Evidence for natural selection of immune genes from *Parachromis managuensis* by transcriptome sequencing

Huan Zhong, Hong Zhang, Zhanyang Tang, Zhongbao Guo, Jinfeng Yan, Jun Xiao, Yongju Luo and Yi Zhou

*Guangxi Key Laboratory of Aquatic Genetic Breeding and Healthy Aquaculture, Guangxi Academy of Fishery Sciences, Nanning, PR China; Guangxi Key Laboratory of Beibu Gulf Marine Biodiversity Conservation, Qinzhou University, Qinzhou, PR China; Department of Cell Biology, School of Life Sciences, Central South University, Changsha, PR China*

**ABSTRACT**

*Parachromis managuensis* is a native cichlid fish from Central America that has been recently introduced to Southern China. The present study aimed to identify the adaptive evolution of *P. managuensis* using transcriptome data and provided genetic information regarding this novel fish in China. We obtained the transcriptome sequences from a mixed cDNA library of *P. managuensis*. A total of 19,419,739 raw reads were obtained from the cDNA library. The de novo assembly generated a total of 102,977 unigenes. The genomic data of Atlantic cod (*Gadus morhua*) and Nile tilapia (*Oreochromis niloticus*) were used to identify orthologs by phylogenetic analysis. Based on the phylogeny, we detected 105 positively selected genes in *P. managuensis* from 6197 orthologous unigenes. The GO annotation revealed that the T-box protein 20 (TBX20) in the category 'Immune system process' was positively selected and four positively selected genes were found in the 'Immune system' pathway by KEGG analysis. Among them, the genes encoding for phosphoinositide 3-kinase adapter protein 1, ras-related protein Rap-1b, complement factor I and probable ATP-dependent RNA helicase DHX58 were associated with immune adaptation. These findings can provide additional information regarding the natural selection of *P. managuensis* and can be used as a representative example for future comparative transcriptome and immunological studies.

**ARTICLE HISTORY**

Received 3 January 2018
Accepted 30 August 2018

**KEYWORDS**

*Parachromis managuensis*; adaptation; positive selection; transcriptome

**Introduction**

Fish farming has been used extensively worldwide over the centuries. Currently, the productivity is persistently growing owing to the rapid development of aquaculture. However, the high-density and large-scale farming models lead to high risks of disease and low resistance to pathogens. Thus, development of disease-resistant fish varieties with high adaptation can be used to improve the production of some species [1]. The processes of breeding cause fish species to adapt to different farming environments which pose certain challenges. The adaptation of teleosts to the environment involves significant changes with regards to their immune system [2], resistance to pathogens [3] and courtship [4] traits. This represents a complicated mechanism underlying adaptation. One possible explanation is attributed to the gene loss and/or duplication in teleosts due to genome duplication, which contributes to significant behavioural complexity [5,6]. For example, the genes associated with viviparity were positively selected and indicate gene loss in platyfish (*Xiphophorus maculatus*) [7]. In order to adapt to the absence of MHC II in the genome, the Atlantic cod (*Gadus morhua*) evolved a compensatory mechanism [8]. Nevertheless, it is difficult to reveal all the evolutionary attributes from thousands of fish species. In addition, the improvement in the genetic traits of the artificial species, for example, via chromosome manipulation and hybridization, contribute substantially to the fish diversity. Fortunately, the application of high throughput sequencing data to the fish genome allows the investigation of the genetic basis that underlies the adaptation process of numerous fish species.

In aquaculture, the process of infection and its spread are affected by multiple environmental factors such as temperature, pH, salinity and ammonia...
content [2]. Meanwhile, the improved germplasm resources may enhance the population’s resistance to the development of disease [9,10]. Furthermore, the genetic variation including gene mutations facilitates speciation [11]. The identification of the key variation among species is essential in deciphering the basis of adaptive biological aspects and in understanding the genetic mechanisms for adaptation [12].

