-stimulated phosphoprotein (VASP)	2.43
OG07302	Na(+)/H(+) exchange regulatory cofactor (NHE-RF2)	2.38
OG08892	Integral membrane protein 2B (IMP2B)	2.20
OG08559	S100 calcium binding protein V2	1.99
OG10756	T cell receptor delta chain	1.85
OG11002	C-C motif chemokine 3 （CCL3）	1.72
OG11268	Protein S100-P	1.72
OG07173	Complement component 8 alpha (C8A)	1.67
OG09053	Nuclear-pore anchor-like	1.52
OG09012	Myozenin-1	1.51
OG12805	Uncharacterized protein LOC103909146 isoform X2	1.42
OG07743	FXYD domain-containing ion transport regulator 5-like isoform X1	1.41
OG04895	Chchd3 protein	1.35
OG09047	Interleukin 2 receptor, gamma (IL2RG)	1.31
OG07177	Etratricopeptide repeat protein 9C	1.28
OG09992	Regulator of cell cycle RGCC isoform X2	1.26
OG06385	Interleukin-12 alpha chain	1.21
OG10887	Pancreatic secretory trypsin inhibitor	1.20
OG08016	ETS domain-containing protein Elk-1 (ELK-1)	1.14
OG11127	Interleukin-34	1.13
OG05509	Cystatin	1.13
OG07729	Interferon receptor 1 （IFNAR1）	1.12
OG12954	Small lysine-rich protein 1	1.12
OG07825	UDP-N-acetylglucosamine transferase subunit ALG14 homolog	1.11
OG08844	Apolipoprotein C Ia	1.10
OG05561	Prostate stem cell antigen	1.10
OG09068	Actin-related protein 2/3 complex subunit 3	1.09
OG09729	Scotin	1.08
OG09300	Transmembrane protein 218	1.07
OG10138	Mitochondrial import inner membrane translocase subunit Tim8 A	1.06
OG09446	CD63 antigen (CD63)	1.06
OG11026	Uncharacterized protein LOC768128 isoform X1	1.05

OG09034	Protein FAM134C	1.05
OG11081	Uncharacterized protein LOC570390	1.04
OG07221	CD302 antigen	1.04
OG03673	C-C motif chemokine 8 (CCL8)	1.03
OG08905	Eukaryotic translation initiation factor 2-alpha kinase 1 (EIF2AK1)	1.03
OG02854	Multidrug resistance-associated protein 1	1.03
OG12777	Mucin-2-like	1.02
OG07502	Asparagine--tRNA ligase synthetase (asnS)	1.02
OG09775	Epithelial membrane protein 2 isoform X1	1.02
OG11590	Nuclear factor, interleukin 3 regulated, member 2	1.01
OG12581	Light harvesting complex protein LHCII-3	1.01
OG07693	Dual specificity tyrosine-phosphorylation-regulated kinase 4 isoform X2	1.01

PSG functions and metabolite pathways

To detect genes that might be involved in the adaptation to diverse aquatic environments in the QTP, we performed GO functional enrichment and KEGG pathways analysis on the 52 PSGs with orthologous unigenes.

Our GO enrichment analysis identified six PSGs (OG03673, OG08892, OG11017, OG06385, OG11002, and OG11127) that were significantly enriched in four molecular functions (cytokine activity, cytokine receptor binding, chemokine activity, and chemokine receptor binding), and one biological process (immune response; Table 4). KEGG enrichment analysis indicated that six PSGs (OG03673, OG09047, OG07729, OG11002, OG11017, and OG08905) were significantly enriched in two KEGG pathways (cytokine-cytokine receptor interaction and measles) (Table 5 and Figure 4). Thus, we identified nine candidate PSGs in the schizothoracine fish that may be involved in the adaptation to diverse aquatic environments. Notably, all nine of the candidate PSGs were related to fish innate and/or adaptive immunity.

