

Transcriptomic Analyses Reveal Novel Genes with Sexually Dimorphic Expression in Yellow Catfish (*Pelteobagrus fulvidraco*) Brain

Jianguo Lu^{1,2,4} · Min Zheng^{1,3} · Jiajia Zheng¹ · Jian Liu² · Yongzhuang Liu² · Lina Peng^{1,5} · Pingping Wang¹ · Xiaofeng Zhang¹ · Qiushi Wang¹ · Peixian Luan¹ · Shahid Mahbooband⁶ · Xiaowen Sun¹

Received: 1 December 2014 / Accepted: 18 May 2015 / Published online: 5 August 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract Yellow catfish (*Pelteobagrus fulvidraco*) is a pivotal freshwater aquaculture species in China. It shows sexual size dimorphism favoring male in growth. Whole transcriptome approach is required to get the overview of genetic toolkit for understanding the sex determination mechanism aiming at devising its monosex production. Beside gonads, the brain is also considered as a major organ for vertebrate reproduction. Transcriptomic analyses on the brain and of different developmental stages will provide the dynamic view necessary for better understanding its sex determination. In this regard, we have performed a de novo assembly of yellow catfish brain transcriptome by high throughput Illumina

sequencing. A total number of 154,507 contigs were obtained with the lengths ranging from 201 to 27,822 bp and N50 of 2,101 bp, as well as 20,699 unigenes were identified. Of these unigenes, 13 and 54 unigenes were detected to be XY-specifically expressed genes (SEGs) for one and 2-year-old yellow catfish, while the corresponding numbers of XX-SEGs for those two stages were 19 and 13, respectively. Our work identifies a set of annotated genes that are candidate factors affecting sexual dimorphism as well as simple sequence repeat (SSR) and single nucleotide variation (SNV) in yellow catfish. To validate the expression patterns of the sex-related genes, we performed quantitative real-time PCR (qRT-PCR) indicating the reliability and accuracy of our analysis. The results in our study may enhance our understanding of yellow catfish sex determination and potentially help to improve the production of all-male yellow catfish for aquaculture.

Jianguo Lu and Min Zheng contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10126-015-9650-z) contains supplementary material, which is available to authorized users.

✉ Jianguo Lu
jianguonk@gmail.com

✉ Xiaowen Sun
sunxw2002@163.com

Keywords Transcriptome · Sexually dimorphic · Yellow catfish · RNA-Seq

Introduction

Yellow catfish (*Pelteobagrus fulvidraco*) is an important freshwater aquaculture species in China. Genetically, yellow catfish harbors an XX/XY system (Liu et al. 2010, 2013). It displays sexual size dimorphism, male exhibits faster growth rate and reaching a larger ultimate size (three times) than female (Park et al. 2004; Liu et al. 2013). Apart from the intriguing reproductive biology, it is economically vital with continuously growing aquaculture in China but limited available genetic resources. Efficient utilization of the natural diversity

- ¹ Heilongjiang River Fisheries Research Institute, Chinese Academy of Fishery Sciences, 43 Songfa Street, Daoli District, Harbin 150070, China
- ² School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
- ³ Department of Civil Engineering, Auburn University, Auburn, AL 36849, USA
- ⁴ National and Local United Engineering Lab for Freshwater Fish Breeding, Harbin, China
- ⁵ Harbin Normal University, Harbin, China
- ⁶ Department of Zoology, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

of trait phenotypes, i.e., utilization of growth trait to produce monosex of yellow catfish, requires the identification of genetic underpinnings of trait differences. To get an overview of the genetic toolkit deployed for the development and maintenance of the differences between sexes, whole transcriptome approach is required (Piferrer et al. 2012).

In our previous study (Lu et al. 2014), we have identified sex-related genes from gonad transcriptomic analysis which has allowed us to profile the expression of a series of genes involved in sex differentiation of female and male yellow catfish. However, sex determination gene and sex determination mechanism have yet to be elucidated. Similar analyses on other tissue parts and of different developmental stages will provide the dynamic view necessary for a better understanding of sex determination.

To our knowledge, beside gonads, the brain is also involved in vertebrate reproduction. It controls growth and reproduction mainly through the brain–pituitary–gonadal axis (Weltzien et al. 2004). The involvement of the brain in gonad development has been established by studies on quail (Munday et al. 2006; Francis and Barlow 1993), indicating the brain may determine the fate of the gonads (Sreenivasan et al. 2008). It is also reported that differential brain gene expression can also lead to different development of the brain in the two sexes; this even occurs before the gonads formed (Dennis 2004; Dewing et al. 2003; Davies and Wilkinson 2006; Sellars et al. 2015). Also, the brain exhibits one of the most complex transcriptomes of all organs in vertebrates; hence, it is a tissue of choice for sequencing a maximum number of transcripts while reducing the need for normalization (Tzika et al. 2011).

So far, many sex-related genes in brain have been detected. In rodents, *Sry* was a male-determining gene expressed in the substantia nigra of the midbrain (Dewing et al. 2006). In zebra finch, the expression level of *trkB* in male brain, which was a Z-linked gene, is found to be higher than in female (Chen et al. 2005). More genes have been identified showing sex-related expression pattern in the brain including *Usp9x/y* (Xu et al. 2002), *Utx/Uty* (Xu et al. 2008), *CHD1Z/W* (Agate et al. 2003, 2004), and sex-specific markers (Xu et al. 2013), whereas their roles in sex determination are yet not known. Not only in rodents and birds, sex-related genes in brain were also found in teleost fish species, including medaka (Maehiro et al. 2014), zebrafish (Sreenivasan et al. 2008; Santos et al. 2008), rainbow trout (Vizziano-Cantonnet et al. 2011; Yano et al. 2014), sharpnose seabream (Manousaki et al. 2014), and goldfish (Parhar et al. 2001).

To gain further insight into molecular mechanism underlying yellow catfish's sexual dimorphism, we extended our studies by assembling brain transcriptome of yellow catfish using in-depth Illumina HiSeq sequencing.

In our study, we sequenced yellow catfish brain transcriptome of two developmental stages and identified the differences in gene expression profiles between sexes. More novel genes were identified related to sex dimorphism in the brain as well as genetic markers. The further identification of the expression profile of genes involved in sex dimorphism may help to illuminate the gene regulatory network controlling sex determination and subsequent maintenance of phenotypic sex and to devise the production of monosex yellow catfish (Liu et al. 2013; Chen et al. 2015).

