Transcriptome Analysis for Identification of Genes Related to Gonad Differentiation, Growth, Immune Response and Marker Discovery in The Turbot (Scophthalmus maximus)

Deyou Ma1,2,3, Aijun Ma1,2*, Zhihui Huang1,2, Guangning Wang1,2, Ting Wang1,2, Dandan Xia1,2, Benhe Ma1,2

1 Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Key Laboratory of Sustainable Development of Marine Fisheries, Ministry of Agriculture, Qingdao Key Laboratory for Marine Fish Breeding and Biotechnology, Qingdao, 266071, China, 2 Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao, 266071, China, 3 Dalian Ocean University, Dalian, 116023, China

* maaj@ysfri.ac.cn

Abstract

Background

Turbot Scophthalmus maximus is an economically important species extensively aquacultured in China. The genetic selection program is necessary and urgent for the sustainable development of this industry, requiring more and more genome background knowledge. Transcriptome sequencing is an excellent alternative way to identify transcripts involved in specific biological processes and exploit a considerable quantity of molecular makers when no genome sequences are available. In this study, a comprehensive transcript dataset for major tissues of S. maximus was produced on basis of an Illumina platform.

Results

Total RNA was isolated from liver, spleen, kidney, cerebrum, gonad (testis and ovary) and muscle. Equal quantities of RNA from each type of tissues were pooled to construct two cDNA libraries (male and female). Using the Illumina paired-end sequencing technology, nearly 44.22 million clean reads in length of 100 bp were generated and then assembled into 106,643 contigs, of which 71,107 were named unigenes with an average length of 892 bp after the elimination of redundancies. Of these, 24,052 unigenes (33.83% of the total) were successfully annotated. GO, KEGG pathway mapping and COG analysis were performed to predict potential genes and their functions. Based on our sequence analysis and published documents, many candidate genes with fundamental roles in sex determination (dmrt1), growth (ghrh, myf5, prl/prlr) and immune response (TLR1/TLR21/TLR22, IL-15/IL-34), were identified for the first time in this species. In
addition, a large number of credible genetic markers, including 21,192 SSRs and 8,642 SNPs, were identified in the present dataset.

Conclusion
This informative transcriptome provides valuable new data to increase genomic resources of *Scophthalmus maximus*. The future studies of corresponding gene functions will be very useful for the management of reproduction, growth and disease control in turbot aquaculture breeding programs. The molecular markers identified in this database will aid in genetic linkage analyses, mapping of quantitative trait loci, and acceleration of marker assisted selection programs.

Introduction
Turbot (*Scophthamus maximus*) is an economically important flatfish widely farmed in Europe and Asia. The intensive culture of turbot has been promoted in the past few years because of great economic value. Turbot production boosted during the last decade in China, which has become the world's largest turbot-producing nation reported by FAO in 2010. However, inbreeding and intensive culture has brought about multiple negative effects on turbot industry. Enormous economic losses resulted from the slow growth and disease outbreaks of fish [1]. Thus, genetic breeding programs in this species are accessible to be carried out. Currently, the main targets of genetic improvement in turbot are controlling sex ratio, increasing growth rate and enhancing disease resistance [2].

Several economic traits are related to sex in aquaculture species. Sexual dimorphism has been observed in growth rate, time and age of maturation, body shape and carcass composition [3]. Turbot exactly exhibits the significantly sexual dimorphism for growth rate in favor of females among aquaculture species [4]. Producing all-female stocks seems promising to increase biomass of turbot for acquiring more interests. Therefore, controlling sex ratio is one of the major targets of genetic improvement in turbot. Understanding the process of gonadal development can offer a powerful support in the control of sex ratios in finfish aquaculture. An undifferentiated bipotential gonad of fish develops into either a testis or an ovary depending on sex determining genes [5], and external factors such as temperature or pH can directly influence gonadal development and then affect sex ratio in some fish [6], while the mechanisms of sex determination and gonad differentiation in turbot are not conclusive [7,8], due to the insufficient genomic information, and consequently need to be further explored. Growth rate from hatching to commercial size is a primary trait of interest in selection programs of most economic fish and has an intrinsic link with productivity and profitability of aquaculture enterprises. Growth of vertebrates (including fish) is primarily controlled by the GH-insulin-like growth factor-I (IGF-I) axis [9]. There are few studies on growth-related genes and corresponding regulation network in turbot. Intensive culture conditions in fish farms aggravate the risk of pathogen infection and the consequent losses of benefit associated with disease outbreaks. Reduction of disease occurrence is a major concern for turbot aquaculture [10]. Obtaining resistant broodstock is a fascinating solution to control diseases in front of the economic cost of vaccines, treatments and the possible generation of resistances against antibiotics. The necessary preparation is a comprehensive understanding of the immune system in economic fish species [11], particularly in turbot. During the last few years, a large number of expressed sequence tags (ESTs) from the turbots challenged with the most common different pathogens...
have pooled into a database relevant to immune response. Despite recent increases in the number of gene sequences for turbot, the available genomic resources are inadequate yet to offer an extensive detection on candidate genes in control of economic traits and the corresponding regulatory pathways in this species.

Transcriptome containing most protein coding genes is a small but essential part of the genome. Sequencing the transcriptome is an attractive alternative for gene discovery in species whose genome is still not available, especially to economic fish. Newly-developed high-throughput sequencing technologies can produce huge transcriptomic data of non-model organisms with low cost and high efficiency [13]. More than 2000 differentially expressed sex-biased genes and several sex-related biological pathways were firstly found in flounder (*Paralichthys olivaceus*) based on an Illumina platform [14]. An overall database from multiple tissues of turbots was reported using a combined strategy involving Sanger and 454 pyrosequencing, in which important genes related to reproduction and disease control were discovered [15]. There are few reports on a whole transcript profile of turbot that has become a highly appreciated aquaculture species since its introduction to China.

The selection on basis of molecular markers including simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs) is another approach to improve the aquaculture production in commercially important fish species [16]. Hundreds of SSRs have been developed and validated in turbot [17], many of which have been already used for genetic linkage mapping [18,19]. Several QTLs significantly associated with sex-determination [20], growth [21] and resistance to pathogens [2,22] were identified through a genome scan using the genetic map in turbot, which implies the existence of genetic factors underlying these characters and supports their application in genetic breeding strategies. Massive molecular markers have been exploited from extensive transcriptomic sequence data with the advances of sequencing technologies in a variety of economic species [23,24]. Hundreds of true SNPs were detected using 454-pyrosequencing method in turbot and most SNP-containing genes were related to immune response and gonad differentiation processes, which could be chosen as candidates to discover the relationship between functional changes and phenotypic changes [25].