*Parachromis managuensis* is a native cichlid species from Central America. This fish species has been recently cultured in southern China and exhibits a high breeding value for ornamental and nutritional purposes [13,14]. The optimum survival temperature for *P. managuensis* is 25°C to 36°C. Under this high temperature, the proliferation of the pathogens is rapidly causing several risks to the growth of the fish. In addition, under high-density culturing conditions, the risk of infection could be higher. Cichlids represent a large number of fish species in evolutionary and genetic studies. In African lakes, the high diversity of cichlid fish has attracted particular research attention. The potential causes of the diversity can be attributed to hybridization [15], genome duplication [16] and/or geographical isolation [17]. Positive selections in fertilization-related genes [18], visual-related genes [19], resistance-related genes [20] and immune-related genes [20] have been revealed in the African cichlid. In contrast to these studies, evidence from other areas is considerably lesser in quantity. *De novo* assembly of RNA sequencing has provided a high throughput approach to study adaptation. Using this technology, several studies have been successfully investigated for the estimation of positive selection as cichlids including model and non-model fish species [20–22].

In order to identify the genes that may have facilitated adaptation to the development of disease resistance in *P. managuensis*, a transcriptome sequencing approach was performed. *P. managuensis* is a non-model fish without reference genome and transcriptome data. Thus, we first reported the *de novo* transcriptome resource of this species. By sequencing the transcriptome of *P. managuensis* and comparing it to the genomic data from Nile tilapia (*Oreochromis niloticus*) and Atlantic cod (*G. morhua*) retrieved from online databases, the orthologs from the three fishes were determined. Branch-site model analysis was employed in order to identify genes that might have undergone positive selection in *P. managuensis*. The function annotation was conducted in order to indicate the accelerated evolution of the positively identified selected genes. The transcriptome dataset and the results analysis provided new information for the annotation of a large number of expressed sequence tags and the investigation of the complex mechanisms underlying the adaptation of *P. managuensis*.

**Materials and methods**

*Fish sampling and transcriptome sequencing*

The experimental *P. managuensis* were obtained from the Genetics and Breeding Center of the Guangxi Academy of Fishery Sciences (Nanning, China). A total of 11 tissues (brain, pituitary, gill, heart, liver, spleen, kidney, muscle, intestine, testis and ovary) were sampled from four individual fishes (two males and two females, one year old). All the fishes were anesthetized by 0.5% pentoxyethanol (v/v, SIGMA) before being sacrificed. The tissues samples were collected and frozen by liquid nitrogen immediately. The remaining samples were stored at −80°C.

Total RNA was extracted according to the product manual of RNAiso Plus (Takara). The RNA purity and quantity were detected by Nanodrop 2000 (Thermo) using the OD_{260/280} value and the RNA integrity was detected by Agilent 2100 (Agilent Technologies). A cDNA library was prepared for transcriptome sequencing. The mRNA transcripts were enriched using Magnetic Beads Oligo(dT) (Life Technologies) and randomly interrupted into segments by fragmentation buffer (life technologies). The first strand cDNA was synthesized using random hexamer primers. Subsequently, deoxynucleoside triphosphates (dNTPs), DNA polymerase I and RNase H (Life Technologies) were added to synthesize the second strand cDNA. The product was purified by AMPure XP beads (Backman Coulter) and subjected to polymerase chain reaction amplification in order to generate the library for sequencing. The quantity was determined by Qubit 2.0 (Thermo) and the cDNA samples were diluted to a concentration of 1.5 ng/μL. The cDNA library was sequenced on an Illumina Hiseq 4000 platform (Illumina) with the paired-end reads method.

**Quality control, raw reads filtering and de novo transcriptome assembly**

The raw reads were determined by CASAVA 1.8 (Illumina) following filtering in order to generate clean reads. The adaptors were removed. Subsequently, the reads with N ratio of more than 10% were trimmed. Finally, the low-quality reads (the reads that had more than 50% bases with a Q_{phred} value of less than 20) were removed. The cut-off
value was $Q_{phred} = -10\log_{10}(e)$, where $e$ represents the sequencing error rate estimated by the CASAVA 1.8 software. De novo assembly was conducted using the Trinity package [23]. The parameters used for the assembly by the Trinity software were $K$-mer coverage =2 and $K=25$ for the generation of unigenes and transcripts.

