Turk. J. Fish.& Aquat. Sci. 20(1), 29-37
http://doi.org/10.4194/1303-2712-v20_1_03

RESEARCH PAPER

Turkish Journal of FISHERIES and AQUATIC SCIENCES

The Complete Mitochondrial Genome of Endemic Freshwater Sculpin *Cottus dzungaricus* (Scorpaeniformes: Cottidae) Revealed by High-Throughput NGS: Genome Structure and Its Phylogenetic Relationships

Tianyan Yang[1], Wei Meng[2], Wang Xuhui[2], Tianxiang Gao[1], Yunguo Liu[2,*]

[1] Zhejiang Ocean University, Fisheries CollegeZhoushan 316004, China.
[2] Xinjiang University, College of life Science and Technology, Urumqi 830046, China.

Article History
Received 02 June 2018
Accepted 20 December 2018
First Online 03 January 2019

Corresponding Author
Tel.: +869918582554
E-mail: yguoliu@163.com

Keywords
Cottus dzungaricus
Complete mitochondrial genome
Next generation sequencing
Phylogenetic relationships

Abstract

Cottus dzungaricus is an indigenous and infrequent cold-water sculpin in Irtysh River China, the classification status was still controversial. In this study, a total of 16,530 bp circular and double-stranded Cottus dzungaricus complete mitochondrial genome was sequenced. The average nucleotide composition showed a slightly higher AT content and strong anti-G bias. Selective pressure tests of 13 protein-coding genes revealed only Cyt b gene experienced the positive selection. The mean genetic distances (0.018-0.123) showed various interspecific divergences within genus Cottus. The NJ phylogenetic tree based on K2-P model fell into two major lineages, one of which was further divided into two sublineages. These three clusters of cottoid fishes might respectively represent different salinity adaption types. Combined with their geographical distribution pattern, we speculated that genus Cottus originated in the Far East and C. dzungaricus was a transitional or intermediate species during the population dispersion. The controversial classification status of C. dzungaricus was also discussed by reconstructing a dendrogram based on mtDNA control region. The conclusions will lay the foundation for clarifying the classification relationship of genus Cottus.

Introduction

Freshwater sculpins *Cottus dzungaricus* belonging to order Osteichthyes, family Scorpaeniformes and genus Cottidae is a scarce cold-water fish that distributes only in Irtysh River China (Guo *et al*., 2012). It was first collected and named as a subspecies of *C. sibiricus* based on morphological characteristic differences by Li and Ho in 1966 (Li *et al*., 1966). In recent years, due to habitat fragmentation and illegal fishing, the population of *C. dzungaricus* has been declining seriously. It has been listed as the class II protected animal in Xinjiang Uyghur Autonomous Region. However, limited studies have been conducted on this aborigine species for a long time because of difficulties in sampling. The molecular genetic studies of *C. dzungaricus* will present great significance for phylogenetics and molecular evolution of sculpin fishes.

Mitochondrial DNA is a useful and popular molecular marker in population genetics, evolutionary biology and phylogenetic studies of species (Avise *et al*., 1987; Harrison, 1989; Saccone *et al*., 2000), since it is characterized by strict maternal inheritance, relative lack of recombination, rapid evolutionary rate compared with nuclear DNA (nDNA), and the ability to provide an abundance of genotype (Rokas Ladoukakis, & Zouros, 2003). Instead of partial sequences, the complete mitochondrial DNA genome is more able to

offer comprehensive and objective data about genetic variations. Nowadays, primer walking sequencing technique combined with LA-PCR (Long and accurate polymerase chain reaction) strategy are widely used in complete mitochondrial genome studies (Lambert, Mariam, & Susan, 2010; Sha Lin, Li, & Huang, 2013). But with the advances in sequencing technology, the recently developed next-generation sequencing (NGS) technologies have gradually led to the proposal of more straightforward integrated pipelines for complete mitogenome sequencing (Timmermans Dodsworth, Culverwell, & Vogler, 2010; Yang & Huang, 2015). Especially for the large scale genome studies, the NGS was the best choice for saving time and efforts.

In this study, the complete mitochondrial genome sequence of *C. dzungaricus* was initially determined by NGS with Illumina HiSeq2500 platform. The structure of mtDNA sequence was described and analyzed. In addition, phylogenetic relationship and divergence time were also discussed to further confirm the taxonomic status of *C. dzungaricus* and provide fundamental genetic information for indigenous fishes in Xinjiang China.

Materials and Methods

Sample Collection

The fresh and alive specimen was collected from Irtysh River China in 2015. The GPS Coordinates of sampling site was 86.20°E, 47.87°N. The pectoral fin was taken and preserved in absolute ethyl alcohol at -20℃. The mitochondrial genomic DNA was extracted by removing nuclear DNA by differential centrifugation using BioVision mitochondrial DNA isolation kit.

Illumina Sequencing and Data Analysis

After being tested qualified, the genomic DNA was first mechanical fragmented by ultrasonic with the average size 450 bp. A sequencing library was generated and paired-end sequenced through a single lane of Illumina Hiseq2500 based on sequencing by synthesis (SBS) technology. The 5.22 Gb raw data with the average length 150 bp were purified (removing adapters, primers, unpaired, short and low-quality reads) and high-quality NGS reads (4.66 Gb, Phred Quality Score Q30≥91.47%) were directly assembled by NOVOPlasty with default parameter settings (Dierckxsens, Mardulyn, & Smits, 2017). The mtDNA genome was annotated and mapped by MitoFish software (Wataru *et al*., 2013).