In this study, the transcriptome of pooled multiple tissues from one male and one female turbots was characterized using an Illumina sequencing platform to maximize the chance of presenting as many transcripts as possible, respectively. The turbot genomic database is enriched by unigenes that were *de novo* assembled and annotated through strict bioinformatic analysis in this study. Many important genes involved in sex-control, growth and immune response were identified. Furthermore, abundant markers including SSRs located within coding regions and SNPs detected amongst deep coverage sequence regions reads were developed. Briefly, the transcriptome offers an invaluable data for further genomic research of flatfishes, as well as to discover new markers potentially useful in promoting turbot molecular breeding progress.

**Materials and Methods**

**Ethics statement**

The animals used in the present study were artificially cultivated, and all experimental treatments are implemented according to the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The study protocol was approved by the Experimental Animal Ethics Committee, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, China.
Sample preparation
Two adult turbots of three years old including one male and one female (total length: ♀, 40.7 cm; ♂, 31.6 cm) used for transcriptome sequencing were purchased from Tian-yuan Fisheries Co. Ltd., Yantai, China. The fish samples were acclimated in the laboratory for one week before the experiment treatment. A variety of tissues including liver, spleen, kidney (inclusive of head kidney), cerebrum, gonads after breeding season and skeletal muscle were dissected from the two samples after euthanasia by immersion in MS-222 buffered solution (3 g/L) on ice. All fresh tissues were frozen in liquid nitrogen immediately and stored at -80°C until RNA extraction within 2 weeks.

RNA isolation, cDNA library construction and Illumina deep sequencing
Total RNA was extracted from each tissue sample using Trizol Reagent (Invitrogen, CA, USA). After checking RNA purity and concentration, the integrity of RNA samples was assessed using the RNA 6000 Pico LabChip with a Bioanalyzer 2100 (Agilent Technologies, CA, USA). mRNA was purified from total RNA that was predigested at 37°C for 1 h using DNase I using Micropoly(A) PuristTM mRNA purification kit (Ambion USA). The eligible mRNA samples (RIN values ≥8) from one turbot individual were pooled in equal amounts to generate one mixed sample. A total amount of 10 μg mRNA per sample was used as input material for preparing one separate Illumina sequencing libraries.

The method of double cDNA synthesis was modified on basis of the published method [26]. Briefly, first strand cDNA was synthesized using GsuI-oligo(dT) and Superscript II reverse transcriptase (Invitrogen, USA). Subsequently, the first strand was lysed from biotin-attached mRNA/cDNA that had been picked out by DynalM280 magnetic bead (Invitrogen) through recognition of biotin linked to mRNA 5’ cap structure. Second strand cDNA synthesis was subsequently performed using Ex Taq polymerase (Takara). The polyA ends and 5’ adaptors were eventually removed by GsuI enzyme.

The double cDNA was cleaved into fragments (300~500 bp) using Fisher ultrasound equipment and then purified by Ampure beads (Agencourt, USA). The generation and amplification of cDNA libraries were carried out by Illumina TruSeq RNA Sample Preparation Kit and TruSeq PE Cluster Kit (Illumina, San Diego, USA) following manufacturer’s recommendations, respectively. Finally, the libraries were sequenced on an Illumina Hiseq 2000 platform in the Chinese National Human Genome Center (Shanghai) and paired-end reads with approximate length of 100 bp were generated.

Bioinformatic analysis

Quality control. Raw reads were produced through base calling and stored in fastq format. The raw data became clean after filtration by removing the adapter sequences, reads with unknown nucleotides (N) more than 10% and low quality sequences (base quality score < Q20). Clean data with high quality were the basis of following analyses.

Transcriptome assembly and gene annotation. De novo assembly of clean reads was carried out using Trinity software (http://trinityrnaseq.sf.net). Trinity was usually consisted of three independent software modules: Inchworm, Chrysalis and Butterfly. Using this software, sequencing data were partitioned into many single de Bruijn graphs (each represented transcriptional complexity for a given gene). Full-length splicing isoforms and transcripts from paralogous genes were obtained after the independent processing of graphs. At this period, the k-mer value was set to 25. The longest transcript from the same component was only preserved as a contig for excluding the interference from alternative splicing of transcripts. The assembled sequences were defined as unigenes.
The prediction of unigenes was performed by mapping to protein-coding sequences using GetORF of EMBOSS [27]. The predicted protein-coding sequences were annotated to the NCBI non-redundant (Nr) protein database and UniProtKB database, using BLASTp with algorithm with an E-value threshold of 1e-5. Gene encoding protein domains were identified by searching against Swiss-Prot and TrEMBL databases through BLASTp program. GO function for all unigenes was classified using Gene2go of GoPipe program [28]. KEGG (Kyoto Encyclopedia of Genes and Genomes) metabolic pathway annotation and COG (Clusters of Orthologous Groups) classification of unigenes were determined by searching against KEGG database and COG database using BLAST algorithm, respectively.

Expression abundance and GO enrichment analysis. Differential expression of unigenes in the two turbot libraries was analyzed using the MA-plot-based method with Random Sampling model (MARS) in DESeq R package [29]. P value was adjusted by means of q value. Q value<0.001&|log2 (fold change)| >1 was set as the threshold for significantly differential expression. GO enrichment analysis of the differentially expressed genes (DEGs) was performed using a hyper geometric distribution test. GO term with false discovery rate (FDR)<0.01 was defined as the term of significantly enriching DEGs.

Markers detection. Molecular markers, including SSRs and SNPs, were detected after mapping all clean reads to the assembled transcripts. The set of unique sequences was searched for SSR markers using MISA (http://pgrc.ipk-gatersleben.de/misa/misa.html). The minimum repeat number used for this search was eight for dinucleotide, five for tri-, four for tetra- and three for penta- and hexanucleotide microsatellites. SSR-containing ESTs were identified as candidates for marker development if they presented enough flanking sequences on either side of the repeats for primer design using Primer 3 (http://primer3.sourceforge.net/releases.php). Putative SNP detection was performed with SOAPSnp software. For identification of potential SNPs, various parameters such as base quality score and read depth were optimized. The following criteria were selected as the final SNP sets: read depth of four and the minimum variant frequency of two, variations compared to the consensus sequence were counted as SNPs. Furthermore, they were considered statistically significant at FDR/tested p-value<0.1.