Materials and Methods

Ethics Statement, Experimental Fish, and Sample Collection

All procedures in our study including the handling and treatment of fish were approved by the Animal Care and Use Committee at Heilongjiang River Fisheries Research Institute (ACUC-HRFRI). We used 40 fishes, including 20 male yellow catfish (10 one-year-old juveniles, 10 two-year-old adults) and 20 female yellow catfish (10 one-year-old juveniles and 10 two-year-old adults). They are full-sibling yellow catfish which came from the National Fish Original Species Farm at Zhaodong City, Heilongjiang Province, China. One-year-old and 2-year-old yellow catfish were collected in August 16, 2012 and August 14, 2013, respectively. The sex of these fish was confirmed anatomically. Then, the experimental fish were euthanized with 250 mg/L tricaine methanesulfonate before sample collection. Brain tissues of male and female yellow catfish from two developmental stages were collected and stored in 1.5 ml RNAlater tube (Qiagen, Hilden, Germany), respectively, and then transferred to a -80°C freezer until prior to RNA extraction.

RNA Extraction

Samples were removed from the -80°C freezer and homogenated using TissueRuptor (Qiagen, Hilden, Germany) to a fine solution. Total RNAs were extracted from each sample using RNeasy Lipid Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions after the RNase-free DNase I (Qiagen, Hilden, Germany) treatment of genomic DNA elimination. The concentration and integrity of RNA were examined using Thermo Scientific™ NanoDrop™ 8000 Spectrophotometer and Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The RNA with OD_{260/280} ≥ 1.8 and RNA integrity number

(RIN) ≥ 7.0 was selected for the following experiment. Equal amounts of the high-quality RNA samples from brain tissue were then pooled together for cDNA synthesis and sequencing.

Library Construction, Illumina Sequencing, and Assembly of cDNA Library

RNA-Seq library preparation and sequencing were carried out by BerryGenomics sequencing company (Beijing, China). cDNA library was prepared with ~ 2.5 μg of total RNA based on the Illumina TrueSeq RNA Sample Preparation Kit (Illumina) protocols. The library was then amplified, and the final library yields ~ 400 ng with average fragment size of ~ 270 bp. The library was sequenced with one lane on an Illumina HiSeq 2000 instrument with 100-bp paired-end reads after KAPA quantitation and dilution. Raw read data of yellow catfish RNA-Seq with SRA format have been uploaded to the NCBI Short Read Archive under the accession number SRR1103702. The clean reads from the four transcriptome were filtered out adaptor-only reads and low-quality reads (reads with Q value ≤ 20). Cleaned reads were used for de novo assembly using the de Bruijn graph approach with Trinity by default parameters.

Transcriptome Construction and Assembly Assessment

To assess the transcriptome of yellow catfish, we compared the transcriptome contigs with Refseq and Ensemble proteins of zebrafish, medaka, stickleback, and tetraodon by BLAST program with default parameters.

Functional Annotation and Gene Ontology Analysis

The assembly RNA-Seq contigs was used for similarity search program against reference protein sequence

including zebrafish, medaka, stickleback, and tetraodon, respectively. The similarity searches were performed using the BLASTX program with the E value cutoff of $1e-10$. Gene annotation was assigned to the RNA-Seq contigs based on the top BLAST hit. Gene ontology (GO) annotation and enrichment analysis was then followed using Blast2GO. The level 2 GO terms associated with transcriptome contigs were retrieved, and then, the annotation result was categorized as biological process, molecular function, and cellular components.

Gene Expression and Differentially Expressed Gene

Audic's method was used to identify differentially expressed genes between two libraries (Audic and Claverie 1997). The threshold for the P value was determined using false discovery rate (FDR) and was widely set at 0.05 (Benjamini and Yekutieli 2001). In this study, “FDR <0.05 ” and “XX_FPKM=0 or XY_FPKM=0” were used to classify as specifically expressed genes (SEGs). “FDR <0.05 ” and “ $|\log_2(\text{XX_FPKM}/\text{XY_FPKM})|>1$ or $|\log_2(\text{XY_FPKM}/\text{XX_FPKM})|>1$ ” were used to classify as differentially expressed genes (DEGs). Those meeting “XX_FPKM <2 and XY_FPKM <2 ” statistical criteria were classified as non-expressed genes (NEGs), whereas all there remaining ones were designated as co-expressed genes (CEGs). In this way, all genes were classified into four types: SEGs, DEGs, CEGs, and NEGs.

SSR and SNV Detection

The SSR detection was applied using msatcommander toolkits (Faircloth 2008). The sequences composed of two or more repeat units with motifs separated by 100 bp were considered to be two or more microsatellites. Only microsatellite sequences with flanking sequences of >50 bp on both sides were collected for future primer designing. The SNP and structure variation detection was employed using bwa/SAMtool (Li et al. 2009).

Experimental Validation Using qRT-PCR

To verify the expression profile of the differentially expressed genes selected using bioinformatics method, we have performed quantitative real-time PCR (qRT-PCR). Total RNAs were reverse-transcribed using reverse transcriptase (Invitrogen, Carlsbad, CA, USA). All primers were designed using Primer Premier 6. The SYBR Green I Master Mix (TaKaRa, Dalian, China) was used for qRT-PCR using an Applied Biosystems Prism 7500 Fast Real-Time PCR system. One denaturation cycle was performed at 95°C for 5 min, and 40 amplification cycles were performed at

Table 1 Statistical summary of yellow catfish transcriptome data

Sequencing	Number of reads	209,599,618
Assembly	Total bases	20.96 Gb
	Cleaned reads	191,071,480
	Number of contigs	154,507
	Average contig length	1,014.05
	N50 value	2,101
Annotation	N90 value	349
	Contigs with blast results	150,291
	Unigenes with blast result	27,226
	Contig with GO terms	64,794
	Unigenes with GO terms	19,039

95 °C for 15 s and 60 °C for 30 s. Results were expressed relative to the expression levels of beta-actin in each sample using the Relative Expression Software Tool (REST) version 2009. Expression differences between groups were assessed for statistical significance using R software. All samples were run in triplicate. The averages of three relative quantities of biological replications were used in a two-tailed Student's *t* test with a 95 % confidence level ($P < 0.05$) to determine the significance with respect to gene expression.

Results

Transcriptome Sequencing and Assembly

To generate a comprehensive reference transcriptome of the yellow catfish and to investigate gene expression variations between sexes in the brain, we prepared four cDNA libraries from male and female pooled brain RNAs of 1-year- and 2-year-old yellow catfish, and then they were sequenced by using Illumina HiSeq 2000 technology. In total, 209,599,618, 100-bp-long paired-end reads were obtained. Then, the low-quality sequences and ambiguous nucleotides were removed, with remaining 191,071,480 clean reads for Trinity de novo assembling (Table 1). A total of 154,507 contigs were assembled, ranging from 201 to 27,822 bp with the average length of 1,014.05 bp, the N50 value of 2,101 bp, and the N90 value of 349 bp (Table 1). The raw transcriptome sequences in this study have been uploaded in the NCBI SRA site with the accession number SRR1103702. In addition, the percentage of GC content was 46.58 %. The transcriptome data analysis workflow was shown in Fig. 1.

