

Article

De Novo Assembly, Characterization and Comparative Transcriptome Analysis of the Mature Gonads in *Spinibarbus hollandi*

Chong Han ¹, Wenwei Huang ¹, Suhan Peng ¹, Jiangwei Zhou ¹, Huawei Zhan ¹, Yuying Zhang ¹, Wenjun Li ¹, Jian Gong ^{2,*} and Qiang Li ^{1,*}

¹ School of life Sciences, Guangzhou University, Guangzhou, 510006, China

² Key Laboratory for Water Quality and Conservation of the Pearl River Delta, Ministry of Education, School of Environmental Science and Engineering, Guangzhou University, Guangzhou 510006, China

* Correspondence: gong_jian@mails.uas.ac.cn (J.G.); lq512328@163.com (Q.L.); Tel.: +86-020-39366937 (J.G.); +86-020-39366913 (Q.L.)

Simple Summary: *Spinibarbus hollandi* has been recognized as an economically important aquaculture species in southeastern China. However, the information on its reproductive regulation is scarce. The long maturity period and low egg laying amount remain a major challenge for large-scale breeding of *S. hollandi*. In the current study, the gonad transcriptomes of *S. hollandi* were first obtained through Illumina Novaseq technology. By transcriptome comparing of ovary and testis, a lot of differential expression genes were identified and most of them were supposed to participate in gonad formation, differentiation and gametogenesis. Moreover, the expression levels of some typical reproductive genes indicated they might play similar roles in gonad differentiation and development of *S. hollandi*. Finally, the gonad transcriptome information can help researchers understand the regulatory functions of sex-related genes in *S. hollandi*.

Abstract: *Spinibarbus hollandi* is an important commercial aquaculture species in southeastern China, but with long maturity period and low egg laying amount. However, there has been little study of its gonad development and reproductive regulation, which limits aquaculture production. Here, for the first time, gonadal transcriptomes of male and female *S. hollandi* were analyzed. A total of 167,152 unigenes were assembled, with only 48,275 annotated successfully. After comparison, a total of 21,903 differentially expressed genes were identified between male and female gonads, of which 16,395 were upregulated and 5508 were downregulated in the testis. In addition, a large number of differentially expressed genes participating in reproduction, gonad formation and differentiation, and gametogenesis were screened out and the differential expression profiles of partial genes were further validated using quantitative real-time PCR. These results will provide basic information for further research on gonad differentiation and development in *S. hollandi*.

Keywords: *Spinibarbus hollandi*; transcriptome; gonad differentiation and development; sex

Citation: Han, C.; Huang, W.; Peng, S.; Zhou, J.; Zhan, H.; Zhang, Y.; Li, W.; Gong, J.; Li, Q. De Novo Assembly, Characterization and Comparative Transcriptome Analysis of the Mature Gonads in *Spinibarbus hollandi*. *Animals* **2023**, *13*, 166. <https://doi.org/10.3390/ani13010166>

Academic Editors: Albamaria Parmeggiani and Alberto Elmi

Received: 1 December 2022

Revised: 26 December 2022

Accepted: 29 December 2022

Published: 31 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Spinibarbus hollandi, belonging to the subfamily Barbinae of Cyprinidae, is an important commercial aquaculture species in southeastern China, which is mainly distributed in Guangdong, Guangxi, Hunan, Hubei, Fujian, Anhui and Taiwan Provinces. *S. hollandi* has high nutritional value and medicinal value, and is popular as a table fish [1]. *S. hollandi* is an omnivorous species, which feeds mainly on aquatic insects, small fish, shrimp, and even blue-green algae and organic detritus [2]. Although *S. hollandi* is easy-to-raise, its growth rate is relatively slow, which limits the aquaculture efficiency and popularization of *S. hollandi* [3]. In recent years, due to its good taste and beautiful appearance,

S. hollandi is increasingly becoming a popular fish with considerable economic value especially in south China. *S. hollandi* is a gonochoristic fish species, where males mature earlier than females. The age of first maturation of *S. hollandi* normally takes 3 to 4 years in south China. In addition, although *S. hollandi* is a cyprinid fish, only a few thousand eggs are laid at one time for female fish [4]. However, dozens to millions of eggs are laid at one time in grass carp and common carp. In short, long maturity period and low egg laying amount are critical factors limiting the yield increase of *S. hollandi*. Thus, it is extremely important to develop our understanding of the reproductive biology of this species.

For fish, gonadal development, differentiation and maturation are more complex than other vertebrates, which are easily influenced by the environmental factors [5]. In recent years, cumulative practices have shown that understanding the mechanisms of reproductive regulation is critically important for reproduction management and breeding practice. However, previous studies mainly focused on genetic resources, rearing conditions, morphology and molecular markers associated with growth of *S. hollandi* [3], whereas less attention was paid to gonadal development, differentiation, maturation and gametogenesis. Thus, to better understand the reproductive mechanisms and facilitate reproductive regulation, it is necessary to explore the functional genes associated with gonad development and differentiation, and gametogenesis.

Despite the rapid development of next generation sequencing (NGS) technology, the available genetic information on *S. hollandi* is rather limited. Only brain and muscle transcriptome data were published to explore the mechanism of starvation response and compensatory growth [6,7]. Up to now, nearly no reproduction-related gene was characterized and fundamental information of their expression is still lacking. Due to low cost and high throughput, NGS-based transcriptome sequencing has been a most effective method to generate a large number of transcripts and gene expression profiles in recent years [8]. It has been widely used for functional gene discovery and analysis of gene expression in teleosts. In addition, using gonad comparative transcriptome analysis, the expression patterns of reproduction-related genes have been revealed in many fish species such as Spotfin porcupine fish (*Diodon hystrix*) [9], spotted scat (*Scatophagus argus*) [10], spotted knifejaw (*Oplegnathus punctatus*) [11], yellow catfish (*Pelteobagrus fulvidraco*) [12], mandarin fish (*Siniperca chuatsi*) [13] and so on. All these studies characterized a lot of candidate reproduction-related genes and provided basic information for reproductive regulation.

In this study, for the first time, Illumina-based transcriptome sequencing and de novo assembly of female and male gonads was carried out in *S. hollandi*. After comparative transcriptome analysis, a large number of sex-biased genes were uncovered and differentially expressed genes involved in reproductive regulation were also identified and discussed. Our data will enrich the currently available genetic and genomic data on *S. hollandi*, and identify candidate genes participating in gonad differentiation, development, maturation, and gametogenesis.

2. Materials and Methods

2.1. Sample Collection

In this study, six 4-year-old *S. hollandi* individuals (three males and three females) were obtained from an aquaculture farm in Shaoguan, Guangdong province. The fish were anesthetized with MS-222 before euthanizing them. Gonads from each individual were collected and stored in liquid nitrogen immediately. All animal handling procedures and experimental protocols were approved by the Experimental Animal Ethics Committee of the Guangzhou University of China (No. GURBBB210405).

2.2. RNA Extraction and Library Construction

Total RNA was extracted from fresh frozen tissue using RNA isolater Total RNA Extraction Reagent (Vazyme, Nanjing, China) according to the manufacturer's instructions.

The RNA concentration and purity were detected preliminarily using the Nanodrop2000 (Thermo Scientific, Wilmington, DE, USA). Then, integrity detection and quantification of RNA was carried out by Agilent 4200 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Eukaryotic transcriptome libraries were constructed using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, Ipswich, MA, USA). Eukaryotic mRNAs were enriched using magnetic beads with Oligo (dT). In a high temperature environment with metal ions, the RNA was fragmented and the first cDNA chain was synthesized with random hexamers, followed by the addition of enzyme, buffer, dNTPs (dATP, dTTP, dGTP, dCTP) to synthesis the second chain of cDNA. The synthesized double-stranded cDNA was purified by magnetic beads and repaired, and additional A was ligated to the tail end of cDNA, and the sequencing connector was connected. The sorted magnetic beads were used for fragment size sorting, and the sorted fragments were enriched by PCR, and finally purified to obtain the final library.

