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