The relative influence of natural selection acting on PCGs can be detected by comparing the rates of non-synonymous substitutions (*Ka*) versus synonymous substitutions (*Ks*), which indicates the net balance between deleterious and beneficial mutations (Yang & Bielawski, 2000; Hurst, 2002; Nielsen, 2005). Eight species of genus *Cottus* that had close relationships (*C. amblystomopsis* KY563345, *C. asper* MF326939, *C. bairdii* KP013090, *C. czerskii* KJ956027, *C. perifretum* MF326940, *C. reinii* AP004442*, C.volki* KY563344 and *C. dzungaricus* in this study) were chosen for selection pressure analysis. We examined three pairs of null and alternative hypothesis models to assess the *Ka/Ks* ratio (or ω, *dN/dS*) for all codon sites by EasyCodeML software (Gao & Chen, 2016) using site model. The likelihood ratio tests (LRTs) were performed to compare the fit of three pairwise models: M0 (one-ratio) vs. M3 (discrete), M1a (nearly neutral) vs. M2a (positive selection), and M7 (β distribution) vs. M8 (β distribution and ω).

In order to explore the evolutionary status of *C. dzungaricus* within family Cottidae, the Neighbor-joining (NJ) phylogenetic tree was constructed by Mega6.0 using Kimura 2-parameter model (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013) based on complete mitochondrial genomes of related species, and *Mesocottus haitej* KF170218, *Icelus spatula* KT004432, *Gymnocanthus herzensteini* KX148474 and *G. intermedius* KX148473 were selected as outgroups. All relative sequences were available from NCBI Genbank and aligned by multiple alignment program MAFFT v7.149 (Katoh & Toh, 2010; Yamada Tomii, & Katoh, 2016). To further explore the relationships between *C. dzungaricus* and *C. sibiricus*, the phylogenetic tree for the mtDNA control region was reconstructed by Maximum Likelihood (ML) method with 1,000 bootstrap replicates using PAUP4.0 (Swofford, 2002). Before tree building, MrModletest2.3 was used to determine the appropriate DNA substitution model and gamma rate heterogeneity using the Akaike Information Criterion (AIC) and the best-fit model HKY+I+G was selected (Nylander, 2004).

Results

Sequence Analysis

The length of *C. dzungaricus* mitochondrial genome was 16, 530 bp in total, containing 13 protein-coding genes (PCGs), 22 tRNA genes, 2 rRNA genes and a control region (Figure 1; Table 1). The raw sequence data have been deposited to GenBank's Sequence Read Archive (SRA) with the submission number SRR7523602, and the NCBI accession number was MH638324. Out of the 37 genes, twenty-eight were encoded on the heavy strand excluding eight tRNA genes and ND6 gene. Ten intergenic spacers (78 bp) were found ranging from 1 to 37 bp in length. Six overlaps (24 bp) were observed, in which the longest overlapping area (10 bp) was between ATP8 and ATP6 genes.

The average A, T, C and G nucleotide composition was 27.0%, 26.3%, 29.7% and 17.0%, respectively, indicating a little higher A+T content. Strong strand-specific compositional bias was known to be a most

remarkable feature of mitochondrial genomes (Alexandre, Nelly, & Jean, 2005), which was measured by the formulas AT skew = (A − T) / (A + T) and GC skew = (G − C) / (G + C), respectively. (Perna & Kocher, 1995). In this study, the overall nucleotide composition skewness revealed negative GC-skews and positive AT-skews of *C. dzungaricus* mitogenome.

As in most vertebrate mtDNA, most PCGs used ATG as initiation codon, except for COI gene starting with GTG. Five genes used TAA as termination codon, while ND6 gene was ended with TAG. Incomplete termination codons containing T (ND2, ND3, ND4, COII and Cyt *b*) or TA (ATP6 and COIII) were also present. A total of 3, 799 amino acids were encoded, and the most frequently used was hydrophobic amino acid Leucine (17.58%), while hydrophilic amino acid Cysteine (0.63%) was the least frequently used one. The values of relative synonymous codon usage (RSCU) of all PCGs codons were calculated excluding the stop codons (Table 2). A total of 30 RSCU values were greater than 1.0, with GCC (A) the highest frequency. The count of NNU and NNA (2042) was higher than that of NNC and NNG (1758), which was consistent with the high A+T content in the third codon position, indicating the strong codon usage bias.

The 12S and 16S ribosomal RNA genes were 947 bp and 1690 bp, respectively. They were located in the typical positions between tRNA-Phe and tRNA-Leu (UUR), and separated by tRNA-Val. The length of 22 tRNA was from 66 to 74 bp, with the total length of 1, 555 bp. Most of them could fold into the typical secondary structure of cloverleaf except tRNA-Ser (AGN) gene because of an absence of the DHU arm. The length of control region (CR) or called displacement loop region (D-loop) was 857 bp. It was located between tRNA-Pro and tRNA-Phe and could be divided into three domains: termination associated sequence (TAS), central conserved sequence block (CSB-F, E and D) and conserved sequence block (CSB-1, 2 and 3) (Table 3).