**Gene annotation**

The unigenes and transcripts were annotated by seven public databases including NCBI (National Center for Biotechnology Information) non-redundant protein (Nr), NCBI nucleotide (Nt), Protein family (Pfam), eukaryotic Ortholog Groups (KOG), Swiss-prot, Kyoto Encyclopedia of Genes and Genomes (KEGG) and GO (Gene Ontology). Nr, Nt and Swiss-prot annotation were carried out by Blast 2.2.28+[24] with an e-value of 1e-5. KOG and KEGG annotations were conducted using Blast 2.2.28+ with e-values of 1e-3 and 1e-10, respectively. Pfam annotation was carried out by the HMMER 3.0 package [25]. The hmmscan e-value was 0.01. GO annotation was used by Blast2GO v2.5 and the e-value was 1e-6 [26].

The coding DNA sequence (CDS) of the unigenes was identified. The unigenes annotated by Nr and Swiss-prot were translated following CDS extraction. The remaining unigenes were identified by estscan 3.0.3 [27]. Following trimming, the identified ORFs were used for orthologous sequence investigation.

**Identification of the orthologous sequences**

In the present study, we used Nile tilapia (O. niloticus) as a related fish type of P. managuensis in order to compare the evolutionary features. Atlantic cod (G. morhua) was used as an outlier species of the group. The CDSs of O. niloticus (ftp://ftp.ensembl.org/pub/release-89/fasta/oreochromis niloticus/dna/) and G. morhua (ftp://ftp.ensembl.org/pub/release-89/fasta/gadus morhua/dna/) were obtained from the Ensembl genome browser. The best reciprocal hit (BRH) method was employed in order to identify the orthologous sequences among the three species. Blastx search was conducted reciprocally between P. managuensis and O. niloticus in order to identify a 1:1 orthologous sequence. Subsequently, the search was carried out for the P. managuensis and the G. morhua. Finally, the 1:1:1 orthologous sequences among these three species were identified.

**Detection of natural selection**

The orthologous sequences were aligned by Prank using codon alignment mode [28,29]. Following alignment, the sequences were trimmed by Gblocks [30]. The aligned sequences that were less than 200 bp were not included in the analysis in order to eliminate false-positive predictions due to short lengths. In addition, the sequences with unexpected stop codons were removed. The detection of natural selection was estimated by CODEML in PAML package using Branch-Site Model [31]. The phylogenetic tree of the three species for the CODEML analysis was used (Figure 1). The orthologous sequences with $P<0.05$ were determined as positively selected genes. The free-ratio model in CODEML was further used in order to estimate the $Ka$, $Ks$ and $Ka/Ks$ ratios of P. managuensis and O. niloticus branches with multiple test correction. The positively selected genes were assigned to GO categories and KEGG categories in order to indicate their associated functions.

**Results and discussion**

**Transcriptome sequencing and de novo assembly**

In the present study, a total of 19,419,739 raw reads were generated from the P. managuensis cDNA library. Low-quality reads were removed in order to obtain high quality clean reads. Finally, 19,050,950 reads with 2.88 Gb clean bases were produced by filtering of raw reads. Using these clean reads, de novo assembly generated 122,128 transcripts with N50 of 1172 bp and mean length of 713 bp. Moreover, the Trinity package generated 102,977 unigenes with N50 of 937 bp and mean length of 620 bp. The length distribution of transcripts (Figure 2(a)) and unigenes (Figure 2(b)) are displayed in Figure 2 and the data are indicated in Table 1. All the raw data are available in the NCBI Sequence Read Archive database (Accession number: SRR5627607).

The transcriptome sequencing profile of P. managuensis provided in this study adds to our knowledge...
of natural selection and genomic resources, which opens up possibilities for further study of cichlids. Cichlid is a large family of fish species that represents a high diversity of teleosts. To date, several fish genomes have been sequenced [8,32–34]. Due to the complexity of the teleosts genome (genome duplication, gene loss and recombination), de novo assembly of genomic sequences faces great challenges. However, the transcriptome studies are extensively conducted due to their simplicity and low cost. Thus, the present results provide a huge amount of transcriptome data about *P. managuensis*.