Gene Identification and Functional Annotation

BLAST-based gene identification was performed to annotate the yellow catfish transcriptome and identify genes with sexually dimorphic expression. BLASTX was used for the assembled contig gene identification by searching against the reference protein sequences, including zebrafish (*Danio rerio*), medaka (*Oryzias latipes*), stickleback (*Gasterosteus aculeatus*), and tetraodon (*Tetraodon nigroviridis*). There were 150,291 contigs corresponding to 27,226 unigenes, with an *E* value cutoff less than $1e-10$. A total number of 64,794 assembled contigs were assigned at least one GO term, corresponding to 19,039 unigenes for describing biological processes, molecular functions, and cellular components (Table 1).

The differentially expressed unique genes were then used as input to perform GO annotation by Blast2GO

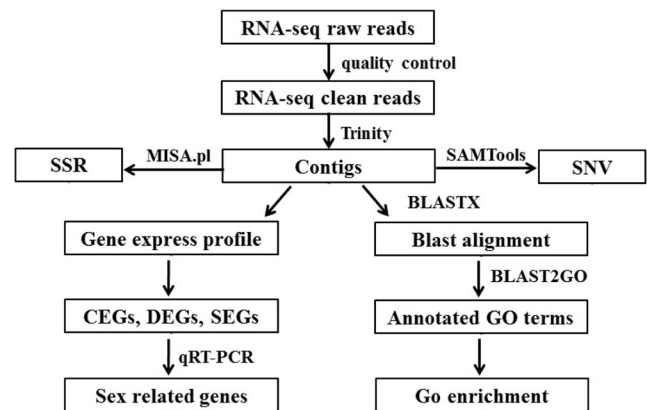


Fig. 1 Yellow catfish transcriptome analysis pipeline

(Conesa et al. 2005). The GO annotations were plotted in Fig. 2. Of these, the number of GO terms in biological process ontology was the most prevalent, followed by the cellular component ontology and the molecular function. Briefly, for biological processes, genes involved in cellular processes and metabolic process were highly represented. For the cellular component, cell and cell part were the most represented categories. Binding was the most represented GO term, followed by molecular transducer activity for molecular functions. Interestingly, within biological processes, a total of 11 unigenes were annotated to reproduction (GO, 0000003), 88 unigenes were annotated to developmental process (GO, 0032502), 12 unigenes were annotated to growth (GO, 0040007), and 11 unigenes were annotated to reproductive process (GO, 0022414).

Then, we performed sex-biased GO term analysis. For male-biased genes, the significant enrichment GO terms were related to nucleoside/ribonucleoside and other complex catabolic process, responding to environment or ion. Meanwhile, the female-biased GO terms were related to DNA modification, meiosis, and cell adhesion. Interestingly, many female-biased GO terms were annotated to female gamete generation (GO, 0007292), oogenesis (GO, 0048477), and germplasm (GO, 0060293). The sex-biased GO terms were shown in Supplementary Tables 1 and 2.

Assessment of Transcriptome Assembly

In order to assess the quality of the yellow catfish transcriptome assembly, the assembled contigs were compared with Refseq proteins of zebrafish, medaka, stickleback, and tetraodon using BLASTX with an *E* value cutoff of $1e-10$. In total, 80,659 (52.2 %), 66,344 (42.9 %), 68,871 (44.6 %), and 65,507 (42.4 %) significant hits were identified with those fish species, respectively (Table 2). In addition, the total numbers of 16,

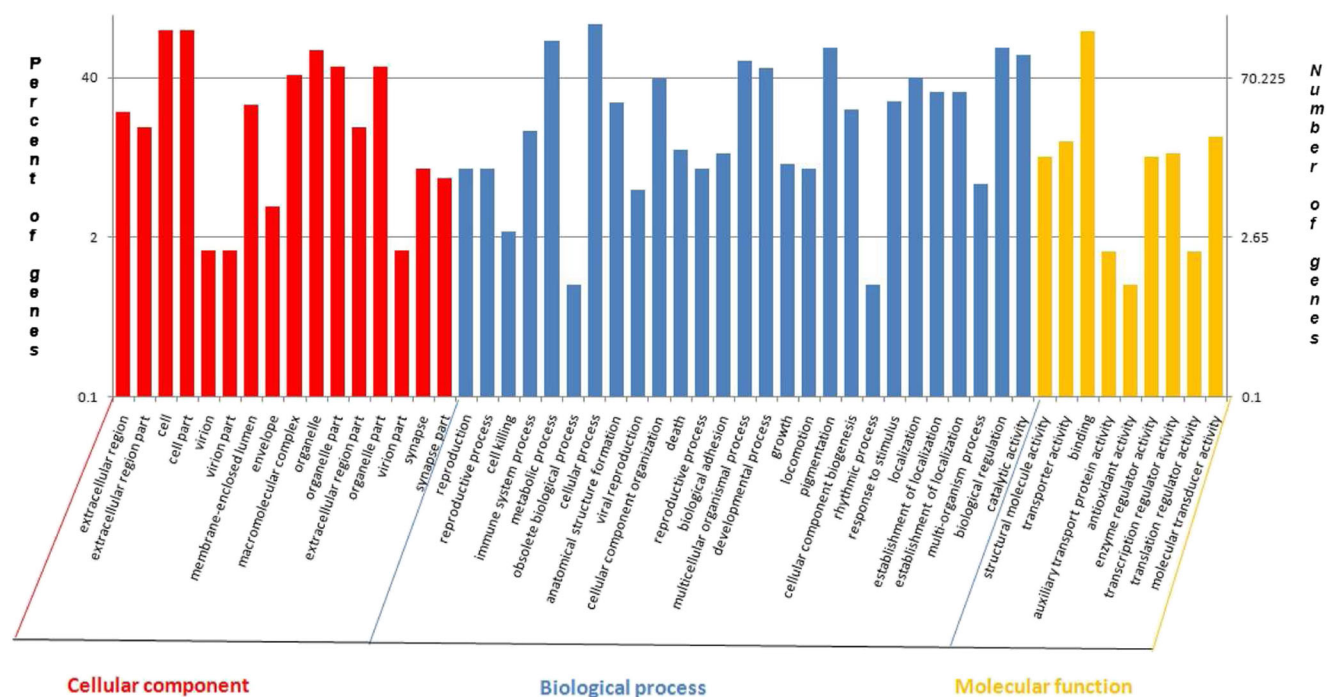


Fig. 2 Gene ontology categories pattern of the differentially expressed unigenes in yellow catfish. Distribution of the GO categories assigned to the differentially expressed unigenes in yellow catfish. Unique transcripts

(unigenes) were annotated in three categories: cellular components, molecular functions, and biological processes

351, 10,989, 12,508, and 11,523 unigenes were retained for zebrafish, medaka, stickleback, and tetraodon after filtering the repeat isogenes. The corresponding uniprotein numbers of these four fish are 20,978 (zebrafish), 15,935 (medaka), 16,918 (stickleback), and 15,579 (tetraodon), respectively (Table 2). We again noted that zebrafish had the highest similarity to yellow catfish, which is consistent with our previous findings (Lu et al. 2014).