The lineage-specific evolution test of the nine candidate PSGs indicated that different candidate PSGs acted more strongly on different specific lineages or species, as compared to other lineages or species (Figure 2). For example, of the 13 species, *S. lantsangensis* possessed the most PSGs (nine), followed by *Ptychobarbus dipogon*, *S. younghusbandi,* and *S. stoliczkai* (eight PSGs), while *S. labiatus* had the fewest PSGs (one).

Table 4 Key PSGs involved in the significantly enriched GO functions.

GO accession	GO terms	GO type	Positively selected genes
GO:0005125	Cytokine activity	MF	OG03673, OG08892, OG11017, OG06385, OG11002
GO:0005126	Cytokine receptor binding	MF	OG11017, OG11002, OG06385, OG11127, OG03673
GO:0008009	Chemokine activity	MF	OG11017, OG11002, OG03673
GO:0042379	Chemokine receptor binding	MF	OG03673, OG11017, OG11002
GO:0006955	Immune response	BP	OG03673, OG08892, OG11017, OG06385, OG11002

Table 5 Key PSGs involved in the significantly enriched KEGG pathways.

	KEGG pathway	ID	Positively selected genes
Immune system	Cytokine-cytokine receptor interaction	ko04060	OG07729, OG11002, OG03673, OG09047, OG11017
Immune system	Measles	ko05162	OG07729, OG08905, OG09047
Immune system	Chemokine signaling pathway	ko04062	OG11002, OG03673, OG11017

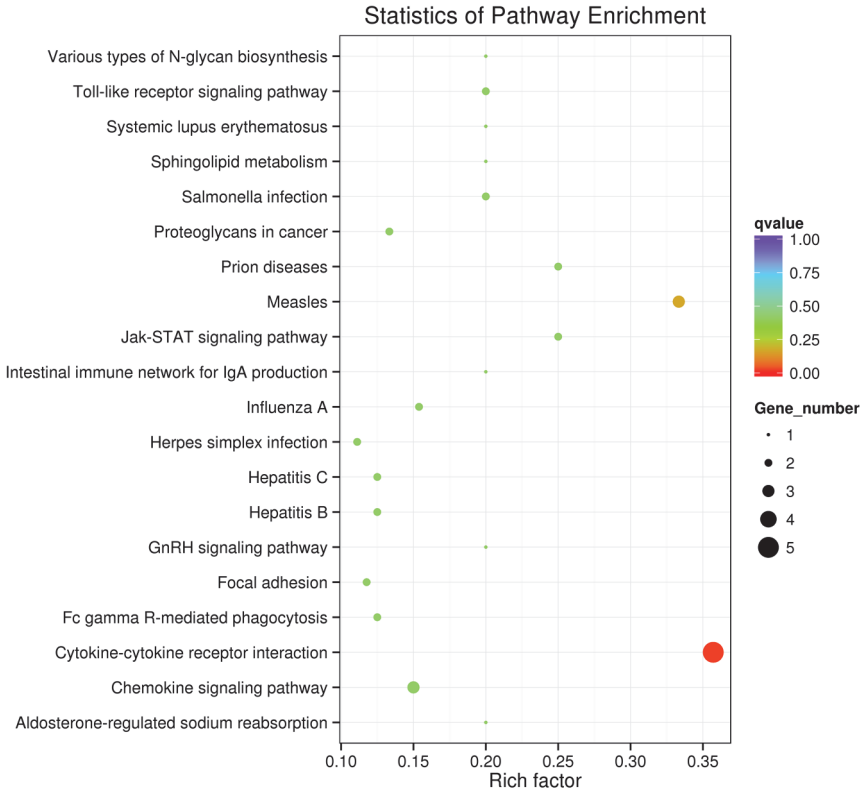

Figure 4 An overview of the KEGG pathways significantly enriched in the 52 PSGs. Specific pathways are plotted along the y-axis, and the x-axis indicates the enrichment factor. The size each colored dot indicates

the number of candidate PSGs associated with each corresponding pathway: pathways with larger-sized dots contain a larger numbers of genes. The color of the each dot indicates the corrected *P*-value for the corresponding pathway.