2.3. Library Sequencing, De Novo Assembly and Annotation

Using the Illumina Novaseq 6000 (Illumina, Inc., San Diego, CA, USA) high-throughput sequencing platform, the qualified libraries were sequenced with the sequencing strategy was Pair-End 150 (PE150). The amount of sequencing data for each library was not less than 6 Gb.

In order to obtain high quality assembly results for subsequent analysis, the raw reads were filtered, and low-quality reads with unknown base N beyond 5% and joint contamination were discarded. The remaining reads, also called “Clean reads”, were saved in FASTQ format. Then, using Trinity software, de novo assembly was carried out [14].

The coding region of Unigene was predicted by three forward and reverse reading frames, which could produce altogether 6 kinds of coding protein sequences. After obtaining the coding protein sequences, the protein sequences were compared with the non-redundant protein sequences (Nr) (<https://www.ncbi.nlm.nih.gov/> (accessed on 20 April 2021)) and the Uniprot protein (<https://www.uniprot.org/> (accessed on 20 June 2021)) database. The encoding method with the best alignment (the largest alignment score) was chosen as the encoding method of the gene.

Using homology searches, unigenes were annotated against five major public databases including Nr database, the Clusters of eukaryotic Orthologous Groups (KOG, <http://www.ncbi.nlm.nih.gov/KOG> (accessed on 27 June 2021)) database, the Kyoto Encyclopedia of Genes and Genomes (KEGG, <http://www.genome.jp/kegg> (accessed on 17 July 2021)) and the Uniprot protein database using BLASTx or Diamond. The NR results were further annotated by Blast2GO [15], and the unigenes were further grouped (GO, <http://geneontology.org> (accessed on 25 July 2021)) according to the classification of biological process, cell composition and molecular function.

2.4. Identification of Differentially Expressed Genes and Enrichment Analysis

The clean reads of each sample are first mapped to the assembled transcripts using hisat2 v2.1.0 [16]. Then, the transcripts per million (TPM), fragments per kilobase of exon model per million mapped fragments (FPKM), coverage and other gene-expression-related values were further calculated using stringtie v1.3.3b according to the comparison results [17]. Finally, using python prepDE.py, the output result of stringtie software was converted into a format recognized by the “edgeR” package (V3.6) [18,19], and the differential gene expression was analyzed using edgeR. The genes with p -value < 0.05 and $|\log_2FC| > 2$ are set as significantly differentially expressed genes (DEGs). The number of differentially expressed genes was counted by statistical significance, and then the clusterProfiler program of the R package [20] was applied against the background of DEGs. Based on Fisher’s exact test and Benjamini correction, the pathways that were significantly enriched in differentially expressed genes compared with the whole genome background with p value < 0.05 were considered to be significantly enriched.

2.5. Validation of DEGs Using Quantitative Real-Time PCR

The quantitative real-time PCR (qRT-PCR) analysis was used to validate the reliability of these DEGs. A total of 14 differentially expressed reproduction-related genes were randomly selected for validation. First, based on the sequences of 14 differentially expressed reproduction-related genes and reference gene β -actin, specific primers were designed by Primer Premier 5.0 (Table 1). Then, the HiScript III RT SuperMix for qPCR (+gDNA wiper) (Vazyme, China) was used for synthesizing cDNA template. Each qRT-PCR was carried out using SYBR Green qPCR Mix (GDSBIO, China) on a LightCycle 480 system. The qRT-PCR reaction included an initial denaturation step at 95 °C for 3 min, followed by 40 cycles of 95 °C for 10 s, 60 °C for 20 s and 72 °C for 20 s, and a final extension at 72 °C for 5 min, ending with a dissociation curve process. Each sample was amplified in triplicate, and the expression of each gene was normalized using β -actin by the comparative CT method ($2^{-\Delta\Delta CT}$) [21] and was shown as mean \pm standard error.

Table 1. Primers of qRT-PCR.

Primer ID	Primer Sequences
GDCB-amh-QP1F	TCCACAGAGTCCCTGCTATCA
GDCB-amh-QP1R	CAATCACCTTTTGCCCCAC
GDCB-hsd11b2-QP1F	AGACAGGCTAAAGGCCGGA
GDCB-hsd11b2-QP1R	TGACGAAGTGTGTTGGTAAGAAGAT
GDCB-fshr-QP1F	CGCCTTCTTCAACCTCAC
GDCB-fshr-QP1R	GTTACTTGGAATCCTTTCTATG
GDCB-era-QP1F	GCATTTTCATTCTGCTCCAGTC
GDCB-era-QP1R	CATTCCTTTGTTGCTCATGTGTCT
GDCB-erbeta1-QP1F	CCGCTGGTGAAGAGGGTAGT
GDCB-erbeta1-QP1R	GCAGATGTAGTCATTGTGTCCTTGTAT
GDCB-erbeta2-QP1F	TGGAGAGCCGTGGGAAGGTT
GDCB-erbeta2-QP1R	TGTTGCTGAGGTGTCGAATATG
GDCB-ar-QP1F	CCCCCAGCAAGGGAATAATATG
GDCB-ar-QP1R	TCAGCGCCCCGAAAGAAA
GDCB-foxk2-QP1F	TGAAGCTGCCGAGAGTGTGC
GDCB-foxk2-QP1R	GGGATGTTGATGGTTAAAGGGG
GDCB-smad4a-QP1F	CCAGGGAGGAGAGAATGAGAGC
GDCB-smad4a-QP1R	CGAGCGTATATGACATGAGGGAA
GDCB-gdf6-QP1F	AGTGTCCAAGTCTGTTCTCTGCT
GDCB-gdf6-QP1R	CTCGTTATGGTGTGTTGCTGC
GDCB-bmp7-QP1F	CACCGCTCTAAAAACCCCA
GDCB-bmp7-QP1R	ATTCTCCCTCACAGTAATACGCA
GDCB-pax7-QP1F	AGGCCAGAGTGCAGGTTTGG
GDCB-pax7-QP1R	TGTGGTAATGTTGTTGAAGGATAGG
GDCB-pax8-QP1F	AAACCAGGAGTCATCGGGGG
GDCB-pax8-QP1R	CCTTTGCTGTCCAGAGGAAGG
GDCB-sox9-QP1F	AGGTCAGAGCTCCGGCTTGTACT
GDCB-sox9-QP1R	TGTGATTGGGTTGGGGAATGG
GDCB-sox5-QP1F	ACCAAAGCCCAAGGACGAGG
GDCB-sox5-QP1R	GCCGGAGGTGATGGATGACA
GDCB-beta-actin-QP1F	GTGTTGGCATAACAGGTCCTTACG
GDCB-beta-actin-QP1R	ACGGACAGGTCATCACCATTG

3. Results

3.1. Overview of Transcriptome Assembly Quality

A sum of six cDNA libraries was constructed in triplicates from testes (JC) and ovaries (LC). After quality control and data filtering, a total of 39.72 Gb clean reads was generated using an Illumina HiSeq platform, with a mean of 6.62 Gb, ranging from 6.0 to 7.28 Gb per sample. The mean values of Q20 and Q30 are higher than 95% (Table 2). After de novo assembly, 167,152 unigenes were assembled, with mean length and N50 of 871 bp and 1198 bp, respectively (Table 3). About 30.13% unigenes were in length of 300–500bp, and 33,368 of them (19.97%) were more than 1000 bp in length (Figure 1).

Table 2. Summary statistics of the gonadal RNA-Seq data.

Sample	Reads Number	Base Number	Total Base	Q20	Q30	GC
JC-K5	23,015,034	3,452,255,100	6,904,510,200	98.33%	95.25%	46.53%
	23,015,034	3,452,255,100		97.96%	94.12%	46.51%
JC-K8	24,277,601	3,641,640,150	7,283,280,300	98.29%	95.16%	46.56%
	24,277,601	3,641,640,150		97.86%	93.88%	46.54%
JC-K9	19,990,173	2,998,525,950	5,997,051,900	98.32%	95.21%	46.21%
	19,990,173	2,998,525,950		98.08%	94.42%	46.18%
LC-K1	23,109,316	3,466,397,400	6,932,794,800	98.11%	94.96%	48.05%
	23,109,316	3,466,397,400		97.85%	94.20%	48.03%
LC-K3	21,181,082	3,177,162,300	6,354,324,600	98.29%	95.23%	48.17%
	21,181,082	3,177,162,300		97.83%	94.01%	48.15%
LC-K9	20,827,911	3,124,186,650	6,248,373,300	98.42%	95.45%	48.03%
	20,827,911	3,124,186,650		98.12%	94.58%	48.00%

Table 3. Summary statistics of the *S. hollandi* gonadal transcriptome assembly.