SSRs and SNV Detection Through RNA Sequencing

For further assessment of the assembly quality and development of new molecular markers, including di-, tri-, tetra-, and penta-nucleotide SSRs with a minimum of

four repetitions for all motifs, the SSRs that were only located in one single read had been filtered. A total number of 90,210 microsatellites were identified from 154,507 contigs. Among these microsatellites, di-nucleotide repeat motifs were the most abundant type (72,143, 80.0 %), followed by tri-, tetra-, and penta-nucleotide repeat motifs (13,676 (15.2 %), 4,273 (4.7 %), and 118 (1.3 %); see Table 3). Eight thousand seven hundred forty-five SSRs are gene associated. These SSRs provide plenty molecular information to design polymorphic primers for further genotyping analysis.

For extended application of the RNA-Seq data, structure variations were discovered using the assembled transcriptome. The short reads of RNA-Seq data were aligned onto the reference transcriptome of yellow catfish, and the total number of insertion and deletion (INDEL) varies from 2,343 to 2,992 from four different

Table 2 The assessment of yellow catfish transcriptome assembly

Fish species	Blast hits number	Unigene number	Uniprotein number
Zebrafish	80,659	16,351	20,978
Medaka	66,344	10,989	15,935
Stickleback	68,871	12,508	16,918
Tetraodon	65,507	11,523	15,579

Unigene number is the number of unique gene after removing duplicated genes. Uniprotein number is the number of unique protein after removing duplicated proteins

Table 3 Statistics of microsatellite identified from yellow catfish transcriptome

Total number of contigs	154,507
Di-nucleotide repeats	72,143
Tri-nucleotide repeats	13,676
Tetra-nucleotide repeats	4,273
Penta-nucleotide repeats	118
Microsatellite total number	90,210

Table 4 Statistics of INDELs and SNPs identified from yellow catfish transcriptome

Sample	INDEL number			SNP number		
	Insertion	Deletion	Subtotal	Synonymous	Non-synonymous	Subtotal
YC1XX ^a	1,661	682	2,343	2,603	18,608	21,211
YC1XY ^b	2,043	949	2,992	2,493	21,148	23,641
YC2XX ^c	1,836	744	2,580	4,604	11,836	16,440
YC2XY ^d	1,877	749	2,626	4,586	12,546	17,132

^a One-year-old female yellow catfish^b One-year-old male yellow catfish^c Two-year-old female yellow catfish^d Two-year-old male yellow catfish

yellow catfish samples. The insertion number varies from 1,661 to 2,043, and the deletion number varies from 682 to 949, respectively. The total number of SNPs varies from 16,440 to 23,641 from four yellow catfish samples. The synonymous number varies from 2,493 to 4,604, and the non-synonymous number varies from 11,836 to 21,148, respectively (Table 4). This availability of microsatellites and structure variants developed in this study serves as precious molecular markers for yellow catfish genetics research.

Sex-Related Gene Expression Profiling in Brain

According to the classification standard described in the method, a total of 27,226 unigenes were classified into four gene expression categories. For 1-year-old yellow catfish, 12,510 unigenes were identified as CEGs in both XX and XY brain transcriptome, and 39 and 19 unigenes were detected to be expressed differentially in XX or XY brain noted as DEGs, respectively (Table 5 and Fig. 3). For 2-year-old adults, 12,503 unigenes were detected as CEGs, and 113 and 186 unigenes were identified as DEGs in XX and XY, respectively. While the numbers of XX-SEGs at 1- and 2-year-olds were 19 and 13, respectively, the corresponding numbers for XY-SEGs were 13 and 54, respectively (Table 5 and Fig. 3).

Table 5 Three types of gene expression profile of yellow catfish

	CEGs	DEGs	SEGs
YC1XX	12,510	39	19
YC1XY		19	13
YC2XX	12,503	113	13
YC2XY		186	54

CEGs co-expressed genes, DEGs differentially expressed genes, SEGs specifically expressed genes

The number of DEGs from two developmental stages was greater than SEGs. The number of DEGs was negatively correlated with that of CEGs expressed at two stages. More SEGs and DEGs were observed in XY brain in 2-year-old yellow catfish. However, in 1-year juvenile, more SEGs were detected in the female. The scatterplots of the gene expression profiles also revealed that there were more up-regulated genes in XY than in XX in 2-year-old but similar gene expression pattern of male and female in 1-year-old (Table 3 and Fig. 4).

Despite the significance in gene expression observed between sexes, CEGs at two developmental stages still made up the majority of all unigenes. However, compared to gonad (Lu et al. 2014), less DEGs and SEGs but more CEGs were found in the brain of yellow catfish.

Identification of Sexually Dimorphic Expression Genes

The brain has pronounced sexual dimorphism in function in mammals. In 1-year-old brain tissue, four sex differentially overexpressed genes were detected in female, including CYP1A (Negrato et al. 2013; Hasselberg et al. 2004), ZP3 (Yuan et al. 2013; Mold et al. 2009), org (oogenesis-related gene), and lyceraldehyde 3-phosphate dehydrogenase (GAPDH). Sperm acrosome-associated 4 genes (SPACA4) (Shetty et al. 2003) as the only DEG was identified to be overexpressed in male (Table 6). The FIGLA gene was specifically expressed in female, whereas prostate stem cell antigen (PSCA) gene was specifically expressed in male (Table 6).