DISCUSSION

A better understanding of fish genetics will provide key insights into the adaptation of fish to diverse aquatic environments in the QTP, and thus increase our knowledge of the molecular mechanisms underlying the speciation, dispersal, and current distribution of the freshwater fish endemic to the QTP. Previous studies, comparing the transcriptomics of single QTP fish species of QTP to other low-land species, have identified a set of candidate genes related to energy metabolism and hypoxia response in freshwater fish in the QTP (Yang et al. 2014; Tong et al. 2017). Here, our aim was to characterize interspecific transcriptomic differences within the subfamily Schizothoracinae, and, more specifically, to identify the key PSGs involved in the significantly enriched functions and metabolite pathways acting on the specific lineages (or species). In this way, we aimed to investigate the genetic basis underlying the adaptation of this subfamily to local aquatic environments. As part of our study, we also sequenced and assembled merged-tissue transcriptomes of 13 schizothoracine species, providing an important genetic resource that was previously lacking.

Transcriptome quality is usually influenced by the method of library construction and the sequencing platform. Our assembly statistics indicated that the species with the greatest number of sequencing reads yielded more assembled transcripts, unigenes, and annotated open reading frames (ORFs). Our merged transcriptomes yielded 95,251–145,805 unigenes for each of the 13 schizothoracine species, which is consistent with a previous report of single-tissue RNA-seq data from *Gymnocypris przewalskii* (Zhang et al. 2015), but far greater than multiple-tissue merged transcriptomes from *Gymnodiptychus pachycheilus* (Yang et al. 2014), *G. przewalskii* (Tong et al. 2017), and glyptosternoid fish (Ma et al. 2015; Kang et al. 2017). The N50 values across the 13 species were similar, ranging between 1,030 and 1,567 kb (mean length ranging from 635–745 bp; Table S1). These N50 values were slightly smaller than those of *G.*

przewalskii (Tong et al. 2017) but consistent with most recent comparative transcriptome studies in fish (Yang et al. 2014; Ma et al. 2015; Kang et al. 2017).

The number of Nr annotations was similar across the 13 species (ranging from 34,532 to 43,254; Table 2), and consistent with estimates for single-tissue RNA-seq data in *G. przewalskii* [17]. However, the number of Nr annotations recovered here was far greater than multiple-tissue merged transcriptomes produced for other fish species (Yang et al. 2014; Ma et al. 2015; Kang et al. 2017; Tong et al. 2017). Out of the 95,251–145,805 assembled unigenes from the 13 species, we annotated 81.43%–90.94% against six public databases. Unigenes without significant hits may be orphan genes, noncoding RNAs, untranslated transcripts, or misassembled transcripts. Thus, there might have been orphan genes specific to the schizothoracine fish (Yang et al. 2014), which might have rapidly evolved, taking indispensable biological roles and contributing to lineage-specific phenotypes and adaptations (Chen et al. 2013). Relatively few shared orthologs were identified because large interspecific comparisons are still somewhat nascent. Even so, we identified 2,064 orthologous genes in 13 schizothoracine species, which were fewer than that identified in two congeneric species of naked carp (Zhang et al. 2015), but the amount was consistent with similar comparative transcriptome studies in other fish species (Yang et al. 2014; Ma et al. 2015; Tong et al. 2017). Thus, we have here obtained a set of high quality transcriptomic resources, which are appropriate for further comparative genomic analyses when genome sequencing data are not available.

The schizothoracine fish (Teleostei: Cyprinidae) are the largest and most diverse taxon of the QTP ichthyofauna (Wu and Wu 1992; Chen and Cao 2000). Previous studies have demonstrated that the uplifting of the QTP from 50 MYA had a profound effect on drainage patterns and changed the environment of the plateau substantially, promoting the speciation of the schizothoracine fish endemic to the region (Li et al. 2013; Qi et al. 2015). Previous studies based on comprehensive transcriptomes have shown that genes associated with energy metabolism and the hypoxia response underwent significantly accelerated or adaptive evolution in the schizothoracine fish, as compared to other teleost fish from the plains (Yang et al. 2014; Tong et al. 2017). Here, we identified 2,064 orthologous genes from the 13