**Functional annotation**

The best-hit search that was conducted in seven public databases (Nr, Nt, Pfam, KOG, Swiss-prot, KEGG and GO) resulted in the annotation of 69,904 unigenes (67.88% of total unigenes) (Table 1). With the exception of the classification ‘others’, the majority of the unigenes were annotated to *O. niloticus* (27.2%), followed by *Neolamprologus brichardi* (11.5%), *Haplochromis burtoni* (10.3%), *Maylandia zebra* (7.4%) and *Pundamilia nyererei* (7.2%) based on Nr annotation (Figure 3). These top five species were from the Cichlidae family.

The annotated unigenes were assigned according to GO terms. The GO categories contain three categories including ‘Biological process’, ‘Cellular component’ and ‘Molecular function’ (Figure 4). In the ‘Biological process’ category, the GO terms ‘cellular process’, ‘metabolic process’ and ‘single-organism process’ included the most abundant unigenes (18,787, 15,451 and 15,274 unigenes, respectively). The top three GO terms in the ‘Cellular component’ category were ‘cell',

| Table 1. Summary of de novo transcriptome of *P. managuensis*. |
|---------------------|---------------------|---------------------|---------------------|---------------------|
| Assembly            |                     |                     |                     |                     |
| Total number of raw reads (counts) | 19,419,739          | Total number of clean reads (counts) | 19,000,950          |
| Total number of unigenes (counts)  | 102,977             | N50 length of unigenes (bp)  | 937                 |
| Mean length of unigenes (bp)      | 620                 | Annotation           |                     |
| Annotated in NR (counts)          | 47,006              | Annotated in NT (counts) | 63,578              |
| Annotated in Pfam (counts)        | 31,660              | Annotated in KOG (counts) | 18,627              |
| Annotated in Swiss-prot (counts)  | 37,974              | Annotated in KEGG (counts) | 25,675              |
| Annotated at least in one database (counts) | 69,904 |
‘cell part’ and ‘membrane’ which contained 10,147, 10,146 and 6952 unigenes, respectively. In the ‘metabolic process’ category, the GO terms ‘binding’, ‘catalytic activity’ and ‘transporter activity’ represented the majority of the unigenes (18,519, 12,435 and 2518, respectively).

The KEGG pathway represents experimental knowledge on the metabolism and cell functions of certain genes. In the first pathway hierarchy, the pathways ‘Organism systems’ (8244 unigenes), ‘Metabolism’ (6995 unigenes), ‘Genetic information processing’ (2570 unigenes), ‘Environmental information processing’ (5598 unigenes) and ‘Cellular processes’ (4232 unigenes) were identified. The top five pathways in the second pathway hierarchy were ‘Signal transduction’ (4036 unigenes), ‘Endocrine system’ (1854 unigenes), ‘Cellular community’ (1556 unigenes), ‘Immune system’ (1527 unigenes) and ‘Transport and catabolism’ (1331 unigenes) (Figure 5).

The enormous numbers of expressed sequence tags were initially generated by the transcriptome sequencing approach with annotation, whereas the identified 6197 orthologous unigenes from the follow-up analysis comprised a pool for additional functional and evolutionary studies in cichlid. The annotation results of *P. managuensis* indicated that the majority of the unigenes were annotated by the cichlid fish species. This outcome was expected since *P. managuensis* was included in the family of Cichlidae. The top three biological process categories that were identified by the GO annotation were ‘Cellular process’, ‘Metabolic process’ and ‘Single-organism process’. These findings were similar to previous reports on the grass carp (*Ctenopharyngodon idellus*) [35], common carp (*Cyprinus carpio*) [36] and red tilapia (*Oreochromis* spp.) [37].