For 2-year-old yellow catfish, we have found 12 sex differentially expressed genes in male, including CYP1A, 20BHS, GTHa, zonadhesin (ZAN) (Hunt et al. 2005), growth hormone (GH), cytochrome P450 arom, leptin receptor (LEPR) (Liu et al. 2010; Gong et al. 2013), Sox9a2 (Nakamoto et al. 2009), insulin-like growth factor binding

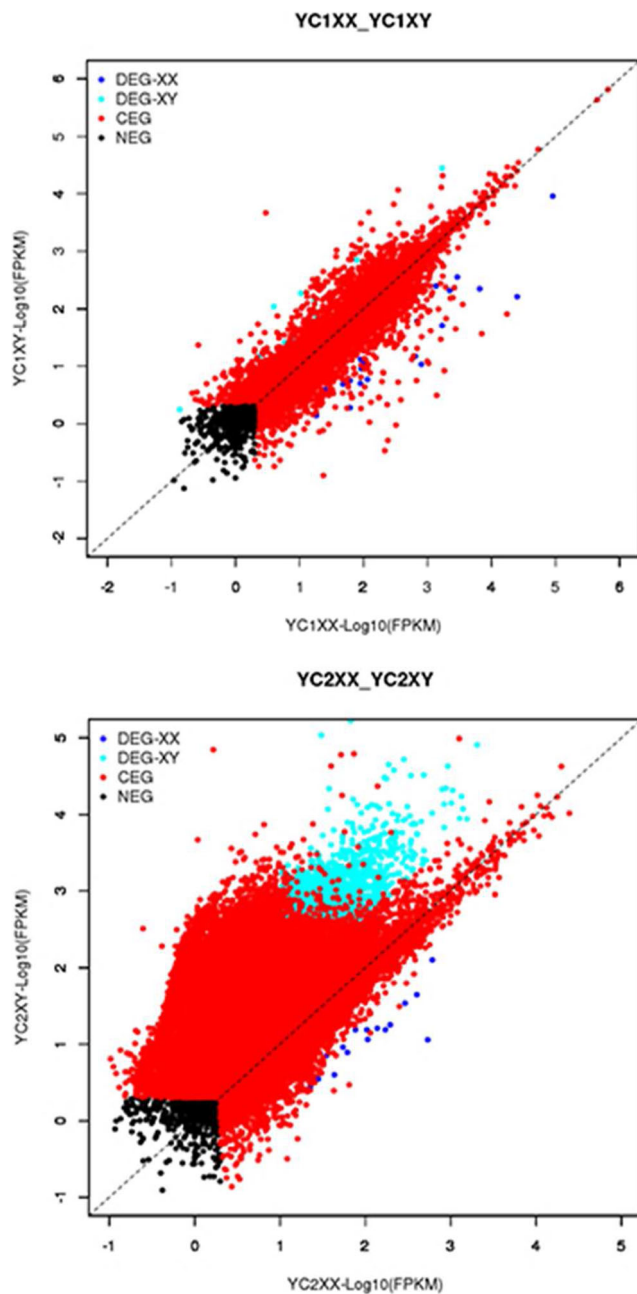


Fig. 3 Gene expression profiling in gonad tissue at two developmental stages. The number of DEGs from two developmental stages was greater than SEGs. The number of DEGs was negatively correlated with that of CEGs expressed at two stages. More SEGs and DEGs were observed in XY brain in 2-year-old yellow catfish. However, in 1-year juvenile, more SEGs were detected in the female

protein 1 (IGFBP1) (Yu et al. 2013), HSP70 (Mayer and Bukau 2005; Tao et al. 2013), RPL15, and myogenin. Only one sex differentially expressed gene was detected in female, which is ppiD (Table 7). GAPDH and GSG1L were found specifically expressed in female, whereas gamete and mating-type-specific protein A (gmsA) and meiosis inhibitor 1 (MEI1) were specifically expressed in male (Table 7).

Validation of RNA-Seq Results by qPCR

In order to validate the sexually dimorphic expressed genes identified by RNA-Seq, we randomly selected eight genes from those with differentially expression patterns and from DEGs and SEGs based on their function for qRT-PCR validation, including oogenesis, folliculogenesis, IGFBP1, MEI1, P450, SOX9a2, GTHa, and CYP1A. Comparing the transcriptome data with the qRT-PCR results from seven selected differentially expressed genes, qRT-PCR results were consistent with RNA-Seq results at two developmental stages (Fig. 4). Overall, the qRT-PCR results indicated the reliability and accuracy of the Trinity reference assembly and the RNA-Seq-based transcriptome expression analysis.

Discussions

The discovered genes provide a baseline for understanding the brain sexual divergence in yellow catfish. In this work, we have sequenced, assembled, and annotated the brain transcriptome of yellow catfish using Illumina HiSeq 2000 sequencing technology. Among those differentially expressed genes, many genes were identified involving sex determination. In addition, a large number of SSRs and SNVs were discovered as genetic markers. We then validated seven selected sexually dimorphic expressed genes by qRT-PCR. The results showed that the sex-related genes were significantly associated with sex determination and sex differentiation indicating the reliability and accuracy of our analysis.

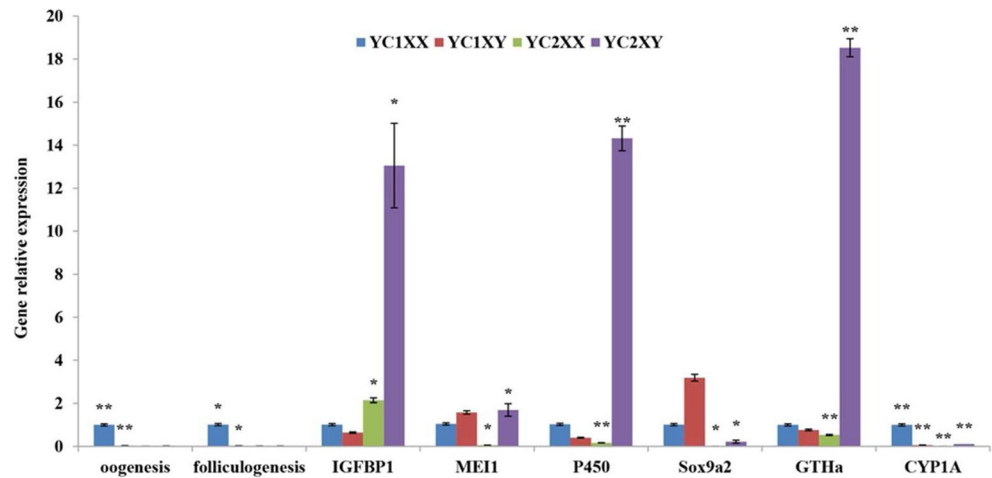
These genes identified are involved in development, metabolism, and biological regulation; some of them have been reported to be associated with sex determination. We identified more XX-related genes in 1-year-old yellow catfish; however, in 2-year-old, most of genes were found to be associated with XY. It is known that the brain of mammals had a remarkable sexual dimorphism in function which is consistent with our study. But the number of identified sex-related genes in the brain was less compared to the gonad, which is consistent with previous studies in other vertebrates (Yang et al. 2006; Mank et al. 2008) and fish species (Manousaki et al. 2014).

