Gene 524 (2013) 1–14

Contents lists available at SciVerse ScienceDirect

Gene

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

Identification of reproduction-related genes and SSR-markers through expressed sequence tags analysis of a monsoon breeding carp rohu, *Labeo rohita* (Hamilton)

Dinesh K. Sahu [a], Soumya P. Panda [a], Sujata Panda [a], Paramananda Das [a], Prem K. Meher [a], Rupenangshu K. Hazra [b], Eric Peatman [c], Zhanjiang J. Liu [c], Ambekar E. Eknath [a], Samiran Nandi [a,*]

[a] Central Institute of Freshwater Aquaculture, Bhubaneswar, Orissa 75 1002, India
[b] Regional Medical Research Centre, Bhubaneswar, Orissa 75 1023, India
[c] Fish Molecular Genetics and Biotechnology Laboratory, Aquatic Genomics Unit, Department of Fisheries and Allied Aquacultures, Auburn University, Auburn, AL 36849, USA

ARTICLE INFO

Article history:
Accepted 25 March 2013
Available online 9 April 2013

Keywords:
Labeo rohita
Expressed sequence tags
Normalized cDNA libraries
Reproduction
EST-SSRs

ABSTRACT

Labeo rohita (Ham.) also called rohu is the most important freshwater aquaculture species on the Indian sub continent. Monsoon dependent breeding restricts its seed production beyond season indicating a strong genetic control about which very limited information is available. Additionally, few genomic resources are publicly available for this species. Here we sought to identify reproduction-relevant genes from normalized cDNA libraries of the brain–pituitary–gonad–liver (BPGL-axis) tissues of adult *L. rohita* collected during *post preparatory* phase. 6161 random clones sequenced (Sanger-based) from these libraries produced 4642 (75.34%) high-quality sequences. They were assembled into 3631 (78.22%) unique sequences composed of 709 contigs and 2922 singletons. A total of 182 unique sequences were found to be associated with reproduction-related genes, mainly under the GO term categories of reproduction, neuro-peptide hormone activity, hormone and receptor binding, receptor activity, signal transduction, embryonic development, cell–cell signaling, cell death and anti-apoptosis process. Several important reproduction-related genes reported here for the first time in *L. rohita* are zona pellucida sperm-binding protein 3, aquaporin-12, spermine oxidase, sperm associated antigen 7, testis expressed 261, progesterone receptor membrane component, Neuropeptide Y and Pro-opiomelanocortin. Quantitative RT-PCR-based analyses of 8 known and 8 unknown transcripts during *preparatory* and *post-spawning* phase showed increased expression level of most of the transcripts during *preparatory* phase (except Neuropeptide Y) in comparison to *post-spawning* phase indicating possible roles in initiation of gonad maturation. Expression of unknown transcripts was also found in prolific breeder common carp and tilapia, but levels of expression were much higher in seasonal breeder rohu. 3631 unique sequences contained 236 (6.49%) putative microsatellites with the AG (28.16%) repeat as the most frequent motif. Twenty loci showed polymorphism in 36 unrelated individuals with allele frequency ranging from 2 to 7 per locus. The observed heterozygosity ranged from 0.096 to 0.774 whereas the expected heterozygosity ranged from 0.109 to 0.801. Identification of 182 important reproduction-related genes and expression pattern of 16 transcripts in *preparatory* and *post-spawning* phase along with 20 polymorphic EST-SSRs should be highly useful for the future reproductive molecular studies and selection program in *Labeo rohita*.

Abbreviations: ESTs, Expressed sequence tags; cDNA, Complementary DNA; CIFA, Central Institute of Freshwater Aquaculture; DSN, Duplex-specific nuclease; BLAST, Basic Local Alignment Search Tool; NR, Non-redundant; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; KAAS, KEGG Automatic Annotation Server; BBH, Bi-directional Best Hit; EC, Enzyme commission; ZFIN, Zebrafish Information Network; GIP, Genetic information processing; EIP, Environmental information processing; ZP-3, Zona pellucida sperm-binding protein 3; GtHα, Gonadotropin alpha subunit; AQP-12, aquaporin-12; SPO, Spermine oxidase; SAA-7, Sperm associated antigen 7; TE-261, Testis expressed 261; PRMC-1, Progesterone receptor membrane component 1; POMC, Pro-opiomelanocortin; NP-Y, Neuropeptide Y; UTRs, Un-translated Regions; ORFs, Open reading frame; SSRs, Simple sequence repeat; HWE, Hardy–Weinberg Equilibrium; Ho, Observed heterozygosity; He, Expected heterozygosity; NA, Number of allele; T, Temperature; BPGL, Brain–Pituitary–Gonad–Liver axis; IGFBP, Insulin like growth factor binding protein; G, Gram; °C, Degree centigrade; mg, Milligram; ng, Nanogram; µl, Microliter; pmol, Picomol; µM, Micromolar; mM, Millimolar; *d*NTP, Deoxy ribose nucleotide Phosphate; mRNA, Messenger ribonucleic acid.

* Corresponding author. Tel.: +91 674 2465446, +91 674 2465421x209; fax: +91 674 2465407.
E-mail address: eurekhain@yahoo.co.in (S. Nandi).

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.03.111

1. Introduction

The Indian major carp *Labeo rohita* (Ham.), commonly known as rohu is the leading candidate species for freshwater aquaculture in the whole sub-continent of South-East Asia including India, Bangladesh, Myanmar, Pakistan, Sri-lanka, Nepal and Thailand. It is a popular table fish because of its unique taste and several other attributes leading to highest consumer preference among carps. Culture has almost taken the shape of an industry in some of these countries such as India and Bangladesh, which require sustained supply of quality seed throughout the year. Although previous research on culture and breeding aspects is available (Routray et al., 2007) several issues related to its gonad maturation, spawning and seed production still remain to be solved. One of the major problems is that it is purely a monsoon breeder (Natarajan and Jhingran, 1963; Quasim and Qayyum, 1962) and cannot be bred in confined pond water without hormonal induction (Bhattacharya, 1999; Chaudhuri and Alikunhi, 1957) thus restricting seed production beyond monsoon season. While these problems are of least concern in other prolific breeders like common carp (*Cyprinus carpio*) and tilapia (*Oreochromis niloticus*) which mature and breed throughout the year under similar environmental conditions, propagation of Indian major carps is a much greater challenge. So addressing these problems requires in depth knowledge about regulation of gonad maturation, spawning and seed production, which is the outcome of both gene (G) and environment (E) interactions. With changing climate patterns and irregular monsoon, seed production has become totally unpredictable, which made it imperative to gain complete control on gonad maturation and breeding for sustained seed supply.

Attempts made to meet the growing demand for carp seed have largely been by environmental manipulations via multiple induced breeding (Gupta et al., 1995), providing improved brood stock diet (Nandi et al., 2001), advancement of gonad maturation and offseason breeding through photo-thermal manipulation (Sarkar et al., 2010), but only a few broods (<5%) showed this type of precocious maturation during brood rearing. These results indicated a strong genetic control for the seasonal nature of reproduction about which very limited information is available so far. Although genetic improvement programs have been undertaken on growth (Gjerde et al., 2002) and disease resistance (Sahoo et al., 2011) few genetic studies have focused on reproduction aspects. The problem has been exacerbated by the lack of genomic resources for this species in public genomics databases. The lack of genome sequence information often limits gene discovery in non-model species. Expressed sequence tags (ESTs) today represent a powerful and efficient tool for rapid identification of the genes that are preferentially expressed in certain tissues or cell types (Adams et al., 1991) and are reported to be helpful for post-transcriptomic large-scale functional genomics particularly to gain new insights into reproductive molecular biology (Cerdà et al., 2008a, 2008b). Genomic resources (i.e. ESTs and type 1 markers) in commercially important fish species are useful for many purposes such as stock identification, stock enhancement, genome mapping, marker assisted breeding, genetic management, and preservation of genetic diversity (Tassanakajon et al., 1997) as well as functional genomics (Gui and Zhu, 2012; Liu, 2007).

Genomic resources are currently available for other commercially important fishes, including rainbow trout (*Oncorhynchus mykiss*) (Gohin et al., 2010; Von-Schalburg et al., 2005), coho salmon (*Oncorhynchus kisutch*) (Luckenbach et al., 2008), tilapia (*Oreochromis mossambicus*) (Chu et al., 2006), Atlantic halibut (*Hippoglossus hippoglossus*) (Mommens et al., 2010), senegalese sole (*Solea senegalensis*) (Cerdà et al., 2008a,b), Atlantic salmon (*Salmo salar*) (Leong et al., 2010) and cod (*Gadus morhua*) (Goetz et al., 2006). The relationship between the transcriptome and physiological indicators of reproduction have been studied for some commercially important fish species e.g. rainbow trout; Hook et al., 2011, coho salmon; Luckenbach et al., 2008 etc. but no similar information is available about reproduction for *L. rohita*.

Therefore, our primary objective here is to develop and identify transcripts of reproduction-related genes, verify their association with reproduction by transcript expression pattern and discover microsatellites within these transcripts derived from reproduction-associated tissues of *L. rohita* which may be utilized further in the future for studying reproductive issues in this species.

2. Materials and methods

2.1. Ethics Statement

This study was approved by the ethical committee of the Central Institute of Freshwater Aquaculture, Bhubaneswar, Orissa, India.

2.2. Animals and tissue collection

Adult males and females of *L. rohita* fishes (800–1200 g) during May–June, (*post preparatory*) were collected from Central Institute of Freshwater Aquaculture (CIFA) farm ponds (Lat.20°1′06″–20°11′45″N, Long.80°50′52″–85°51′35″E). The fishes were euthanized with MS-222 at 300 mg/L before dissection. Liver, brain, pituitary, ovary and testis samples were collected from a minimum of five fishes each, quickly frozen in liquid nitrogen and stored at −80 °C, until used for RNA extraction.

2.3. cDNA synthesis, normalization and sequencing

Total RNA was isolated from 50 to 100 mg of different tissue samples, following the Guanidium Thiocyanate method (Chomczynski and Sacchi, 1987) using the Trizol-reagent (Invitrogen, Carlsbad, CA) according to manufacturer's instructions. RNA was freed from genomic DNA contamination with DNase (NEB, Ipswich, MA) treatment, and the integrity was checked in a 1% denaturing gel. Those RNA samples showing clear separation of the 28S and 18S bands in the gel were taken for further analysis. The quality and quantity of the RNA in each preparation were checked by UV-spectrophotometry (Varian Cary 50 Bio) and RNA samples showing A260/A280 absorption ratios greater than 1.8 were accepted for further work.

Normalized cDNA libraries were constructed for rohu liver, brain, pituitary, ovary, and testis from the pooled RNA for each tissue. First strand cDNA was synthesized with 2 μg total RNA from each sample at 42 °C using MMLV-based Mint reverse transcriptase (Schmidt and Muller, 1999) as per the MINT protocol (Evrogen, Moscow, Russia). The second strand was synthesized using Advantage 2 polymerase mix (Clonetech, Mountain View, CA). Double stranded cDNA products were purified using the QIAquick PCR Purification Kit (QIAGEN) and approximately 1000 ng purified cDNA product was aliquotted and ethanol precipitated for subsequent normalization. The normalization was carried out following the Trimmer direct protocol of cDNA normalization kit (Evrogen, Moscow, Russia) using Duplex-Specific Nuclease (DSN) enzyme (Zhulidov et al., 2004). The re-hybridization was performed by incubation at 68 °C for 7 h for brain, 6 h for liver, ovary and testis respectively, and 5 h for pituitary cDNA. Both normalized and non-normalized cDNAs were amplified for same numbers of cycles using advantage 2 polymerase mix and normalization efficiency was verified by electrophoresis in 1.5% agarose gel. PCR products generated from the normalized cDNA of all selected tissues were cloned in TOPO-TA cloning Vector System (Invitrogen, Carlsbad, CA) and transformed into chemically competent TOP-10 strain of *Escherichia coli*. The transformed colonies (ampicillin resistant) were picked up for different tissues (i.e. 2500 for liver and brain each; 1500 for pituitary, ovary and testis each, respectively) and were preserved in 15% glycerol at −80 °C until use. Plasmid DNA was prepared from the overnight culture (3 ml) by alkaline lysis using the vacuum manifold platform (Eppendorf, Hauppauge, NY) according to the manufacturer's protocol. Plasmid DNA concentrations were determined

using a UV-spectrophotometer (Varian-Cary 50 Bio) and samples were stored at −20 °C until sequencing. Sequencing reactions were carried out in 10 μl reaction volumes using 75 ng of plasmid DNA templates and M13 (-40) forward universal primer along with the ABI Big dye terminator v3.1 cycle sequencing kit (ABI, Foster City, CA) according to the manufacturer's protocol (Sanger et al., 1977). The reaction products were purified by EDTA–ethanol precipitation and analyzed on an ABI 3130xl Genetic Analysis System.