schizothoracine fish. Of these, we identified 52 genes as candidate genes. These candidate genes have probably undergone positive selection along the whole schizothoracine fish lineage. We then found that nine candidate PSGs were significantly enriched in several key GO functions and metabolite pathways. Interestingly, all of the functions and metabolite pathways significantly enriched in the schizothoracine fish were associated with the immune system, which was inconsistent with previous studies of the schizothoracine fish (Yang et al. 2014; Tong et al. 2017). This discrepancy may be because previous studies focused on comparative transcriptomics between the schizothoracine fish and other teleost fishes from the plains, instead of interspecies comparisons within schizothoracine fish.

Fish immune function is known to be associated with habitat structure (Bowden 2008; Diepeveen et al. 2013). During the colonization of ecologically novel drainages or upon the exploitation of vacant ecological niches, selection on immune-related genes can be particularly strong as fishes encounter different pathogens (Birrer et al. 2012). Here, we identified nine candidate PSGs that were significantly enriched into five GO functions and two KEGG pathways. All of the enriched GO functions were related to the immune system, including cytokine activity, cytokine receptor binding, chemokine activity, chemokine receptor binding, and immune response. These results were supported by our KEGG pathway analysis.

Cytokines are a family of low-molecular-weight proteins that are secreted by activated immune-related cells upon induction by various parasitic, bacterial, or viral pathogens (Salazar-Mather and Hokeness 2006). Cytokines regulate the immune response as autocrines or paracrines by binding to the appropriate receptors (Zhu et al. 2013a). Cytokines are divided into interferons (IFNs), interleukins (ILs), tumor necrosis factors (TNFs), colony-stimulating factors, and chemokines; cytokines are secreted by various types of immune cells, including macrophages, lymphocytes, granulocytes, dendritic cells, mast cells, and epithelial cells (Salazar-Mather and Hokeness 2006). Therefore, cytokines have been catalogued as key regulators of the immune response, acting as a bridge between innate and adaptive responses (Alejo and Tafalla 2011). Here, seven of the nine candidate PSGs that were significantly enriched in GO functions and KEGG pathways were either cytokines or cytokine receptors, including interferon receptor 1 (IFNAR1, OG07729), C-C motif chemokine 25 (CCL25,

OG11017), C-C motif chemokine 3 (CCL3, OG11002), C-C motif chemokine 8 (CCL8, OG03673), interleukin-34 (IL-34, OG11127), interleukin 2 receptor, gamma (IL2RG, OG09047), and interleukin-12 alpha chain (IL-12, OG06385).

Chemokines are early-response, chemotactic members of cytokine family secreted by infected tissue cells (Salazar-Mather and Hokeness 2006; Alejo and Tafalla 2011; Bird and Tafalla 2015). Chemokines recruit monocytes, neutrophils, and other effector cells from blood vessels to the infection site (Rajesh 2016). In fish, CCL3, CCL8, and CCL25 have been shown to play important roles in the recruitment of monocytes and leukocytes, as well as in the induction of the inflammatory response to microbial invasion (de Oliveira et al. 2013; de Oliveira et al. 2015).

Interleukins, the largest group of cytokines, are associated with both pro- and anti-inflammatory functions (Rajesh 2016). For example, IL-12 is produced primarily by antigen presenting cells (APC), such as macrophages, dendritic cells, and B cells; however, this cytokine is indispensable for macrophage, neutrophil, and lymphocyte recruitment to infected tissues and for the activation of these cells as pathogen eliminators (Svanborg et al. 1999). In fish, IL-34 expression is sensitive to inflammatory stimuli and may regulate macrophage biology (Wang et al. 2013).

Interferons (IFNs) are secreted proteins that may induce an antiviral state in cells, and that, in vertebrates, play a major role in antiviral defense (Samuel 2001). During viral infections, IFNs bind to the IFN receptors (IFNARs), which trigger signal transduction through the JAK-STAT signal transduction pathway, resulting in the expression of Mx and other antiviral proteins (Samuel 2001).