The sex-related differentially expressed genes were discussed as follows:

The Genes Overexpressed in Male Brains

SPACA4 also known as SAMP14, found only expressed in human testis (Shetty et al. 2003), was overexpressed in the 1-year-old male yellow catfish. PSCA is firstly

Fig. 4 Validation of gonad tissue transcriptome results by qRT-PCR using eight selected differentially expressed genes in two developmental stages. qRT-PCR fold changes are relative to control samples and normalized by changes in beta-actin values. The averages of three relative quantities of biological replications were used in a two-tailed Student's *t* test with a 95 % confidence level ($P<0.05$) to determine the gene expression significance



identified as a sex-related gene, specifically expressed in the male brain of 1-year-old yellow catfish in our study.

In 2-year-old yellow catfish, ZAN encodes a sperm protein that is involved in the binding of egg and sperm (Hunt et al. 2005; Hardy and Garbers 1995). LEPR was reported to be highly expressed in testes of zebrafish (Liu et al. 2010). Sox9a2 as a member of Sox9, found overexpressed in male of yellow catfish, has also been reported as a sex-related gene and overexpressed in male rainbow trout (Vizziano et al. 2007) and channel catfish (Sun et al. 2013). IGFBP1 as an insulin-like growth factor binding protein 1, and insulin-like growth factor (IGF), is very important to growth, development, and metabolic regulation in vertebrate. It is reported that IGFBP is critical to gonad development (Yi et al. 2001).

The expression of GH had no difference between male and female in 1-year-old yellow catfish; however, in 2-year-old, the expression in male was found higher than female, which may explain the sex growth dimorphism phenomenon that male grows faster than female in 2-year-old (Liu et al. 2013). RPL15 is firstly identified as a sex-related gene in brain as well as in gonad in yellow catfish, which was previously reported as a molecular marker in teleost phylogenesis (Schulze et al. 2005).

The Genes Overexpressed in Female Brains

At the developmental stage of 1-year-old, FIGLA, as a folliculogenesis-specific basic helix-loop-helix gene, is identified as a female-related gene. It is considered critical to the development of female germline (Soyal et al. 2000; Pala et al. 2009) and also found to play an important role in XX medaka individuals (Matsuda 2003). ZP3, as a member of zona pellucid (ZP) family, was identified as a female-specific gene in zebrafish (Jiang et al. 2012) and Nile tilapia (Eshel et al. 2014).

In 2-year-old yellow catfish, ppiD identified as DEG and GAPDH and germ cell-specific gene 1 identified as SEGs in female yellow catfish have not been reported as for its role in sex determination.

Compared to our previous study (Lu et al. 2014), CYP1A, GAPDH, GH, HSP70, and RPL15 identified in this study were found as sex-related genes in the brain as well as in the gonad, while other sex-related genes are considered as novel genes firstly found in yellow catfish. In terms of SOX9 family, SOX9a1 and Sox9a2 were identified in the gonad of yellow catfish, whereas only sox92a was found here. IGFBP1 were found in the gonad; however, in the brain, IGFBP1 were

Table 6 Differentially expressed genes and specifically expressed genes list in 1-year-old yellow catfish

Sequence ID	FPKM1 (YC1XX)	FPKM2 (YC1XY)	Gene description	Enrichment
comp1014_c0_seq1	2.50E-7	1,331.06	Cytochrome P450 1A (CYP1A)	♀DEG
comp42555_c1_seq1	29,897.90	167.78	Zona pellucida sperm-binding protein 3-like (ZP3)	♀DEG
comp347596_c0_seq1	8,982.63	337.89	Oogenesis-related gene (org)	♀DEG
comp627131_c0_seq1	3,050.23	314.02	Glyceraldehyde 3-phosphate dehydrogenase (GAPDH)	♀DEG
comp157281_c0_seq1	14.52	282.80	Sperm acrosome membrane-associated protein 4-like (SPACA4)	♂DEG
comp368208_c0_seq1	1,383.14	0	Folliculogenesis-specific basic helix-loop-helix (FIGLA)	♀SEG
comp521437_c0_seq1	0	2,120.74	Prostate stem cell antigen precursor-like (PSCA)	♂SEG

FPKM1 gene expression read value in 1-year-old female yellow catfish, FPKM2 gene expression read value in 1-year-old male yellow catfish

Table 7 Differentially expressed genes and specifically expressed genes list in 2-year-old yellow catfish

Sequence ID	FPKM1 (YC2XX)	FPKM2 (YC2XY)	Gene discription	Enrichment
comp19602_c0_seq4	22.61	2.77	Peptidyl-prolyl cis-trans isomerase D (ppiD)	♀DEG
comp3119_c0_seq1	54.36	4,762.43	Cytochrome P450 1A (CYP1A)	♂DEG
comp14556_c0_seq1	96.95	5,826.18	20 beta-hydroxysteroid dehydrogenase (20BHSD)	♂DEG
comp95107_c0_seq1	16.04	769.38	Gonadotropin alpha subunit (GTHa)	♂DEG
comp10436_c0_seq1	170.31	3,218.14	Zonadhesin-like (ZAN)	♂DEG
comp7131_c0_seq1	225.28	4,158.40	Growth hormone (GH)	♂DEG
comp17721_c0_seq1	95.6	1,732.57	Cytochrome P450 arom (CYP)	♂DEG
comp15039_c0_seq1	38.77	664.98	Leptin receptor (LEPR)	♂DEG
comp4302_c0_seq1	72.28	1,034.19	HMG box transcription factor Sox9a2 (Sox9a2)	♂DEG
comp2531_c0_seq1	404.08	5,427.10	Insulin-like growth factor binding protein 1 (IGFBP1)	♂DEG
comp31477_c1_seq3	31.17	362.04	Heat shock protein 70 (HSP70)	♂DEG
comp11313_c0_seq1	100.56	796.7	Ribosomal protein L15 (RPL15)	♂DEG
comp7379_c0_seq1	23.84	161.26	Myogenin	♂DEG
comp666188_c0_seq1	24.17	0	GAPDH protein	♀SEG
comp207731_c0_seq1	13.02	0	Germ cell-specific gene 1-like protein-like (GSG1L)	♀SEG
comp261806_c0_seq1	0	307.89	Gamete and mating-type-specific protein A-like (gmsA)	♂SEG
comp143101_c0_seq1	0	8.42	Meiosis inhibitor protein 1-like (MEI1)	♂SEG

FPKM1 gene expression read value in 2-year-old female yellow catfish, *FPKM2* gene expression read value in 2-year-old male yellow catfish

identified. Further investigation is warranted to study the functions of these genes in terms of sex determination.

Conclusions

This is the first report of yellow catfish brain transcriptome using RNA-Seq technology. In this study, we have assembled and annotated the brain transcriptome and have identified novel sexually dimorphic expressed genes of the yellow catfish. These genes may be candidate genes involved in sex determination. The validation by qRT-PCR indicated the reliability and accuracy of our analysis. In addition, we have also identified many SSRs and SNVs. These genetic markers will assist our understanding of the sex determination mechanism of yellow catfish. Our analysis of sex-related gene expression reveals differences that are likely to be pertinent for the mechanism underlying sex dimorphism.