SSR validation and polymorphism evaluation

Genomic DNA was isolated from 90 randomly selected turbots using TIANamp marine animals DNA kit (TIANGEN, Beijing, China) according to the protocols. The integrity of DNA samples was checked using 1% agarose gel electrophoresis and their purity and concentration were assessed by Nanodrop 1000. All samples with the final concentration of 40 ng/μl were reserved at -20°C for upcoming analysis. The annealing temperature of primers was initially tested for amplification using a pool DNA samples. PCR amplifications were carried out using Master-cycler gradient thermal cycler (Eppendorf) in a final volume of 15 μl. Each reaction tube contains 1.4 μl 10× PCR buffer, 1.2 μl of dNTP (2.5 mM), 0.6 μl of each primer (10 μmol), and 1μl of genomic DNA (40 ng/ul), 0.2 μl of rTaq DNA polymerase (5 U/ ul, Takara), 10 μl of ddH2O. The PCR reaction program was: DNA denaturation at 95°C for 5 min; 30 cycles of 95°C for 45 s, 57~60°C for 50 s, 72°C for 50 s; and 72°C for 10 min as a final extension. Amplification products were resolved in 8% denaturing polyacrylamide gel, visualized by silver-staining to determine allele sizes using a 50-bp DNA ladder as a reference marker.

There were 17 pairs of SSR primers for the assessment of genetic diversity in turbot progenies. The number of alleles (Na), polymorphism information content (PIC), expected and observed heterozygosities (He and Ho, respectively) were calculated with the software PopGene32 (version 1.32).
SNP validation

Genomic DNA of 96 individuals obtained from eight selected families was extracted from tail fins, using TIANamp marine animals DNA kit (TIANGEN, Beijing, China) following the manuscripts. DNA samples were prepared at -20°C with the concentration of 40 ng/μl. In total, 147 SNPs with the coverage ≥500 were chosen as candidates to validate the putative SNPs identified in transcripts, using high resolution melting (HRM) technology. Primers of high quality were designed using Primer premier 5.0 software and synthesized in Sangon (Shanghai, China), where the unlabeled oligonucleotides as the internal temperature controls for genotyping by amplicon melting [30] were produced. HRM genotyping was performed on LightScanner device using primers with distinct and single amplified products and LC Green Plus dye. The genotyping data were analyzed by PopGene32 (version 1.31).

Results and Discussion

Illumina sequencing, reads assembly and gene annotation

Total RNA was isolated from multiple tissues of two adult turbots (one male and one female), including liver, spleen, kidney, cerebrum, gonad and skeletal muscle, for achieving a full-scale *S. maximus* transcriptome. RNA samples were pooled with equal quantities to construct two cDNA libraries and sequenced by two Illumina platforms, respectively. This mixing strategy was commonly reported in some similar studies [23,31,32]. In total, 44,219,773 clean reads of 100 bp in length were received from the two libraries of one male and one female turbot samples after trimming adaptors and low-quality sequences (Table 1). The remaining reads were assembled into 106,643 contigs, of which 71,107 were left as unigenes by eliminating redundancies. As indicated in Table 1, the length distribution of all unigenes was 201–17,407 bp with a mean length of 892 bp. The final assembled sequences and detailed gene annotations were presented in S1 File and S1 Table, respectively.

Unigene annotation

The alignment of non-redundant unigenes was performed with public Nr and Swiss-Prot databases to estimate their putative function. Totally 24,052 unigenes, which took up an approximate proportion of 33.83%, were annotated to the know sequence databases with significant blast scores. As shown in Fig 1, more than half of annotated sequences (50.25%, 12,087) had an E-value from 9E-10 to 1E-110, while 29.38% (7,162) with the E-value to be zero. Nearly two-third (66.17%) of *S. maximus* unigenes were not annotated to any sequences in the reference databases. The low annotation ratio seems unsurprising in non-model organisms without published genomes, especially aquaculture varieties [33–35]. Previous studies on transcriptome analyses indicate that unannotated sequences mainly represent transcripts of spanning only untranslated mRNA regions, chimeric sequences derived from assembly errors [36] and containing non-conserved protein regions [37]. Some may also be components of novel genes.

Table 1. Summary of Illumina transcriptome, assembly and annotation for *Scophthalmus maximus*.

| Raw results (after trimming) | Assembly results | Annotation results |
|-----------------------------|------------------|-------------------|
| Clean bases (G)             | 8,844            | 106,643           | 24052 |
| Read pairs                  | 44,219,773       | 71,107            | 37058 |
| Read length (bp)            | 100              | 201–17,407        | 16540 |
|                            |                  | KEGG hits         | 11938 |

doi:10.1371/journal.pone.0149414.t001
specific to this species, which are likely to be matched to certain genome sequences in the near future.

The result of main species distribution matched against Nr database (Fig 2) showed that 28.74% of the annotated unigenes shared similar sequences with *Oreochromis niloticus*, whose draft genome was published in 2014. The following species were *Neolamprologus brichardi* (10.29%), *Maylandia zebra* (10.27%), *Haplochromis burtoni* (8.38%), *Pundamilia nyererei* (7.48%), *Poecilia formosa* (6.98%), *Takifugu rubripes* (5.77%), *Xiphophorus maculatus* (3.80%), *Oryzias latipes* (3.56%), *Dicentrarchus labrax* (3.41%), *Tetraodon nigroviridis* (1.75%), and others (5.10%). As expected unigenes of turbot transcriptome matched well to proteins of other fish including species with reported genome. The ratio of unigene annotation to *S. maximus* was only 0.33%, which could be due to its few genomic sequences submitted in Genbank.

**Classification of COG, GO and KEGG**

Sequences of assembled unigenes were also subjected to BLASTp searching against databases of COG, GO and KEGG. Summary of statistical results were shown in Table 1.

COG database provides the classification of orthologous gene products. Unigene annotations of COG were selected for checking the completeness of our transcriptome library and the effectiveness of the annotation process. The possible functions of unigenes were predicted and classified by searching their predicted CDSs of unigenes against COG database (Fig 3). Possible functions of 37,058 unigenes were clustered into 25 COG categories, and the top five were 'signal transduction mechanisms' (7,236), 'general function prediction only' (4,521), 'function unknown' (3,143), 'transcription' (3,034), and 'posttranslational modification, protein turnover, chaperone' (2,432), while the three smallest clusters were 'defense mechanisms' (242), 'nuclear structure' (228) and 'cell motility' (93).