	Number
Gene Number (#):	167,152
Total length (nt):	145,709,933
Average length (nt):	871
Max length (nt):	18,177
Min length (nt):	280
N50	1198

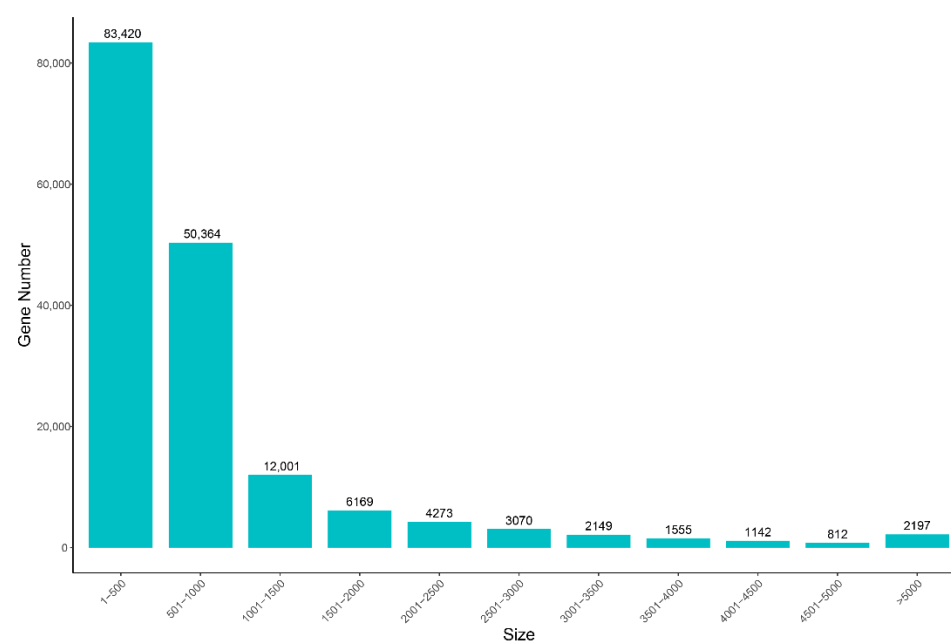


Figure 1. Length distribution of all assembled unigenes of the *S. hollandi* gonadal transcriptome. X-axis: size of unigene; Y-axis: number of unigenes.

3.2. Unigene Annotation

After annotation, a total of 48,275 unigenes were successfully annotated. The highest percentage of annotated unigenes was in the Uniprot database (47,343; 98.07%) and second is the Nr database (46,151; 95.60%), while the lowest was in the GO database (23,186; 48.03%) (Figure 2A). The distribution of BLASTx top-hit species showed that *Onychostoma macrolepis* (12,937; 28.03%) has the biggest number of homologous genes to *S. hollandi*, followed by *Cyprinus carpio* (5924; 12.84%), *Sinocyclocheilus rhinoceros* (5554; 12.03%) and *Sinocyclocheilus anshuiensis* (4699; 10.18%) (Figure 2B).

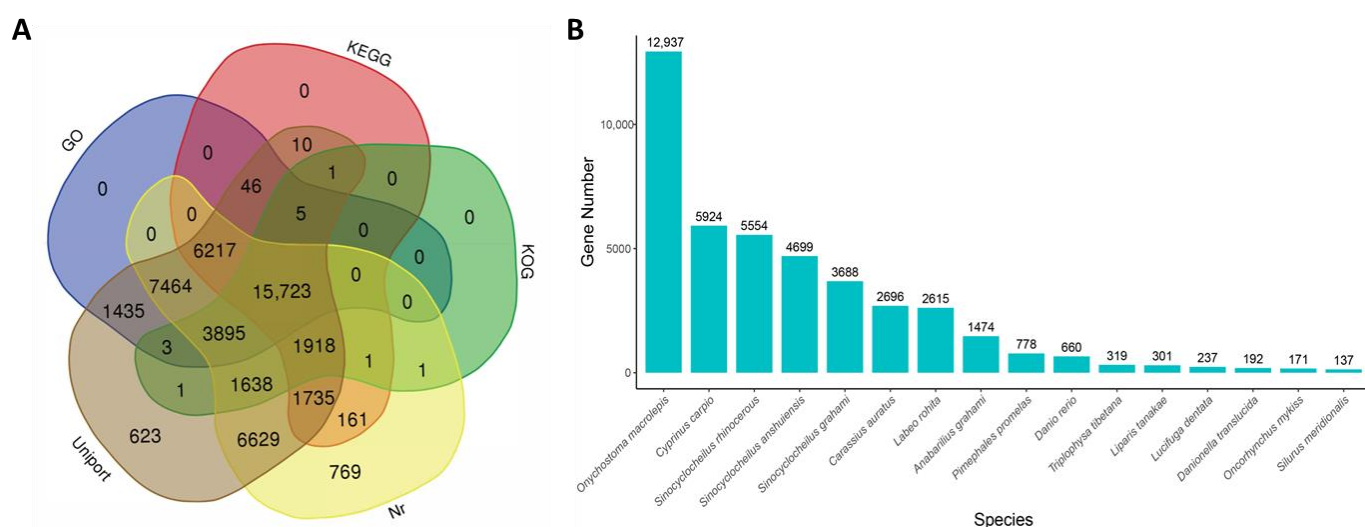


Figure 2. Venn diagrams of functional annotation against five major databases and species distribution of annotated homologous genes against NR database.

In addition, all unigenes were further annotated in GO, KOG and KEGG databases to get knowledge of their functional classification. A total of 23,186 (48.03%) unigenes were grouped into three categories of GO. In the biological processes category, cellular (19,188) was the most representative item. In the category of cellular component, the most represented term was cellular anatomical entity (22,121). In addition, in the molecular function category, the most representative terms were binding (17,708) and catalytic activity (11,582) (Figure 3A). After KEGG annotation, a sum of 25,817 (53.48%) unigenes was divided into five different functional categories. The largest distribution was “signal transduction” (3943 unigenes), “global and overview maps” (2552 unigenes) and “immune system” (2030 unigenes) (Figure 3B). According to KOG annotation, a total of 23,186 (48.03%) unigenes were grouped into 25 families, and the richest distribution was “General function prediction only” (5349 unigenes), followed by “Signal transduction mechanisms” (3452 unigenes); the smallest distribution was “Nuclear structure”, with only 8 unigenes (Figure 3C).