2.4. Sequence analysis and EST processing

The sequence analyses were carried out for all the clones, except those sequences shorter than 140 bp which were not included in the analysis. The EGassembler bioinformatics pipeline (Masoudi-Nejad et al., 2006) was used with default analysis parameters for cleaning and removing low quality sequence stretches, followed by the detection and removal of repetitive elements from sequences. Vector sequences were trimmed using VecScreen of NCBI (http://www.ncbi.nlm.nih.gov/VecScreen/UniVec.html) before assembly and searching for the homologies against the database. The sequences were analyzed for contigs and singletons by cap3 software (Huang and Madan, 1999) using 90% minimum match and 80 bp minimum overlap length as parameters. The homologies were searched using the Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) of NCBI against the non-redundant (nr) and EST-other nucleotide database from NCBI through the Blast2go program (http://www.blast2go.org/) of gene ontology (GO).

2.5. EST functional assignment

Sequences were functionally characterized by comparison against Gene Ontology (GO) databases. Putative functions of the unique sequences were extracted and assigned with GO hierarchical terms of their homologous genes from the UniProtKB/SwissProt and the non-redundant (nr) protein database in Blast2go program (Conesa et al., 2005). GO terms were assigned to the clusters in the three main GO categories: molecular function, biological process and cellular component at level 2. Each cluster was further assigned to one or more sub-categories based on all the GO terms assigned to it. The final annotation file was produced after gene-ID mapping, GO term assignment, annotation augmentation and generic GO-Slim process. Pathway analyses of unique sequences were carried out based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa et al., 2002) using the online (http://www.genome.jp/tools/kaas/) KEGG Automatic Annotation Server (KAAS) (Moriya et al., 2007) by using Bi-directional Best Hit (BBH) method. Enzyme commission (EC) numbers were obtained and used to putatively map protein sequence to a specific biochemical pathway.

2.6. Identification of gene orthologues involved in reproduction

To identify the genes and transcripts in rohu that play important roles in reproductive processes, the ESTs associated with GO terms under reproduction (GO:0000003), neuro-peptide hormone activity (GO:0005184), receptor binding (GO:0005102), receptor activity (GO:0004872), signal transduction (GO:0007165), embryonic development or egg hatching (GO:0009792), cell–cell signaling (GO:0007267), cell death (GO:0008219) and anti-apoptosis (GO:0006916) were selected.

2.7. Putative ORF sequences searching

Unidentified sequences (no match found in BLASTX and BLASTN) among unique sequences were analyzed by star-orf software (http://web.mit.edu/star/orf/runapp.html) with the parameter of 80 bp minimal ORF length to search for a putative ORF. Only those sequences clearly containing the start codon, coding sequence, stop codon and poly (A) tail were considered a full-length cDNA, and were searched in all the possible frames against the protein database using BLASTP tool of NCBI.

2.8. Expression analysis of selected gene orthologues using quantitative real-time PCR during preparatory and post-spawning phases

Relative expression level of 16 transcripts (8 each of known and unknown genes with putative ORFs) were measured by real-time PCR in brain, liver, pituitary, ovary and testis tissues collected from 15 individuals during initiation of gonad maturation (*preparatory*) and resting phase (*post-spawning*) each, taking β-actin as a reference gene (GenBank accession no. EU184877). Similarly, brain, liver, pituitary and ovary from common carp (*C. carpio*) and brain, liver, and ovary tissues from tilapia (*O. niloticus*) were also collected during *preparatory* phase for comparison with rohu. RNA was extracted for each tissue as described earlier, pooled into three sets and first strand cDNA was synthesized by annealing 2 μg of RNA with 0.5 μg of oligo $d(T)_{18}$ primer in a total reaction volume of 20 μl using M-MLV reverse transcriptase (Finnzymes, Vantaa, Finland) and incubated at 37 °C for 1 h. The reaction was stopped by heating at 95 °C for 5 min. Validation of transcript specific primers (Supplementary Table 3) was checked by normal PCR and band intensities for different tissues were observed in agarose gel electrophoresis with β-actin as control (data not shown). The real-time PCR amplifications were carried out using a Light Cycler® 480 (Roche, Germany) in a final volume of 20 μl containing 0.5 μl cDNA, 10 pmol of each primer and 10 μl SYBR Green master mix (Roche, Germany) together with a negative control with no template. The PCR was initiated with an activation step at 95 °C for 5 min, followed by 40 cycles of: 15 s at 95 °C, 15 s at the Tm specific for the primer pairs used, and 35 s at 72 °C with a single fluorescence measurement. After the amplification phase, a melting curve cycle was set at 95 °C for 5 s, 67 °C for 1 min with acquisitions 5 per °C from 97 to 65 °C and a continuous measurement to confirm later about the amplification of a single product. Real time PCR was repeated twice for each tissue for each sample with three replicates. The crossing point, *Cp* values were acquired for both the target and reference gene using software version LCS480 1.5.0.39 of Light Cycler® 480 (Roche, Germany). The relative level of each transcript in different tissue was calculated by normalization of the value with the corresponding reference and compared among them using *Cp* values for brain cDNA as positive calibrator (Pfaffl et al., 2002). Comparison of relative expression level of each transcript between the two reproductive phases in individual tissue as well as between the two species was analyzed in REST 2009 software and the whisker-box plots were extracted with 2000 time iterations (http://www.REST.de.com).

2.9. Identification of microsatellite-containing ESTs and analysis of microsatellite polymorphism

The sequences of individual clones were examined for the presence of repeat motifs using the repeat finder (http://www.imtech.res.in/raghava/srf/) software keeping the parameters as: at least 5 repeats for di-, and four repeats each for tri, tetra, penta and hexa-nucleotides. Primers were designed at the flanking regions of the repeat sequences using Primer 3 (v.0.4.0) software (http://frodo.wi.mit.edu/). Genomic DNA was isolated from 36 unrelated rohu individuals collected from three different geographically isolated rivers (12 each) in India (i.e. Mahanadi, Krishna and Kaveri). PCR was performed using 25 ng genomic DNA, 2.5 pmol of each forward and reverse primer, 200 μM of each *d*NTP, 1.5 mM $MgCl_2$ and 0.25 units *Taq* DNA polymerase (Bangalore Genei). Amplification was carried out with a touchdown PCR profile on a GeneAmp 9700 thermal cycler (Applied Biosystem) with the following temperature setting for all primers: Initial denaturation of 94 °C for 5 min (1 cycle) followed by 45 s at 94 °C (denaturation), 1.30 min annealing at required temperature and 2 min extension at 72 °C for each cycle. Annealing temperature in touch down procedure

Table 1
Basic characteristics of the five cDNA libraries used in this study.

	Liver	Brain	Pituitary	Ovary	Testis	Assembled data from all five tissues
No. of clones sequenced	1598	1935	1097	945	586	6161
High quality sequences	1363	1461	740	518	560	4642
Average length	500	500	500	500	500	500
Contigs	169	413	50	27	13	709
Singletons	944	636	583	443	465	2922
Unique sequences	1113	1049	633	470	488	3631
Annotated sequences (E-value < 10^{-6})	478	338	76	85	167	978
Un-annotated sequences (Only BLASTN E-value < 10^{-6})	446	506	241	204	183	1696
Unidentified sequences	189	205	316	181	138	957

Tissues	NCBI GenBank accession numbers of rohu ESTs
Liver:	JK729643–JK729652, JK714437–JK715554, GR977101–GR977067 and GR958186–GR957987.
Brain:	JK757806–JK757949, GR426801–GR426902, GR463794–GR463893, GR881193–GR881290, GR977102–GR977152, HO758290–HO758779 and HO762029–HO762504.
Pituitary:	JK671301–JK672040.
Ovary:	JK627102–JK627585 and JK649715–JK649748.
Testis:	JK729653–JK730212.

ranged from 60 °C to 55 °C with successive decrement of 1 °C. Two cycles of amplification were set for each annealing temperature (60–56 °C) but 25 rounds of amplification were performed at the final annealing temperature (55 °C) making the total number of cycles 35. PCR products were then dried in a DNA Plus (vacuum concentrator) and analyzed on a 6% denaturing PAGE gel along with standard marker, ΦX 174/*Hae*III. For studying microsatellite polymorphism, silver stained gels were visualized using Gel scanner and fragment size was analyzed in Quantity One™ software (BioRad). Since the microsatellites are co-dominant markers, allele frequencies were estimated by direct count either from the stained gel or from scanned document. Observed heterozygosity (H_O), expected heterozygosity (H_e), tests for Hardy–Weinberg Equilibrium and pair wise linkage dis-equilibrium were calculated by Fisher's exact test using GDA software (Lewis, and Zaykin, 2001).

3. Results

3.1. Normalized cDNA libraries and EST sequence analysis

6161 random clones were sequenced of which 1363 clones of liver, 1461 clones from brain, 740 clones of pituitary, 518 clones of ovary and 560 clones from testis, totaling 4642 sequences (75.34%) were retained after trimming vector, adaptor and bad sequences. Cleansed sequences were submitted to the dbEST database at NCBI and accession numbers were obtained. Each cDNA library had minimum average insert size of 500 bp, ranging from 140 to 1300 bp. Clustering and assembly produced 709 contigs (19.52%) and 2922 singletons (80.47%) together which constituted 3631 unique sequences (78.22%) (Table 1). The length of contigs ranged between 140 bp to 1700 bp. The numbers of ESTs in a contig ranged from a minimum of 2 to a maximum of 83 sequences (Fig. 1). Tissue participation in contigs formation showed that, single tissue contigs were obtained maximally from brain, whereas double tissue contigs in brain and liver, and only eight contigs were made of ESTs from five different tissues (Supplementary Table 1). BLAST search analysis in databases showed that 80.45% of sequences from brain, 83.01% of sequences from liver, and 71.72% of sequences from the testis cDNA libraries had significant matches, whereas 49.93% of sequences from pituitary and 38.52% of ovary cDNA library sequences did not match with any sequences in public databases. 978 (26.93%) and 1696 (46.70%) out of the total 3631 unique sequences had significant BLASTX and BLASTN matches with a cut-off e-value more than 10^{-6} with the nr database and were called known and un-annotated sequences, respectively. Finally, the remaining 957 (26.35%) sequences had no match with the nr database and were called un-identified sequences (Table 1). Statistical distribution showed that the majority of BLASTX hit sequences of *L. rohita* ESTs produced similarities with the cyprinid species i.e. *Danio rerio* in ZFIN databases, followed by *C. carpio*, *Ctenopharyngodon idella*, *Carassius auratus* and a significant number of sequences were also similar with other fish species such as *O. niloticus*, *Ictalurus punctatus* and *S. salar*.

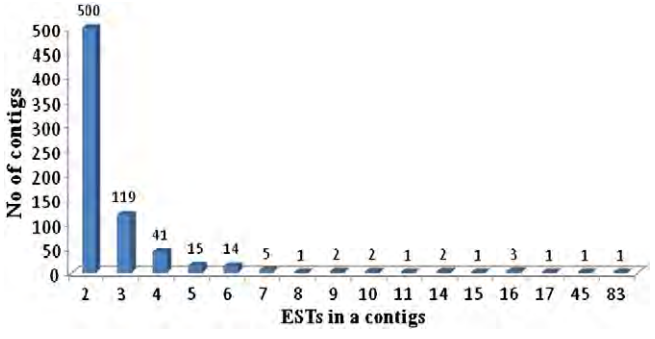


Fig. 1. Distribution of 4642 ESTs in 709 contigs: Arrangement of 4642 transcripts from brain, liver pituitary, ovary and testis in 709 contigs.