The schizothoracine fish are widely distributed in drainages throughout the QTP (Wu and Wu 1992; Chen and Cao 2000). Although there is little information about fish pathogens in the QTP, the different aquatic environments of the QTP may have different microbial populations and, thus, potentially unique pathogens. Even in the same drainage system, pathogens may be unique to a given micro-ecosystem due to differences in dissolved oxygen, ultraviolet radiation, and physical and chemical properties of the water. Thus, different schizothoracine fish are

inevitably under selective pressures from new pathogens during the colonization of ecologically novel drainages or upon the exploitation of vacant ecological niches. Here, our lineage-specific test of the evolution of the nine candidate PSGs indicated that species from the same drainage shared more of the same PSGs. However, these PSGs also had species-specific features due to ecological differences among drainages, as well as among micro-habitats in the same drainage (e.g., benthic and pelagic). For example, of the four species collected from the Tsangpo River (*S. younghusbandi*, *P. dipogon*, *Schizothorax macropogon,* and *Oxygymnocypris stewartii*), three (*S. younghusbandi*, *P. dipogon,* and *S. macropogon*) are benthic (Wu and Wu 1992; Chen and Cao 2000; Qi et al. 2012). However, *S. younghusbandi* and *P. dipogon* share only eight PSGs, while *S. macropogon* shares only six. *O. stewartii*, a pelagic fish, shares four PSGs with *S. younghusbandi*, *P. dipogon,* and *S. macropogon*. Similar patterns were observed in the fish species from the Yellow River.

Fish are armed with less-developed adaptive immune systems as compared to those or mammals (Meng et al. 2012). The innate immune system plays an important role in fish, acting to rapidly eliminate pathogens, including bacteria and parasites, as the first line of defense against infection (Zhu et al. 2013b; Tong et al. 2017). The previous comparative transcriptomics of the schizothoracine fish and other teleost fish from the plains (Yang et al. 2014; Tong et al. 2017), together with our findings based on interspecific transcriptomic comparisons within the schizothoracine fish, suggested that genes associated with energy metabolism and the hypoxia response in the schizothoracine have been subjected to accelerated or adaptive evolution. These genes may have been critical for the adaptation of the schizothoracine fish to the cold and hypoxic environment of the uplifted QTP. In addition, the adaptive evolution of immune genes may have allowed the fish species to colonize ecologically novel environments or to exploit vacant ecological niches during speciation.

In summary, we assembled the transcriptomes of 13 schizothoracine fish. The genetic adaptation of these fish to local aquatic environments during speciation was investigated, based on an interspecific comparative analysis of the transcriptomes. We identified nine PSGs (out of 52) that were significantly enriched in immune system-related functions and metabolite pathways, which indicated that species-specific differences may have evolved in response to

different ecological environments and to different living habits (e.g., benthic or pelagic). Our results suggested that the adaptive evolution of immune genes allowed new schizothoracine species to colonize ecologically novel environments or to exploit vacant ecological niches during speciation.

ACKNOWLEDGMENTS

This work was supported by grants from the National Natural Science Foundation of China (31460094) and the Natural Science Foundation of Qinghai Science & Technology Department in China (2015-ZJ-901). We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

LITERATURE CITED

Alejo, A., and C. Tafalla, 2011 Chemokines in teleost fish species. *Dev Comp Immunol* 35 (12):1215-1222.

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang *et al.*, 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25 (17):3389-3402.

Baack, E.J., and M.L. Stanton, 2005 Ecological factors influencing tetraploid speciation in snow buttercups (Ranunculus Adoneus): niche differentiation and tetraploid establishment. *Evolution* 59 (9):1936-1944.

Bird, S., and C. Tafalla, 2015 Teleost Chemokines and Their Receptors. *Biology (Basel)* 4 (4):756-784.

Birrer, S.C., T.B. Reusch, and O. Roth, 2012 Salinity change impairs pipefish immune defence. *Fish Shellfish Immunol* 33 (6):1238-1248.