Selection Pressure

In order to test for the possibility of selection mode on mtDNA protein-coding sequences, we compared the average pairwise *Ka/Ks* ratio in 13 PCGs for *C. dzungaricus.* The reverse complement sequence of ND6 gene was used to infer the correct amino acid sequence. The LRTs under the first and second models showed that no positive site was detected for all PCGs, while the third model (M7 vs. M8) suggested that positive or called Darwinian selection pressure acted on six genes (Cyt *b*, ND2, ND4, ND4L, ND5 and ND6), but only Cyt *b* gene was significant (*P*=0.0004). But the posterior probabilities of

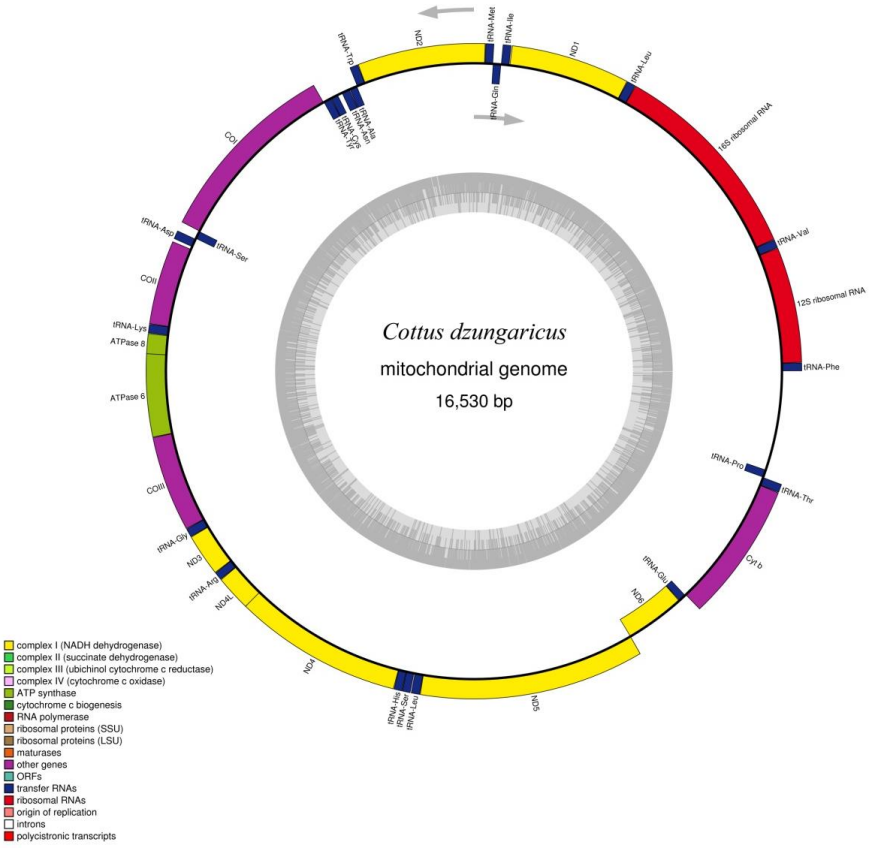

Figure 1. The mitochondrial genome map of *C. dzungaricus.*

these genes were all lower than 95% level. The results of site model on six genes presented in Table 4.

Phylogenetic Relationships

All sculpins examined in this study fell into two major lineages (A and B) strongly supported by a bootstrap value of 100% (Figure 2). One lineage comprised *C. amblystomopsis*, *C. reinii*, *C. hangiongensis* and *C.poecilopus*. The remainders constituted another lineage that could be further divided into two sublineages (A1 and A2). *C. dzungaricus* had close relationships to *C. perifretum* and *C. rhenanus*. Remarkably, the only two recognized species of genus *Comephorus*: *C. baikalensis* (Big Baikal oilfish) and *C. dybowskii* (Little Baikal oilfish) clustered among fishes of genus *Cottus*.

The controversial classification status of *C. dzungaricus* was explored by reconstructing the phylogenetic tree of *C. sibiricus* and relative species (Figure 3). Two divergent clades of Siberian sculpins were existed: one was from Irtysh River China and the other was from Yakutsk, as well as other sculpins mixed between them.

Discussions

Sequence Difference and Selective Pressure

The structure and characteristic of complete mitochondrial genome of *C. dzungaricus* was circular with a similar number and arrangement of genes as other reported sculpins (Hwang, Byeon, & Lee, 2013;