3.2. Gene ontology annotation and KEGG pathway analysis

Gene ontology (GO) annotation of 3631 unique sequences of *L. rohita* showed the highest number of annotations from the biological process category 1120 (41.51%), followed by cellular component 847 (31.39%) and molecular function category 731 (27.09%), making a total of 2698 GO terms (Fig. 2). Among the biological process category, 304 genes and 278 genes were related to metabolic and cellular processes respectively; under the cellular component category, 368 genes were from cell, and 232 genes corresponded to cytoplasm whereas under the molecular function category, the majority of genes (320) were involved in the binding process. Individual tissue level analysis showed that a large number of genes from brain (29.93%), pituitary (25.30%) and testis (28.07%) under the biological process category, were involved in cellular process, whereas in liver (30.21%) and ovary (25.30%) a greater number of genes were responsible for the metabolic process (Fig. 3). KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis was performed on all unique sequences of *L. rohita* as alternative approach for functional

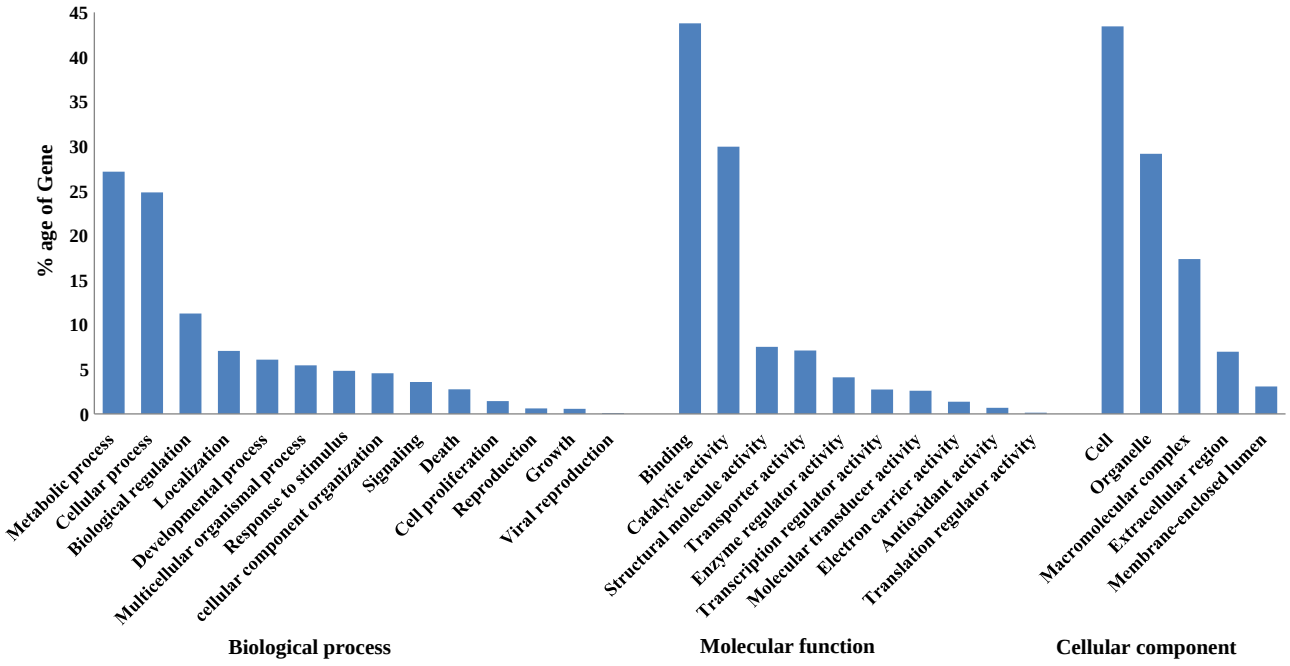


Fig. 2. Percentages of annotated *Labeo rohita* sequences assigned with GO terms according to level 2 categories: GO-terms were processed by Blast2Go and categorized at level 2 under three main categories. Each of the three GO categories is presented including (left to right): biological process, molecular function and cellular component.

categorization and annotation. Several KEGG pathways were represented by more than 30 unique sequences. A total of 873 unique sequences were assigned with 624 enzyme codes (KO numbers) or Enzyme Commission (EC) numbers (Table 2). Among these KEGG annotated sequences, 264 (30.24%) were classified under metabolism, 185 (21.19%) under genetic information processing (GIP), 73 (18.14%) under environmental information processing (EIP), 100 (11.45%) under cellular processes and 251 (28.75%) sequences under organismal systems.

3.3. Identification of reproduction-relevant genes and putative ORFs

A total of 182 reproductive genes were identified (Supplementary Table 2), among which 47 were mainly associated with reproduction-related proteins, 5 related to neuro-peptide hormone activity, 22 belonging to hormone and receptor binding related proteins, 11 receptor activity related proteins, 18 signal transduction related proteins, 43 embryonic development or egg hatching related proteins, 4 related to cell–cell signaling, and 32 each related to cell death and

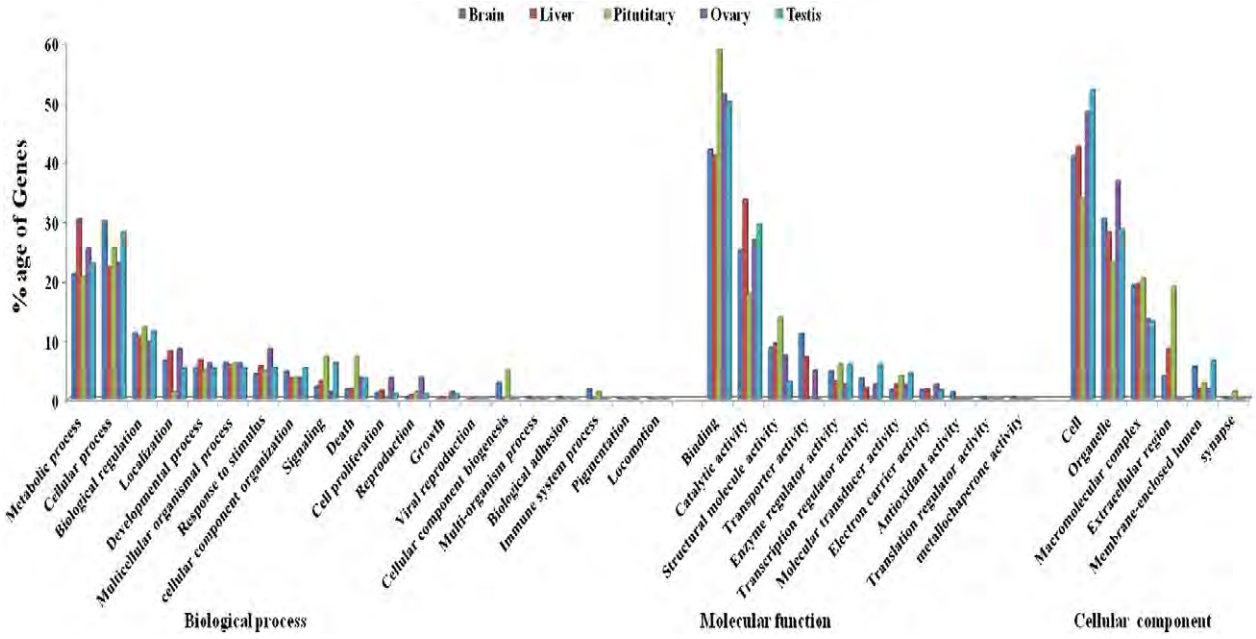


Fig. 3. Comparison of the GO categories of genes tissue wise: Comparison analysis of GO-terms processed by Blast2Go and categorized at level 2 under three main categories (cellular component, molecular function and biological process) between brain, liver, pituitary, ovary and testis.

Table 2
Details of enzymes found in KEGG pathway database from rohu libraries.

KEGG categories represented	Number of KO	Unique sequences
Metabolism	**174**	**264**
Carbohydrate metabolism	40	64
Energy metabolism	38	56
Lipid metabolism	18	27
Amino acid metabolism	21	21
Nucleotide metabolism	19	26
Metabolism of cofactors and vitamins	16	17
Glycan biosynthesis and metabolism	07	20
Metabolism of other amino acids	06	13
Xenobiotics biodegradation and metabolism	06	17
Biosynthesis of secondary metabolites	02	02
Metabolism of terpenoids and polyketides	01	01
Genetic information processing	**142**	**185**
Transcription	23	28
Translation	67	94
Folding, sorting and degradation	42	53
Replication and repair	10	10
Environmental information processing	**51**	**73**
Membrane transport	01	01
Signal transduction	32	44
Signaling molecules and interaction	18	28
Cellular processes	**78**	**100**
Transport and catabolism	29	34
Cell motility	08	09
Cell growth and death	21	30
Cell communication	20	27
Organismal systems	**179**	**251**
Immune system	51	65
Endocrine system	29	33
Circulatory system	19	29
Digestive system	31	51
Excretory system	05	09
Nervous system	34	50
Sensory system	01	02
Development	07	07
Environmental adaptation	02	05
Total	624	873

anti-apoptosis process. Among the important gene orthologues, the reproduction-related genes found and reported for the first time in *L. rohita* are zona pellucida sperm-binding protein 3 (ZP-3), gonadotropin alpha subunit (GtHα), stathmin 1 oncoprotein 18, aquaporin-12 (AQP-12), vitellogenin isomers (1, 2, & 3), vitelline membrane outer layer protein 1 homolog, pre-albumin, spermine oxidase (SPO), sperm associated antigen 7 (SAA-7), testis expressed 261(TE-261), isotocin-i precursor, melanin-concentrating hormone, progesterone receptor membrane component 1 (PRMC-1), pro-opiomelanocortin (POMC), Neuropeptide Y (NP-Y) and clusterin.

Out of the 957 unidentified sequences, 150 (15.67%) sequences were found with a putative ORF with a minimum of 80 to a maximum of 182 amino acid length, and the remaining 84.32% of sequences did not contain any ORF. These ORF-less ESTs were likely ESTs sequenced within the un-translated regions (UTRs) or within intron-retaining cDNAs. Out of 150 putative ORFs, 20 sequences started with methionine and finished with a stop codon.

3.4. Expression analysis of reproduction-relevant transcripts during preparatory and post-spawning phases

Real-time RT-PCR was performed for 16 unigenes, including 8 known function category genes ZP-3, SAA-7, NP-Y, PRMC-1, APQ-12, SPO, POMC and TE-261; and 8 unknown transcripts with putative ORFs (Rh21T, Rh1231L, Rh1160O, Rh1069P, Rh1399P, Rh1930B, Rh619B and Rh1313O). Results showed clear differences in the level of expression among different tissues as well as in the same tissue between the two different phases of reproduction (i.e. *preparatory* and *post-spawning* phase) in rohu. Comparison of relative transcript level for the known genes in brain with other tissues during *preparatory* phase (Fig. 4), revealed that the expression ratio of ZP-3 was significantly higher in testis ($p < 0.042$) and ovary ($p < 0.001$) and lower in liver ($p < 0.05$). Similarly, SAA-7 level was also higher both in testis (by a factor of 274) and ovary (by a factor of 11) compared to brain, thus was significantly different in testis ($p < 0.013$), but not in ovary. Expression level of NP-Y was similar in testis and pituitary with brain but significantly lower in ovary ($p < 0.001$) and liver ($p < 0.077$). PRMC-1 level was significantly higher in testis ($p < 0.001$), ovary ($p < 0.046$) and liver ($p < 0.001$) where as APQ-12 was significantly higher in liver ($p < 0.001$), testis ($p < 0.001$) and pituitary ($p < 0.001$) as compared to brain. In contrast, SPO level did not differ among the tissues except reduction in liver ($p < 0.049$). POMC level was significantly higher (by a factor of 642) in pituitary ($p < 0.001$) and lower in ovary ($p < 0.001$) while TE-261 expression was significantly higher in testis ($p < 0.001$), ovary ($p < 0.047$) and liver ($p < 0.001$).

Relative expression analysis in the same tissue for comparison between *preparatory* and *post-spawning* phase (Supplementary Fig. 1) showed that expression ratio of all known genes including ZP-3($p < 0.001$), SAA-7($p < 0.021$), NP-Y($p < 0.001$), PRMC-1($p < 0.045$), APQ-12($p < 0.001$), SPO($p < 0.001$), POMC($p < 0.001$) and TE-261($p < 0.001$) respectively were up regulated in testes, while only SAA-7 was up-regulated ($p < 0.034$) in ovary and APQ-12 was up-regulated ($p < 0.001$) in pituitary during *preparatory* phase as compared to respective tissue levels in *post-spawning* phase. On the other hand SAA-7, NP-Y, PRMC-1, APQ-12, POMC, and TE-261 in brain; ZP-3, PRMC-1, APQ-12, SPO and POMC in liver; NP-Y, SAA-7, PRMC-1, APQ-12, SPO, POMC and TE-261 in ovary and SAA-7, PRMC-1 in pituitary, respectively was down regulated in *preparatory* phase and the expression levels were statistically different from the corresponding tissue levels during the *post-spawning* phase. Though expression level of NP-Y was observed up regulated in pituitary particularly during the *post-spawning* phase, the difference was not statistically significant over the *preparatory* phase.