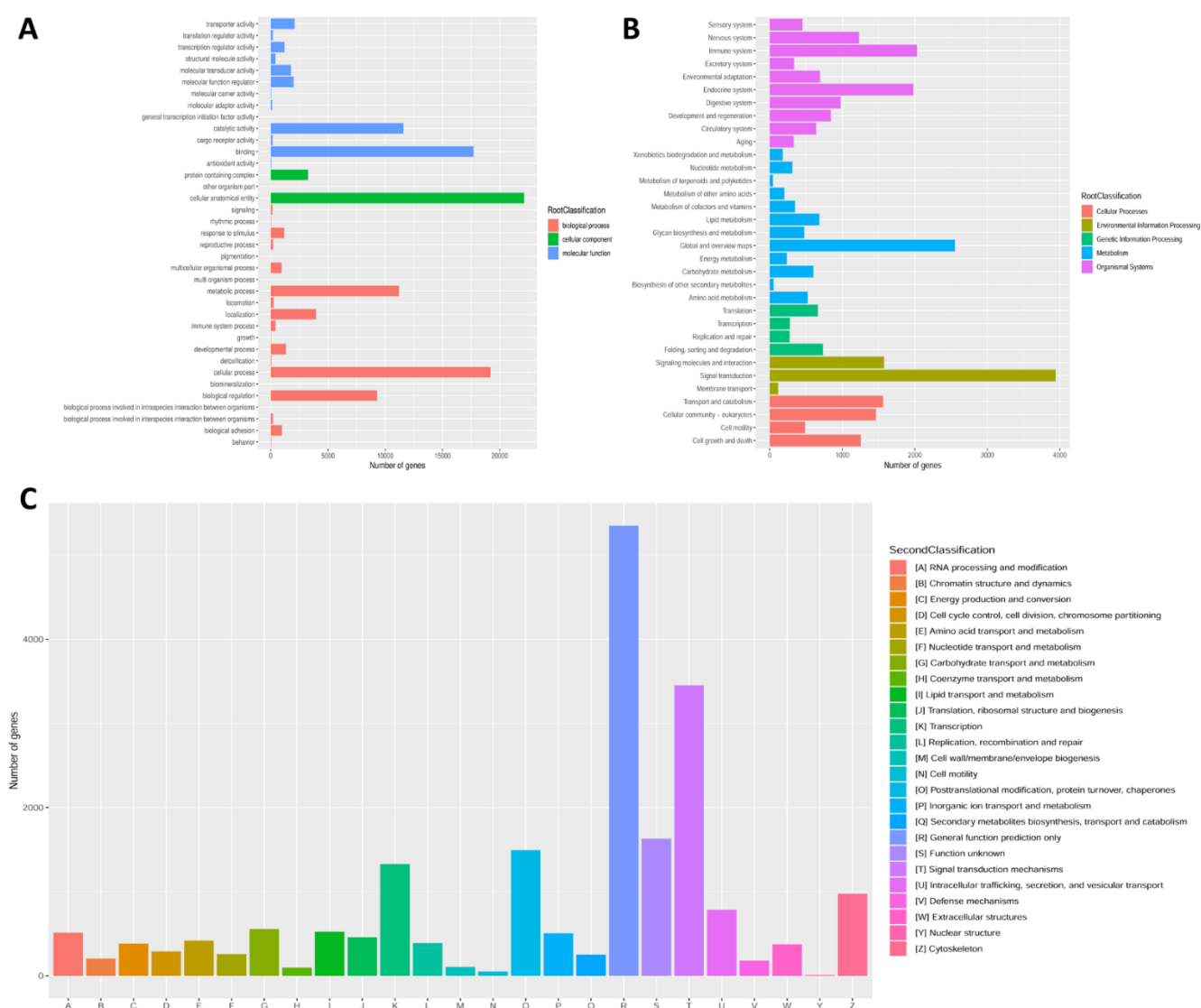


Figure 3. Gene function classification based on GO (A), KEGG (B) and KOG (C) databases.

3.3. Differential Expression Analysis

The levels of gene expression were normalized using the TPM values. A total of 21,903 DEGs were identified between the ovary and testis samples, with 16,395 (74.85%) DEGs significantly high expressed in the testes and 5508 (25.15%) high expressed in the ovaries (Figure 4A). A volcano plot shows the DEG information (Figure 4B). The KEGG enrichment analysis revealed that the metabolic pathway was most representative, followed by the MAPK signaling pathway and the neuroactive ligand-receptor interaction pathway (Figure 5).

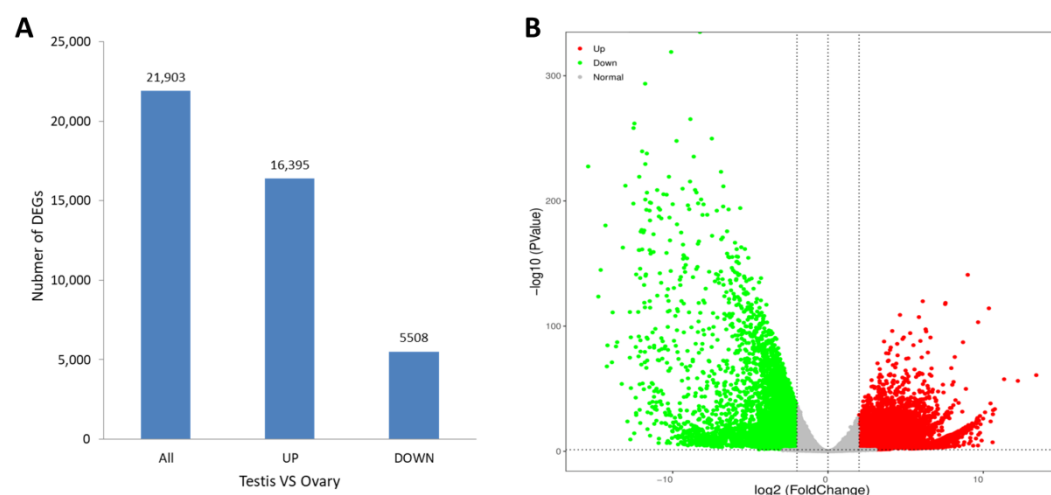


Figure 4. (A) Differentially expressed genes in males (testis) and females (ovary) of *S. hollandi*. (B) Volcano plot of DEGs in testis versus ovary. Red and green keys and dots represent upregulated and downregulated genes, respectively, in males, and vice versa.

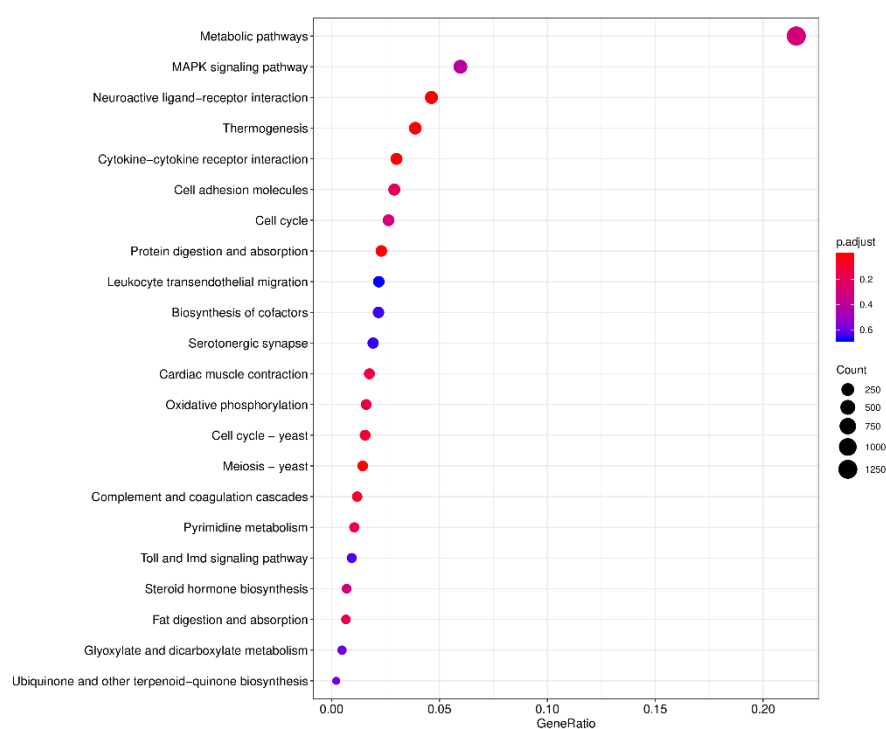


Figure 5. Top 22 KEGG enrichment analysis. The horizontal axis is the ratio of the number of differential genes annotated to the KEGG pathway to the total number of differential genes. The vertical axis is the enriched KEGG pathways. The size of the dots represents the number of genes annotated on the KEGG pathway term. The colors key, from red to blue, represent significant enrichment.

In addition, based on the GO and KEGG annotation, numerous genes related to reproduction and gonad development and differentiation were identified (Table 4). Anti-Mullerian hormone (*amh*), SRY-box transcription factor 9 (*sox9*), double sex- and mab-3-related transcription factor 2 (*dmrt2*) and androgen receptor (*ar*) were highly expressed in the testis, while cytochrome P450 (*cypa19a*), catenin beta-1 (*ctnnb1*) and bone morphogenetic protein 15 (*bmp15*) expressed highly in the ovary.

