Meta Gene 7 (2016) 28–33

Contents lists available at ScienceDirect

Meta Gene

journal homepage: www.elsevier.com/locate/mgene

# *De novo* assembly and characterization of tissue-specific transcriptome in the endangered golden mahseer, *Tor putitora*

Ashoktaru Barat, Rohit Kumar, Chirag Goel, Atul Kumar Singh, Prabhati Kumari Sahoo *

*ICAR-Directorate of Coldwater Fisheries Research, Bhimtal, Nainital 263136, Uttarakhand, India*

ARTICLE INFO

*Article history:*
Received 31 August 2015
Accepted 7 November 2015
Available online 11 November 2015

*Keywords:*
Golden mahseer
Endangered
Organ-specific transcriptome
SSR markers

ABSTRACT

The golden mahseer (*Tor putitora*) graces most of the Himalayan Rivers of India and neighboring South Asian countries. Despite its several importance as a research model, as food, and in sport fishing, knowledge on transcriptome database is nil. Therefore, it was targeted to develop reference transcriptome databases of the species using next-generation sequencing. In the present study, 100,540,130 high-quality paired-end reads were obtained from six cDNA libraries of spleen, liver, gill, kidney, muscle, and brain with 28.4 GB data using Illumina paired-end sequencing technology. Tissue-specific transcriptomes as well as complete transcriptome assembly were analyzed for concise representation of the study. In brief, the *de novo* assembly of individual tissue resulted in an average of 31,829 (18,512–46,348) contigs per sample, while combined transcriptome comprised 77,907 unique transcript fragments (unigenes) assembled from reads of six tissues. Approximately 75,407 (96.8%) unigenes could be annotated according to their homology matches in the nr, SwisseProt, GO, or KEGG databases. Comparative analysis showed that 84% of the unigenes have significant similarity to zebra fish RefSeq proteins. Tissue-specific-dominated genes were also identified to hypothesize their localization and expression in individual tissue. In addition, 2485 simple sequence repeats (SSRs) were detected from 77,907 transcripts in the combined transcriptome of the golden mahseer. This study has generated organ-specific transcriptome profiles, which will be helpful to understand the local adaptation, genome evolution, and also future functional studies on immune system of the golden mahseer.



## 1. Introduction

The golden mahseer is one of the most important coldwater fish species. It has great importance as a research model, as food, and in sport fishing. However, its slow growth and reproductive dysfunction in captive condition is a major concern for domestication and aquaculture practice. The population in the wild is also reported to be declining due to several human and environmental factors and therefore, declared as an endangered fish (http://www.iucnredlist.org/). Due to its ecological and economic importance particularly in the Himalayan regions, several attempts were made to propagate and rehabilitate the species in rivers and lakes by artificial breeding using sexually matured fishes from wild (Pandey et al., 1998). Assessment of genetic vigor and population genetic variability was carried out using mitochondrial gene and microsatellite markers in geographically isolated populations from the Ganga, the Indus, and the Brahmaputra basins (Sati et al., 2015). To study the genetic profiles, polymorphic microsatellite markers from a partial genomic library were identified (Sahoo et al., 2013). Complete mitochondrial genome of this species has also been reported (Sati et al., 2014) for the proper phylogeny of the species in the family Cyprinidae.

The information on genetic profiles is still not enough to study the complete genome evolution, environmental challenges that species is facing, the immunology, the developmental biology, etc., of this species. The emergence of next-generation sequencing technologies has made possible to generate significant gene/genetic resources through transcriptome profiling of non-model species. The availability of annotated gene collections provide the baseline for future functional genomics applications. Hence, the major objective of this study was to create a reference set of transcriptome using multiple organs that would be helpful to understand the local adaptation, genome evolution, and also future functional studies on immune system of the golden mahseer.

## 2. Materials and method

### 2.1. Fish samples

All protocols were performed in accordance with Animal welfare Act approved by ICAR-DCFR Ethical body. Live fishes were collected from the River Kosi (Ramnagar, 29.40°N 79.12°E) using cast net and transported to wet laboratory. The fishes were acclimatized for

* Corresponding author.
*E-mail address:* psahooin@yahoo.co.in (P.K. Sahoo).

http://dx.doi.org/10.1016/j.mgene.2015.11.001

1 month with proper feeding and aeration. After acclimatization, 5 individuals (size range 50–100 cm) were sacrificed and different organs (spleen, liver, gill, kidney, muscle, and brain) were collected in RNAlater® (Ambion Inc., Austin, Texas) and stored at −80 °C.