Again, during comparison of relative transcript level for the unknown transcripts in brain with other tissues in the preparatory phase (Fig. 5), no statistical variation could be found for Rh21T level among different tissues. Rh1231L level was comparatively higher in ovary (by a factor of 21) but the difference was not significant ($p > 0.947$) over brain. Expression ratio of Rh1160O and Rh1069P were significantly higher only in ovary ($p < 0.027$ and $p < 0.001$) as compared to brain but were not different in other tissues. Rh1399P level was significantly higher in ovary ($p < 0.046$) and pituitary ($p < 0.001$) than in brain while Rh1930B was found to be high in ovary (by a factor of 18) and liver (by a factor of 4) but was not statistically different. Expression levels of both Rh619B and Rh1313O transcripts were significantly higher in ovary ($p < 0.001$ and $p < 0.016$) as well as in pituitary ($p < 0.001$ and $p < 0.031$), respectively.

Expression ratio analysis (Supplementary Fig. 2), showed that Rh1231L, Rh1399P and Rh21T were up-regulated ($p < 0.001$, 0.001 and 0.037) in testes; while only Rh1069P was up-regulated ($p < 0.031$) and Rh1231L, Rh1930B, Rh1313O and Rh21T were down-regulated ($p < 0.001$) in ovary, during the *preparatory* phase and the levels were different from the *post-spawning* phase. No statistical difference could be observed for the transcript Rh1160O between the two phases. Almost all the transcripts (except Rh1313O) were down-regulated in brain during the preparatory phase and the levels were statistically different from the *post-spawning* phase. Expression levels of majority of the transcripts in pituitary (except Rh21T) and in liver (except Rh1930B and Rh21T) were not different between the *preparatory* and *post-spawning* phases. Rh21T was down-regulated ($p < 0.043$) in pituitary and Rh1930B and Rh21T

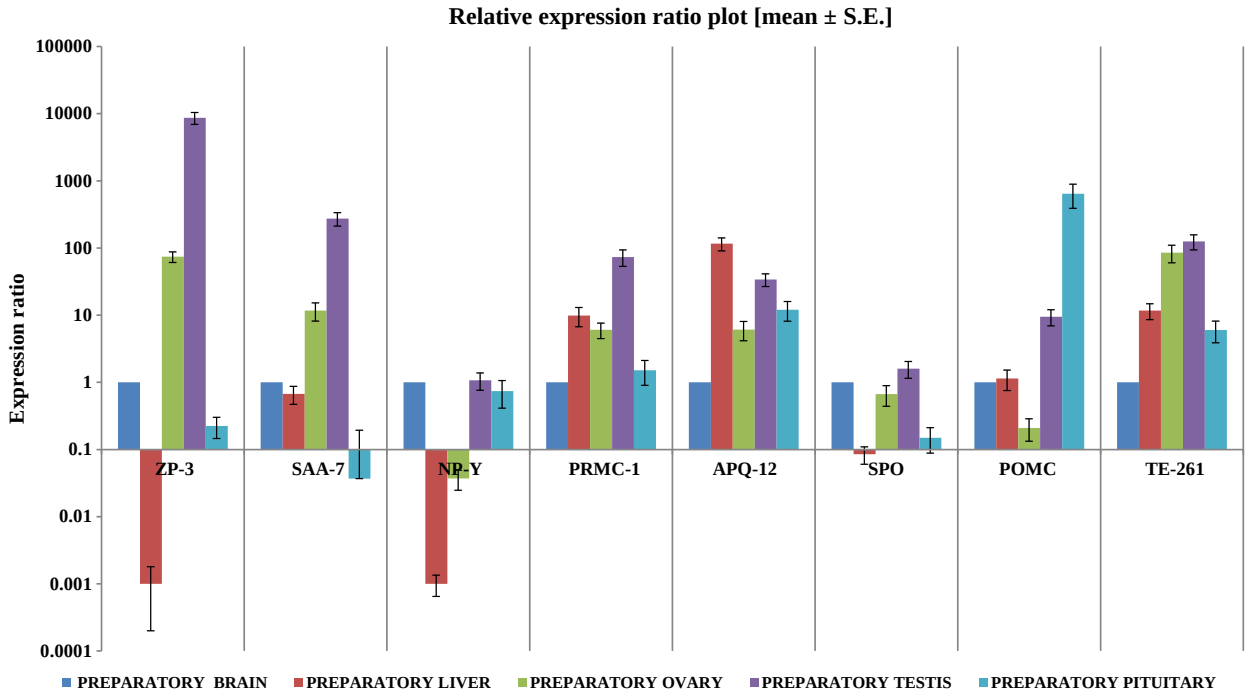


Fig. 4. Comparative analysis of relative-expression levels of 8 known transcripts during the *preparatory* phase among different tissues of *Labeo rohita* with beta actin as reference and brain as positive calibrator (expression ratio is plotted in Log scale).

were down-regulated ($p < 0.001$) in liver during the *preparatory* phase and were different from the *post-spawning* phase.

Expression study of these unknown transcripts in other two prolific breeders i.e. common carp and tilapia in similar tissues (except in tilapia pituitary) confirms that they also possess similar sequences. Expression levels were compared between common carp and rohu (Fig. 6A) as well as between tilapia and rohu (Fig. 6B) while all in the *preparatory* phase. The results showed that transcript levels of Rh1231L, Rh11600, Rh1069P, Rh619B and 1930B in rohu ovary ($p < 0.001$); Rh21T and Rh13130 in rohu pituitary ($p < 0.001$), Rh11600 and Rh619B in rohu liver ($p < 0.001$) and Rh11600 ($p < 0.001$) and Rh21T ($p < 0.003$) in rohu brain respectively, were higher as compared to respective tissues in common carp. On the other hand Rh1069P and Rh619B levels were found more ($p < 0.001$) in common carp brain than in rohu. Comparison with tilapia (Fig. 6B), showed that all unknown transcript levels were higher ($p < 0.001$) in rohu

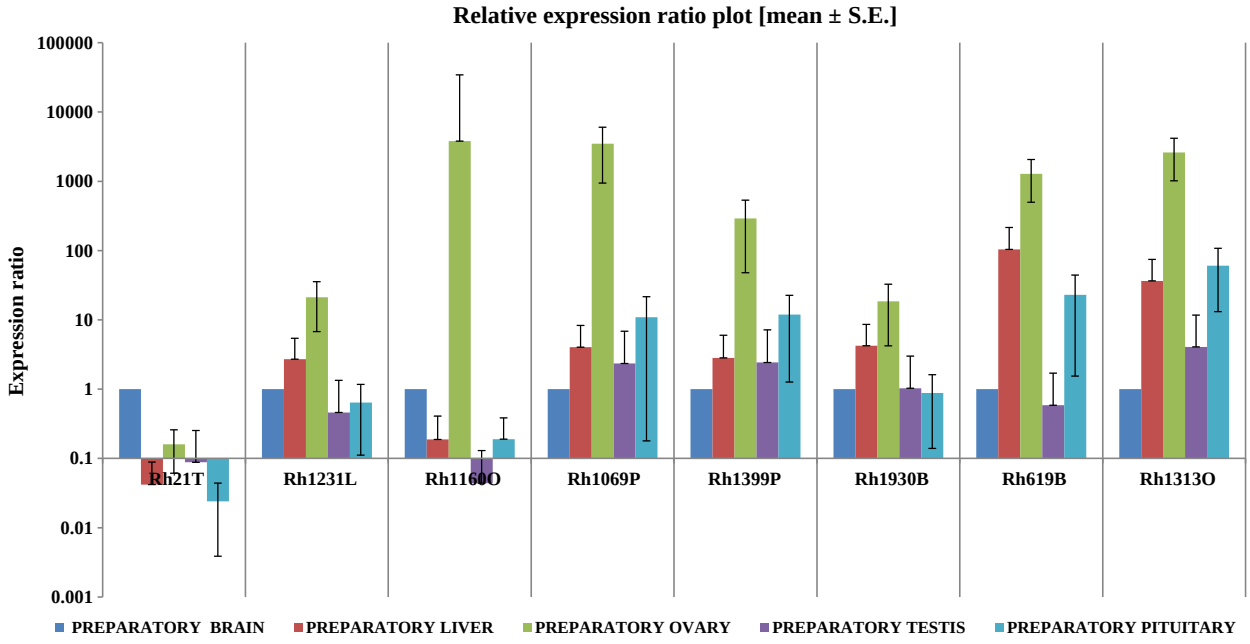
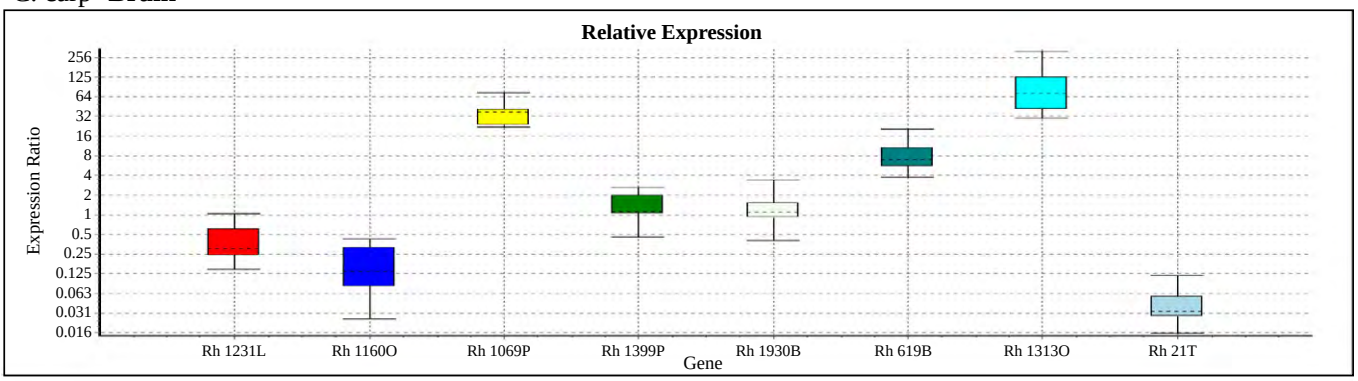


Fig. 5. Comparative analysis of relative-expression levels of 8 un-known transcripts during the *preparatory* phase among different tissues of *Labeo rohita* with beta actin as reference and brain as positive calibrator (expression ratio is plotted in Log scale).

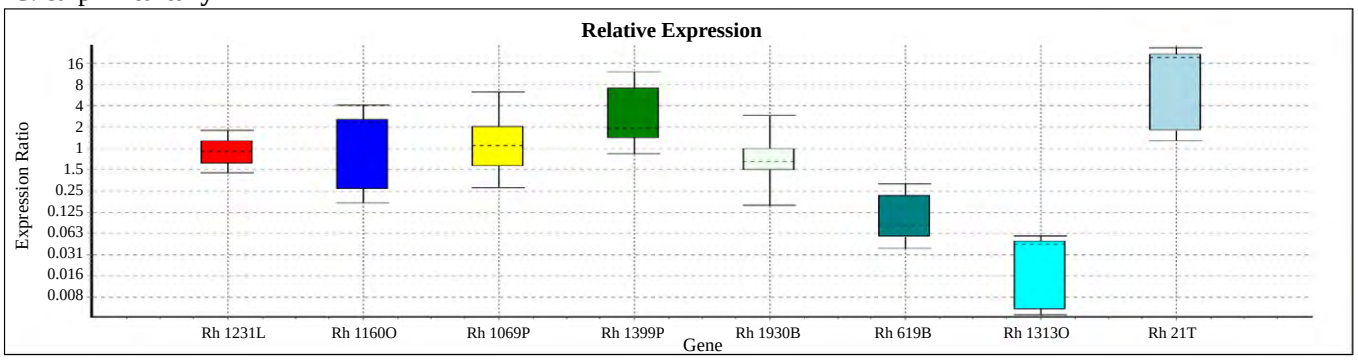
ovary, and Rh1231L was higher ($p < 0.001$) in rohu liver, whereas transcript Rh1160O ($p < 0.001$) and Rh 21T ($p < 0.001$) were higher in tilapia brain than in rohu. In BLASTP analysis, one of the highly expressed putative ORF Rh1160O was partially matched with cytochrome c oxidase subunit III (amino acid position 31–77, Identities < 21%, Positives < 45%, GENE ID: 13080099) with common carp and fibronectin type III domain-containing protein 4 like (amino acid position 4–33, Identities < 35%, Positives < 71%, GENE ID: 100695625) with tilapia.