Table 4. The patterns of differentially expressed genes related to reproduction, gonad development and differentiation in the gonads of *S. hollandi* (JC_vs_LC means testis vs ovary).

Unigene ID	log2FC (JC_vs_LC)	PValue (JC_vs_LC)	FDR (JC_vs_LC)	Nr Annotation	Gene Name
unigene074455	7.572	4.92×10 ⁻⁹	1.83×10 ⁻⁰⁸	Growth differentiation factor 6	<i>gdf6</i>
unigene027408	7.318	2.71×10 ⁻⁹	1.04×10 ⁻⁰⁸	Neuropeptide Y receptor type 2	<i>npy2r</i>
unigene016644	7.159	7.36×10 ⁻¹⁴	4.50×10 ⁻¹³	Cytochrome P450 11B	<i>cyp11b1</i>
unigene058283	6.803	4.63×10 ⁻⁰⁸	1.52×10 ⁻⁰⁷	Forkhead box protein K2	<i>foxk2</i>
unigene065602	6.718	3.09×10 ⁻⁰⁴	5.54×10 ⁻⁰⁴	Paired box protein Pax-6	<i>pax6</i>
unigene047253	6.481	4.79×10 ⁻⁰⁶	1.16×10 ⁻⁰⁵	Fibroblast growth factor 20	<i>fgf20</i>
unigene082603	6.325	2.40×10 ⁻⁰⁵	5.23×10 ⁻⁰⁵	Bone morphogenetic protein 7	<i>Bmp7</i>
unigene100115	6.112	1.68×10 ⁻⁰⁴	3.17×10 ⁻⁰⁴	Forkhead box G1-like protein	<i>foxg1</i>
unigene034505	6.052	1.43×10 ⁻⁰⁵	3.24×10 ⁻⁰⁵	Catenin delta-2	<i>ctnnd2</i>
unigene022426	5.941	1.44×10 ⁻¹⁹	1.38×10 ⁻¹⁸	Paired box protein Pax-2a	<i>pax2</i>
unigene141021	5.878	2.57×10 ⁻¹²	1.37×10 ⁻¹¹	Nanos homolog 2	<i>nanos2</i>
unigene021432	5.874	5.49×10 ⁻⁰⁵	1.12×10 ⁻⁰⁴	Paired box protein Pax-7	<i>pax7</i>
unigene008910	5.867	3.93×10 ⁻⁰⁵	8.24×10 ⁻⁰⁵	Forkhead box protein L1	<i>foxl1</i>
unigene106729	5.856	5.75×10 ⁻¹⁰⁸	1.08×10 ⁻¹⁰⁵	Anti-Mullerian hormone	<i>amh</i>
unigene035421	5.798	7.51×10 ⁻¹³	4.21×10 ⁻¹²	Insulin-like growth factor-I	<i>igf1</i>
unigene086594	5.661	1.96×10 ⁻¹⁰	8.53×10 ⁻¹⁰	Protein Wnt-6	<i>wnt6</i>
unigene130369	5.617	2.78×10 ⁻⁰³	4.27×10 ⁻⁰³	Transforming growth factor beta re- ceptor type 3	<i>tgfbr3</i>
unigene123135	4.990	7.83×10 ⁻³³	1.54×10 ⁻³¹	Transcription factor SOX9	<i>sox9</i>
unigene089170	4.878	1.56×10 ⁻¹³	9.30×10 ⁻¹³	Protein Wnt-5a	<i>wnt5</i>
unigene025970	4.539	6.90×10 ⁻¹¹	3.16×10 ⁻¹⁰	Protein Wnt-2b	<i>wnt2</i>
unigene009697	4.487	1.83×10 ⁻¹⁰	8.00×10 ⁻¹⁰	Gonadotropin-releasing hormone II receptor	<i>gnrhr</i>
unigene125833	4.461	1.52×10 ⁻¹⁹	1.45×10 ⁻¹⁸	Sperm-associated antigen 1	<i>spag1</i>
unigene159834	4.418	1.43×10 ⁻⁰⁶	3.77×10 ⁻⁰⁶	Follicle stimulating hormone recep- tor	<i>fshr</i>
unigene018630	4.344	6.47×10 ⁻²³	7.59×10 ⁻²²	Mothers against decapentaplegic homolog 4a	<i>smad4a</i>
unigene143320	4.016	5.73×10 ⁻⁰⁶	1.38×10 ⁻⁰⁵	Forkhead box D2-like protein	<i>foxd2</i>
unigene036991	3.766	1.74×10 ⁻⁰⁴	3.27×10 ⁻⁰⁴	Protein Wnt-3a	<i>wnt3</i>
unigene131983	3.725	1.54×10 ⁻⁰⁹	6.08×10 ⁻⁰⁹	Bone morphogenetic protein 6 iso- form X1	<i>bmp6</i>
unigene072066	3.692	1.06×10 ⁻⁰⁴	2.07×10 ⁻⁰⁴	Forkhead box P1-B-like isoform X1	<i>foxp1</i>
unigene021595	3.626	2.58×10 ⁻⁰⁶	6.54×10 ⁻⁰⁶	Bone morphogenetic protein recep- tor type-2	<i>bmpr2</i>
unigene094351	3.575	1.65×10 ⁻⁰⁹	6.47×10 ⁻⁰⁹	Paired box protein Pax-3	<i>pax3</i>
unigene077500	3.532	5.58×10 ⁻⁰⁸	1.81×10 ⁻⁰⁷	Paired box protein Pax-8	<i>pax8</i>
unigene114563	3.461	6.23×10 ⁻⁰⁶	1.49×10 ⁻⁰⁵	Forkhead box protein O6	<i>foxo6</i>
unigene020954	3.447	2.05×10 ⁻⁰⁷	6.11×10 ⁻⁰⁷	Transcription factor Sox-1a	<i>sox1</i>
unigene151019	3.352	6.41×10 ⁻⁰⁸	2.06×10 ⁻⁰⁷	Protein Wnt-11	<i>wnt11</i>
unigene157484	3.304	7.84×10 ⁻³³	1.54×10 ⁻³¹	Fibroblast growth factor receptor 2	<i>fgfr2</i>
unigene060599	3.220	2.39×10 ⁻¹⁸	2.10×10 ⁻¹⁷	Transcription factor SOX-5	<i>sox5</i>
unigene060624	3.212	9.24×10 ⁻¹¹	4.17×10 ⁻¹⁰	Growth-hormone-releasing hor- mone receptor	<i>ghrhr</i>
unigene071307	3.210	6.50×10 ⁻¹⁷	5.12×10 ⁻¹⁶	Transforming growth factor beta-2	<i>tgfb2</i>
unigene119381	3.181	1.92×10 ⁻¹⁰	8.37×10 ⁻¹⁰	Insulin-like growth-factor-binding 3	<i>igfb3</i>
unigene002746	3.074	6.17×10 ⁻²⁰	6.04×10 ⁻¹⁹	Corticosteroid 11-beta-dehydrogen- ase isozyme 2	<i>hsd11b2</i>
unigene073844	2.890	4.90×10 ⁻¹³	2.79×10 ⁻¹²	Estrogen receptor beta 2	<i>erb2</i>