### 2.2. Isolation and qualitative and quantitative analyses of total RNA

Total RNA was isolated from 6 tissues (spleen, liver, gill, kidney, muscle, and brain) from 5 individuals each using solution based RaFlex Total RNA isolation Kit (GeNei, Bangalore, India) as per manufacturer's recommended protocol. RNA was DNAse treated at 30 °C for 30 min to remove genomic DNA contamination. The quality of total RNA was checked on 1.2% denatured agarose gel for the presence of intact 28S and 18S rRNA bands. Total RNA was quantified using Qubit®3.0 fluorometer (Thermo Fisher Scientific, Waltham, USA).

### 2.3. Illumina MiSeq 2X150 PE library preparation

The RNA samples of each tissue from 5 individuals were pooled separately to minimize the variation among individual fish. Library preparation was started with mRNA isolation and enrichment using TruSeq RNA Library Preparation Kit v2 (Illumina, San Diego, USA) as per manufacturer's recommended protocol. The mRNA is fragmented enzymatically. The fragments were then subjected to first and second strand cDNA conversion, followed by end repair, adapter ligation, and finally ended by index PCR amplification of adaptor-ligated library. Quantitative and qualitative estimation of library was performed using HT DNA High Sensitivity Assay Kit (Caliper Life Sciences, Hopkinton, Massachusetts, USA) on Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, USA).

### 2.4. Cluster generation and sequencing run

After obtaining the Qubit concentration for the library and the mean peak size, 10pM of library was loaded into the reagent cartridge of MiSeq reagent kits v2 (Illumina, San Diego, USA) 300 cycles PE. Cluster generation was carried out by hybridization of template molecules onto the oligonucleotide-coated surface of the flow cell. Immobilized template copies were amplified by bridge amplification to generate clonal clusters. The kit reagents were used in binding of samples to complementary adapter oligos on paired-end flow cell. The adapters were designed to allow selective cleavage of the forward strands after re-synthesis of the reverse strand during sequencing. The copied reverse strand was then used to sequence from the opposite end of the fragment. The libraries were sequenced on Illumina Miseq 500 platform using 2 × 150 PE chemistry at Xcelris Genomics Ltd., Ahmedabad.

### 2.5. Sequence filtering and de novo assembly

The raw reads generated from each 6 samples were subjected to quality check using parameters like per sequence analysis, per base analysis, and overpresentation analysis using Fastx toolkit and CLC Genomics Workbench v.7.5.2 (CLC Bio, Aarhus, Denmark). Further, the reads were separately trimmed for ambiguity, low quality, and PCR duplicates. The *de novo* assembly of the clean reads of all 6 organs was carried out separately using Trinity v2.0.5 (Grabherr et al., 2011) with default parameters. The assembled transcript contigs were validated using CLC Genomics workbench v.7.5.2 by mapping high-quality reads back to the assembled transcript contigs with the settings mismatch cost of 2, an insert cost of 3, a minimum contig length of 200 bp, a similarity of 0.8, and a trimming quality score of 0.05. The high-quality data of all six samples were also concatenated into single file as a total transcriptome of *Tor putitora* using Trinity v2.0.5.