A

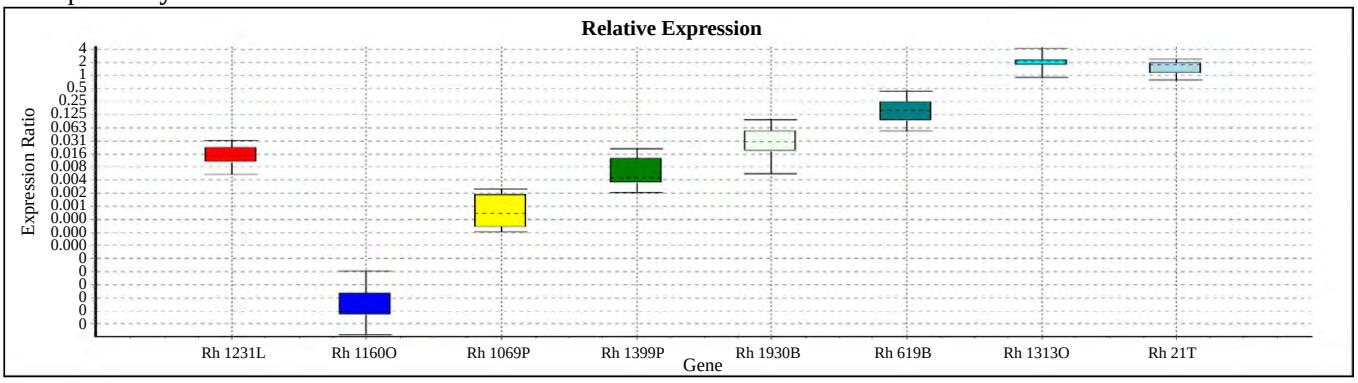
C. carp **Brain**



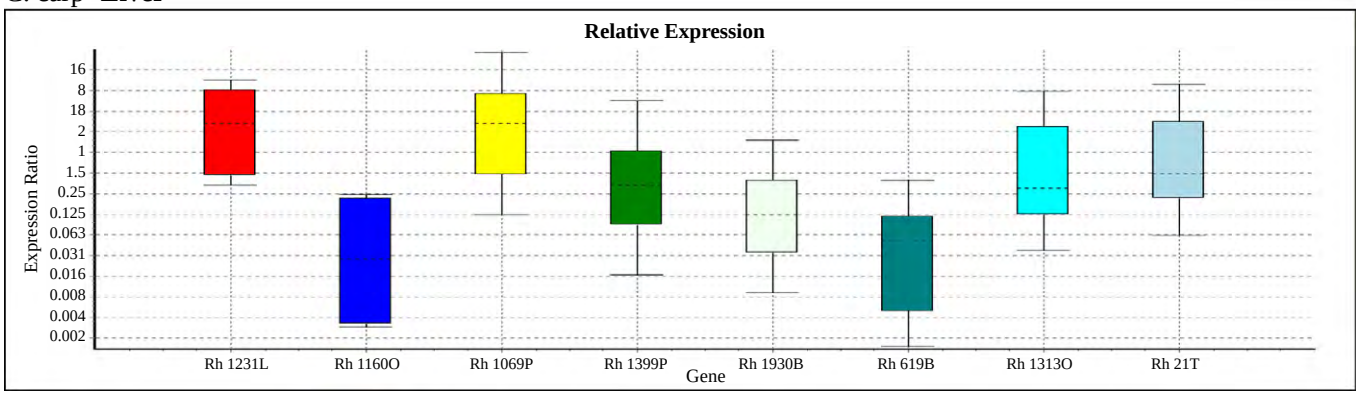
C. carp **Pituitary**



C. carp **Ovary**



C. carp **Liver**



B

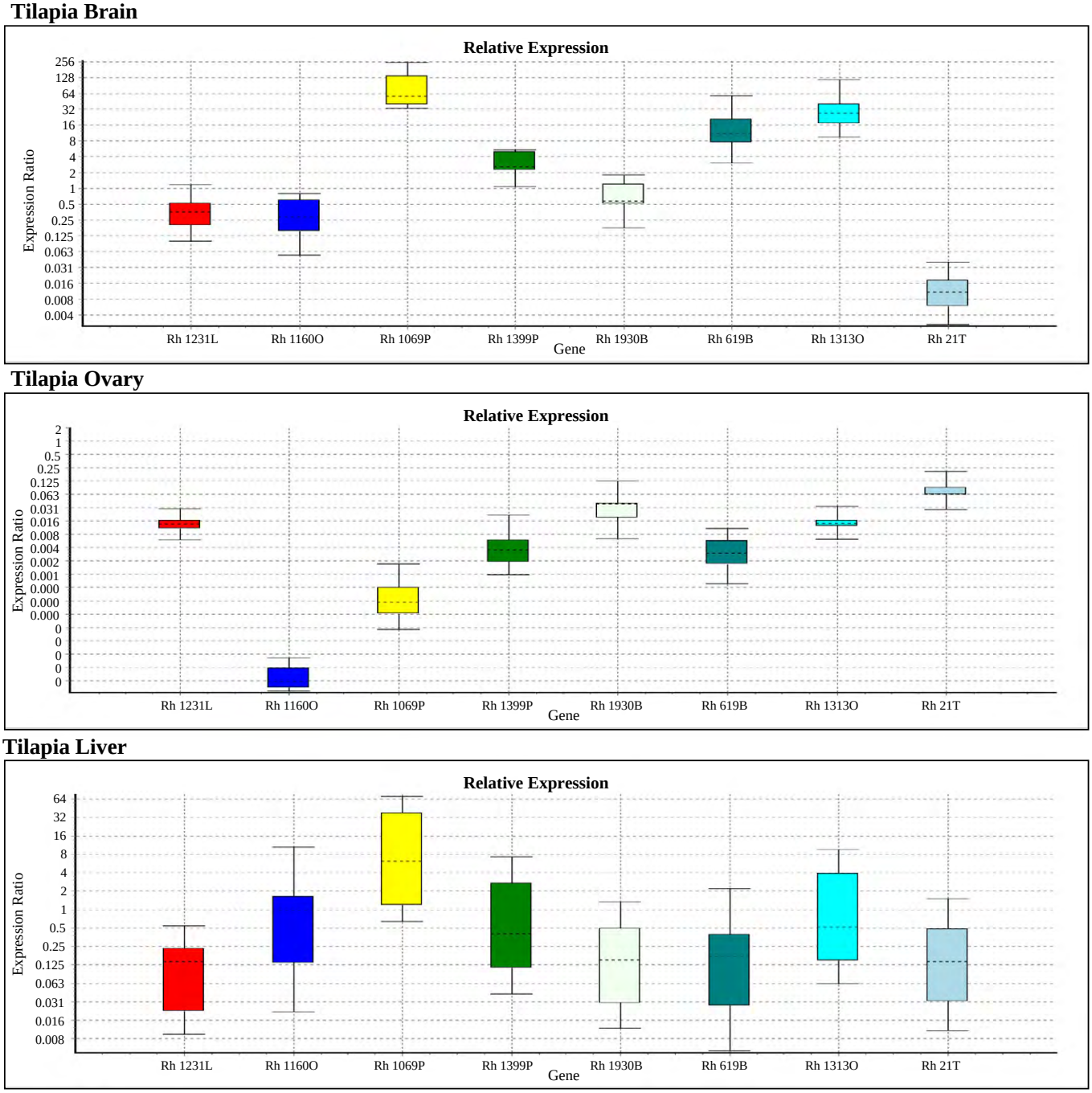


Fig. 6. Comparison of tissue expression ratio of 8 un-known transcripts between two species during *preparatory* phase with beta actin as reference (whisker-box plots). a) Rohu with common carp in brain, pituitary, ovary and liver tissues. b) Rohu with tilapia in brain, ovary and liver tissues.

3.5. Analysis of EST-SSRs, polymorphism and cross species amplification

236 (6.49%) microsatellite-containing sequences were found in 3631 unique rohu sequences of which dinucleotides (154), trinucleotides (65), tetra-nucleotides (14) and penta-nucleotides (03) were the major motifs, respectively. AG repeats were the most frequent motif (28.16%), followed by TG (26.76%) repeats (Fig. 7). Out of these, 129 EST-SSRs contained sufficient flanking sequences of good quality and were chosen for polymerase chain reaction (PCR) analysis in *Labeo rohita*. 52 of the 129 ESTs were annotated sequences (with known function), 73 sequences un-annotated and two un-identified sequences (Supplementary Table 4). Screening among reproduction-related ESTs (182) for microsatellites showed 10 sequences, 3 had di- and 7 had tri repeats, respectively (Table 3).

Out of 129 EST-SSRs 87 (67%) primer pairs successfully amplified scorable products and 20 loci showed polymorphism in 36 unrelated individuals of *L. rohita* with allele number ranging from 2 to 7 (Table 4). Out of 20 polymorphic transcripts, 11 transcripts are un-annotated. The observed heterozygosity ranged from 0.096 to 0.774 whereas the expected heterozygosity ranged from 0.109 to 0.801. Fourteen (PD3, PD11, PD15, PD16, PD19, PD26, PD27, PD42, PD48, PD49, PD86, PD87, PD98 and PD99) of the 20 loci deviated from HWE

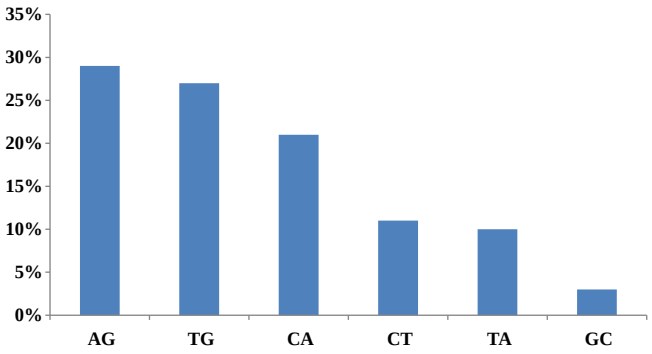


Fig. 7. Distribution of di-nucleotide motifs in microsatellites.

($p < 0.05$). However no linkage disequilibrium was observed among the microsatellite loci. As far as cross-species amplification was concerned, 16 loci were amplified in *Catla catla*, 15 in *Labeo gonius*, 14 each in *Labeo fimbriatus* and *Cirrhinus mrigala*, 13 in *Labeo bata* and 10 in *Labeo kalbasu* (Table 5). The best cross-species amplification was observed in *Catla catla*.

4. Discussion

The reproductive cycle of Indian major carps are reported to be divided into four phases, *preparatory* period (February–April), *pre-spawning* or *post-preparatory* period (May–June), *spawning* period (July–August), and *post-spawning* period (September–January) (Sundararaj and Vasal, 1976) and at each phase gonads show a discrete change (Chakraborti and Bhattacharya, 1984). However ovarian recrudescence, responsiveness and stimulatory effect to both steroidogenic and gametogenic functions of the gonad normally occurs during the *pre-spawning/post preparatory* period. The tissues were collected for EST library preparation during *post preparatory* phase and should represent most of the reproduction-related genes. Normal transcriptomic patterns along the brain-pituitary-gonad-liver (BPGL) axis should be better characterized to identify potential control points for reproductive performance. This paper describes an initial attempt to develop and analyze reproduction-relevant ESTs derived from the BPGL tissues of *L. rohita* by Sanger sequencing. Many gene/protein orthologues were identified for the first time in this species, which will be beneficial as a reference set for the production of large-scale transcriptome by next generation sequencing in the future. Clustering and assembly of 4642 high-quality sequences produced 709 contigs (19.52%) and 2922 singletons (80.47%) (Table 1). Presence of smaller numbers of contigs may be due to reduction in complementary overlapping and redundant (repetitive) sequences resulting from normalization. Presence of 70.52% of contigs made up of two ESTs while 6.90% of ESTs made of five or more ESTs also reflect the normalization effect (Fig. 1). Smaller numbers of contigs (24.93%) were also reported in sheepshead minnow (*Cyprinodon variegates*) in which normalized libraries were prepared for larvae (Pirooznia et al., 2010). Although large numbers of sequences from liver showed matches with public databases, pituitary and ovary sequences produced a limited percentage of matching, indicating the possibility of recovering more unknown sequences from these tissues of *Labeo rohita*. A similar trend was observed in ovary from zebrafish also (Li et al., 2004).