### Detection of positively selected genes

Following BRH searching among the *P. managuensis*, *O. niloticus* and *G. morhua* fish species, 6197 orthologous unigenes were identified. Although 102,977 unigenes were obtained from *P. managuensis*, only 6197 orthologous unigenes were identified by BRH searching among the *P. managuensis, O. niloticus* and *G. morhua* species. These findings may be attributed to the relative far phylogenetic distance between the cichlid and the *G. morhua*. These orthologs are sufficient to provide the basis for further analysis of the evolutionary features in *P. managuensis*. For example, 7012 putative orthologs from the transcriptome of two toad-headed Agamas (genus *Phrynocephalus*) revealed the molecular evolution basis with regards to the adaptation to high elevations of reptiles [38]. Using 1216 alignments of fully overlapping sequences, the positive selection of eastern African cichlid fishes *Astatotilapia burtoni* and *Ophthalmotilapia ventralis* was investigated by comparative analyses [39].

A total of 105 positively selected genes were identified in *P. managuensis* using a branch-site model. These positively selected genes were assigned to 17 GO categories (Figure 6(a)). The T-box protein 20 (TBX20) in the ‘immune system process’ category was identified. TBX20 is essential for heart development and mutations of this gene can result in heart disease; TBX20 also regulates haematopoietic and/or lymphoid organ development [40]. Additionally, 14 genes with
the classification ‘response to stimulus’ were identified as positively selected genes (Table 2).

The KEGG pathway analysis indicated that the positively selected genes were included in 20 pathways (Figure 6(b)). Among them, the category ‘Signal transduction’ included the majority of the genes, whereas four positively selected genes were noted in the ‘Immune system’ pathway including the genes encoding for phosphoinositide-3-kinase 3 (PIK3AP1), Ras-related protein Rap-1A (RAP1A), complement factor I (CFI) and ATP-dependent RNA helicase DHX58 (DHX58) (Table 2). Phosphoinositide 3-
kinase adapter protein 1 and/or B-cell adapter for phosphoinositide 3-kinase is an Ablinteractor 1-regulated substrate of the Abl kinase which contributes to B-cell development [41]. Ras-related protein Rap-1b is a member of the RAS-like small GTP-binding protein superfamily. Ras-related protein Rap-1b is responsible for the maintenance of appropriate endothelial cell polarity and vascular lumen [42]. Complement factor I is a protein of the complement system that regulates complement activation and plays a significant role in the innate immune response [43]. The ATP-dependent RNA helicase DHX58 exerts specific functions in the innate antiviral immunity [44]. The results suggested that these immune genes were under natural selection in the *P. managuensis* species. The data further demonstrated that 14 ‘response to stimulus’ genes that were identified by GO annotation, were positively selected, which indicated some adaptation of this species via natural selection.

*P. managuensis* exhibits similar appearance, behaviour and trophic niche with other cichlid fishes. Cichlid comprises the most species-rich group among teleosts, which has rapidly evolved [45]. This renders cichlid a suitable fish model for the investigation of fish speciation [46]. Cichlid fishes from Africa have been investigated for several years [47]. However, *P. managuensis* is a species from Central America that was not included in these studies. In the present study, we found that the genes included in the ‘response to stimulus’ and ‘immune system’ categories were under natural selection. These findings are similar to reports in other cichlid fishes [47]. Further studies should examine the variations between native fish in Central America and the introduced fish in China. The adaptation of the introduced *P. managuensis* to the new environment can be studied by the investigation of the corresponding evolution attributes.

**Conclusions**

In the present study, we used a transcriptome approach and a branch-site model to evaluate the genetic evolution pattern of *P. managuensis*. A total of 105 positively selected genes were identified in *P. managuensis*. Several gene functional categories were associated with resistance to pathogen infection and the immune system. The data may add insight into the adaptation of cichlids and provide useful information for future evolutionary and immunological studies.

**Disclosure statement**

The authors declare that they have no competing interests.

**Funding**

This work was supported by the Natural Science Foundation of China under grant number 31460688 and 31672627, the Scientific Research and Technological Development Project of Guangxi under grant number GuiKeGong1598006-5-10, the Natural Science Foundation of Guangxi under grant number 2015GXNSFBA139049, the Guangxi Science and Technology Infrastructure Project under grant number 15-235-10B, Public Welfare of Institute Research Funding in Guangxi under grant number CXIF-2016-14, the Scientific Research Fund of Guangxi Education Department under grant number ZD2014139 and the Beibu...
Gulf Marine Biological Resources Development and Protection of Key Laboratory under grant number 2017ZB07.