Table 1. The mitochondrial genome characteristics of *C. dzungaricus*

	Start	End	Length (bp)	Strand	Start coden	End coden	No. of proteins	Intergenic (+) Overlap (-)
tRNA-Phe	1	68	68	+				0
12S rRNA	69	1015	947	+				0
tRNA-Val	1016	1087	72	+				0
16S rRNA	1088	2777	1690	+				0
tRNA-Leu(UUR)	2778	2851	74	+				0
ND1	2852	3826	975	+	ATG	TAA	324	0
tRNA-Ile	3831	3900	70	+				4
tRNA-Gln	3900	3970	71	-				-1
tRNA-Met	3970	4039	70	+				-1
ND2	4040	5084	1045	+	ATG	T--	348	0
tRNA-Trp	5085	5155	71	+				0
tRNA-Ala	5157	5225	69	-				1
tRNA-Asn	5227	5299	73	-				1
tRNA-Cys	5337	5402	66	-				37
tRNA-Tyr	5403	5472	70	-				0
COI	5474	7024	1551	+	GTG	TAA	516	1
tRNA-Ser(UCN)	7025	7095	71	-				0
tRNA-Asp	7099	7171	73	+				3
COII	7194	7884	691	+	ATG	T--	230	22
tRNA-Lys	7885	7958	74	+				0
ATPase 8	7960	8127	168	+	ATG	TAA	55	1
ATPase 6	8118	8800	683	+	ATG	TA-	227	-10
COIII	8801	9585	785	+	ATG	TA-	261	0
tRNA-Gly	9586	9658	73	+				0
ND3	9659	10007	349	+	ATG	T--	116	0
tRNA-Arg	10008	10076	69	+				0
ND4L	10077	10373	297	+	ATG	TAA	98	0
ND4	10367	11747	1381	+	ATG	T--	460	-7
tRNA-His	11748	11816	69	+				0
tRNA-Ser(AGN)	11817	11884	68	+				0
tRNA-Leu(CUN)	11889	11961	73	+				4
ND5	11962	13800	1839	+	ATG	TAA	612	0
ND6	13797	14318	522	-	ATG	TAG	173	-4
tRNA-Glu	14319	14387	69	-				0
Cyt *b*	14392	15532	1141	+	ATG	T--	380	4
tRNA-Thr	15533	15604	72	+				0
tRNA-Pro	15604	15673	70	-				-1
Control region	15674	16530	857	+				0

"+" means heavy strand, "-" means light strand
Negative numbers indicate overlapping nucleotide.

Han, Li, Zhao, & Xu, 2016; Balakirev, Saveliev, & Ayala, 2016; Fast, Aguilar, Nolte, & Sandel, 2017), but slightly different from the sequence (Genbank no. KM093860) uploaded into NCBI database (Ao *et al*., 2014), which was obtained by the conventional Sanger dideoxy sequencing method.
It was found by sequence contrast that mutation sites o ccurred among three codons, but didn't alter the number of amino acids. It only changed the termination codons on some PCGs (ATP6, ND2, ND3 and COIII). For the tissue that was not derived from the same specimen, it was uncertain whether the difference resulted from sampling or sequencing method.

The similarity between two sequences was 99% by BLAST and the K2-P pairwise distance was 0.002. Meanwhile, no frameshift mutation or abnormal termination codons in the protein-coding region were observed. The results confirmed that high-throughput DNA sequencing technology was applicable to the complete mitochondrial genome analysis. In addition, it could reduce experiment complexity, shorten time consumed and implement work efficiency.

The positive selection sites were detected in Cyt *b*, ND2, ND4, ND4L, ND5 and ND6 genes. As the center of cellular energy metabolism, mitochondrial DNA evolves faster than nuclear genetic markers (Boursot & Bonhomme, 1986; Delsuc, Stanhope, & Douzery, 2003), which possibly led these loci to be susceptible to selective pressure. Within all positive selection sites, LRT *P*-value was only significant (P<0.01) in Cyt *b* gene. The positive diversifying selection was dominated in this gene, and diversity at the amino acid level was favored, likely due to the fitness advantage provided by the mutations (Seibert, Howell, Hughes, & Hughes, 1995).

Genetic Distance and Phylogenetics

The K2-P genetic distance among genus *Cottus* varied from 0.018 (*C. perifretum* vs. *C. rhenanus*) to 0.123 (*C. hangiongensis* vs. *C. reinii*; *C. hangiongensis* vs. *C. amblystomopsis*). Anthropogenic changes within the past 200 years had allowed the hybridization between them in European natural water bodies (Stemshorn, Reed, Nolte, & Tautz, 2011). In this study, the genetic distance between *C. perifretum* and *C. rhenanus* was the lowest, reflecting closed relationships of them genetically, and it could be also proved in the cladogram. The large span of genetic distances showed various interspecific divergence within this genus, some were small and some were larger. Therefore, it was necessary

Table 2. Codon usage in *C. dzungaricus* mitochondrial PCGs

Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU
UUU(F)	111	0.97	GUG(V)	28	0.5	GCA(A)	85	0.98	GAG(E)	17	0.34
UUC(F)	119	1.03	UCU(S)	57	1.36	GCG(A)	8	0.09	UGU(C)	10	0.83
UUA(L)	82	0.74	UCC(S)	74	1.77	UAU(Y)	29	0.54	UGC(C)	14	1.17
UUG(L)	26	0.23	UCA(S)	50	1.2	UAC(Y)	78	1.46	UGA(W)	99	1.65
CUU(L)	187	1.68	UCG(S)	10	0.24	CAU(H)	34	0.62	UGG(W)	21	0.35
CUC(L)	170	1.53	CCU(P)	74	1.37	CAC(H)	76	1.38	CGU(R)	14	0.75
CUA(L)	148	1.33	CCC(P)	94	1.74	CAA(Q)	87	1.78	CGC(R)	16	0.85
CUG(L)	55	0.49	CCA(P)	39	0.72	CAG(Q)	11	0.22	CGA(R)	34	1.81
AUU(I)	131	1.02	CCG(P)	9	0.17	AAU(N)	26	0.45	CGG(R)	11	0.59
AUC(I)	126	0.98	ACU(T)	53	0.68	AAC(N)	89	1.55	AGU(S)	11	0.26
AUA(M)	86	1.16	ACC(T)	139	1.78	AAA(K)	66	1.83	AGC(S)	49	1.17
AUG(M)	62	0.84	ACA(T)	106	1.36	AAG(K)	6	0.17	GGU(G)	37	0.59
GUU(V)	62	1.1	ACG(T)	14	0.18	GAU(D)	18	0.48	GGC(G)	91	1.46
GUC(V)	60	1.06	GCU(A)	80	0.92	GAC(D)	57	1.52	GGA(G)	67	1.08
GUA(V)	76	1.35	GCC(A)	174	2.01	GAA(E)	83	1.66	GGG(G)	54	0.87