unigene076208	2.764	1.34×10^{-06}	3.55×10^{-06}	Follitropin subunit beta	<i>fshb</i>
unigene145623	2.763	7.92×10^{-09}	2.87×10^{-08}	Insulin-like growth-factor-binding protein 1	<i>igfbp1</i>
unigene144860	2.759	2.51×10^{-07}	7.39×10^{-07}	Forkhead box protein O3	<i>foxo3</i>
unigene052054	2.730	6.22×10^{-11}	2.86×10^{-10}	Fibroblast growth factor receptor 4	<i>fgfr4</i>
unigene126929	2.604	6.48×10^{-06}	1.55×10^{-05}	Bone morphogenetic protein 2	<i>bmp2</i>
unigene146529	2.383	7.29×10^{-07}	2.01×10^{-06}	Transcription factor sox7	<i>sox7</i>
unigene156535	2.380	2.36×10^{-13}	1.38×10^{-12}	Estrogen receptor beta 1	<i>erb1</i>
unigene057687	2.372	8.63×10^{-07}	2.35×10^{-06}	stAR-related lipid transfer protein 13	<i>stard13</i>
unigene031051	2.351	1.86×10^{-05}	4.13×10^{-05}	Forkhead box protein D1	<i>foxd1</i>
unigene013811	2.265	3.43×10^{-03}	5.18×10^{-03}	Androgen receptor protein	<i>ar</i>
unigene101954	2.264	2.05×10^{-04}	3.80×10^{-04}	Gonadotropin-releasing hormone	<i>gnrh</i>
unigene100104	2.236	3.13×10^{-06}	7.84×10^{-06}	double sex- and mab-3-related transcription factor 2	<i>dmrt2</i>
unigene008405	2.207	1.52×10^{-03}	2.43×10^{-03}	Spermatogenesis-associated protein 16	<i>sgap16</i>
unigene105851	2.174	3.57×10^{-03}	5.37×10^{-03}	Paired box protein Pax-4	<i>pax4</i>
unigene040229	2.160	1.35×10^{-08}	4.77×10^{-08}	Growth/differentiation factor 10	<i>gdf10</i>
unigene122511	2.117	1.60×10^{-06}	4.18×10^{-06}	Transcription factor SOX-4	<i>sox4</i>
unigene054659	-2.416	3.36×10^{-13}	1.94×10^{-12}	Spermatid perinuclear RNA-binding protein isoform X3	<i>strbp</i>
unigene147986	-2.463	3.91×10^{-09}	1.47×10^{-08}	Cytochrome P450 aromatase	<i>cyp19a1a</i>
unigene017060	-2.563	1.13×10^{-31}	2.10×10^{-30}	TGF-beta receptor type-2-like isoform X1	<i>tgfbr2</i>
unigene124450	-2.947	2.27×10^{-38}	5.69×10^{-37}	Catenin beta-1	<i>ctnnb1</i>
unigene062930	-2.964	5.89×10^{-11}	2.72×10^{-10}	Transcription factor Sox-3	<i>sox3</i>
unigene014765	-3.027	4.13×10^{-10}	1.73×10^{-09}	Spermatogenesis-associated protein 13	<i>sgap13</i>
unigene000100	-3.070	8.27×10^{-07}	2.26×10^{-06}	Sperm-associated antigen 5	<i>spag5</i>
unigene127688	-3.105	3.71×10^{-44}	1.16×10^{-42}	sTART domain-containing protein 10	<i>stard10</i>
unigene152354	-3.674	8.97×10^{-10}	3.63×10^{-09}	Paired box protein Pax-1	<i>pax1</i>
unigene051443	-3.717	2.71×10^{-18}	2.37×10^{-17}	Double sex- and mab-3-related transcription factor A1	<i>dmrta1</i>
unigene143561	-3.915	1.79×10^{-71}	1.39×10^{-69}	Growth/differentiation factor 9	<i>gdf9</i>
unigene125277	-3.925	5.06×10^{-13}	2.88×10^{-12}	Estradiol 17-beta-dehydrogenase 2	<i>hsd17b2</i>
unigene069571	-3.965	1.60×10^{-88}	1.96×10^{-86}	Bone morphogenetic protein 15	<i>bmp15</i>
unigene086306	-4.644	2.71×10^{-18}	2.37×10^{-17}	Nuclear receptor subfamily 5 group A member 2	<i>nr5a2</i>
unigene007641	-5.956	5.52×10^{-41}	1.55×10^{-39}	Forkhead box protein P3	<i>foxp3</i>
unigene060440	-5.993	3.07×10^{-50}	1.20×10^{-48}	Transcription factor Sox-19b	<i>sox19b</i>
unigene150909	-6.870	4.59×10^{-170}	3.12×10^{-167}	Transcription factor Sox-21-A	<i>sox31a</i>
unigene158966	-7.139	8.15×10^{-47}	2.82×10^{-45}	Protein Wnt-8a	<i>wnt8</i>
unigene043515	-11.722	3.97×10^{-142}	1.52×10^{-139}	Zona pellucida sperm-binding protein 4	<i>zp4</i>
unigene055600	-12.153	5.03×10^{-220}	1.32×10^{-216}	Zona pellucida sperm-binding protein 3	<i>zp3</i>
unigene077416	-14.328	3.82×10^{-181}	3.10×10^{-178}	Growth/differentiation factor 3	<i>gdf3</i>

3.4. Validation of Transcriptomic Data

A total of fourteen DEGs related to sex differentiation and gonadal development were analyzed from the transcriptome data, including *amh*, *hsd11b2*, *fshr*, *era*, *erb1*, *erb2*, *ar*, *smad4a*, *gdf6*, *bmp7*, *pax7*, *pax8*, *sox5* and *sox9*. In general, the results of qRT-PCR and RNA-seq analysis were consistent (Figure 6), and the expression of all tested genes was higher in testes than in ovaries, which indicated the gene expression profiles quantified by transcriptomic analysis were reliable and accurate.

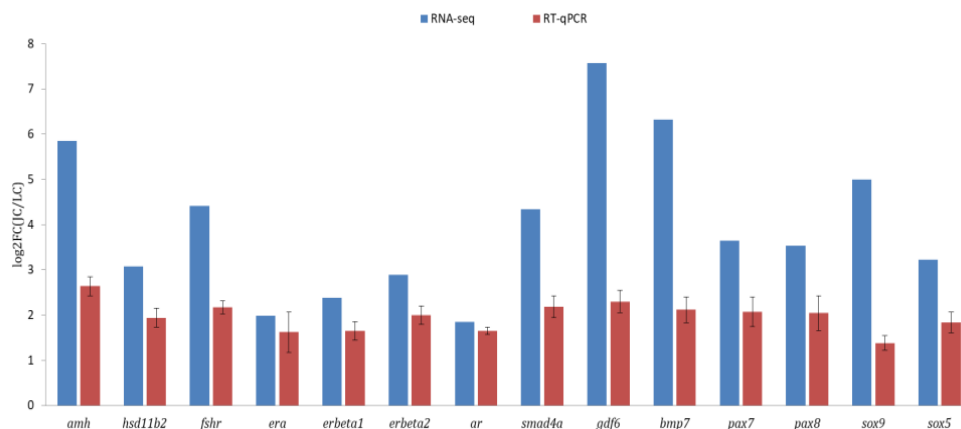


Figure 6. Validation of gonad transcriptome data by qRT-PCR.

4. Discussion

Gonad development, differentiation and maturity are extremely complicated biological process, which contains a number of functional genes. In recent years, transcriptome sequencing has been demonstrated to be a highly effective method to obtain comprehensive information of gene expression for specific tissue. *S. hollandi* is an important commercial aquaculture species with high nutritional value and medicinal value. Its long sexual maturity time and low egg laying amount seriously hinder the development of the breeding industry. However, up to now, the molecular mechanisms of gonadal differentiation and development have not been revealed in *S. hollandi*. In this study, comparative transcriptome analysis of ovary and testis was first used to identify sex-related DEGs in *S. hollandi*.

4.1. DEGs Involved in Steroid Synthetic Pathway

Sex steroid hormones are critically important for gonadal differentiation, development and maturation in fish species, which mainly includes androgen and estrogen. In fish, the 11-ketotestosterone (11-KT) and 17-estradiol (E2) are the major androgen and estrogen that are essential to testicular and ovarian development [22]. In fact, sex steroid hormone synthesis needs a lot of steroid-metabolizing enzymes, which are encoded by many genes such as *cyp19a1a*, *cyp11b2*, *hsd11b1*, *cyp11a1*, *hsd17b* and so on.