References

[1] Fjalestad KT, Gjedrem T, Gjerde B. Genetic improvement of disease resistance in fish: an overview. Aquaculture. 1993;111(1):65–74.
[2] Bowden TJ. Modulation of the immune system of fish by their environment. Fish Shellfish Immunol. 2008;25(4):373–383.
[3] Zhu Z, Wang R, Ren L, et al. Characterization of the CCR3 and CCR9 genes in miyuki croaker and different selection pressures imposed on different domains between mammals and teleosts. Dev Comp Immunol. 2013;41(4):631–643.
[4] Candolin U, Heuschele J. Is sexual selection beneficial during adaptation to environmental change? Trends Ecol Evol. 2008;23(8):446–452.
[5] Vollf J. Genome evolution and biodiversity in teleost fish. Heredity. 2005;94(3):280–294.
[6] Taylor JS, Braasch I, Frickey T, et al. Genome duplication, a trait shared by 22,000 species of ray-finned fish. Genome Res. 2003;13(3):382–390.
[7] Schartl M, Walter RB, Shen Y, et al. The genome of the platyfish, Xiphophorus maculatus, provides insights into evolutionary adaptation and several complex traits. Nat Genet. 2013;45(5):567–572.
[8] Star B, Nederbragt AJ, Jentoft S, et al. The genome sequence of Atlantic cod reveals a unique immune system. Nature. 2011;477(7363):207–210.
[9] Knibb W. Genetic improvement of marine fish—which method for industry? Aquac Res. 2000;31(1):11–23.
[10] Zhu LY, Nie L, Zhu G, et al. Advances in research of fish immune-relevant genes: a comparative overview of innate and adaptive immunity in teleosts. Dev Comp Immunol. 2013;39(1):39–62.
[11] Schluter D. Evidence for ecological speciation and its alternative. Science. 2009;323(5915):737–741.
[12] Barrett RD, Hoekstra HE. Molecular spandrels: tests of selection pressures imposed on different domains between mammals and teleosts. Dev Comp Immunol. 2013;41(4):631–643.
[13] Liu LH, He JZ, Li NQ, et al. Complete mitochondrial sequence of a flatfish provides insights into ZW sex coloration. Proc Natl Acad Sci USA. 2003;100(24):14074–14079.
[14] Gerrard DT, Meyer A. Positive selection and gene conversion in SPP120, a fertilization-related gene, during the East African cichlid fish radiation. Mol Biol Evol. 2007;24(10):2286–2297.
[15] Sugawara T, Terai Y, Okada N. Natural selection of the rhodopsin gene during the adaptive radiation of East African Great Lakes cichlid fishes. Mol Biol Evol. 2002;19(10):1807–1811.
[16] Elmer KR, Fan S, Gunter H, et al. Rapid evolution and selection inferred from the transcriptomes of sympatric crater lake cichlid fishes. Mol Ecol. 2010;19(s1):197–211.
[17] Yang LD, Wang Y, Zhang ZL, et al. Comprehensive transcriptome analysis reveals accelerated genomic evolution in a Tibet fish, Gymnodiptychus pachycheilus. Genome Biol Evol. 2014;7(1):251–261.
[18] Wang Y, Yang LD, Wu B, et al. Transcriptome analysis of the plateau fish (Triplophysa dalaica): implications for adaptation to hypoxia in fishes. Gene. 2015;565(2):211–220.
[19] Grabherr MG, Haas BJ, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29(7):644–652.
[20] Camacho C, Coulouris G, Avagyan V, et al. BLAST+: architecture and applications. BMC Bioinf. 2009 [cited 2018 Mar 19];10(1):421. [9 p.]. DOI:10.1186/1471-2105-10-421.
[21] Mistry J, Finn RD, Eddy SR, et al. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 2013 [cited 2018 Mar 19];41(12):e121. [10 p.]. DOI:10.1093/nar/gkt263.
[22] Conesa A, Götz S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. Int. J Plant Genomics. 2008 [cited 2018 Mar 19];2008:619832. [12 p.]. DOI:10.1155/2008/619832.
[23] Iseli C, Jongeneel CV, Bucher P. ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. Proc Int Conf Intell Syst Mol Biol. 1999;99:138–148.
[24] Löytynoja A, Goldman N. An algorithm for progressive multiple alignment of sequences with insertions. Proc Natl Acad Sci USA. 2005;102(30):10557–10562.
[25] Löytynoja A, Goldman N. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. Science. 2008;320(5883):1632–1635.
[26] Zhang J, Nielsen R, Yang Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol. 2000;17(4):540–552.
[27] Wang YP, Lu Y, Zhang Y, et al. The draft genome of the grass carp (Ctenopharyngodon idellus) provides insights into its evolution and vegetarian adaptation. Nat Genet. 2015;47(6):625–631.
[28] Chen SL, Zhang GJ, Shao CW, et al. Whole-genome sequence of a flatfish provides insights into ZW sex
[34] Xu P, Zhang XF, Wang XM, et al. Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. Nat Genet. 2014;46(3):253–260.