Gene Ontology (GO) is an international classification system of standardizing gene function across species to comprehensively profile characteristics of genes, gene products and sequences
In this study, 16,540 unigenes were categorized by GO analysis (Table 1). Second-level GO terms were used to classify the involvement terms of unigenes in three main categories (cellular component, molecular function and biological process) and each unigene was assigned to one or more GO term. In this study, 14,633 unigenes are involved in cellular component categories, among which, 'cell' (14,155 unigenes; 30.86%), 'intracellular' (10,740; 23.03%), 'cytoplasm' (7,486; 16.05%) and 'membrane' (7,050; 15.12%) comprised the largest proportion (Fig 4A). Further, 11,462 unigenes are involved in 25 level-2 terms of molecular function category, and 'binding' (12,114; 27.22%), 'protein binding' (7,650; 17.19%), 'catalytic activity' (5,651; 12.70%) and 'nucleic acid binding' (2,743; 6.16%) were the most abundant (Fig 4B). Additionally, 13,676 unigenes are involved in various biological process categories, as shown in Fig 4C, the top four terms were 'cellular process' (10,468; 19.81%), 'metabolic process' (7,213; 13.65%), 'regulation of biological process' (6,003; 11.36%), and 'macromolecule metabolic process' (5,287; 10.01%). In summary, these GO terms cover a majority of the overall assignments in S. maximus transcriptomic dataset. It is easily understood that genes encoding these functions may be well annotated in the database as a result of their conservation across different species.

KEGG pathway-based analysis facilitated to further study complicated metabolic pathways and biological behaviors of genes [39]. A total of 11,938 unigenes were consequently classified into specific pathways (Fig 5), among which most fell into 'human diseases' (3,516), 'organism
system’ (2,530), and ‘metabolism’ (2,198), followed by ‘cellular processes’ (1,288) and ‘environmental information processing’ (1,284), while least were assigned to ‘genetic information processing’ (1,122). Predominant subcategories of all the pathways were ‘infectious diseases’ (1,439), ‘cancers’ (1,085) and ‘signal transduction’. These annotations offer a valuable resource for investigating specific processes, functions, and pathways in flatfish research.

Identification of genes related to sex determination, growth and immunity

Genes related to sex determination and gonad differentiation. As one of the most promising aquaculture species in Europe and China, turbot shows extreme differential growth rates between sexes. Compared with males, significantly faster growth rate and later sexual maturity of females make all-female population production desirable for turbot industry [40]. Several studies revealed the sex ratio of turbot is determined by major genetic sex-related factors [41] and limitedly influenced by environmental factors like temperature [42]. Genes involved in sex differentiation and gonad development play obvious roles in controlling sex ratio of this species. It seems reasonable that elucidating the mechanisms of sex determination and gonad differentiation is a prior goal to boost turbot production, but delayed due to insufficient known sex-related genes and explicit biological pathways in turbot.

In this study, more than 44 million clean reads were produced using Solexa technology and a large number of unigenes (71,107) were strictly annotated after de novo assembly. The selection criterion on sex-associated genes was mainly referred to the comprehensively gonadal transcriptome information of turbot [15] and olive flounder [14]. Table 2 shows 40 relevant
genes including both well-known genes in SD of fish and novel genes identified for the first time in turbot. Our result shows some genes (ar, mis, sox9, sox6) proved to play roles in testicular development, consistent with the report of Ribas et al [15]. The sox genes encode an important family of transcription factors with highly conserved HMG domain [43], involved in a variety of developmental processes including sex determination and differentiation. Sox9 is an important member of SoxE group and attracts extensive attention, because sox9 enables to lead male development even in the absence of sex-determining region of the Y chromosome (sry), and plays a critical role in the male sex-determining pathway as the downstream gene directly
regulated by SRY in vertebrates [44,45]. Several annotated genes including sox6b (sox6 homologue), ar, hsp90α, dnali1 and ropn1l, are considered to be correlated with turbot testicular development according to the male-specific expression proofs in flounder [14]. The dmrt genes have attracted considerable interest recently because of their involvement in sex determination and differentiation among animal phyla. Dmrt1 with a highly conserved zinc finger-like DNA-binding (DM), is confirmed as the master regulator of male gonad differentiation [46], while its action mechanisms have yet to be elucidated in the turbot. Recently, dmrt1 was validated in the Z chromosome of half-smooth tongue sole (Cynoglossus semilaevis) that belongs to the same order with turbot [47]. This finding will intrigue researchers to seek the more reliable proof through in-depth study of drmt1 on the sex ratio to define the sex determination pattern, based on its firstly identified coding sequence of turbot in this study.
Another group of identified genes in turbot is involved in ovarian development. The female-biased genes found in the present transcriptome, such as *cyp19a*, *zpc5*, *zar1*, *gtc*, *gdf9*, *star*, *start-5/7*, and *gtc*, are well-established in the previous study [15]. Most of them are responsible for the synthesis of female hormones [48]. The transcripts of *sox17* and *sox19* genes were