Among them, the *cyp19a1a* gene is the most essential regulator, which can directly catalyze the switching of androgens to estrogens in the ovary. In *Oreochromis niloticus* and *Danio rerio*, knock out of the *cyp19a1a* gene caused a drop in E2 levels and further induced female-to-male sex reversal [23,24]. The *Hsd11b2* gene encodes an important enzyme that turns testosterone into 11-KT, which is important in testis maintenance. In *Danio rerio*, the reproductive capability of *hsd11b2* homozygous mutation adult males is almost completely abrogated [25]. The *Cyp11b1* gene is also involved in the synthesis of 11-KT, encoding the critical enzyme that turns testosterone to 11-KT [26]. In the sea bass, *cyp11b1* not only was predominantly expressed in the first stages of spermatogenesis but most likely was also expressed in spermatogonia [26]. In addition, high expression of *hsd11b2* and *cyp11b1* genes in the testis was also found in mandarin fish [27]. In *S. hollandi*, the

expression of *cyp11b1* and *hsd11b2* genes was higher in the testis, while the *cyp19a1a* gene expressed highly in the ovary, indicating the expression pattern was similar among different fish species [9,10,27]. Thus, these results indicated that these DEGs play important roles in gonad development and reproduction in fish.

4.2. DEGs Involved in Gonad Differentiation and Development

In recent years, more and more genes were demonstrated to participate in the sex determination and gonad differentiation of teleosts. The upstream regulatory genes are diverse, but the downstream genes are much conserved. The common regulatory gene downstream of most sex-related genes is the *cyp19a1a* gene, which can directly catalyze the conversion of androgens to estrogens.

Anti-Müllerian hormone (*amh*), a transforming growth factor from the TGF- β family, controls the degeneration of female primordial germ cell ducts in mammals and promotes male development in vertebrates. Previous studies have demonstrated *amh* mutation could lead to male-to-female sex reversal in zebrafish [28] and tilapia [29]. Moreover, direct evidence revealed that *amh* could inhibit the transcription of *cyp19a1a* through *Amhr2/Smads* signaling [30]. *Bmp15* and *gdf9*, as oocyte-produced signals, participate in maintaining adult female sex differentiation and, in zebrafish, females deficient in *bmp15* begin development normally but switch sex during the mid- to late-juvenile stage, and become fertile males [31]. In *S. hollandi*, the results revealed that the expression of *amh* was higher in the testis, indicating the *Amh* pathway was of great importance in testicular differentiation regulation. In addition, *bmp15* and *gdf9* also showed ovary-biased expression in *S. hollandi*, suggesting a conserved role of *bmp15* and *gdf9* in teleosts.

Sox family genes are widely involved in the sex determination and differentiation process of vertebrates, and are a family of transcription factors that are closely related to the mammalian sex-determining region of the Y genes, namely *sry* [32]. Previous study has found that *sox5* could downregulate the activity of *dmy* (main sex determining gene) by binding to its promoter, and females with the *sox5* mutation will reverse to males in medaka [33]. *Sox9* is known to have an essential role in Sertoli cell differentiation; conditional *sox9* knockout testes fail to maintain germ cells in mice [34]. Previous evidence has proved that *sox9* is capable of interfering with WNT/ β -catenin signaling and repressing the ovarian-differentiating genes including follistatin, Iroquois-related homeobox 3 and *foxl2* [35]. As in the medaka and mouse, *sox5* and *sox9* also showed high expression in the testis of *S. hollandi*, suggesting an important function of these genes in vertebrates.

4.3. DEGs Involved in Gametogenesis and Gamete Maturation

In aquaculture practice, it is extremely important to obtain parental fish with well-developed gametes. Many genes are involved in the regulation of gamete development and maturation. Here, the expression of many genes involved in gametogenesis and gamete maturation was also characterized, such as *spata16*, *zp3* and *zp4*. *Spata16*, namely spermatogenesis-associated protein 16, is essential for spermiogenesis [36] and male fertility; homozygous mutation of *spata16* leads to male infertility [37]. *Zp3* is a major kind of female-specific factor involved in the regulation of the reproductive process, forming an extracellular matrix surrounding oocytes which mediates sperm binding. Similar to *zp3*, *zp4* also surround oocytes and may act as a sperm receptor. In this study, the expression of *zp3* and *zp4* was higher in the ovary, indicating that they might participate in ovarian folliculogenesis in *S. hollandi*. In addition, *spata13*, *spag5* and *strbp* also showed higher expression in the ovaries, suggesting an important role in oogenesis and oocyte maturation. In all, similar expression patterns of these reproduction-related genes were also found in *Diodon hystrix* [9], *Scatophagus argus* [10], *Patinopekten yessonsis* [38] and so on, indicating a conserved role of these genes in teleosts.

5. Conclusions

This work is the first study on the gonad transcriptome of *S. hollandi*. Based on testis and ovary transcriptome data, a number of 167,152 unigenes were identified. Comparative transcriptome analysis identified a large scale of DEGs involved in gonadal development, differentiation and gametogenesis. For these DEGs related to gonad development and reproduction, similar expression profiles were found, suggesting their conserved roles in gonad development and gametogenesis. Our results will provide a valuable resource for further research on sex determination and gonad development.

Author Contributions: Conceptualization, J.G. and Q.L.; methodology, C.H.; software, C.H.; validation, W.H., S.P. and J.Z.; formal analysis, C.H.; investigation, W.H., S.P. and J.Z.; resources, H.Z. and Y.Z.; data curation, C.H.; writing—original draft preparation, C.H.; writing—review and editing, J.G. and Q.L.; visualization, C.H. and W.L.; supervision, J.G. and Q.L.; project administration, J.G. and Q.L.; funding acquisition, Q.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Project of Science and Technology of Guangzhou city grant number 201804010486.

Institutional Review Board Statement: All procedures in this experiment were approved by the Institutional Review Committee of Bioethics and Biosafety of Guangzhou University on 5 April 2021 (No. GURBBB210405).

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data are available from the SRA (<http://www.ncbi.nlm.nih.gov/sra/>, accessed on 12 October 2021) data repository (accession number: PRJNA907069).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cai, Z.D.; Lin, G.; Ni, J.Y.; Ye, Z.; Tang, Z.B.; Zhang, S. Study on the reproductive biology of *Spinibarbus caldwelli*. *Guangxi Agric. Sci.* **2007**, *2*, 24–28. (In Chinese with English abstract).
2. Yang, S.D.; Lin, T.S.; Liou, C.H.; Peng, H.K. Influence of dietary protein levels on growth performance, carcass composition and liver lipid classes of juvenile *Spinibarbus hollandi* (Oshima). *Aquat. Res.* **2003**, *34*, 661–666.
3. Yang, Y.; Lan, Z.; Shu, H.; Zhou, H.; Jiang, X.; Hou, L.; Gu, P. Association between expression levels and growth trait-related SNPs located in promoters of the MC4R and MSTN genes in *Spinibarbus hollandi*. *Genes Genom.* **2018**, *40*, 1119–1125.
4. Jian, X.; Peizhen, Z.; Caiyan, W.; Jianrong, C.; Liangming, Z.; Qinhong, L.; Yuxuan, W. The gonadal development and the annual change of ovary in *Spinibarbus hollandi*. *Chin. J. Zool.* **2004**, *39*, 7–10.
5. Heule, C.; Salzburger, W.; Böhne, A. Genetics of sexual development: An evolutionary playground for fish. *Genetics* **2014**, *196*, 579–591.
6. Yang, Y.; Zhou, H.; Hou, L.; Xing, K.; Shu, H. Transcriptional profiling of skeletal muscle reveals starvation response and compensatory growth in *Spinibarbus hollandi*. *BMC Genom.* **2019**, *20*, 938.
7. Yang, Y.; Zhou, H.; Shu, H.; Zhong, D.; Zhang, M.; Xia, J.H. Effect of fasting and subsequent refeeding on the transcriptional profiles of brain in juvenile *Spinibarbus hollandi*. *PLoS ONE* **2019**, *14*, e0214589.
8. Garber, M.; Grabherr, M.G.; Guttman, M.; Trapnell, C. Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat. Methods* **2011**, *8*, 469–477.
9. Chen, H.; Li, Z.; Wang, Y.; Huang, H.; Yang, X.; Li, S.; Yang, W.; Li, G. Comparison of gonadal transcriptomes uncovers reproduction-related genes with sexually dimorphic expression patterns in *Diodon hystrix*. *Animals* **2021**, *11*, 1042.
10. Mustapha, U.F.; Peng, Y.X.; Huang, Y.Q.; Assan, D.; Zhi, F.; Shi, G.; Huang, Y.; Li, G.L.; Jiang, D.N. Comparative transcriptome analysis of the differentiating gonads in *Scatophagus argus*. *Front. Mar. Sci.* **2022**, *9*, 962534.
11. Du, X.; Wang, B.; Liu, X.; Liu, X.; He, Y.; Zhang, Q.; Wang, X. Comparative transcriptome analysis of ovary and testis reveals potential sex-related genes and pathways in spotted knifejaw *Oplegnathus punctatus*. *Gene* **2017**, *637*, 203–210.
12. Lu, J.; Luan, P.; Zhang, X.; Xue, S.; Peng, L.; Mahbooband, S.; Sun, X. Gonadal transcriptomic analysis of yellow catfish (*Pelteobagrus fulvidraco*): Identification of sex-related genes and genetic markers. *Physiol. Genom.* **2014**, *46*, 798–807.
13. Zhu, Q.; Han, C.; Liu, S.; Ouyang, H.; Liu, D.; Zhang, Z.; Huang, J.J.; Han, L.Q.; Li, S.S.; Li, G.F.; et al. Development and gene expression analysis of gonad during 17 α -methyltestosterone-induced sex reversal in mandarin fish (*Siniperca chuatsi*). *Aquat. Rep.* **2022**, *23*, 101049.
14. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323.