Bowden, T.J., 2008 Modulation of the immune system of fish by their environment. *Fish Shellfish Immunol* 25 (4):373-383.

Cai, Q., X. Qian, Y. Lang, Y. Luo, J. Xu *et al.*, 2013 Genome sequence of ground tit Pseudopodoces humilis and its adaptation to high altitude. *Genome Biol* 14 (3):R29.

Cao, W.X., Y.Y. Chen, Y.F. Wu, and S.Q. Zhu, 1981 Origin and evolution of schizothoracine fishes in relation to the upheaval of the Xizang Plateau, in Collection in Studies on the Period, Amplitude and Type of the Uplift of the Qinghai-Xizang Plateau pp. 118-130 in *Studies on the Period, Amplitude and Type of the Uplift of the Qinghai–Xizang Plateau*, edited by C.A.o.S. The Team of the Comprehensive Scientific Expedition to the Qinghai-Xizang Plateau. Science Press, Beijing.

Castresana, J., 2000 Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17 (4):540-552.

Chen, S., B.H. Krinsky, and M. Long, 2013 New genes as drivers of phenotypic evolution. *Nat Rev Genet* 14 (9):645-660.

Chen, Y.F., and W.Y. Cao, 2000 Schizothoracinae, pp. 273-390. in *Fauna Sinica, Osteichthyes, Cypriniformes III.*, edited by P.Q. Yue. Science Press, Beijing.

Collin, H., and L. Fumagalli, 2011 Evidence for morphological and adaptive genetic divergence between lake and stream habitats in European minnows (Phoxinus phoxinus, Cyprinidae). *Mol Ecol* 20 (21):4490-4502.

de Oliveira, S., A. Lopez-Munoz, F.J. Martinez-Navarro, J. Galindo-Villegas, V. Mulero *et al.*, 2015 Cxcl8-l1 and Cxcl8-l2 are required in the zebrafish defense against Salmonella Typhimurium. *Dev Comp Immunol* 49 (1):44-48.

de Oliveira, S., C.C. Reyes-Aldasoro, S. Candel, S.A. Renshaw, V. Mulero *et al.*, 2013 Cxcl8 (IL-8) mediates neutrophil recruitment and behavior in the zebrafish inflammatory response. *J Immunol* 190 (8):4349-4359.

Diepeveen, E.T., O. Roth, and W. Salzburger, 2013 Immune-related functions of the Hivep gene family in East African cichlid fishes. *G3 (Bethesda)* 3 (12):2205-2217.

Edgar, R.C., 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32 (5):1792-1797.

Funk, D.J., K.E. Filchak, and J.L. Feder, 2002 Herbivorous insects: model systems for the comparative study of speciation ecology. *Genetica* 116 (2-3):251-267.

Gotz, S., J.M. Garcia-Gomez, J. Terol, T.D. Williams, S.H. Nagaraj *et al.*, 2008 High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 36 (10):3420-3435.

Grabherr, M.G., B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson *et al.*, 2011 Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29 (7):644-652.

Guan, L., W. Chi, W. Xiao, L. Chen, and S. He, 2014 Analysis of hypoxia-inducible factor alpha polyploidization reveals adaptation to Tibetan Plateau in the evolution of schizothoracine fish. *BMC Evol Biol* 14:192.

Kang, J., X. Ma, and S. He, 2017 Evidence of high-altitude adaptation in the glyptosternoid fish, Creteuchiloglanis macropterus from the Nujiang River obtained through transcriptome analysis. *BMC Evol Biol* 17 (1):229.

Kirkpatrick, M., and N. Barton, 2006 Chromosome inversions, local adaptation and speciation. *Genetics* 173 (1):419-434.

Li, L., C.J. Stoeckert, Jr., and D.S. Roos, 2003 OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 13 (9):2178-2189.