| Gene Name | Annotation | Query Name | Length | Expression quantity (♂/♀) | Q-value | Trend |
|-----------|------------|------------|--------|---------------------------|---------|-------|
| *dmtr1*   | Doublesex and mab-3-related transcription factor 1 | comp30078_c2_seq1 | 566    | 6.43                      | 1.36E-10 | up    |
| *fshr*    | Follicle stimulating hormone receptor II | comp34552_c0_seq1 | 3389   | 2.39                      | 2.73E-70 | up    |
| *sox6*    | SRY-box containing protein 9 | comp36634_c7_seq1 | 1879   | -0.20                     | 0.47    |       |
| *ar*      | Androgen receptor | comp34574_c1_seq1 | 411    | -0.51                     | 0.69    |       |
| *gsf*     | gonadal soma derived factor 1 | comp38118_c0_seq1 | 2047   | 1.33                      | 1.4E-160 | up    |
| *mdm1*    | meiotic nuclear division protein 1 homolog | comp34969_c0_seq1 | 1598   | -0.75                     | 2.7E-07 |       |
| *mis*     | Müllerian inhibiting substance | comp15802_c0_seq1 | 3092   | 0.22                      | 0.47    |       |
| *hsp90a*  | Cytosolic heat shock protein 90 alpha | comp18173_c0_seq1 | 294    | 0.07                      | 1.00    |       |
| *dnal1*   | Axonemal dynein light intermediate polypeptide 1 | comp21140_c0_seq1 | 1007   | 8.30                      | 5.61E-32 | up    |
| *ropn1l*  | Ropporin-1-like protein-like isoform X2 | comp28799_c0_seq1 | 1720   | 3.26                      | 3.45E-19 | up    |
| *cyp19a*  | P450 aromatase | comp20558_c0_seq1 | 2026   | -7.64                     | 5.1E-45 | down  |
| *zpc5*    | Zona pellucida glycoprotein 5 | comp20390_c0_seq1 | 1801   | -12.65                    | 0       | down  |
| *zar1*    | Zygote arrest protein 1-like | comp20043_c0_seq1 | 1446   | -12.54                    | 0       | down  |
| *gtc*     | Gonadotropin alpha | comp28377_c1_seq1 | 830    | -11.45                    | 0       | down  |
| *gdf9*    | Growth differentiation factor 9 | comp30868_c0_seq1 | 3680   | -5.76                     | 0       | down  |
| *star*    | Steroidogenic acute regulatory protein | comp23962_c0_seq1 | 967    | 0.32                      | 1       |       |
| *start-5* | STAR-related lipid transfer protein 5-like | comp35171_c2_seq1 | 5165   | -1.50                     | 0       | down  |
| *start-7* | STAR-related lipid transfer protein 7-like | comp37065_c0_seq1 | 3481   | -2.00                     | 0       | down  |
| *sox17*   | SRY-box containing protein 17 | comp26556_c1_seq1 | 266    | 1.07                      | 0.48    |       |
| *sox19*   | SRY-box containing protein 19 | comp34381_c0_seq1 | 2961   | -2.04                     | 0       | down  |
| *foxi2*   | Forkhead transcription factor L2 | comp34087_c3_seq1 | 1401   | -5.00                     | 2.68E-46 | down  |
| *vtgr*    | Vitellogenin receptor isoform 1 | comp36664_c17_seq1 | 3808   | -3.35                     | 0       | down  |
| *zp4*     | Zona pellucida sperm-binding protein 4-like | comp30721_c0_seq1 | 2238   | -9.38                     | 3.1E-122 | down  |
| *zp3*     | Zona pellucida sperm-binding protein 3-like | comp29565_c0_seq1 | 1247   | -13.66                    | 0       | down  |
| *zp*      | Zona pellucida sperm-binding protein | comp28111_c0_seq1 | 1390   | -12.61                    | 0       | down  |
| *fem1c*   | Protein fem-1 homolog C-like | comp35674_c0_seq1 | 4180   | -2.02                     | 1.2E-151 | down  |
| *fzd3*    | Frizzled-3-like | comp36698_c1_seq1 | 3514   | -1.84                     | 5.9E-209 | down  |
| *42sp43*  | P43 SS RNA-binding protein-like | comp20349_c2_seq1 | 1515   | -8.56                     | 0       | down  |
| *gtf3a*   | Transcription factor IIIA-like | comp17728_c0_seq1 | 1282   | -14.12                    | 0       | down  |
| *wee2*    | Wee1-like protein kinase 2-like isoform X1 | comp36348_c0_seq1 | 1904   | -7.57                     | 0       | down  |
| *ovca2*   | Ovarian cancer-associated gene 2 protein homolog | comp31026_c0_seq1 | 1363   | -5.28                     | 2.2E-289 | down  |
| *ctss*    | Cathepsin S-like | comp38095_c0_seq1 | 2038   | -8.92                     | 0       | down  |
| *vasa*    | Vasa | comp36885_c2_seq1 | 2658   | -0.21                     | 0       |       |
| *dmrt2*   | Doublesex and mab-3-related transcription factor 2 | comp30078_c0_seq1 | 1113   | -4.68                     | 9.19E-07 | down  |
| *dmrt3*   | Doublesex and mab-3-related transcription factor 3 | comp29037_c2_seq1 | 1392   | -1.55                     | 0.00073 | down  |
| *dmrt5*   | Doublesex and mab-3-related transcription factor 5 | comp55674_c0_seq1 | 291    | -1.51                     | 0.519288 | down  |
| *sox8a*   | SRY-box containing protein 8a | comp36634_c2_seq1 | 2437   | -1.54                     | 2.35E-13 | down  |
| *shbg*    | Sex hormone binding globulin | comp17530_c0_seq1 | 2212   | -1.33                     | 4.1E-62 | down  |
| *cyp1b*   | Steroid 11-beta-hydroxylase | comp248709_c0_seq1 | 422    | 1.49                      | 0.23    |       |

Genes with Italics bold are novel.

doi:10.1371/journal.pone.0149414.t002
significantly upregulated in the differentiation of the ovary in sea bass *Dicentrarchus labrax* [49,50], implicating their roles in ovarian development of fish. *Foxl2* has been validated to be expressed exclusively in female, and joins in ovarian development as the encoding gene of *cyp19a* activating transcription factor [51]. The remaining genes are considered to be engaged in oogenesis (*wec2*), oocyte differentiation and development (*zp*, 42sp43) and vitellogenesis (*vtgr*), justified by Fan et al. [14]. Moreover, *vasa* expressed in the germ cell were identified, which is involved in germ cell determination and development and plays an important role in formation of the primordial germ cell and migration to the germinal ridge [52]. The expression patterns of the above mentioned sex-biased genes support their roles in the turbot gonadal differentiation and development.

In this study, the differential expression of some sex-biased genes (including *dmrt2*, *dmrt3*, *sox8a*) was not consistent with their putative functions as reported in other teleosts. *Dmrt2-3* showed strong male-specific gonadal expression in adult testis of the medaka (*Oryzias latipes*) [53], while both male and female gonadal expression in developing germ cells of medaka [53], zebrafish (*Danio rerio*) [54] and swamp eel (*Monopterus albus*) [55]. The significantly male-biased expression pattern of *Sox8a* (*Sox8* homologue) in *Paramisgurnus dabryanus* [56], *Epinephelus coioides* [57] and *P. olivaceus* [14], suggesting this gene could be essential for differentiation of testis in fish. However, these genes could be assigned as female-biased class on the view of their expression favor in the female adult turbot transcriptome. Therefore, it is intriguing to probe into the diverse roles of these disputable genes in gonad differentiation as well as somite development [58].