Acknowledgments This research was supported by grants from the National Science Foundation of China (31302177), Heilongjiang River Fisheries Research Institute Foundation (HSY201301), and National Science Foundation of Heilongjiang Province (LC201419).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Agate RJ, Grisham W, Wade J, Mann S, Wingfield J, Schanen C, Palotie A, Arnold AP (2003) Neural, not gonadal, origin of brain sex differences in a gynandromorphic finch. *Proc Natl Acad Sci U S A* 100:4873–4878
- Agate RJ, Choe M, Arnold AP (2004) Sex differences in structure and expression of the sex chromosome genes CHD1Z and CHD1W in zebra finches. *Mol Biol Evol* 21:384–396
- Audic S, Claverie JM (1997) The significance of digital gene expression profiles. *Genome Res* 7:986–995
- Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29:1165–1188
- Chen X, Agate RJ, Itoh Y, Arnold AP (2005) Sexually dimorphic expression of *trkB*, a Z-linked gene, in early posthatch zebra finch brain. *Proc Natl Acad Sci U S A* 102:7730–7735
- Chen X, Mei J, Wu J, Jing J, Ma W, Zhang J, Dan C, Wang W, Gui JF (2015) A comprehensive transcriptome provides candidate genes for sex determination/differentiation and SSR/SNP markers in yellow catfish. *Mar Biotechnol (NY)* 17:190–198
- Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676
- Davies W, Wilkinson LS (2006) It is not all hormones: alternative explanations for sexual differentiation of the brain. *Brain Res* 1126:36–45
- Dennis C (2004) Brain development: the most important sexual organ. *Nature* 427:390–392
- Dewing P, Shi T, Horvath S, Vilain E (2003) Sexually dimorphic gene expression in mouse brain precedes gonadal differentiation. *Brain Res Mol Brain Res* 118:82–90
- Dewing P, Chiang CW, Sinchak K, Sim H, Fernagut PO, Kelly S, Chesselet MF, Micevych PE, Albrecht KH, Harley VR, Vilain E (2006) Direct regulation of adult brain function by the male-specific factor SRY. *Curr Biol* 16:415–420

- Eshel O, Shirak A, Dor L, Band M, Zak T, Markovich-Gordon M, Chalifa-Caspi V, Feldmesser E, Weller JI, Seroussi E, Hulata G, Ron M (2014) Identification of male-specific amh duplication, sexually differentially expressed genes and microRNAs at early embryonic development of Nile tilapia (*Oreochromis niloticus*). *BMC Genomics* 15:774
- Faircloth BC (2008) msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol Ecol Resour* 8:92–94
- Francis RC, Barlow GW (1993) Social control of primary sex differentiation in the Midas cichlid. *Proc Natl Acad Sci U S A* 90:10673–10675
- Gong Y, Luo Z, Zhu QL, Zheng JL, Tan XY, Chen QL, Lin YC, Lu RH (2013) Characterization and tissue distribution of leptin, leptin receptor and leptin receptor overlapping transcript genes in yellow catfish *Pelteobagrus fulvidraco*. *Gen Comp Endocrinol* 182:1–6
- Hardy DM, Garbers DL (1995) A sperm membrane protein that binds in a species-specific manner to the egg extracellular matrix is homologous to von Willebrand factor. *J Biol Chem* 270:26025–26028
- Hasselberg L, Meier S, Svardal A, Hegelund T, Celander MC (2004) Effects of alkylphenols on CYP1A and CYP3A expression in first spawning Atlantic cod (*Gadus morhua*). *Aquat Toxicol* 67:303–313
- Hunt PND, Wilson MD, Von Schalburg KR, Davidson WS, Koop BF (2005) Expression and genomic organization of zonadhesin-like genes in three species of fish give insight into the evolutionary history of a mosaic protein. *BMC Genomics* 6. doi:10.1186/1471-2164-6-165
- Jiang N, Jin X, He J, Yin Z (2012) The roles of follistatin 1 in regulation of zebrafish fecundity and sexual differentiation. *Biol Reprod* 87:54
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing, S (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079
- Liu Q, Chen Y, Copeland D, Ball H, Duff RJ, Rockich B, Londraville RL (2010) Expression of leptin receptor gene in developing and adult zebrafish. *Gen Comp Endocrinol* 166:346–355
- Liu H, Guan B, Xu J, Hou C, Tian H, Chen H (2013) Genetic manipulation of sex ratio for the large-scale breeding of YY super-male and XY all-male yellow catfish (*Pelteobagrus fulvidraco* (Richardson)). *Mar Biotechnol* (NY) 15:321–328
- Lu J, Luan P, Zhang X, Xue S, Peng L, Mahbooband S, Sun X (2014) Gonadal transcriptomic analysis of yellow catfish (*Pelteobagrus fulvidraco*): identification of sex-related genes and genetic markers. *Physiol Genomics* 46:798–807
- Maehiro S, Takeuchi A, Yamashita J, Hiraki T, Kawabata Y, Nakasone K, Hosono K, Usami T, Paul-Prasanth B, Nagahama Y, Oka Y, Okubo K (2014) Sexually dimorphic expression of the sex chromosome-linked genes *cntfa* and *pdlim3a* in the medaka brain. *Biochem Biophys Res Commun* 445:113–119
- Mank JE, Hultin-Rosenberg L, Webster MT, Ellegren H (2008) The unique genomic properties of sex-biased genes: insights from avian microarray data. *BMC Genomics* 9:148
- Manousaki T, Tsakogiannis A, Lagnel J, Sarropoulou E, Xiang JZ, Papandroulakis N, Mylonas CC, Tsigenopoulos CS (2014) The sex-specific transcriptome of the hermaphrodite sparid sharpnose seabream (*Diplodus puntazzo*). *BMC Genomics* 15. doi:10.1186/1471-2164-15-655
- Matsuda M (2003) Sex determination in fish: lessons from the sex-determining gene of the teleost medaka, *Oryzias latipes*. *Develop Growth Differ* 45:397–403
- Mayer MP, Bukau B (2005) Hsp70 chaperones: cellular functions and molecular mechanism. *Cell Mol Life Sci* 62:670–684
- Mold DE, Diniz AE, Sambandan DR (2009) Regulation of zebrafish zona pellucida gene activity in developing oocytes. *Biol Reprod* 81:101–110
- Munday PL, Buston PM, Warner RR (2006) Diversity and flexibility of sex-change strategies in animals. *Trends Ecol Evol* 21:89–95
- Nakamoto M, Muramatsu S, Yoshida S, Matsuda M, Nagahama Y, Shibata N (2009) Gonadal sex differentiation and expression of *Sox9a2*, *Dmrt1*, and *Foxl2* in *Oryzias luzonensis*. *Genesis* 47:289–299
- Negrato E, Vascellari M, Capolongo F, Binato G, Da Dalt L, Boscolo Papo M, Gioacchini G, Carnevali O, Bertotto D, Radaelli G, Pascoli F (2013) Expression of 8-OHDG in *Zosterisessor ophiocephalus* from the Venetian lagoon, Italy. *Eur J Histochem* 57:e8.
- Pala I, Scharl M, Thorsteinsdottir S, Coelho MM (2009) Sex determination in the *Squalius alburnoides* complex: an initial characterization of sex cascade elements in the context of a hybrid polyploid genome. *PLoS One* 4:e6401
- Parhar IS, Tosaki H, Sakuma Y, Kobayashi M (2001) Sex differences in the brain of goldfish: gonadotropin-releasing hormone and vasotocinergic neurons. *Neuroscience* 104:1099–1110
- Park IS, Kim JH, Cho SH, Kim DS (2004) Sex differentiation and hormonal sex reversal in the bagrid catfish *Pseudobagrus fulvidraco* (Richardson). *Aquaculture* 232:183–193
- Piferrer F, Ribas L, Diaz N (2012) Genomic approaches to study genetic and environmental influences on fish sex determination and differentiation. *Mar Biotechnol* 14:591–604
- Santos EM, Kille P, Workman VL, Paull GC, Tyler CR (2008) Sexually dimorphic gene expression in the brains of mature zebrafish. *Comp Biochem Physiol A Mol Integr Physiol* 149:314–324
- Schulze SR, Sinclair DA, Fitzpatrick KA, Honda BM (2005) A genetic and molecular characterization of two proximal heterochromatic genes on chromosome 3 of *Drosophila melanogaster*. *Genetics* 169:2165–2177
- Sellers MJ, Trewin C, McWilliam SM, Graves RS, Hertzler PL (2015) Transcriptome profiles of *Penaeus (Marsupenaeus) japonicus* animal and vegetal half-embryos: identification of sex determination, germ line, mesoderm, and other developmental genes. *Mar Biotechnol* (NY) 17:252–265
- Shetty J, Wolkowicz MJ, Digilio LC, Klotz KL, Jayes FL, Diekmann AB, Westbrook VA, Farris EM, Hao ZL, Coonrod SA, Flickinger CJ, Herr JC (2003) SAMP14, a novel, acrosomal membrane-associated, glycosylphosphatidylinositol-anchored member of the Ly-6/urokinase-type plasminogen activator receptor superfamily with a role in sperm-egg interaction. *J Biol Chem* 278:30506–30515
- Soyal SM, Amleh A, Dean J (2000) FIGalpha, a germ cell-specific transcription factor required for ovarian follicle formation. *Development* 127:4645–4654
- Sreenivasan R, Cai M, Bartfai R, Wang X, Christoffels A, Orban L (2008) Transcriptomic analyses reveal novel genes with sexually dimorphic expression in the zebrafish gonad and brain. *PLoS One* 3:e1791
- Sun F, Liu S, Gao X, Jiang Y, Perera D, Wang X, Li C, Sun L, Zhang J, Kaltenboeck L, Dunham R, Liu Z (2013) Male-biased genes in catfish as revealed by RNA-Seq analysis of the testis transcriptome. *PLoS One* 8:e68452
- Tao WJ, Yuan J, Zhou LY, Sun LN, Sun YL, Yang SJ, Li MH, Zeng S, Huang BF, Wang DH (2013) Characterization of gonadal transcriptomes from Nile tilapia (*Oreochromis niloticus*) reveals differentially expressed genes. *PLoS One* 8:e63604
- Tzika AC, Helaers R, Schramm G, Milinkovitch MC (2011) Reptilian-transcriptome v1.0, a glimpse in the brain transcriptome of five divergent Sauropsida lineages and the phylogenetic position of turtles. *EvoDevo* 2:19
- Vizziano D, Randuineau G, Baron D, Cauty C, Guiguen Y (2007) Characterization of early molecular sex differentiation in rainbow trout, *Oncorhynchus mykiss*. *Dev Dyn* 236:2198–2206
- Vizziano-Cantonnet D, Anglade I, Pellegrini E, Guiguen MM, Fostier A, Guiguen Y, Kah O (2011) Sexual dimorphism in the brain aromatase