Li, Y., Z. Ren, A.M. Shedlock, J. Wu, L. Sang *et al.*, 2013 High altitude adaptation of the schizothoracine fishes (Cyprinidae) revealed by the mitochondrial genome analyses. *Gene* 517 (2):169-178.

Ma, X., W. Dai, J. Kang, L. Yang, and S. He, 2015 Comprehensive Transcriptome Analysis of Six Catfish Species from an Altitude Gradient Reveals Adaptive Evolution in Tibetan Fishes. *G3 (Bethesda)* 6 (1):141-148.

Malinsky, M., R.J. Challis, A.M. Tyers, S. Schiffels, Y. Terai *et al.*, 2015 Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science* 350 (6267):1493-1498.

Marra, N.J., V.P. Richards, A. Early, S.M. Bogdanowicz, P.D. Pavinski Bitar *et al.*, 2017 Comparative transcriptomics of elasmobranchs and teleosts highlight important processes in adaptive immunity and regional endothermy. *BMC Genomics* 18 (1):87.

Meng, F., Y. Sun, X. Liu, J. Wang, T. Xu *et al.*, 2012 Analysis of C3 suggests three periods of positive selection events and different evolutionary patterns between fish and mammals. *PLoS One* 7 (5):e37489.

Nosil, P., L.J. Harmon, and O. Seehausen, 2009 Ecological explanations for (incomplete) speciation. *Trends Ecol Evol* 24 (3):145-156.

Qi, D., Y. Chao, S. Guo, L. Zhao, T. Li *et al.*, 2012 Convergent, parallel and correlated evolution of trophic morphologies in the subfamily schizothoracinae from the Qinghai-Tibetan plateau. *PLoS One* 7 (3):e34070.

Qi, D., T. Li, X. Zhao, S. Guo, and J. Li, 2006 Mitochondrial cytochrome b sequence variation and phylogenetics of the highly specialized Schizothoracine fishes (Teleostei: Cyprinidae) in the Qinghai-Tibet plateau. *Biochem Genet* 44 (5-6):270-285.

Qi, D.L., S.C. Guo, Y. Chao, Q.H. Kong, C.Z. Li *et al.*, 2015 The biogeography and phylogeny of

schizothoracine fishes (Schizopygopsis) in the Qinghai-Tibetan Plateau. *Zoologica Scripta* 44 (5):523-533.

Qu, Y., H. Zhao, N. Han, G. Zhou, G. Song *et al.*, 2013 Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan plateau. *Nat Commun* 4:2071.

Rajesh, P., Prasanth, B. Venkatesh, K. Mukesh, P. Jesu, A., 2016 Innate and adaptive immune molecules of striped murrel Channa striatus. *Reviews in Aquaculture* 0:1-24.

Rocha, L.A., D.R. Robertson, J. Roman, and B.W. Bowen, 2005 Ecological speciation in tropical reef fishes. *Proc Biol Sci* 272 (1563):573-579.

Saitoh, K., T. Sado, R.L. Mayden, N. Hanzawa, K. Nakamura *et al.*, 2006 Mitogenomic evolution and interrelationships of the Cypriniformes (Actinopterygii: Ostariophysi): the first evidence toward resolution of higher-level relationships of the world's largest freshwater fish clade based on 59 whole mitogenome sequences. *J Mol Evol* 63 (6):826-841.

Salazar-Mather, T.P., and K.L. Hokeness, 2006 Cytokine and chemokine networks: pathways to antiviral defense. *Curr Top Microbiol Immunol* 303:29-46.

Samuel, C.E., 2001 Antiviral actions of interferons. *Clin Microbiol Rev* 14 (4):778-809, table of contents.

Svanborg, C., G. Godaly, and M. Hedlund, 1999 Cytokine responses during mucosal infections: role in disease pathogenesis and host defence. *Curr Opin Microbiol* 2 (1):99-105.

Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar, 2013 MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30 (12):2725-2729.

Tatusov, R.L., N.D. Fedorova, J.D. Jackson, A.R. Jacobs, B. Kiryutin *et al.*, 2003 The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4:41.