15. Conesa, A.; Gotz, S.; Garcia-Gomez, J.M.; Terol, J.; Talon, M.; Robles, M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **2005**, *21*, 3674–3676.
16. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360.
17. Perte, M.; Perte, G.M.; Antonescu, C.M.; Chang, T.C.; Mendell, J.T.; Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **2015**, *33*, 290–295.
18. Anders, S.; McCarthy, D.J.; Chen, Y.; Okoniewski, M.; Smyth, G.K.; Huber, W.; Robinson, M.D. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat. Protoc.* **2013**, *8*, 1765–1786.
19. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140.
20. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. Cluster profiler: An R package for comparing biological themes among gene clusters. *OMICS* **2012**, *16*, 284–287.
21. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{(-Delta Delta C(T))} Method. *Methods* **2001**, *25*, 402–408.
22. Ijiri, S.; Kaneko, H.; Kobayashi, T.; Wang, D.S.; Sakai, F.; Paul-Prasanth, B.; Nagahama, Y. Sexual dimorphic expression of genes in gonads during early differentiation of a teleost fish, the Nile tilapia *Oreochromis niloticus*. *Biol. Reprod.* **2008**, *78*, 333–341.
23. Zhang, X.; Li, M.; Ma, H.; Liu, X.; Shi, H.; Li, M.; Wang, D. Mutation of *foxl2* or *cyp19a1a* results in female to male sex reversal in XX Nile tilapia. *Endocrinology* **2017**, *158*, 2634–2647.
24. Lau, E.S.W.; Zhang, Z.; Qin, M.; Ge, W. Knockout of zebrafish ovarian aromatase gene (*cyp19a1a*) by TALEN and CRISPR/Cas9 leads to all-male offspring due to failed ovarian differentiation. *Sci. Rep.* **2016**, *6*, 37357.
25. Theodoridi, A.; Dinarello, A.; Badenetti, L.; Pavlidis, M.; Dalla Valle, L.; Tsalafouta, A. Knockout of the *hsd11b2* Gene Extends the Cortisol Stress Response in Both Zebrafish Larvae and Adults. *Int. J. Mol. Sci.* **2021**, *22*, 12525.
26. Viñas, J.; Piferrer, F. Stage-specific gene expression during fish spermatogenesis as determined by laser-capture microdissection and quantitative-PCR in sea bass (*Dicentrarchus labrax*) gonads. *Biol. Reprod.* **2008**, *79*, 738–747.
27. Zhu, Q.; Han, C.; Peng, C.; Zhou, X.; Wang, C.; Han, L.; Li, S.S.; Li, G.F.; Lin, H.R.; Zhang, Y. Identification of potential sex-related genes in *Siniperca chuatsi*. *J. Oceanol. Limnol.* **2021**, *39*, 1500–1512.
28. Lin, Q.; Mei, J.; Li, Z.H.; Zhang, X.; Zhou, L.; Gui, J.F. Distinct and cooperative roles of *amh* and *dmrt1* in self-renewal and differentiation of male germ cells in zebrafish. *Genetics* **2017**, *207*, 1007–1022.
29. Li, M.; Sun, Y.; Zhao, J.; Shi, H.; Zeng, S.; Ye, K.; Jiang, D.; Zhou, L.; Sun, L.; Tao, W.; et al. A Tandem Duplicate of Anti-Müllerian Hormone with a Missense SNP on the Y Chromosome Is Essential for Male Sex Determination in Nile Tilapia, *Oreochromis niloticus*. *PLoS Genet.* **2015**, *11*, e1005678.
30. Liu, X.; Dai, S.; Wu, J.; Wei, X.; Zhou, X.; Chen, M.; Tan, D.; Pu, D.; Li, M.; Wang, D. Roles of Anti-Müllerian hormone (Amh) and its duplicates in sex determination and germ cell proliferation of Nile tilapia. *Genetics* **2021**, *220*, iyab237.
31. Dranow, D.B.; Hu, K.; Bird, A.M.; Lawry, S.T.; Adams, M.T.; Sanchez, A.; Amatruda, G.F.; Draper, B.W. Bmp15 is an oocyte-produced signal required for maintenance of the adult female sexual phenotype in zebrafish. *PLoS Genet.* **2016**, *12*, e1006323.
32. Wei, L.; Yang, C.; Tao, W.; Wang, D. Genome-wide identification and transcriptome-based expression profiling of the Sox gene family in the Nile Tilapia (*Oreochromis niloticus*). *Int. J. Mol. Sci.* **2016**, *17*, 270.
33. Scharf, M.; Schories, S.; Wakamatsu, Y.; Nagao, Y.; Hashimoto, H.; Bertin, C.; Mourot, B.; Schmidt, C.; Wilhelm, D.; Centanin, L.; et al. Sox5 is involved in germ-cell regulation and sex determination in medaka following co-option of nested transposable elements. *BMC Biol.* **2018**, *16*, 16.
34. Barrionuevo, F.; Georg, I.; Scherthan, H.; Lécureuil, C.; Guillou, F.; Wegner, M.; Scherer, G. Testis cord differentiation after the sex determination stage is independent of *Sox9* but fails in the combined absence of *Sox9* and *Sox8*. *Dev. Biol.* **2009**, *327*, 301–312.
35. Li, Y.; Zheng, M.; Lau, Y.F.C. The sex-determining factors SRY and SOX9 regulate similar target genes and promote testis cord formation during testicular differentiation. *Cell Rep.* **2014**, *8*, 723–733.
36. Fujihara, Y.; Oji, A.; Larasati, T.; Kojima-Kita, K.; Ikawa, M. Human globozoospermia-related gene Spata16 is required for sperm formation revealed by CRISPR/Cas9-mediated mouse models. *Int. J. Mol. Sci.* **2017**, *18*, 2208.
37. Dam, A.H.; Kosciński, I.; Kremer, J.A.; Moutou, C.; Jaeger, A.S.; Oudakker, A.R.; Tournaye, H.; Charlet, N.; Largier-Tourenne, C.; Bokhoven, H. Homozygous mutation in SPATA16 is associated with male infertility in human globozoospermia. *Am. J. Hum. Genet.* **2007**, *81*, 813–820.
38. Yang, D.; Yin, C.; Chang, Y.; Dou, Y.; Hao, Z.; Ding, J. Transcriptome analysis of male and female mature gonads of Japanese scallop *Patinopekten yessoensis*. *Genes Genom.* **2016**, *38*, 1041–1052.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.