**Growth related genes.** Growth rate of cultured fish from hatching to commercial size is one of the most important factors in the success of aquaculture. A variety of genes involved in regulating growth were identified from our sequence database (see Table 3), based on three principal search strategies: 1) associations between genes in the somatotropic axis and growth,

| Gene Name | Annotation | Query Name | Length | Expression quantity (♂/♀) | Q-value | Trend |
|-----------|------------|------------|--------|----------------------------|---------|-------|
| gh        | Growth hormone precursor | comp17564_c0_seq1 | 775 | -10.66 | 4.01E-123 | down |
| ghr       | Growth hormone receptor | comp34224_c0_seq1 | 1177 | -0.78 | 0.04 | |
| ghhr      | Growth hormone releasing hormone | comp374142_c0_seq1 | 325 | 2.66 | 0.10 | |
| ghih      | Growth hormone inhibiting hormone-like | comp23996_c0_seq1 | 1004 | 0.32 | 0.59 | |
| sstr      | Somatostatin receptor | comp24799_c1_seq1 | 1202 | -1.13 | 0.0039 | |
| igf       | Insulin-like growth factors | comp38041_c8_seq1 | 3754 | -0.88 | 1.85E-45 | |
| igfr      | IGF receptor | comp33325_c0_seq1 | 3008 | -0.44 | 0.040 | |
| igfbp     | IGF binding protein | comp32674_c4_seq1 | 1454 | -1.43 | 4.03E-51 | down |
| mstn      | Myostatin | comp20346_c1_seq1 | 2996 | 0.16 | 0.60 | |
| myf5      | Myogenic regulatory factor 5 | comp37586_c17_seq1 | 1393 | -2.65 | 1.21E-34 | down |
| fabp      | Fatty acid-binding protein | comp38300_c0_seq1 | 606 | -4.13 | 0 | down |
| htr       | 5-Hydroxytryptamine receptor | comp26484_c1_seq1 | 1741 | 1.66 | 0.48 | |
| prl       | Prolactin | comp38254_c0_seq1 | 747 | -15.12 | 2.99E-188 | down |
| prlr      | Prolactin receptor | comp37125_c1_seq1 | 6331 | -2.36 | 0 | down |

doi:10.1371/journal.pone.0149414.t003
2) controlling growth at the muscle tissue level, 3) other candidate genes related to growth. A total of 60 assembled sequences, partial encoding regions of 17 genes or gene families, were verified in this study, which have been identified previously to have roles in growth of fish and other species.

Growth hormone-releasing hormone (GHRH), growth hormone inhibiting hormone (GHIH or somatostatin), growth hormone (GH), insulin-like growth factors (IGF-I and -II), and relevant carrier proteins and receptors, main components of the somatotropic axis, are widely accepted to play a critical role in regulating the formation of skeletal muscles in finfish [9]. As the main regulator of postnatal somatic growth, GH has been proved to play a vital role in stimulating anabolic processes such as cell division, skeletal growth and protein synthesis. Polymorphisms in the piscine GH gene have shown association with growth performance of *Salmo salar* [59] and *P. olivaceus* [60]. The biological actions of GH on target cells, including transmembrane signal transduction and subsequently transcriptional induction of many genes (e.g. IGF-I), are mediated by its receptor GHR. The study on Atlantic salmon [61] provides strong evidence that the expression of GHR regulates the production of IGF-I in fish, which has a pivotal role in growth determination [62]. Therefore, the GHR gene should be further examined as a possible candidate for growth improvement in finfish. GHRH codes for a peptide important in upregulating the GH expression [63]. Six signal nucleotide polymorphisms (SNPs) in the 5’UTR of cattle GHRH gene were revealed association with growth traits including weight improvement [64]. A SNP in the GHRH fourth intron out of Arctic charr was related to a significant increase (9.4%) in growth rate of early life stages [65]. However, there are few studies on the variability within this gene of flatfishes reported despite the importance of GHRH in the somatotropic hormonal axis.

Myostatin (MSTN, also known as GDF-8), a member of the transforming growth factor-β (TGF-β) superfamily, functions as a negative regulator of skeletal muscle development and growth [66]. Suppression of MSTN in the transgenic fishes resulted in the increase of muscle production [67]. Recently, MSTN has become a focal gene in the polymorphism detection and association studies towards selective breeding for growth traits in livestock [68] and some aquaculture species, such as the mollusk [69] and genetically improved farmed tilapia [70]. Myf-5 is a key member of the myogenic regulatory factors (MRFs) with a characteristic basic helix-loop-helix (bHLH) domain. A mutation of a single-base pair in one intron of the myf5 gene associating with increases of the cattle body weight [71] suggests that this gene is a potential candidate for marker-assisted selection of economic varieties. The reported studies of myf5 mainly focused on molecular structure, dynamic expression, and promoter analysis in some fish [72,73] excluding *S. maximus*. Reports on polymorphisms of this gene have not been found as yet.

Fatty acid-binding proteins (FABPs), belonging to a superfamily of intracellular lipid-binding proteins, occur ubiquitously in tissues of vertebrates and invertebrates with distinct expression patterns for the individual FABPs. These proteins have multiple proposed roles, such as promoting the cellular uptake and transfer of fatty acids (FAs), targeting FAs to specific metabolic pathways, and involving in the regulation of gene expression and cell growth [74]. There are various members of the FABP family, of which liver (L-), intestinal (I-) and heart (H-) are the dominating types [75]. As shown in Table 3, transcripts of the three FABPs were all detected in the present transcriptome. To date, the progress of fish *fabps* is limited to the expression patterns in Atlantic salmon [76] and the promoter function in zebrafish [77]. Consequently, there exists a broad area of interest to explore the functions of FABP family in turbot *S. maximus*. SNP in the gene encoding 5-hydroxytryptamine receptor has shown the significant association with growth trait in crustaceans [78], while its functions and polymorphisms in fish are hardly ever to be investigated.
Prolactin (PRL) is an important regulator with multiple biological functions through binding to its receptor (PRLR) in fish [79]. Studies of calcium uptake in larval tilapia [80] indicate that PRL could be involved in regulating calcium balance as has been suggested in adult fish, while more evidence needs to be revealed about the role of PRL in larval calcium balance, because calcium accretion is important for numerous processes, particularly skeletal formation a process initiating soon after hatch in fish. Few reports in fish support a somatotropic action for PRL. It was suggested to influence growth of Mozambique tilapia with stimulating liver IGF-I production [81]. Additionally, transcripts of PRL and PRLR have been identified in association with post-hatching development of fish larvae [82]. Regarding the effect of PRL in fish immune function, it is increasingly clear that PRL is an important modulator mediated by PRLR [83], enhancing mitosis as well as phagocytic activity of leucocytes [84], and stimulating immunoglobulin production [85]. Immunoregulation in larval fish is crucial for commercial interest to be attained in aquaculture. Overall, PRL exerts multiple functions with receptor PRLR in fish, and probably also in fish larvae, and may have a different spectrum of activities in different species. Further work is required to fully characterise their activities.