### 2.6. CDS prediction, functional annotation, and simple sequence repeat (SSRs) identification

The transcript contigs were used to identify the coding region using TransDecoder on default parameters to identify the coding sequences from assembled transcript contig of 6 organs, i.e., spleen, liver, gill, kidney, muscle, and brain separately (Supplementary File 1, CDS length distribution). The functional annotation of CDS generated from each sample and concatenated transcripts of all six organs (total transcriptome) were performed separately by aligning the contigs to non-redundant database of NCBI using BLASTx with an e-value threshold of 1e-06. Gene ontology (GO) annotation is used by Blast2Go Pro software suite v3.0. The GO sequence distribution was analyzed for all the three GO domains, i.e., biological processes, molecular functions, and cellular components. Each annotated sequence was assigned to detailed GO terms. KEGG pathway analysis was performed using KEGG automatic annotation server (KAAS) (www.genome.jp/tools/**kaas**/). All the CDS were compared against the KEGG database using BLASTx with threshold bit-score value of 60 (default). Transcripts exclusive to each organs and those common in between organs were identified using In-house Pearl Script. We also identified simple sequence repeat (SSR) from the CDS of total transcriptome (merged CDS of all six tissues) using pearl script MISA (http://pgrc.ipk-gatersleben.de/misa/). The parameters were adjusted for identification of perfect di-, tri-, tetra-, and penta-nucleotide motifs with a minimum of 8, 5, 5, and 5 repeats, respectively. Only microsatellite sequences with flanking sequences longer than 50 bp on both sides were collected for primer designing.

## 3. Results and discussion

### 3.1. Sequence and assembly

The transcriptome sequencing using 2 × 150 bp chemistry on the Illumina Miseq500 generated a total of 100,540,130 high-quality reads containing 28,705,168,638 bases across six organs (spleen, liver, gill, kidney, muscle, and brain) yielding 28.4 GB data. The raw data of each sample were deposited separately at the NCBI Sequence Read Archive (SRA) under the study accession number as shown in Table 1. The reads statistics of six organs are also shown in Table 1.

*De novo* assemblies of each library resulted in an average of 31,829 (18,512–46,348) contigs per sample. Brain tissue had the highest contig count, totaling approximately 60.3 Mbp. The lowest contig count was observed in muscle tissue; with approximately 42.4 Mbp. The longest contig length was 11,907 bp, and the contig N50 was 1206 bp. The assembly statistics of each tissue was shown in Table 1.

### 3.2. CDS prediction, functional annotation, and SSR identification

The contigs were used to identify the coding regions using TransDecoder of all six tissues. The CDS predictions were 23,832 in spleen, 26,454 in liver, 25,489 in gill, 28,937 in kidney, 25,790 in muscle, and 22,819 in brain, respectively. BLASTx hit resulted in the annotation of 22,989, 25,639, 24,581, 28,116, 24,668, and 22,151 CDS in spleen, liver, gill, kidney, muscle, and brain, respectively. The CDS length distributions among six tissues were shown in Supplementary File 1, CDS length distribution. When the CDS of all six tissues were concatenated, it was observed that 257 transcripts were between the smallest group of 200 to 300 bp, and the highest numbers of 27,624 transcripts were ≥1000 bp in length. The next highest 11,507 transcripts were between 500 and 600 bp (Fig. 1).

A *de novo* assembly was also prepared from all six tissues to create a reference transcriptome that resulted in 77,907 transcripts. Functional annotation was performed by aligning them to non-redundant database of NCBI using BLASTx with an e-value threshold of 1e-06. The BLASTx hit resulted in the annotation of 75,407 CDS and 2500 CDS of no BLAST hit. Majority of the sequences showed matches with the eukaryotes, i.e.,

**Table 1**
Read and assembly statistics of golden mahseer's transcriptome

| Description | Spleen | Liver | Gill | Kidney | Muscle | Brain |
|---|---|---|---|---|---|---|
| NCBI SRA accession no. | SRX767362 | SRX767364 | SRX768559 | SRX768561 | SRX770050 | SRX800776 |
| Sample ID | SAMN03196595 | SAMN03196596 | SAMN03196597 | SAMN03196598 | SAMN03196599 | SAMN03196600 |
| No. of high-quality reads | 18,456,821 | 15,355,597 | 20,875,681 | 17,011,668 | 14,847,540 | 13,992,823 |
| Number of contigs | 25,733 | 38,343 | 36,683 | 25,357 | 46,348 | 18,512 |
| Total transcriptome length (in base pairs) | 48,421,544 | 47,919,997 | 49,432,894 | 59,609,973 | 42,439,395 | 60,329,006 |
| Max transcript size (in base pairs) | 9129 | 9694 | 11,455 | 16,876 | 76,237 | 14,184 |
| N50 value | 1878 | 1493 | 1989 | 2342 | 894 | 3249 |