Table 3. Sequence features of the control region in *C. Dzungaricus*

Items	Location	Sequence (5'-3')
TAS	15711-15722	TACATATATGTA
CSB-F	15996-16020	TAACGGTTATTGAAGGTGAGGGACA
CSB-E	16028-16035	GTGGGGGT
CSB-D	16050-16071	TATTCCTGGCATTTGGTTCCTA
CSB-1	16323-16346	CATACTTGTATCTCAAGAGCATAA
CSB-2	16404-16426	CCCCCCTACCCCCCTAAAACTCC
CSB-3	16444-16463	CTGAAAACCCCCCGGAAACA

Table 4. Results of positive selection tests for Cyt *b*, ND2, ND4, ND4L, ND5 and ND6 genes

Gene	Model	np	Ln L	Estimates of parameters	Model compared	LRT *P*-value	Positive selection sites
Cyt *b*	M7	17	-3322.499039	p = 0.04066, q = 0.04066	M7 vs. M8	0.000421215	Not allowed
	M8	19	-3314.726671	p_0 = 0.98791, p = 0.04273, q = 0.62638 (p_1 = 0.01209), ω = 1.00000			320 V 0.857
ND2	M7	17	-3658.046246	p = 0.12170, q = 1.29100	M7 vs. M8	0.637847533	Not allowed
	M8	19	-3657.596590	p_0 = 0.99452, p = 0.14538, q = 1.73811 (p_1 = 0.00548), ω = 1.00000			222 N 0.720,280 F 0.576
ND4	M7	17	-4731.832858	p = 0.03999, q = 0.27261	M7 vs. M8	0.092310905	Not allowed
	M8	19	-4729.450265	p_0 = 0.98147, p = 0.03937, q = 0.29323 (p_1 = 0.01853), ω = 1.00000			20 P 0.943,422 T 0.516
ND4L	M7	17	-827.451966	p = 0.03856, q = 0.22230	M7 vs. M8	0.055937740	Not allowed
	M8	19	-824.568450	p_0 = 0.91954, p = 2.27550, q = 99.00000 (p_1 = 0.08046), ω = 1.00000			53 T 0.617,61 M 0.571,8
ND5	M7	17	-5861.966355	p = 0.04123, q = 0.28307	M7 vs. M8	0.998777748	Not allowed
	M8	19	-5861.967578	p_0 = 0.99999, p = 0.04125, q = 0.28332 (p_1 = 0.00001), ω = 1.00000			540 V 0.800
ND6	M7	17	-1614.432691	p = 0.04205, q = 0.27912	M7 vs. M8	0.831430141	Not allowed
	M8	19	-1614.248083	p_0 = 0.98540, p = 0.01958, q = 0.11668 (p_1 = 0.01460), ω = 1.00000			89 A 0.794