It is well-reasoned that the different expression abundance of growth related genes are mainly responsible for the marked sex difference in growth rate of turbot. The expression of many selected genes in Table 3 was detected to be significantly downregulated in the male transcriptome, compared with that of the female, including gene of GH precursor, igfbp2, myf5, fabp, htr, prl and prlb, all of which have been proved to have positive effects on growth of vertebrates. The result may provide supplementary evidence on the validation of the chosen growth genes.

**Candidate genes related to immune response.** The innate and adaptive immunity compose the immune defense system of fish, and the former plays a major role in immune response than that of the latter in fish [86]. The knowledge of the relevant genes for immune response in *S. maximus* has greatly increased recently with the effort of high-throughput sequencing. Pereiro et al. [87] detected the antiviral transcripts of immune-related tissues from infected turbots using 454-pyrosequencing and provided a rich source of data to increase the knowledge of *S. maximus* immune transcriptome. Consequently, a large number of contigs and singletons involved in innate and acquired immunity were discovered after combining the Sanger and pyrosequencing data [15]. Here, a significant number of genes detected in our transcriptome (see S2 Table) were confirmed to be the main components of the immune pathways (complement, Toll-like receptor signaling, B cell receptor signaling, T cell receptor signaling and programmed cell death), agreed with the previous two results. Certainly, several interesting genes related to innate and acquired immunity were presented in Table 4.

Recognition of pathogen-associated molecular patterns (PAMPs) mediated by pattern-recognition receptors (PRRs) is critical to the initiation of innate immune responses. PRRs sense the conserved molecular structure (PAMPs) of a pathogen and induce subsequent host immunity through multiple signaling pathways for eradicating the pathogen [88]. Toll-like receptors (TLRs) are the first characterized PRRs, sharing structural and functional similarities from Drosophila to humans [89]. They are believed to play a crucial role in host defense of pathogenic microbes in innate immune system through recognizing PAMPs expressed on infectious agents. Three crucial members of TLR group (*TLR1, TLR21* and *TLR22*) were identified in the genomic database of turbot. TLR1 plays an essential role in pathogen recognition and activation of innate immunity [90]. In recent years, *TLR1* has been characterized in a number of fish, such as *Tetraodon nigroviridis* [91], *Epinephelus coioides* [92], and *P. olivaceus* [93]. TLR22 occurs exclusively in aquatic animals with similar functions of mammalian TLR3, and supervises the infection of dsRNA virus to alert the immune system for antiviral protection in fish [94]. *TLR22* evolves functional diversification and adaptation of the response to different
PAMPs while the information of TRL21 is still rare in the model zebrafish [95], not to mention turbot. It is evident that TNF receptor-associated factor 3 (TRAF3) plays multiple roles in mammalian T and B lymphocytes [96], such as antagonizing the effects of TRAF2 in NF-κB activation, but the function of TRAF3 in the two piscine immune cells needs to be investigated. Interleukin-34 (IL-34) is a cytokine that promotes the differentiation and viability of monocytes and macrophages through the colony-stimulating factor-1 receptor [97]. Genes encoding two interleukin receptors essential for ILs mediated signal transduction were first detected in the turbot cDNA library. C9 encodes the final component of the complement system, which participates in the formation of the membrane attack complex (MAC) that assembles on bacterial membranes to form a pore, permitting disruption of bacterial membrane organization. C-type lectin is a large family that includes all known collectins and selectins in animals. The collectins with carbohydrate recognition domains, are secreted proteins that play important roles in the innate immune system by binding to carbohydrate antigens on microorganisms, facilitating their recognition and removal. The selectin P may play a role in inflammatory response. The encoding genes of these C-type lectins found in the present database will be useful to investigate their roles in host defense of S. maximus. Many cytokines that play pivotal roles in the mammalian specific immunity have been validated in teleost fish, including transforming growth factor-β (TGF-β), and a number of IL involved in adaptive immune responses. IL-15 was discovered in relevance to the adaptive immune response of pufferfish [98] and rainbow trout [99]. All isoforms of the TGF-β family have been identified in a variety of teleosts, such as Danio rerio [100], Oncorhynchus mykiss [101], Morone saxatilis [102]. However, information on biological functions of these molecules in fish immune system remains limited. The identified nucleotide sequences of IL-15, receptors of other ILs, TGF-β and related receptors in this study will support the foundation to make their functions in turbot adaptive immunity clear.

**Molecular markers**

**SSR characterization and polymorphism evaluation.** SSRs have been widely used in construction of genetic linkage, QTL analysis and assessment of genetic diversity in aquaculture

| Gene Name                  | Annotation                          | Query Name |
|----------------------------|-------------------------------------|------------|
| Toll-like receptor signaling pathway | TLR1 | Toll-like receptor 1 | comp32484_c3_seq1 |
|                            | TOLLIP | Toll-interacting protein | comp21400_c0_seq1;comp23286_c0_seq1;comp23286_c1_seq1 |
|                            | TRAF3 | PREDICTED: TNF receptor-associated factor 3-like | comp22885_c1_seq1;comp23509_c0_seq1 |
|                            | TLR21 | Toll-like receptor 21 | comp32484_c3_seq1 |
|                            | TLR22 | Toll-like receptor 22 | comp36502_c11_seq1 |
|                            | IL-34 | Interleukin-34 | comp30033_c0_seq1 |
|                            | IL18R1 | Interleukin-18 receptor 1 | comp3046_c0_seq1;comp22532_c0_seq1 |
|                            | IL1R1 | Interleukin-1 receptor 1 | comp28977_c1_seq1 |
| Apoptosis or programmed cell death | TRAF2 | PREDICTED: TNF receptor-associated factor 2-like | comp26413_c0_seq1;comp26467_c3_seq1;comp26467_c5_seq1 |
| Complement pathway | C9 | PREDICTED: Complement component C9-like | comp37672_c0_seq1 |
| Other immune molecules | SLEP | Selectin P | comp35827_c0_seq1 |
|                            | COLEC | Collectin 10/11/12 | comp31521_c1_seq1;comp31124_c0_seq1;comp30265_c0_seq1;comp31642_c0_seq1 |

PLOS ONE | DOI:10.1371/journal.pone.0149414 February 29, 2016 16 / 26
species, due to the distinct advantages of high variability, abundance, neutrality and co-dominance [103]. An important emerging application of high-throughput Illumina-Solexa is the identification of molecular markers from genomic DNA. Our search revealed 21,192 SSRs were contained in ESTs of the transcriptomic dataset, of which 39.58% were di-nucleotide repeats, followed by 38.89% tri-nucleotide repeats and 21.53% tetra/penta/hexa-nucleotide-repeats (Fig 6). The most abundant SSR repeat types of animals are generally believed to be di-nucleotide repeats [36,41], and our findings are coincident with this conclusion. In the di-nucleotide repeats motifs, (GT/TG)n and (AC/CA)n were the two predominant types with frequencies

![Fig 6. Frequency distribution of SSRs by motif length found in Scophthalmus maximus.](https://doi.org/10.1371/journal.pone.0149414.g006)

![Fig 7. Polyacrylamide gel electrophoresis for one SSR marker (comp15993_c0_seq1) in the 30 individuals.](https://doi.org/10.1371/journal.pone.0149414.g007)
of 42.72% and 29.94%, respectively. Among the 20 types of tri-nucleotide repeats, (GGA/GAG/AGG)n, (GCA/CAG/AGC)n, (TCC/CCT/CTC)n and (AAG/AGA/GAA)n were the leading types with a combined frequency of 46.46%.