Tong, C., T. Fei, C. Zhang, and K. Zhao, 2017 Comprehensive transcriptomic analysis of Tibetan Schizothoracinae fish Gymnocypris przewalskii reveals how it adapts to a high altitude aquatic life. *BMC Evol Biol* 17 (1):74.

Tong, C., C. Zhang, R. Zhang, and K. Zhao, 2015 Transcriptome profiling analysis of naked carp (Gymnocypris przewalskii) provides insights into the immune-related genes in highland fish. *Fish Shellfish Immunol* 46 (2):366-377.

Tonnis, B., P.R. Grant, B.R. Grant, and K. Petren, 2005 Habitat selection and ecological speciation in Galapagos warbler finches (Certhidea olivacea and Certhidea fusca). *Proc Biol Sci* 272 (1565):819-826.

Wang, T., T. Kono, M.M. Monte, H. Kuse, M.M. Costa *et al.*, 2013 Identification of IL-34 in teleost fish: differential expression of rainbow trout IL-34, MCSF1 and MCSF2, ligands of the MCSF receptor. *Mol Immunol* 53 (4):398-409.

Wu, Y.F., and C.Z. Wu, 1992 *The fishes of the Qinghai – Xizang plateau*. Chengdu: Science and Technology Press.

Xie, C., X. Mao, J. Huang, Y. Ding, J. Wu *et al.*, 2011 KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res* 39 (Web Server issue):W316-322.

Xu, J., J.T. Li, Y. Jiang, W. Peng, Z. Yao *et al.*, 2017 Genomic Basis of Adaptive Evolution: The Survival of Amur Ide (Leuciscus waleckii) in an Extremely Alkaline Environment. *Mol Biol Evol* 34 (1):145-159.

Yang, J., X. Chen, J. Bai, D. Fang, Y. Qiu *et al.*, 2016 The Sinocyclocheilus cavefish genome provides insights into cave adaptation. *BMC Biol* 14:1.

Yang, L., Y. Wang, Z. Zhang, and S. He, 2014 Comprehensive transcriptome analysis reveals accelerated genic evolution in a Tibet fish, Gymnodiptychus pachycheilus. *Genome Biol Evol* 7 (1):251-261.

Yang, Z., 2007 PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24 (8):1586-1591.

Yang, Z., W.S. Wong, and R. Nielsen, 2005 Bayes empirical bayes inference of amino acid sites under positive selection. *Mol Biol Evol* 22 (4):1107-1118.

Yang, Z.H., and R. Nielsen, 2002 Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol* 19: 908–917.

Young, M.D., M.J. Wakefield, G.K. Smyth, and A. Oshlack, 2010 Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol* 11 (2):R14.

Zhang, J.Z., Nielsen, R., Yang, Z.H., 2005 Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol* 22:2472–2479.

Zhang, R., A. Ludwig, C. Zhang, C. Tong, G. Li *et al.*, 2015 Local adaptation of Gymnocypris przewalskii (Cyprinidae) on the Tibetan Plateau. *Sci Rep* 5:9780.

Zhang, R., Z. Peng, G. Li, C. Zhang, Y. Tang *et al.*, 2013 Ongoing speciation in the Tibetan plateau Gymnocypris species complex. *PLoS One* 8 (8):e71331.

Zhou, W.W., Y. Wen, J. Fu, Y.B. Xu, J.Q. Jin *et al.*, 2012 Speciation in the Rana chensinensis species complex and its relationship to the uplift of the Qinghai-Tibetan Plateau. *Mol Ecol* 21 (4):960-973.

Zhu, L.Y., L. Nie, G. Zhu, L.X. Xiang, and J.Z. Shao, 2013a Advances in research of fish immune-relevant genes: a comparative overview of innate and adaptive immunity in teleosts. *Dev Comp Immunol* 39 (1-2):39-62.

Zhu, Z., Y. Sun, R. Wang, and T. Xu, 2013b Evolutionary analysis of TLR9 genes reveals the positive selection of extant teleosts in Perciformes. *Fish Shellfish Immunol* 35 (2):448-457.