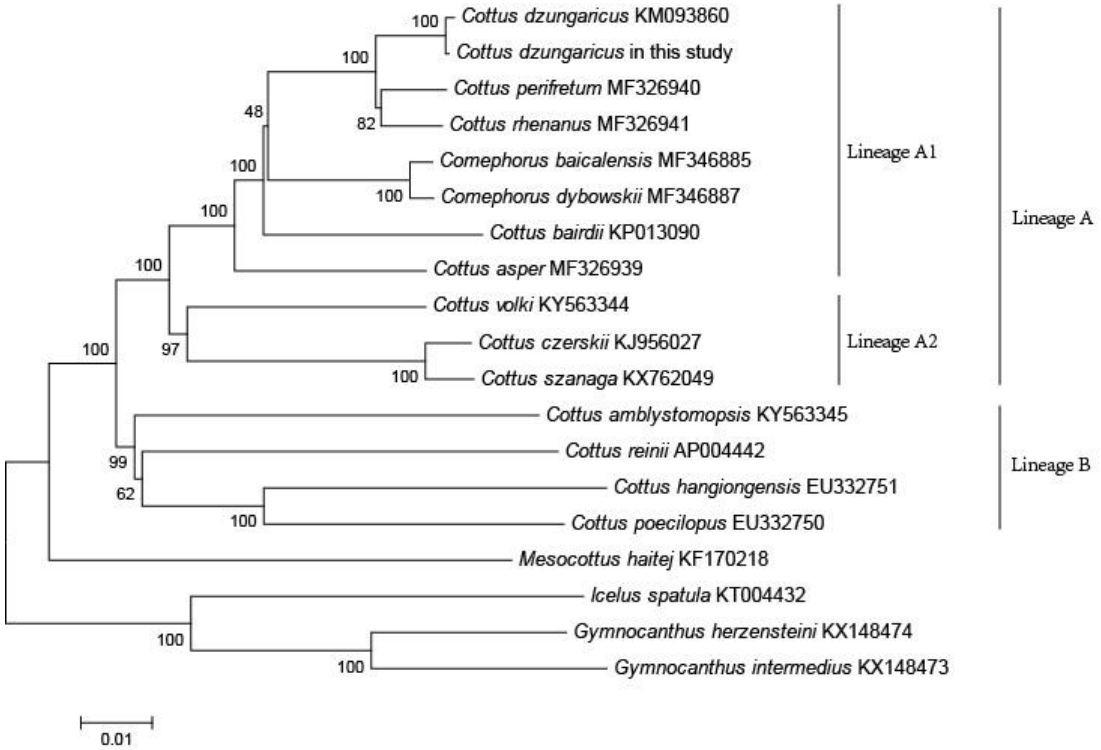

Figure 2. NJ phylogenetic tree of *C. dzungaricus* based on K2P model.

to carry out species identification supplemented by nucleotide DNA markers analysis.

It is known that the divergence of homologous genes from different species is in well congruent with phylogenetic relationship of them. From the dendrogram we found that cottoid fishes distributing in Japan, Korea and Russia in lineage B were the earliest divergent species of genus *Cottus*, then species in lineage A2 which mainly distributed in Mongolia and Russia, and species spreading over Eurasia and North America in lineage A1 suggested the latest speciation. Hence we speculated that fishes of *Cottus* originated from water systems in the Far East, gradually spread westward to Mongolia and Russia, and finally passed through Europe to North America.

C. amblystomopsis (amphidromous) and *C. hangiongensis* (catadromous) in lineage B had strong salinity adaptability. There were two main forms of *C. asper*: the inland form and the coastal form, and the latter lived in rivers and swam down into brackish estuaries to breed (Arne, 1964). The others in lineage A were all freshwater species. The

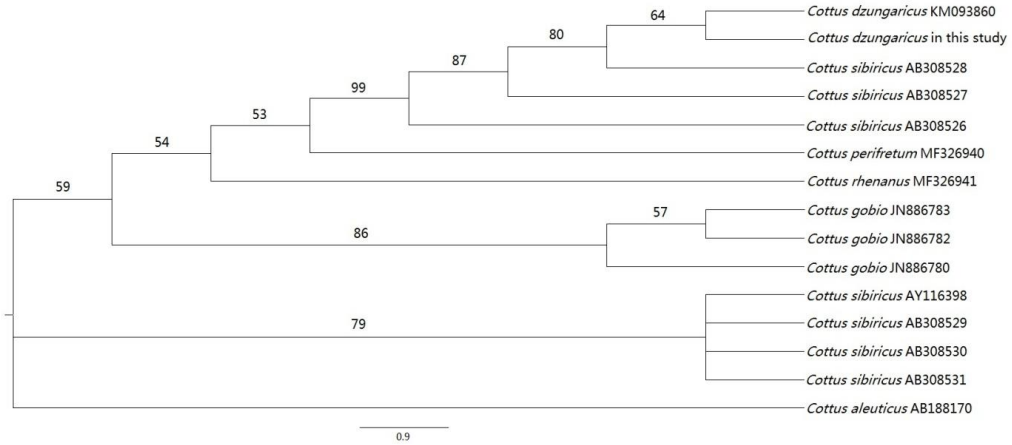

Figure 3. The reconstructed phylogenetic tree of *C. sibiricus* and relative species based on mtDNA control region.

ecological habits and biological characteristics suggested a reduced salinity adaptability of them. The speciation time of *C. dzungaricus* in Xinjiang was relatively later. It had close relationships with cottoid fishes in Europe and North America, but distant relationships with species in the Far East. With that in mind, we considered *C. dzungaricus* was a transitional or intermediate species of *Cottus* that spread from the Far East to North America.

Two sister species of cottoid fish from Lake Baikal *Comephorus baicalensis* and *Comephorus dybowskii* were located among genus *Cottus* in the phylogenetic tree, it was speculated that these two genera might be closely related phylogenetically. We conjectured that genus *Comephorus* derived from genus *Cottus* during recent stage of evolutionary history.

Taxonomic Status

There have always been many disputes in the species status of *C. dzungaricus* over a long period of time. As early as 1966, the sculpins collected from upper reaches of Irtysh River and its tributary (altitude 800-1200m) had been recorded by the name of *C. sibiricus altaicus* by Chinese scholars Li and Ho firstly. They were distinguished from type species by the morphometric characteristics such as the posterior pair of nostrils, longer pectoral fin, and the number of bone prickles (Li *et al*., 1966). Then, Swiss ichthyologist Maurice Kottelat promoted its subspecies taxonomic category to the level of species and renamed it as *C. dzungaricus* considering *C. sibiricus altaicus* was the junior primary homonym of *C. poecilopus altaicus* Kaschenko, 1899 (Kottelat, 2006). Bogdanov and Knizhin (2007) discovered a rather wide range of variation in biological traits and exterior morphometric characters of *C. sibiricus* from three large river basins of Siberia (the Lena, Yenisei and Ob basins). Their conclusions offered further evidences for Kottelat that in spite of internal heterogeneity, there were no reasons for subdivision of the Siberian sculpin *C. sibiricus* into geographic forms (subspecies). Recent comparative researches on morphological data combined with molecular evidence demonstrated that differences between *C. sibiricus* and *C. dzungaricus* lay within the limits of intraspecific variation. They supported to return *C. dzungaricus* to its subspecific level and rename *C. sibiricus dzungaricus* (Sideleva, 2017). In the present study, consulting the research results of Yokoyama, Sideleva, Shedko, and Goto (2008), we reconstructed the phylogenetic tree of *C. sibiricus* and relative species based on mtDNA control region, and found apparent two divergent clades of Siberian sculpins in phylogenetic tree: Irtysh River clade and Yakutsk clade, with some other cottoid species locating between them. The individual of *C. dzungaricus* in this study clustered with *C. sibiricus* in Irtysh River. From the data of mtDNA control region, we speculated differentiation between them might reach the rank of species, and supported the name of Xinjiang's cottoid fish as *C. dzungaricus*. For lacking of complete mitogenome information of *C. sibiricus*, the results needed to be further clarified by more sufficient morphology and molecular data of holotypes.