A battery of different gene/protein orthologues were identified for *L. rohita* in which large numbers (182) were related to reproduction-relevant genes and categorized into different sub-categories (Supplementary Table 2). Among reproduction-relevant genes identified, Insulin-like growth factor binding protein (IGFBP) was reported to have an expression profile similar to zinc finger proteins and RAD1, an ATPase associated with DNA repair, during the reproductive cycle in the liver in rainbow trout (Hook et al., 2011). During ovarian oocyte growth (vitellogenesis) cytochrome isomers produced high energy, selenoprotein increased antioxidant protection, heat shock protein 90 and Ras homolog member play roles in intracellular signaling pathways, beta 1 integrin play role in cell-to-cell and cell-to-matrix interactions (Sequeira et al., 2009).

Other genes/protein orthologues related to reproduction (Supplementary Table 2) includes i.e. cyclin b2, ribonucleoside-diphosphate reductase subunit m2 in testis; zona pellucida sperm-binding protein 3 in ovary; histone h2a deubiquitinase, 40s ribosomal protein s20 in liver; ribosomal protein s6, cyclin-a cdk2-associated protein p19 in brain; serine arginine-rich splicing factor, cyclin-dependent kinase inhibitor 3 from mixed tissue (contigs) were found in *Labeo rohita*. Many of these are well-characterized transcripts and several recent and informative papers have been published detailing the functions of these genes and their protein products in fish oocytes and embryos (Bobe and Labbe, 2010; Cerdà et al., 2008a,b; Chu et al., 2006; Gohin et al., 2010; Li et al., 2004; Luckenbach et al., 2008; Mommens et al., 2010; Von-Schalburg et al., 2005).

Another important protein vitellogenin (Vg) a precursor of yolk proteins is synthesized in the liver in maturing female fish under estradiol-17 β (E2) stimulation, transported to the ovary through bloodstream and incorporated into the growing oocytes by receptor-mediated endocytosis, followed by a proteolysis into smaller yolk proteins (Mommsen and Walsh, 1988). Three vitellogenin isoforms (vitellogenin 1, 2 and 3) have been observed in liver of *L. rohita* in our study; however, two forms of vitellogenin protein were reported earlier in the serum of estradiol-17 beta-treated *L. rohita* (Nath et al., 1992). The transcript for gonadotropin alpha subunit, common in both the GtHs the GtH I (FSH like) and II (LH like) was found both in pituitary (Melamed et al., 1998) as well as in the ovary of *post preparatory* phase of *L. rohita*. Isotocin-I precursor (Hiraoka et al., 1993), galanin message-associated-peptide, opioid growth factor, Neuropeptide Y (Söderberg et al., 1994) and c-type natriuretic peptide-4 act as neurohypophyseal hormones. Similarly, somatolactin precursor is responsible for the hormone activity (Iraqi et al., 1993), and thioredoxin helps in regulation of the biological process and cellular homeostasis. Identification of cholesterol-7-alfa-mono-oxygenase and catechol-o-methyltransferase from steroid hormone biosynthesis and sterol esterase and delta-24-sterol reductase from steroid biosynthesis in KEGG pathways analysis (Table 2) suggest the presence of both pathways in *L. rohita*. Steroids are essential for the regulation of spermatogenesis and oocyte maturation and their production is regulated by pituitary gonadotropins (Marín Juez et al., 2011).

Expression analyses of reproduction-relevant genes during *preparatory* and *post-spawning* stages are reported here for the first time in *L. rohita*. Although all these transcripts (ESTs) were collected during the *post preparatory* phase, generally the *preparatory* stage is considered as the initiation of gonad maturation in Indian carps (Manna and Bhattacharya, 1993) and accumulation of any transcript in this stage

Table 3
List of SSRs from reproductive relevant expressed sequence tags identified in *Labeo rohita*.

Sl. no.	Gene identity	Repeat region	Accession number
Polymorphic			
1.	Catechol-O-methyltransferase a	$(TG)_5$	HO758525
Monomorphic			
2.	Zgc:153925 protein	$(AAC)_4$	JK729777
3.	bcl2 adenovirus e1b 19 kDa-interacting protein 3	$(GCA)_6$	GR881241
4.	Vesicle-associated membrane protein 3	$(TCA)_4$	GR958070
5.	Thioredoxin	$(TGG)_4$	JK714776
6.	Vitelline membrane outer layer protein 1 homolog	$(TGT)_4$	GR958073
7.	Phosducin-like protein 3	$(GGA)_4$	JK714842
8.	Nuclear receptor corepressor hdac3 complex subunit tblr1	$(CAT)_5$	JK627268
9.	Guanine nucleotide-binding protein subunit beta-2-like 1	$(CA)_9$	JK671933
10.	Ribosomal l24 domain-containing protein 1	$(TG)_5$	Contig30

Table 4
Locus, GenBank accession number, repeat type, primer sequence, product size, annealing temperature (T), Number of allele (NA), observed heterozygosity (Ho), expected heterozygosity (He) and test for departure from HWE (*P*HW) of 20 *Labeo rohita* microsatellite loci.

Sl. no.	Locus	GenBank accession no.	SSR motif	Primer pair sequence (5′–3′)	Product Size (bp)	T (°C)	NA	Ho	He	*P*HW
01	PD03	HO758528	[TG]5	F:TGCTGCCTTCCTGTTTATTAGC R:AGCACACAAGCAAACCCTAGAT	156	58	2	0.774	0.482	0.000
02	PD11	HO758317	[AAT]6	F:CTAGGCTTCTGATGGCTGAAAT R:AGTAGCTCATCTCGTTGCCTTT	168	58	7	0.714	0.801	0.011
03	PD15	HO758308	[CT]16	F:TCGTATTAGTGGTCGTGTCTCG R:TCCTCTCAAAGGAACAGAATGC	156	60	3	0.379	0.662	0.000
04	PD16	HO758364	[CAG]5	F:GCAGCTCTACTTGTTTGTCACG R:GGTCATTCCACTTACCATGGAC	135	60	6	0.636	0.764	0.0162
05	PD19	GR958186	[GAT]5	F:GATACCAGACTGCCTGCCTATT R:GCAAACAACTTCAGGCTCTTCT	199	60	2	0.156	0.289	0.0278
06	PD26	JK714479	[AT]6	F:GAGGCTCCAAAGAGAAATATGG R:TACAGGAAGACCACTTGATTCG	245	58	3	0.096	0.537	0.0000
07	PD27	JK714495	[ATT]6	F:AGTAGCTCATCTCGTTGCCTTT R:CTAGGCTTCTGATGGCTGAAAT	170	58	3	0.485	0.629	0.0228
08	PD41	HO758779	[TC]8	F:CAGTGGTATCAACGCAGAGTGT R:GAACAGAAATGCTGAGATGCAG	140	58	3	0.533	0.550	0.7459
09	PD42	HO758312	[AAT]4	F:GGCTTCTGATGGCTGAAATACT R:AGTAGCTCATCTCGTTGCCTTT	171	60	3	0.600	0.675	0.0084
10	PD43	HO758525	[TG]5	F:CGGAGTGAAAGAGAAAGGAAAG R:AGCGATAGAGGACGTACAGGAC	130	58	2	0.250	0.221	1.0000
11	PD48	HO758564	[TG]5	F:CGGAGTGAAAGAGAAAGGAAAG R:AGCGATAGAGGACGTACAGGAC	130	58	2	0.294	0.505	0.0146
12	PD49	HO758654	[AGAA]26	F:CACCACAGCACTTAACACTTACTG R:AGAGTGATCCTCAGGGATTTCA	150	60	3	0.354	0.459	0.0006
13	PD84	HO758664	[GAT]4	F:ATGGACTCTCTGGATGGATACG R:GTTTACTTTGGCAACGGTCACT	151	60	3	0.285	0.299	0.1028
14	PD86	JK714585	[AAAAG]4	F:GAAATGGCTGAGACAGACAGTG R:CCAAAGAGACTTGATCCTTACG	183	60	4	0.117	0.552	0.0000
15	PD87	GR977079	[TCG]4	F:GCTTTGCAAACAACTCAGCTC R:AAGGAGGAAAAGCAGATGGTG	192	58	4	0.558	0.526	0.0284
16	PD98	JK714636	[TGG]4	F:ACCACTTGCAGTGAGTTGTACC R:CCGGGGAGACATTTAAACACT	215	58	3	0.147	0.332	0.0015
17	PD99	JK714758	[TG]5	F:TTGTTCCAGATCCAGAAGTGTG R:TTCTCCTCACCAACTCTTCCTC	179	60	3	0.242	0.428	0.0021
18	PD103	JK714764	[AC]5	F:GAGCTTTTTATCCCACCCATTC R:TGAACTCCACACAGGCTGTACT	163	58	2	0.114	0.109	0.0815
19	PD109	JK714807	[AC]8	F:ACGTGACACCATGTACAGGAAC R:TGGAGATGGATGAGTGAAACTG	231	58	2	0.892	0.607	0. 7451
20	PD112	JK714857	[TC]10	F:GCTTGGGCTAATACCAAGTGTT R:GCTGTTATAGGGAATGACTTCTGG	238	58	3	0.814	0.645	0.9990

Fourteen (PD3, PD11, PD15, PD16, PD19, PD26, PD27, PD42, PD48, PD49, PD86, PD87, PD98 and PD99) of the 20 loci deviated from HWE ($p < 0.05$).

indicates its role in maturation initiation. On the other hand, the *post-spawning* phase is considered as a resting period with almost no gonad activity. Again events in initiation of gonad maturation are highly important to study for a monsoon breeder which occurs with an increase in water temperature and photoperiod during March–April. It is also of special interest and pre-requisite for the successful attempts for year-round seed production by artificial simulation of the natural environment under controlled conditions (Sarkar et al., 2010). Expression levels of selected transcripts were evaluated and compared during *preparatory* and *post-spawning* phase. ZP-3 is a vital constitutive part of the oocyte and is essential for oocyte fertilization. This structure binds spermatozoa and is required to initiate the acrosome reaction (Conner et al., 2005). Higher expression levels of ZP-3 were found both in testis and ovary in the *preparatory* phase in our study. It was earlier reported to be highly expressed in the liver and ovary in *Gilthead seabream* (Modig et al., 2006), and in the ovary and kidney of females but could be barely detected in the testis and kidney of males in *Cynoglossus semilaevis* (Sun et al., 2010). PRMC-1 has been described as a possible participant in the progestin-induced oocyte maturation in fish. PRMC-1 was found both in ovary and testis, while up regulated in testis but down regulated in ovary during *preparatory* phase in rohu; however presence of PRMC-1 mRNA in the oocyte and its strong expression in ovarian tissue was confirmed earlier in rainbow trout. The expression level was reported to be highest during vitellogenesis and a decrease in expression during late vitellogenesis of the reproductive cycle (Mourot et al., 2006). Similarly, Pro-opiomelanocortin is a precursor polypeptide expressed in pituitary gland and is necessary for adrenal function and energy homeostasis (Raffin-Sanson et al., 2003) and a higher level of POMC in *preparatory* pituitary suggests similar function in rohu also. Spermine oxidase is a FAD-dependent enzyme that specifically oxidizes spermine (Spm) (Cervelli et al., 2012) and high level of SPO in *preparatory* testis confirms its dominant role in rohu. Aquaporins (APQs) are a super family of integral membrane proteins that facilitate the rapid and yet highly selective flux of water and other small solutes across biological membranes and are strongly associated with the

Table 5
Cross-species amplification for the twenty microsatellite loci developed in *Labeo rohita*.