mammals, teleost fishes, and few prokaryotes. Among the teleost fishes, most sequences showed similarity match with *Danio rerio* (84%), *Oreochromis niloticus* (2%), and *Cyprinus carpio* (1%), followed by other fishes (5%). A gene ontology (GO) sequence distribution helps in specifying all the annotated nodes comprising of GO functional groups. CDS associated with similar functions are assigned to same GO functional group. The term within themselves could be assigned to one or more transcripts in the transcriptome. The GO annotations resulted a total of 190,679 GO term assignment to the annotated transcriptome comprising 69,372 (36.38%) assignments from biological process, 70,391 (36.92%) from molecular function, and 50,916 (26.7%) from cellular components ontology (Supplementary File 1, GO distribution). There was not much difference in the expression profiles in association with GO classification when compared to entire transcriptome as well as six individual tissues respectively (Supplementary File 1, GO distribution). In GO distribution, the transcripts are involved with only 22 biological process terms, 12 molecular function terms, and 8 cellular component terms. The mapped CDS fall in to six broad categories such as metabolism, genetic information processing, environmental information processing, cellular processes, oraganismal system, and diseases (Supplementary File 1, pathway distribution).

In case of individual tissues, the highest assignments of GO terms mapped were cellular and metabolic process in association with biological process, binding and catalytic activity associated with molecular function, and cell and organelle in association with cellular component process (Fig. 2). When we compared the GO distribution among six tissues, the apparent contribution of liver in metabolic process (20.46% of its total biological process) and brain in cellular process (23%) was found. Further, the involvement of muscle in response to stimuli (3.7%) suggested more exposure of muscle to external environment than other organs, particularly gills (3.0%). Two GO terms, i.e., binding activity (48.16%) and catalytic activity (38.3%), in the molecular function process were significantly overexpressed in the kidney than other organs. The highest assignments in case of cellular component category were cell (61.4%) and synapse (1.2%) in brain and organelle (26.0%) in liver (Supplementary File 1, GO distribution).

The transcripts involved in KEGG pathway were distributed under six categories, metabolism, genetic information processing, environmental information processing, cellular process, organismal system, and diseases. Tissue-specific total numbers of CDS were mapped in KEGG pathway: 9696 CDS (42.17%) in spleen; 13,002 (50.71%) in liver; 11,289 (45.9%) in gill; 13,745 (48.88%) in kidney; 11,305 (45.8%) in muscle; and 10,442 (47.1%) in brain. It was observed that liver is the major site of metabolism, with 2743 CDS involved in the mechanism and 1571 CDS involved in genetic information processing. It was also observed that lipid metabolism (425 CDS) was the major pathways under the metabolism category in liver (Supplementary File 1, KEGG annotation). In kidney, 2082 CDS were annotated to be involved in environmental information processing, 1845 CDS in cellular process, 3127 CDS in organismal system, and 2933 CDS in diseases (Supplementary File 1, KEGG annotation). The associations of 1533 CDS in signal transduction under environmental information processing pathway, 783 CDS in immune system under organismal system, and 1177 CDS in infectious diseases under diseases pathway in the kidney tissue were also significant. Comparatively, most of the transcripts were associated at least in four major KEGG pathways in the kidney tissue than other tissues.

### 3.3. Tissue-specific transcripts' profiles

The functionally annotated transcripts were considered for common and exclusive transcript study. The hit accessions of all annotated CDS from all six tissues were extracted and were analyzed as unique using in-house Perl script. Concisely, 1377 unique sequences from spleen,