Acknowledgments

This study was supported by the Scientific Research Startup Foundation of Zhejiang Ocean University (2016-2017), the National Natural Science Foundation of China (31560719), and the Natural Science Foundation of Xinjiang Uygur Autonomous Region (2015211C261).

References

Alexandre, H., Nelly, L., & Jean, D. (2005). Evidence for Multiple Reversals of Asymmetric Mutational Constraints during the Evolution of the Mitochondrial Genome of Metazoa, and Consequences for Phylogenetic Inferences.

Systems Biology, 54(2), 277-298. https://dx.doi.org/10.1080/10635150590947843

Ao, M., Xie, P., Zhao, G., Guo, F.F., Niu, J.G., & Ma, X.F. (2014). The complete mitochondrial genome of *Cottus sibiricus altaicus* (Scorpaeniformes: Cottidae). Mitochondrial DNA, 2, 1-2. https://dx.doi.org/10.3109/19401736.2014.945573

Arne, B. (1964). The effect of salinity on the iodine metabolism and standard metabolic rates of coastal and inland prickly sculpins *Cottus asper* Richardson. (Master Thesis). The University of British Columbia, Vancouver, Canada.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., & Saunders, N.C. (1987). Intraspecific phylogeography: The Mitochondrial DNA Bridge Bewteen Population Genetics and Systematics. Annual Review of Ecology and Systematics, 18, 489-522. https://doi.org/0066-4162/87/1120-0489$02.00

Balakirev, E.S., Saveliev, P.A., & Ayala, F.J. (2016). Complete mitochondrial genome of the Amur sculpin *Cottus szanaga* (Cottoidei: Cottidae). Mitochondrial DNA Part B: Resources, 1, 737-738. https://doi.org/10.1080/23802359.2016.1233470

Bogdanov, B.E., & Knizhin, I. B. (2007). Variability of siberian sculpin *Cottus sibiricus* (Cottidae, Sorpaeniformes). Journal of Ichthyology, 47(4), 271-278. https://doi.org/10.1134/S0032945207040030

Boursot, P., & Bonhomme, F. (1986). Génétique et évolution du génome mitochondrial des Métazoaires. Génétique sélection evolution, 18(1), 73-98. https://doi.org/10.1186/1297-9686-18-1-73

Delsuc, F., Stanhope, M. J., & Douzery, E.J. (2003). Molecular systematics of armadillos (Xenarthra, Dasypodidae): contribution of maximum likelihood and Bayesian analyses of mitochondrial and nuclear genes. Molecular Phylogenetics and Evolution, 28(2), 261-75. https://doi.org/10.1016/S1055-7903(03)00111-8

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

Fast, K., Aguilar, A., Nolte, A.W., & Sandel, M.W. (2017). Complete mitochondrial genomes for *Cottus asper*, *Cottus perifretum*, and *Cottus rhenanus* (Perciformes, Cottidae). Mitochondrial DNA Part B: Resources, 2, 666-668. https://doi.org/10.1080/23802359.2017.1375870

Gao, F., & Chen, J. (2016). EasyCodeML: an interactive visual tool for CodeML analysis. Retrieved from http://blog.sciencenet.cn/blog-460481-903608.html

Guo, Y., Zhang, R.M., Cai, L.G., Ma, Y.W., Niu, J. G., Tursun, T.V.D., … Xie, C. G. (2012). Fish Fauna of Xinjiang (pp. 212-213). Urumqi, China, Xinjiang Science and Technology Press., 220 pp.(In Chinese)

Han, X., Li, C., Zhao, S., & Xu, C. Z. (2016). The complete mitochondrial genome of Cherskii's sculpin (*Cottus czerskii*) (Scorpaeniformes: Cottidae). Mitochondrial DNA Part A: DNA Mapping, Sequencing, and Analysis, 27(4), 2629-2630. https://doi.org/10.3109/19401736.2015.1041127

Harrison, R.G. (1989). Animal Mitochondrial DNA as a Genetic Marker in Population and Evolutionary Biology. Trends in Ecology and Evolution, 4(1), 6-11. https://doi.org/10.1016/0169-5347(89)90006-2

Hurst, L. D. (2002). The Ka/Ks ratio: diagnosing the form of sequence evolution. Trends in Genetics, 18(9), 486-487. https://doi.org/10.1016/S0168-9525(02)02722-1

Hwang, D.S., Byeon, H.K., & Lee, J.S. (2013). Complete mitochondrial genome of the freshwater sculpin *Cottus hangiongensis* (Scorpaeniformes, Cottidae). Mitochondrial DNA, 24(5), 490-491.https://doi.org/10.3109/19401736.2013.773987

Katoh, K., & Toh, H. (2010). Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics, 26(15), 1899-1900. https://doi.org/10.1093/bioinformatics/btq224

Kottelat, M. (2006). Fishes of Mongolia. A Check-List of the Fishes Known to Occur in Mongolia with Comments on Systematics and Nomenclature (pp. 59-60). Washington D.C., USA, The World Bank.103 pp.