Sl. no	Species	No. of loci with successful cross amplification	Annealing temperature (°C)
01	*Cirrhinus mrigala*	14	55–60
02	*Catla catla*	16	55–60
03	*Labeo rohita*	18	55–60
04	*Labeo bata*	13	55–60
05	*Labeo gonius*	15	55–60
06	*Labeo fimbriatus*	14	55–60
07	*Cirrhinus reba*	12	55–60
08	*Labeo kalbasu*	10	55–60

various functions in the female and male reproductive systems (Huang et al., 2006) as well as egg hydration in teleosts (Cerdà, 2009). In mammals, 13 aquaporin subfamilies are present whereas 18 aquaporin genes with homologies are reported in zebrafish and other teleosts (Madsen, 2012). Aquaporin-1b transcripts were detected exclusively in *Heteropneustes fossilis* ovary and brain and showed significant seasonal variations; in the ovary, aquaporin 1b was maximally expressed during the *pre-spawning* period, whereas in the brain the highest expression was detected during *spawning* (Chaube et al., 2011). Itoh et al. (2005) reported that the APQ-12 was specifically expressed in the pancreas in xenopus. In rohu, APQ-12 was highly expressed in liver (equivalent to hepato-pancreas in carp), followed by testis and pituitary and significant up regulation of it was noticed in pituitary in *preparatory* phase. Teleosts generally lack a hypophyseal portal system (Levavi-Sivan et al., 2010) and increase in APQ-12 in pituitary with initiation of gonad activity may be a mechanism to increase flow of solutes. However, its specific role in rohu reproduction needs further study. NP-Y expression level in rohu was higher in brain, ovary and pituitary during *post-spawning* phase as compared to preparatory phase. Mazumdar et al. (2007) reported that NP-Y immune-reactivity in olfactory neurons and pituitary of *Clarias batrachus* gradually augmented during resting through *pre-spawning* phases, attaining a peak in the spawning phase. It remains to be seen if the expression pattern is similar in rohu also and ovarian NP-Y was not reported earlier. Sperm associated antigens (SAA) are a testis-specific component that is expressed during germ cell differentiation (Wu et al., 2010), and variances with wide diversity of SAA were observed in different species. In rohu higher expression levels of SAA-7 and TE-261 were observed in testis during *preparatory* phase but no such report for SAA-7 and TE-261 expression is found in any other fish species. Therefore, the general trend of increased expression level of almost all known transcripts studied in testis, ZP-3, PRMC-1 in ovary and POMC, APQ-12 in pituitary during *preparatory* phase of *L. rohita* may confirm their important roles in initiation of gonad maturation, while increased NP-Y level in brain, ovary and pituitary in *post-spawning* phase indicates its possible role during resting of gonads.

15.67% of sequences with putative ORFs among un-identified sequences are quite significant for the possible discovery of new genes from *L. rohita*. Among these Rh1160O, Rh1069P, Rh1399P, Rh619B and Rh1313O transcripts were found accumulated in higher levels in ovary and Rh1399P, Rh619B and Rh1313O transcripts in pituitary during the *preparatory* phase. Searching of similar sequences in the present database of common carp and tilapia was not successful, and always showing only a few bases matching both at BLASTN, tBLASTN as well as in BLASTX and BLASTP, but expression of all these transcripts were also observed in these prolific breeders. Level of expression was much higher in the seasonal breeder rohu in comparison to carp and tilapia which indicates that there may be some possible role of these transcripts in reproduction of rohu. Many unknown transcripts or novel sequences (34.7%) were also reported during transcriptomic analyses in zebrafish gonad and brain (Sreenivasan et al., 2008). However, functional validation of these un-known transcripts and their exact role in *L. rohita* reproduction remains to be confirmed in future.

In this study 52 out of 129 SSR-containing ESTs were annotated sequence (with known function) and the repeats were located either at 3′UTR or 5′UTR, as well as within the coding region. This has great significance for development of type I marker as ESTs are considered an important resource for discovery of polymorphic markers by microsatellite tagging (Liu et al., 1999). These type I markers in reproduction-related genes should be useful for *L. rohita* genetic research, genome mapping and for the identification of trait-associated alleles for marker-assisted selection (Liu and Cordes, 2004). In *L. rohita* AG/TG was the most frequent motif for microsatellites found in this experiment (Fig. 7). Robinson et al. (2012) reported TG as the common repeat motif in *L. rohita* microsatellites in genes with putative health and stress related functions. 129 EST-SSRs of 236 identified contained sufficient flanking sequences, out of which 42 primer pairs failed to amplify any PCR product. This may be due to the primer sequences spanning introns and/or containing mutations and/or indels (insertion or deletion) (Gao et al., 2010). Out of the 20 polymorphic loci (Table 4), fourteen deviated from the Hardy–Weinberg Equilibrium (HWE) ($p < 0.05$). This result was likely due to the limited sample size, and/or presence of over-dominant selection (Sui et al., 2009). Microsatellite mutation rates have been reported as high as 10^{-2} per generation (Crawford and Cuthbertson, 1996), and are believed to be caused by polymerase slippage during DNA replication, resulting in differences in the number of repeat units (Levinson and Gutman, 1987). The reason for better cross-species amplification (Table 5) in *Catla catla* with rohu primers cannot be explained properly with the present state of knowledge. *Labeo kalbasu* was reported more genetically similar to *L. rohita* than *Catla catla* or *Cirrhinus mrigala* as estimated by RAPD analysis (Barman et al., 2003). Singh et al. (2008) tested cross-amplification of SSR markers and assessed their suitability for population genetic analysis in *Labeo kalbasu*. Patel et al. (2009) during cross-species amplification tests with rohu primers observed greater numbers of microsatellite loci amplified among the *Labeo* genus than with *Catla catla* or *Cirrhinus mrigala*. Conservation of flanking regions is a general property of microsatellite loci and has been specifically reported in fishes (Rico et al., 1996). This fact permits homologous microsatellites to be amplified in closely related species using the same set of primers (cross amplification). This strategy of cross-species amplification may provide a cost-effective alternative to microsatellite isolation and development, which is time consuming and labor intensive.

5. Conclusion

Identification of 182 reproduction-relevant genes from 4642 high-quality *L. rohita* EST sequences from reproduction-related tissues represents a significant advancement of genomic resources in this species. Tissue expression analyses of 8 known genes (ZP-3, SAA-7, NP-Y, PRMC-1, APQ-12, SPO, POMC and TE-261) and 8 unknown genes with putative ORFs (Rh21T, Rh1231L, Rh1160O, Rh1069P, Rh1399P, Rh1930B, Rh619B and Rh1313O) during *preparatory* and *post-spawning* phases indicated the potential roles of these transcripts in initiation of gonad maturation of the monsoon breeder carp *L. rohita*. Expressions of unknown transcripts were also observed in prolific breeder common carp and tilapia which were not reported earlier. Identification of 20 polymorphic EST-SSR developed in these transcripts will be highly useful in application to the selective breeding program for *L. rohita* in the near future.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2013.03.111.

Acknowledgments

This work was funded by the Department of Biotechnology, Government of India (Project Reference: BT/PR9916/AAQ/03/360/2007; 07.07.2008), supported by Indian Council of Agricultural Research (ICAR), and Central Institute of Freshwater Aquaculture (CIFA), Government of India. We would like to thank DG, ICAR and Director, CIFA for their encouragement. Special thanks to Dr P. Routary, Carp Breeding Unit, CIFA for providing fish and to Madhusmita Panda and Debashree Sahu for their help.

References

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. Science 252, 1651–1656.

Barman, H.K., Barat, A., Yadav, B.M., Banerjee, S., Meher, P.K., Reddy, P.V.G.K., Jana, R.K., 2003. Genetic variation between four species of Indian major carps as reveled by random amplified polymorphic DNA assay. Aquaculture 217, 115–123.

Bhattacharya, S., 1999. Recent advances in the hormonal regulation of gonadal maturation and spawning in fish. Curr. Sci. 3, 342–349.

Bobe, J., Labbe, C., 2010. Egg and sperm quality in fish. Gen. Comp. Endocrinol. 165, 535–548.

Cerdà, J., 2009. Molecular pathways during marine fish egg hydration: the role of aquaporins. J. Fish Biol. 75 (9), 2175–2196.