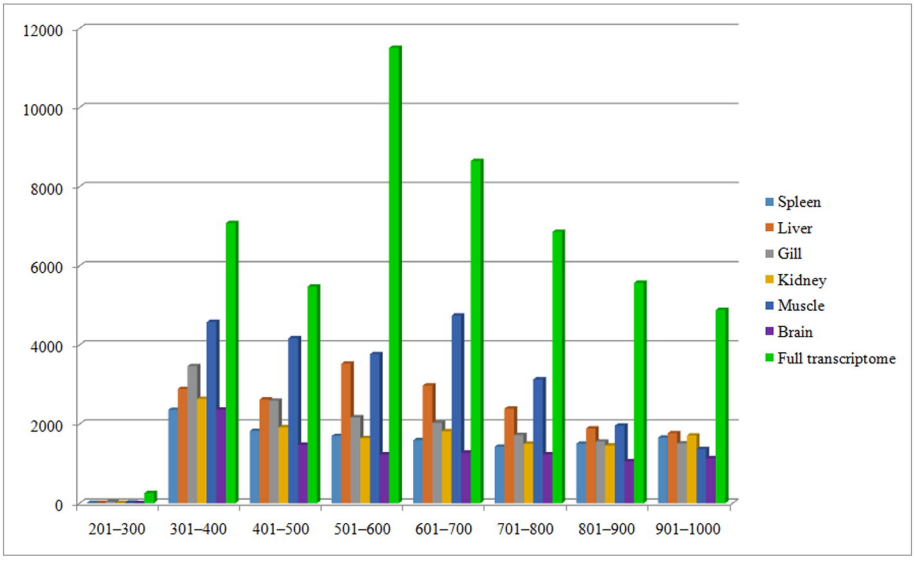


**Fig. 1.** CDS length distribution of tissue-specific as well as concatenated transcriptome of golden mahseer.

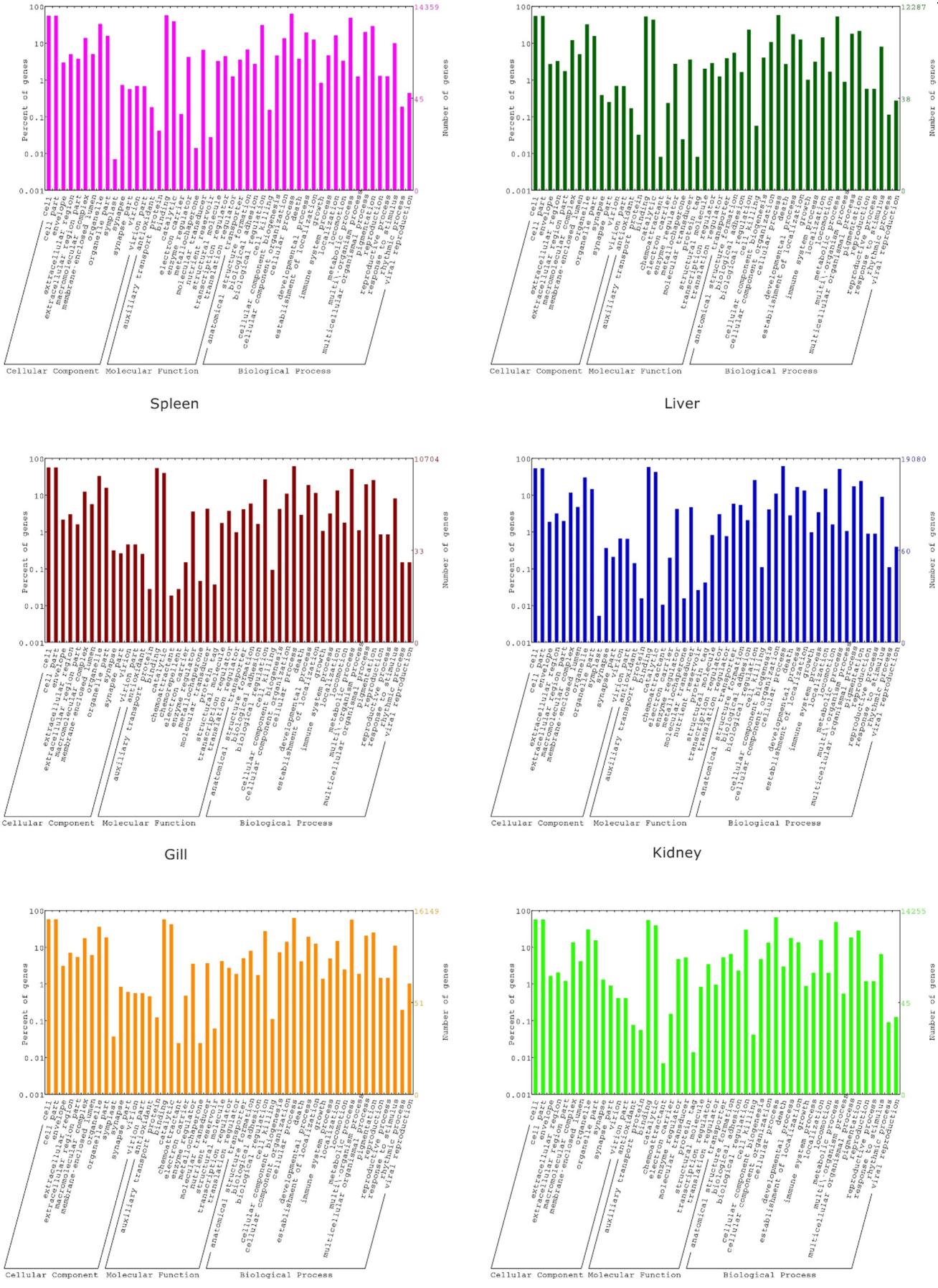


**Fig. 2.** Gene ontology classification of transcripts of six different tissues viz. spleen, liver, gill, kidney, muscle, and brain.

1175 from liver, 981 from gill, 837 from kidney, 2011 from muscle, and 412 from brain were identified (Supplementary File 2). We observed different interleukins, toll like receptors, immunoglobulins, MHC class 1 antigen, semaphorins, and plexins among most dominant transcripts (2.1%) in the spleen that are mostly involved in antigens processing and immune system process. Spleen is the organ that shows its efficiency in trapping and processing antigens for blood filtering in fishes (Zapata et al. 1996, Zapata et al., 2006). Therefore, the expression of genes related to immune system and antigen processing was evident in spleen. However, semaphorins act as axon guidance factors for neural development, but it was also observed to regulate both immune cell interactions and immune cell trafficking during physiological and pathological immune response (Takamatsu et al., 2010).

Apolipoproteins (0.17%) and molecular chaperones like *Dnajc* and *HSP 27* (0.51%) overexpressed in liver. Apolipoproteins in liver compensate the metabolic demand through lipid mobilization and transport (Huth and Place, 2013) and may be the possible reason of its overexpression. The KEGG pathway mapping of liver's transcripts also showed the association of highest number of transcripts in metabolism. The Dnaj proteins also known as HSP 40 are common heat shock proteins that assist in folding or unfolding of other proteins or macromolecule structures in cells (Ellis, 1987). They are found to be induced under several stress conditions *i.e.*, thermal stress, hypoxia and diseases (Song et al., 2014). However, it was also suggested by Roberts et al. (2010) that they play major role in immune responses. Therefore, the overexpression of these HSPs in liver might be an act of normal defense mechanism.