- expression and activity, and in the central expression of other steroidogenic enzymes during the period of sex differentiation in monosex rainbow trout populations. *Gen Comp Endocrinol* 170: 346–355
- Weltzien FA, Andersson E, Andersen O, Shalchian-Tabrizi K, Norberg B (2004) The brain-pituitary-gonad axis in male teleosts, with special emphasis on flatfish (*Pleuronectiformes*). *Comp Biochem Physiol A Mol Integr Physiol* 137:447–477
- Xu J, Burgoyne PS, Arnold AP (2002) Sex differences in sex chromosome gene expression in mouse brain. *Hum Mol Genet* 11:1409–1419
- Xu J, Deng X, Watkins R, Disteche CM (2008) Sex-specific differences in expression of histone demethylases Utx and Uty in mouse brain and neurons. *J Neurosci* 28:4521–4527
- Xu D, Lou B, Xu H, Li S, Geng Z (2013) Isolation and characterization of male-specific DNA markers in the rock bream *Oplegnathus fasciatus*. *Mar Biotechnol* (NY) 15:221–229
- Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L, Drake TA, Lusis AJ (2006) Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res* 16: 995–1004
- Yano A, Nicol B, Jouanno E, Guiguen Y (2014) Heritable targeted inactivation of the rainbow trout (*Oncorhynchus mykiss*) master sex-determining gene using zinc-finger nucleases. *Mar Biotechnol* (NY) 16:243–250
- Yi Z, Hathaway MR, Dayton WR, White ME (2001) Effects of growth factors on insulin-like growth factor binding protein (IGFBP) secretion by primary porcine satellite cell cultures. *J Anim Sci* 79:2820–2826
- Yu WK, Shi YF, Fong CC, Chen Y, Van De Merwe JP, Chan AKY, Wei F, Bo J, Ye R, Au DWT, Wu RSS, Yang MS (2013) Gender-specific transcriptional profiling of marine medaka (*Oryzias melastigma*) liver upon BDE-47 exposure. *Comp Biochem Physiol D Genomics Proteomics* 8:255–262
- Yuan C, Zhang YY, Hu GJ, Li M, Zheng Y, Gao JC, Yang YP, Zhou Y, Wang ZZ (2013) Expression of two zona pellucida genes is regulated by 17 alpha-ethinylestradiol in adult rare minnow *Gobiocypris rarus*. *Comp Biochem Physiol C Toxicol Pharmacol* 158:1–9