Cerdà, J., Bobe, J., Babin, P.J., Admon, A., Lubzens, E., 2008a. Functional genomics and proteomic approaches for the study of gamete formation and viability in farmed finfish. Res. Fish. Sci. 16, 54–70.
Cerdà, J., Mercadé, J., Lozano, J.J., Manchado, M., Tingaud-Sequeira, A., Astola, A., Infante, C., Halm, S., Viñas, J., Castellana, B., Asensio, E., Cañavate, P., Martínez-Rodríguez, G., Piferrer, F., Planas, J.V., Prat, F., Yúfera, M., Durany, O., Subirada, F., Rosell, E., Maes, T., 2008b. Genomic resources for a commercial flatfish, the Senegalese sole (*Solea senegalensis*): EST sequencing, oligo microarray design, and development of the Soleamold bioinformatic platform. BMC Genomics 9, 508.
Cervelli, M., Amendola, R., Polticelli, F., Mariottini, P., 2012. Spermine oxidase: ten years after. Amino Acids 42, 441–450.
Chakraborti, P., Bhattacharya, S., 1984. Plasma thyroxine levels in freshwater perch: influence of season, gonadotropins, and gonadal hormones. Gen. Comp. Endocrinol. 53 (2), 179–186.
Chaube, R., Chauvigné, F., Tingaud-Sequeira, A., Joy, K.P., Acharjee, A., Singh, V., Cerdà, J., 2011. Molecular and functional characterization of catfish (*Heteropneustes fossilis*) aquaporin-1b: changes in expression during ovarian development and hormone-induced follicular maturation. Gen. Comp. Endocrinol. 170 (1), 162–171.
Chaudhuri, H., Alikunhi, K.H., 1957. Observation on the spawning of Indian major carps by hormone injection. Curr. Sci. 26, 381–382.
Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162 (1), 156–159.
Chu, S.L., Weng, C.F., Hsiao, C.D., Hwang, P.P., Chen, Y.C., Ho, J.M., Lee, S.J., 2006. Profile analysis of expressed sequence tags derived from ovary of tilapia, *Oreochromis mossambicus*. Aquaculture 251, 537–548.
Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21 (18), 3674–3676.
Conner, S.J., Lefièvre, L., Hughes, D.C., Barratt, C.L., 2005. Cracking the egg: increased complexity in the zona pellucida. Hum. Reprod. 20 (5), 1148–1152.
Crawford, A.M., Cuthbertson, R.P., 1996. Mutations in sheep microsatellites. Genome Res. 6, 876–879.
Gao, X.G., Li, H.J., Li, Y.F., Sui, L.J., Zhu, B., Liang, Y., Liu, W.D., He, C.B., 2010. Sixteen polymorphic simple sequence repeat markers from expressed sequence tags of the Chinese mitten crab *Eriocheir sinensis*. Int. J. Mol. Sci. 11, 3035–3038.
Gjerde, B., Reddy, P.V.G.K., Mahapatra, K.D., Saha, J.N., Jana, R.K., Meher, P.K., Sahoo, M., Lenka, S., Govindassamy, P., Rye, M., 2002. Growth and survival in two complete diallele crosses with five stocks of Rohu carp (*Labeo rohita*). Aquaculture 209, 103–115.
Goetz, F.W., McCauley, L., Goetz, G.W., Norberg, B., 2006. Using global genome approaches to address problems in cod mariculture. ICES J. Mar. Sci. 63, 393–399.
Gohin, M., Bobe, J., Chesnel, F., 2010. Comparative transcriptomic analysis of follicleenclosed oocyte maturational and developmental competence acquisition in two non-mammalian vertebrates. BMC Genomics 11, 18.
Gui, J.F., Zhu, Z.Y., 2012. Molecular basis and genetic improvement of economically important traits in aquaculture animals. Chin. Sci. Bull. 57, 1751–1760.
Gupta, S.D., Rath, S.C., Dasgupta, S., Tripathi, S.D., 1995. A first report on quadruple spawning of *Catla catla* (Ham.). Vet. Arhiv. 65 (5), 143–148.
Hiraoka, S., Suzuki, M., Yanagisawa, T., Iwata, M., Urano, A., 1993. Devergence of gene expression in neurohypophysial hormone precursors among salmonids. Comp. Endocrinol. 92, 292–301.
Hook, S.E., Nagler, J.J., Cavileer, T., Verducci, J., Liu, Y., Hayton, W., Schultz, I.R., 2011. Relationships between the transcriptome and physiological indicators of reproduction in female Rainbow trout over an annual cycle. Environ. Toxicol. Chem. 30, 309–318.
Huang, X., Madan, A., 1999. CAP3: a DNA sequence assembly program. Genome Res. 9, 868–877.
Huang, H.F., He, R.H., Sun, C.C., Zhang, Y., Meng, Q.X., Ma, Y.Y., 2006. Function of aquaporins in female and male reproductive systems. Hum. Reprod. Updat. 12 (6), 785–795.
Iraqi, F., Gong, Z., Hew, C.L., Crim, L.W., 1993. Isolation and characterization of somatolactin genes from two cold water marine teleosts, lumpfish (*Cyclopterus lumpus*) and halibut (*Hippoglossus hippoglossus*). Mol. Mar. Biol. Biotechnol. 2, 96–103.
Itoh, T., Rai, T., Kuwahara, M., Ko, S.B., Uchida, S., Sasaki, S., Ishibashi, K., 2005. Identification of a novel aquaporin, AQP12, expressed in pancreatic acinar cells. Biochem. Biophys. Res. Commun. 330, 832–838.
Kanehisa, M., Goto, S., Kawashima, S., Nakaya, A., 2002. The KEGG databases at Genome Net. Nucleic Acids Res. 30 (1), 42–46.
Leong, J.S., Jantzen, S.G., Von-schalburg, K.R., Cooper, G.A., Messmer, A.M., Liao, N.Y., Munro, S., Moore, R., Holt, R.A., Jones, S.J.M., Davidson, W.S., Koop, B.F., 2010. Salmo salar and Esox lucius full-length cDNA sequences reveal changes in evolutionary pressures on a post-tetraploidization genome. BMC Genomics 11, 279.
Levavi-Sivan, B., Bogerd, J., Mañanós, E.L., Gómez, A., Lareyre, J.J., 2010. Perspectives on fish gonadotropins and their receptors. Gen. Comp. Endocrinol. 165 (3), 412–437.
Levinson, G., Gutman, G.A., 1987. High frequency of short frameshifts in poly-CA/GT tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. Nucleic Acids Res. 15, 5323–5338.
Lewis, P.O., Zaykin, D., 2001. Genetic Data Analysis: computer program for analysis of allelic data, Ver 1.0. http://hydrodyction.eeb.uconn.edu/people/plewis/software.php.
Li, Y., Chia, J.M., Bartfai, R., Christoffels, A., Yue, G.H., Ding, K., Ho, M.Y., Hill, J.A., Stupka, E., Orban, L., 2004. Comparative analysis of the testis and ovary transcriptomes in zebrafish by combining experimental and computational tools. Comp. Funct. Genom. 5, 403–418 (1).
Liu, Z.J., 2007. Aquaculture Genome Technologies. Blackwell Publishing, Ames, IA, Oxford, UK.
Liu, Z.J., Cordes, J.F., 2004. DNA marker technologies and their applications in aquaculture genetics. Aquaculture 238, 1–37.
Liu, Z., Karsi, A., Dunham, R.A., 1999. Development of polymorphic EST markers suitable for genetic linkage mapping of catfish. Mar. Biotechnol. 1, 437–0447.
Luckenbach, J.A., Iliev, D.B., Goetz, F.W., Swanson, P., 2008. Identification of differentially expressed genes during primary and early secondary oocyte growth in coho salmon. Oncorhynchus kisutch. Reprod. Biol. Endocrinol. 6, 2.
Madsen, S.S., 2012. Aquaporins in fishes — expression, localization, and functional dynamics. Front. Physiol. http://dx.doi.org/10.3389/fphys.2012.00434.
Manna, P.R., Bhattacharya, S., 1993. [^{125}I] gonadotropin binding to the ovary of an Indian major carp, *Catla catla*, at different stages of reproductive cycle. J. Biol. Sci. 18 (3), 361–372.
Marín Juez, R., Castellana, B., Manchado, M., Planas, J.V., 2011. Molecular identification of genes involved in testicular steroid synthesis and characterization of the response to gonadotropic stimulation in the Senegalese sole (*Solea senegalensis*) testis. Gen. Comp. Endocrinol. 172, 130–139.
Masoudi-Nejad, A., Tonomura, K., Kawashima, S., Moriya, Y., Suzuki, M., Itoh, M., Kanehisa, M., Endo, T., Goto, S., 2006. EGassembler: online bioinformatics service for large-scale processing, clustering and assembling ESTs and genomic DNA fragments. Nucleic Acids Res. 34, 459–462.
Mazumdar, M., Sakharkar, A.J., Singru, P.S., Subhedar, N., 2007. Reproduction phase-related variations in neuropeptide Y immunoreactivity in the olfactory system, forebrain, and pituitary of the female catfish, *Clarias batrachus* (Linn.). J. Comp. Neurol. 504 (5), 450–469.
Melamed, P., Rosenfeld, H., Elizur, A., Yaron, Z., 1998. Endocrine regulation of gonadotropin and growth hormone gene transcription in fish. Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol. 119 (3), 325–338.
Modig, C., Modesto, T., Canario, A., Cerdà, J., von-Hofsten, J., Olsson, P.E., 2006. Molecular and functional characterization of catfish (*Heteropneustes fossilis*) aquaporin-1b: changes in expression during ovarian development and hormone-induced follicular maturation. Biol. Reprod. 75 (5), 717–725.
Mommens, M., Fernandes, J.M.O., Bizuayehu, T.T., Bolla, S.L., Johnston, I.A., Babiak, I., 2010. Maternal gene expression in Atlantic halibut (*Hippoglossus hippoglossus* L.) and its relation to egg quality. BMC Res. Notes 3, 138.
Mommsen, T.P., Walsh, T.J., 1988. Vitellogenesis and oocyte assembly. : Fish Physiology. Academic Press, San Diego.
Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C., Kanehisa, M., 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 35, 182–185.
Mourot, B., Nguyen, T., Fostier, A., Bobe, J., 2006. Two unrelated putative membrane-bound progestin receptors, progesterone membrane receptor component 1 (PGMRC1) and membrane progestin receptor (mPR) beta, are expressed in the rainbow trout oocyte and exhibit similar ovarian expression. Reprod. Biol. Endocrinol. 4, 6.
Nandi, S., Chattopadhyay, D.N., Verma, J.P., Sarkar, S.K., Mukhopadhyay, P.K., 2001. Effect of dietary supplementation of fatty acids and vitamins on the breeding performance of the carp *Catla catla*. Reprod. Nutr. Dev. 41 (4), 365–375.
Natarajan, A.V., Jhingran, A.G., 1963. On the biology of *Catla catla* (Ham.) from the River Yamuna. Proc. Natn. Inst. Sci. India 29B, 326–355.
Nath, P., Bhakta, M., Mitra, K., 1992. Demonstration of two forms of vitellogenin in serum of estradiol-17 beta-treated Indian major carp, *Labeo rohita*. Indian J. Exp. Biol. 30 (6), 464–469.
Patel, A., Das, P., Barat, A., Meher, P.K., Jayasankar, P., 2009. Utility of cross-species amplification of 34 rohu microsatellite loci in *Labeo bata*, and their transferability in six other species of the cyprinidae family. Aquac. Res. 41, 590–593.
Pfaffl, M.W., Horgan, G.W., Dempfle, L., 2002. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res. 30 (9), 36.
Pirooznia, M., Pozhitkov, A., Perkins, E.J., Deng, Y., Brouwer, M., 2010. Generation, analysis and functional annotation of expressed sequence tags from the sheepshead minnow (Cyprinodon variegatus). BMC Genomics 2, 11.
Quasim, S.Z., Qayyum, A., 1962. Spawning frequencies and breeding season of some freshwater fishes with special reference to those occuring in the plains of northern India. Indian J. Fish. 8, 24–43.
Raffin-Sanson, M.L., Keyzer, Y.D., Bertagna, X., 2003. Proopiomelanocortin, a polypeptide precursor with multiple functions: from physiology to pathological conditions. Eur. J. Endocrinol. 149, 79–90.
Rico, C., Rico, I., Hewitt, G., 1996. 470 million years of conservation of microsatellite loci among fish species. Proc. R. Soc. Lond. B Biol. Sci. 263, 549–557.
Robinson, N., Sahoo, P.K., Baranski, M., Das-Mahapatra, K., Saha, J.N., Das, S., Mishra, Y., Das, P., Barman, H.K., Eknath, A.E., 2012. Expressed sequences and polymorphisms in Rohu carp (*Labeo rohita*, Hamilton) revealed by mRNA-seq. Mar. Biotechnol. 10, 1007.
Routray, P., Verma, D.K., Sarkar, S.K., Sarangi, N., 2007. Recent advances in carp seed production and milt cryopreservation. Fish Physiol. Biochem. 2007 (33), 413–427.
Sahoo, P.K., Rauta, P.R., Mohanty, B.R., Mahapatra, K.D., Saha, J.N., Rye, M., Eknath, A.E., 2011. Selection for improved resistance to Aeromonas hydrophila in Indian major carp *Labeo rohita*: survival and innate immune responses in first generation of resistant and susceptible lines. Fish Shellfish Immunol. 31 (3), 432–438.
Sanger, F., Nicklen, S., Coulson, A., 1977. DNA sequencing with chain terminating inhibitor. Proc. Natl. Acad. Sci. 74, 5463–5467.
Sarkar, S.K., Saha, A., Dasgupta, S., Nandi, S., Verma, D.K., Routray, P., Davaraj, C., Mohanty, J., Sarangi, N., Eknath, A.E., Ayyapan, S., 2010. Photothermal manipulation of reproduction in Indian major carp: a step forward for offseason breeding and seed production. Curr. Sci. 7, 960–965.
Schmidt, W.M., Muller, M.W., 1999. Cap Select: A highly sensitive method for 5′ cap dependent enrichment of full length cDNA in PCR mediated analysis of mRNAs. Nucleic Acids Res. 27, 21.
Sequeira, T.A., Chauvigné, F., Lozano, J., Agulleiro, M.J., Asensio, E., Joan Cerdà, J., 2009. New insights into molecular pathways associated with flatfish ovarian

development and atresia revealed by transcriptional analysis. BMC Genomics 10, 434.
Singh, R.K., Lal, K.K., Mohindra, V., Punia, P., Lakra, W.S., 2008. Cross-species amplification of microsatellite loci in the cyprinid fish *Labeo calbasu* (Hamilton, 1822). Acta Zool. Sin. 55, 937–940.
Söderberg, C., Pieribone, V.A., Dahlstrand, J., Brodin, L., Larhammar, D., 1994. Neuropeptide role of both peptide YY and neuropeptide Y in vertebrates suggested by abundant expression of their mRNAs in a cyclostome brain. J. Neurosci. Res. 37 (5), 633–640.
Sreenivasan, R., Cai, M., Bartfai, R., Wang, X., Alan Christoffels, A., Orban, L., 2008. Transcriptomic analyses reveal novel genes with sexually dimorphic expression in the zebrafish gonad and brain. PLoS One 3 (3), 1791.
Sui, L.Y., Zhang, F.M., Wang, X.M., Bossier, P., Sorg-Hanfling, B., 2009. Genetic diversity and population structure of the Chinese mitten crab *Eriocheir sinensis* in its native range. Mar. Biol. 156, 1573–1583.
Sun, Y., Yu, H., Zhang, Q., Qi, J., Zhong, Q., Chen, Y., Li, C., 2010. Molecular characterization and expression pattern of two zona pellucida genes in half-smooth tongue sole (*Cynoglossus semilaevis*). Comp. Biochem. Physiol. B Biochem. Mol. Biol. 155 (3), 316–321.
Sundararaj, B.I., Vasal, S., 1976. Photoperiod and temperature control in the regulation of reproduction in the female catfish *Heteropneustes fossilis*. J. Fish. Res. Board Can. 33 (4), 959–973.
Tassanakajon, A., Pongsomboon, S., Rimphanitchayakit, V., Jarayabhand, P., Boonsaeng, V., 1997. Random amplified polymorphic DNA (RAPD) markers for determination of genetic variation in wild populations of the black tiger prawn (*Penaeus monodon*) in Thailand. Mol. Mar. Biol. Biotechnol. 6, 110–115.
Von-Schalburg, K.R., Rise, M.L., Brown, G.D., Davidson, W.S., Koop, B.F., 2005. A comprehensive survey of the genes involved in maturation and development of the rainbow trout ovary. Biol. Reprod. 72, 687–699.
Wu, H., Chen, Y., Miao, S., Zhang, C., Zong, S., Koide, S.S., Wang, L., 2010. Sperm associated antigen 8 (SPAG8), a novel regulator of activator of CREM in testis during spermatogenesis. FEBS Lett. 584 (13), 2807–2815 (Jul 2).
Zhulidov, P.A., Bogdanova, E.A., Shcheglov, A.S., Vagner, L.L., Khaspekov, G.L., Kozhemyako, V.B., Matz, M.V., Meleshkevitch, E., Moroz, L.L., Lukyanov, S.A., Dmitry, A., Shagin, D.A., 2004. Simple cDNA normalization using kamchatka crab duplex-specific nuclease. Nucleic Acids Res. 32.