Out of 981 exclusive transcripts in gill tissue, transcripts like actins (0.71%) and claudin (0.81%) were highly expressed. Actin 2 and actin-related protein 10 were overexpressed among actin group of transcripts. The similar results have been obtained by Cui et al. (2014) in the transcriptome analysis of gills of *Takigugu rubripes*. Cui et al. (2014) stated that actin forms microfilaments which contribute in the processes such as intracellular transport of vesicles and organelles. Eight different isoforms of claudins were also found to be overexpressed in the present study. Krause et al. (2008) reviewed the structure and functions of claudins as a transmembrane protein of tight junctions, which enable the vital functions of epithelial organs. Cui et al. (2014) also observed six highly expressed claudin genes in gill transcriptome.

In the exclusive transcript analysis of kidney tissue, 837 transcripts were identified, out of which 21 (2.5%) transcripts were dominated as different solute carrier family genes (Supplementary File 2). These solute carrier families include zinc transporter, neurotransmitter transporter, sodium-glucose transporter, anion exchanger, acetyl CoA transporter, etc., which play a major role in renal organic anion transport. The kidney functions to filter and resorb the inorganic and organic ions, amino acids, bicarbonate, carbohydrates, lipids, and sometimes water itself (Dantzler, 1988) to minimize the ion loss. Therefore, it is evident the overexpression of solute carrier family genes in kidney to maintain the physiological balance in the aquatic environment.

Among 2011 exclusive transcripts in muscle, eleven (0.54%) transcripts were annotated to myosin genes (myosin IIIa, myosin VIIa, myosin IG, myosin heavy chain, myosin binding protein, tropomyosin alpha-1 chain, etc.) and four (0.19%) integrin genes. Fishes remain most active in their aquatic environment, and they need stronger muscle function. The myosin is the primary regulator of muscle strength and contractility (Rooij et al., 2009). Garcia de la serrana et al. (2012) also observed some isoforms of myosin heavy chain, myosin light chain, and myosin binding protein in the muscle transcriptome of the Gilthead seabream (*Sparus aurata*).