[35] Li GX, Zhao YL, Wang J, et al. Transcriptome profiling of developing spleen tissue and discovery of immune-related genes in grass carp (*Ctenopharyngodon idella*). Fish Shellfish Immunol. 2017;60:400–410.

[36] Jiang Y, Zhang S, Xu J, et al. Comparative transcriptome analysis reveals the genetic basis of skin color variation in common carp. PLoS One. 2014 [cited 2018 Mar 19];2014;9(9):e108200. [10 p.]. DOI:10.1371/journal.pone.0108200

[37] Zhu WB, Wang LM, Dong ZJ, et al. Comparative transcriptome analysis identifies candidate genes related to skin color differentiation in red tilapia. Sci Rep. 2016 [cited 2018 Mar 19];2016;6:31347. [12 p.]. DOI: 10.1038/srep31347

[38] Joger U, Yang W, Qi Y, et al. Exploring the genetic basis of adaptation to high elevations in reptiles: a comparative transcriptome analysis of two toad-headed agamas (Genus *Phrynocephalus*). PLoS One. 2014 [cited 2018 Mar 19];2014;9(11):e112218. [8 p.]. DOI:10.1038/srep31347

[39] Baldo L, Santos ME, Salzburger W. Comparative transcriptomics of eastern African Cichlid fishes shows signs of positive selection and a large contribution of untranslated regions to genetic diversity. Genome Biol Evol. 2011;3:443–455.

[40] Cai C-L, Zhou W, Yang L, et al. T-box genes coordinate regional rates of proliferation and regional specification during cardiogenesis. Development. 2005;132(10):2475–2487.

[41] Maruoka M, Suzuki J, Kawata S, et al. Identification of B cell adaptor for PI3-kinase (BCAP) as an Abi interactor 1-regulated substrate of Abi kinases. FEBS Lett. 2005;579(14):2986–2990.

[42] Pannekoek W-J, van Dijk JJJ, Chan OYA, et al. Epac1 and PDZ-GEF cooperate in Rap1 mediated endothelial junction control. Cell Signal. 2011;23(12):2056–2064.

[43] Roversi P, Johnson S, Caesar JJ, et al. Structural basis for complement factor I control and its disease-associated sequence polymorphisms. Proc Natl Acad Sci USA. 2011;108(31):12839–12844.

[44] Huang T, Su J, Heng J, et al. Identification and expression profiling analysis of grass carp *Ctenopharyngodon idella* LGP2 cDNA. Fish Shellfish Immunol. 2010;29(2):349–355.

[45] Sturmbauer C, Husemann M, Danley PD. Explosive speciation and adaptive radiation of East African cichlid fishes. In: Frank E. Zachos, Jan Christian Habel, editors. Biodiversity hotspots. Heidelberg: Springer; 2011. p. 333–362.

[46] Kocher TD. Adaptive evolution and explosive speciation: the cichlid fish model. Nat Rev Genet. 2004;5(4):288–298.

[47] Seehausen O. African cichlid fish: a model system in adaptive radiation research. Proc Biol Sci. 2006;273(1597):1987–1998.