Lambert, M.S., Mariam, T.T., & Susan, F.H. (2010). Primer Walking. Whitefish, USA, Betascript Publishing., 116 pp.

Li, S.Z., Dai, D.Y., Zhang, S.Y., Ma, G.Z., He, Z.W., & Gao, S.D. (1966). Notes on a Collection of Fishes from North Sinkiang, China. Acta Zoologica Sinica, 18(1), 41-56. (In Chinese)

Nielsen, R. (2005). Molecular signatures of natural selection. Annual Review of Genetics, 39(39), 197-218. https://doi.org/10.1146/annurev.genet.39.073003.112420

Nylander, J.A.A. (2004). MrModeltest v2. Program distributed by the author. https://www.researchgate.net/publication/285805344_MrModeltest_V2_Program_Distributed_by_the_Author

Perna, N.T., & Kocher, T.D. (1995). Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. Journal of Molecular Evolution, 41(3), 353-358. https://doi.org/10.1007/BF01215182

Rokas, A., Ladoukakis, E. L., & Zouros, E. (2003). Animal mitochondrial DNA recombination revisited. Trends in Ecology and Evolution, 18(8), 411-417. https://doi.org/10.1016/S0169-5347(03)00125-3

Saccone, C., Gissi, C., Lanave, C., Larizza, A., Pesole, G., & Reyes, A. (2000). Evolution of the mitochondrial genetic system: an overview. Gene, 261(1), 153-159. https://doi.org/10.1016/S0378-1119(00)00484-4

Seibert, S.A., Howell, C.Y., Hughes, M.K., & Hughes, A.L. (1995). Natural Selection on the gag, pol, and env Genes of Human Immunodeficiency Virus 1 (HIV-1). Molecuar Biology and Evolution, 12(5), 803-813. https://doi.org/10.1093/oxfordjournals.molbev.a040257

Sha, M., Lin, L.L., Li, X.J., & Huang, Y. (2013). Strategy and methods for sequencing mitochondrial genome. Chinese Journal of Applied Entomology, 5, 293-297. (In Chinese)

Sideleva, V.G. (2017). Redescription of *Cottus sibiricus* and an estimation of differences between this species and *Cottus dzungaricus* (Cottidae). Journal of Ichthyology, 57(6), 827-834. https://doi.org/10.1134/S003294521706011X

Stemshorn, K.C., Reed, F.A., Nolte, A.W., & Tautz, D. (2011). Rapid formation of distinct hybrid lineages after secondary contact of two fish species (*Cottus* sp.). Molecular Ecology, 20(7), 1475-1491. https://doi.org/10.1111/j.1365-294X.2010.04997.x

Swofford, D.L. (2002). PAUP*: Phylogenetic Analysis Using Parsimony. Sunderland, MA, USA, Sinauer Associates Inc., 210 pp.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution, 30, 2725-2729. https://doi.org/10.1093/molbev/mst197

Timmermans, M.J.T.N., Dodsworth, S., Culverwell, C.L., & Vogler, A. (2010). Why barcode? High-throughput multiplex sequencing of mitochondrial genomes for molecular systematics. Nucleic Acids Research, 38(21), e197. https://doi.org/10.1093/nar/gkq807

Wataru, I., Tsukasa, F., Ryota, I., Koichiro, Y., Yasunobu, M., Takashi, P.S.,Tetsuya, S., Kohji, M.,Hirohiko, T., Masaki, M., & Mutsumi, N. (2013). MitoFish and MitoAnnotator: A Mitochondrial Genome Database of Fish with an Accurate and Automatic Annotation Pipeline. Molecular Biology and Evolution, 30(11), 2531-3540. https://doi.org/10.1093/molbev/mst141

Yamada, K.D., Tomii, K., & Katoh, K. (2016). Application of the MAFFT sequence alignment program to large data–reexamination of the usefulness of chained guide trees. Bioinformatics, 32(21), 3246-3251. https://doi.org/10.1093/bioinformatics/btw412

Yang, J., & Huang, Y. (2015). Strategy of high-throughput sequencing technology in the mitochondrial genome sequencing. Chinese Bulletin of Life Sciences, 27, 1-6. (In Chinese)

Yang, Z., & Bielawski, J.P. (2000). Statistical methods for detecting molecular adaptation. Trends in Ecology and Evolution, 15(12), 496-503. https://doi.org/10.1016/S0169-5347(00)01994-7

Yokoyama, R., Sideleva, V.G., Shedko, S.V., & Goto, A. (2008). Broad-scale phylogeography of the Palearctic freshwater fish *Cottus poecilopus* complex (Pisces: Cottidae). Molecular Phylogenetics and Evolution, 48(3), 1244-1251. https://doi.org/10.1016/j.ympev.2008.02.002