The transcriptome of brain tissue comprised 412 exclusive transcripts: two synapsin IIa, three neural adhesion molecules, two potassium voltage-gated channel, and six protocadherin, and its isoforms were dominant among the exclusive transcripts. Synapsins are involved in the regulation of neurotransmitter release at synapses (Ferreira and Rapoport, 2002). The neural adhesion molecule observed in the present study, normally expresses on the surface of neurons and has a role in cell–cell adhesion, neurite outgrowth, synaptic plasticity, etc. (Senkov et al., 2006). The potassium voltage gate channel Zerg24 (*kcnh2*) plays important role in ion channel and found to be expressed in the brain of zebra fish also (Leong et al., 2010). Frank and Kemler (2002) reviewed the role of protocadherins in tissue morphogenesis, the formation of neuronal circuits, the modulation of synaptic transmission, and the generation of specific synaptic connection. Chen et al. (2008) observed that the transcripts involved in protection and repair of neural tissue were highly represented in brain tissue transcriptome of *D. mawsoni*.

### 3.4. Identification of simple sequence repeats

A total of 2485 microsatellites were identified from 77,907 unique sequences of merged transcriptome using perl script MISA (http://pgrc.ipk-gatersleben.de/misa/), and primers were designed using Primer3 v.2.3.6. Out of these 2485 sequences, 360 contained di-nucleotide, 2104 tri-nucleotide, 19 tetra-nucleotide, and 2 penta-nucleotide repeats. There are about 32 microsatellite containing transcripts for which the primer could be designed successfully (Supplementary File 1, SSR markers). These microsatellites markers in the transcriptome of *Tor putitora* are linked with protein coding genes and would be useful for genetic linkage map.

## 4. Conclusion

As far as we are aware, this is the first non-model fish species of Himalayan Rivers, in which the transcriptome profile of different organs and a reference data bases of full transcriptome is studied. We have generated the transcriptomes of six organs (23,832, 26,454, 25,489, 28,937, 25,790, and 22,819 CDS in spleen, liver, gill, kidney, muscle, and brain, respectively) and a concatenated full transcriptome containing 77,907 contigs. We also identified organ-specific transcripts and 2485 CDS-associated microsatellites. The information generated in the present study using GO annotation, KEGG pathway, comparative study on organ-specific genes, and gene-associated microsatellites would be helpful for evolutionary studies, mitigation of environmental, and climate challenges and rehabilitation of this species in the Himalayan water bodies. These data are also to be used as reference transcriptome to study other coldwater fishes.

## Conflict of interest

The authors report no conflict of interest on content of the manuscript.

## Acknowledgments

The authors are thankful to the Indian Council of Agricultural Research (ICAR) and the Department of Biotechnology (DBT), Government of India, New Delhi, for laboratory facilities and financial support to carry out this research activity. The authors are also grateful to District Forest Officer, Ramnagar, Government of Uttarakhand, and Mr. Vijay K Singh, Technical Officer, DCFR, Bhimtal, for sampling of fishes under study.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.mgene.2015.11.001.

## References


Chen, Z., Cheng, C.-H.C., Zhang, J., Cao, L., Chen, L., Zhou, L., Jin, Y., Ye, H., Deng, C., Dai, Z., Xu, Q., Hu, P., Sun, S., Shen, Y., Chen, L., 2008. Transcriptomic and genomic evolution under constant cold in Antarctic notothenioid fish. PNAS 105 (35), 12944–12949.

Cui, J., Wang, H., Liu, S., Qiu, X., Jiang, Z., Wang, X., 2014. Transcriptome analysis of the gill of *Takifugu rubripes* using Illumina sequencing for discovery of SNPs. Comp. Biochem. Physiol. Part D Genomics Proteomics 10, 44–51.

Dantzler, W.H., 1988. Comparative Physiology of the Vertebrate Kidney. Springer-Verlag, Berlin.

Ellis, J., 1987. Proteins as molecular chaperones. Nature 328, 378–379.

Ferreira, A., Rapoport, M., 2002. The synapsins: beyond the regulation of neurotransmitter release. CMLS, cell. Mol. Life Sci. 59, 589–595.

Frank, M., Kemler, R., 2002. Protocadherin. Curr. Opin. Cell Biol. 14, 557–562.

Garcia De la Serrana Castillo, D., Estevez, A., Andree, K., Johnston, I.A., 2012. Fast skeletal muscle transcriptome of the gilthead sea bream (*Sparus aurata*) determined by next generation sequencing. BMC Genomics 13, 181.

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

Huth, T.J., Place, S.P., 2013. De novo assembly and characterization of tissue specific transcriptomes in the emerald notothen, *Trematomus bernacchii*. BMC Genomics 14, 805.

Krause, G., Winkler, L., Mueller, S.L., Haseloff, R.F., Piontek, J., Blasig, I.E., 2008. Structure and function of claudins. Biochim. Biophys. Acta 1778, 631–645.

Leong, I.U.S., Skinner, J.R., Shelling, A.N., Love, D.R., 2010. Identification and expression analysis of *kcnh2* genes in the zebrafish. Biochem. Biophys. Res. Commun. 396, 817–824.

Pandey, A.K., Patiyal, R.S., Upadhyay, J.C., Tyagi, M., Mahanta, P.C., 1998. Induced spawning of the endangered golden mahseer, *Tor putitora*, with ovaprim at the state fish farm near Dehradun. Indian J. Fish. 45 (4), 457–459.

Roberts, R.J., Agius, C., Saliba, C., Bossier, P., Sung, Y.Y., 2010. Heat shock proteins (chaperones) in fish and shellfish and their potential role in relation to fish health: a review. J. Fish Dis. 33, 789–801.

Rooij, E.-v., Quiat, D., Johnson, B.A., Sutherland, L.B., Qi, X., Richardson, J.A., Kelm Jr., R.J., Olson, E.N., 2009. A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. Dev. Cell 17, 662–667.

Sahoo, P.K., Kumar, R., Sati, J., Matura, R., Barat, A., 2013. Isolation and characterization of novel microsatellite markers in endangered mahseer, *Tor putitora* (Hamilton 1822) (family: Cyprinidae, Pisces). Conserv. Genet. Resour. 5 (3), 759–761.

Sati, J., Goel, C., Kumar, R., Ali, S., Patiyal, R.S., Singh, V.K., Sahoo, P.K., Barat, A., 2014. Complete mitochondrial genome organization of *Tor putitora*. Mitochondrial DNA 25 (4), 278–279.

Sati, J., Kumar, R., Sahoo, P.K., Patiyal, R.S., Ali, S., Barat, A., 2015. Genetic characterization of golden mahseer (*Tor putitora*) populations using mitochondrial DNA markers. Mitochondrial DNA 26 (1), 68–74.

Senkov, O., Sun, M., Weinhold, B., Gerardy-Schahn, R., Schachner, M., Dityatev, A., 2006. Polysialylated neural cell adhesion molecule is involved in induction of long-term potentiation and memory acquisition and consolidation in a fear-conditioning paradigm. J. Neurol. Sci. 26 (42), 10888–109898.

Song, L., Zhang, J., Li, C., Yao, J., Jiang, C., Li, Y., Liu, S., Liu, Z., 2014. Genome-wide identification of Hsp40 genes in channel catfish and their regulated expression after bacterial infection. PLoS One 9 (12), e115752.

Takamatsu, H., Takegahara, N., Nakagawa, Y., Tomura, M., Taniguchi, M., Friedel, R.H., Rayburn, H., Tessier-Lavigne, M., Yoshida, Y., Okuno, T., et al., 2010. Semaphorins guide the entry of dendritic cells into the lymphatics by activating myosin II. Nat. Immunol. 11, 594–600.

Zapata, A., Diez, B., Cejalvo, T., Gutiérrez-de Frías, C., Cortés, A., 2006. Ontogeny of the immune system of fish. Fish Shellfish Immunol. 20 (2), 126–136.

Zapata, A.G., Chiba, A., Varas, A., 1996. Cells and tissues of the immune system of fish. In: Iwama, G., Nakanishi, T. (Eds.), The Fish Immune SystemOrganism. Pathogen and Environment. Academic Press, London, pp. 1–62.