In this study, 2,357 (11.22%) SSR-containing sequences enabled the design of primers (S3 Table), which are the highly desirable development of SSRs for this species. One hundred SSRs were randomly selected for primer synthesis and identification, among which, 70 received clearly amplified target products in PCR amplification. Of these, seventeen were available for the further genetic linkage map construction based on a turbot family through polymorphism validation in a population of 30 individuals (Fig 7). Using the 17 primer pairs, we described the genetic structure characterization of a turbot family including 90 progenies (Table 5). A total of

| Transcripts ID | Repeat motif | Primer sequences (5'-3') | Length | Ta | Na | He | Ho | PIC |
|----------------|--------------|--------------------------|--------|----|----|----|----|-----|
| comp15993_c0_seq1 | (GACCTG)3 | F: ACTGGAGACGGTCCGATACCTG  R: ACTCATGGTGTGTCAGGTC | 129 | 58 | 5 | 0.75 | 0.79 | 0.65 |
| comp18360_c0_seq1 | (CCTGGA)3 | F: ACTCTTGGCGCTGATGAAAG  R: GCTCTGGGCTGATGTCGTC | 125 | 58 | 2 | 0.5 | 0.39 | 0.27 |
| comp20204_c0_seq1 | (CTC)6 | F: GTAGCTGCTGCTGAGTGAAG  R: GAGGATGAGGAGGAGGAGG | 132 | 57 | 2 | 0.5 | 0.41 | 0.27 |
| comp20308_c0_seq1 | (GAC)5 | F: CATGCACAGGACAGGACAG  R: ACTGTGCTGCTGTCGTC | 233 | 57 | 2 | 0.5 | 0.58 | 0.33 |
| comp20617_c0_seq1 | (GGA)5 | F: CCTTCACCCTTCACCCTTCGC  R: TATCCACCACAGCAGCAGC | 120 | 57 | 2 | 0.5 | 0.35 | 0.37 |
| comp21529_c1_seq1 | (GCT)6 | F: TTCTTTTTTCCCCAGGAGCTG  R: GCCCGTTGGATCAGAGCTGTTC | 120 | 58 | 2 | 0.5 | 0.43 | 0.28 |
| comp21594_c1_seq1 | (CGC)6 | F: TTCGCCACACAGCGAGAATCTC  R: TCCGACCTGTCTCCACATGTC | 213 | 57 | 2 | 0.5 | 0.53 | 0.31 |
| comp29608_c1_seq1 | (GCTA)5 | F: AGACCAGATCCTGGATGACAC  R: TCAGGTGGTTCCGTTCTTC | 148 | 57 | 2 | 0.5 | 0.48 | 0.37 |
| comp30754_c0_seq1 | (AGTTTC)6 | F: CAACTCCCTCCTGCTGCTGCTC  R: AGGAACGATGAGGAGGCTG | 135 | 59 | 2 | 0.5 | 0.42 | 0.28 |
| comp31253_c0_seq1 | (TGC)6 | F: GTAGGCGCCATAGTTGAGCTG  R: TCCCAGCACCAGACAGCAG | 143 | 58 | 2 | 0.5 | 0.46 | 0.29 |
| comp31355_c0_seq1 | (AAT)7 | F: AGACCAAACTCAAACCATACGC  R: TGGAGAACAGGACACAGCAG | 178 | 58 | 2 | 0.5 | 0.35 | 0.25 |
| comp34582_c0_seq1 | (CGTG)5 | F: GTCTCTGCTGACCTGCTGCTC  R: GCGAGCACTGCTCAGCACAA | 187 | 59 | 2 | 0.5 | 0.53 | 0.31 |
| comp36059_c15_seq1 | (CGTG)6 | F: AAGGAGGGGGTGAGGACAGGTC  R: GCCAGCTGAGAGGAGGAGG | 207 | 58 | 2 | 0.75 | 0.64 | 0.5 |
| comp36274_c2_seq1 | (TGG)7 | F: GTCTCTGTGCTGCAGAGACG  R: ACTCTTGACACGACACACAC | 126 | 58 | 2 | 0.5 | 0.49 | 0.3 |
| comp36960_c0_seq1 | (AG)9 | F: GTCCGCTGTGCTCCAGACATG  R: TGGACACTCTGAGGACGCTC | 123 | 58 | 2 | 0.5 | 0.51 | 0.3 |
| comp37121_c6_seq1 | (GAG)6 | F: GGAGAGACACGAGGAGACAC  R: GTGTTCTTGATGCCGACAG | 174 | 58 | 2 | 0.5 | 0.29 | 0.22 |
| comp37181_c7_seq1 | (GCTC)5 | F: AAGAGGCGGCTGCTGAGAG  R: GTGAGGCGGTCTCCCGTCAG | 182 | 58 | 4 | 1 | 1 | 0.7 |

Ta annealing temperature, He expected heterozygosity, Ho observed heterozygosity, Na observed number of alleles, PIC polymorphic information content. 

doi:10.1371/journal.pone.0149414.t005
39 alleles were identified, with an average of 2.29 alleles per locus. The polymorphic information content (PIC) ranged from 0.22 to 0.7 with an average of 0.35, indicating that these identified EST-SSRs were at least moderate polymorphic.

**SNP characterization and validation.** Putative SNPs were detected from alignments of multiple sequences during contig assembly. In this study, a total of 8,642 SNPs were obtained, of which 4,894 were transitions (Ts) and 3,748 were transversions (Tv), giving a mean Ts: Tv ratio of 1.306:1 across the turbot transcriptome (Fig 8). The four transitions were the most common SNP types, and GC/CG transversions were the least SNP types on account of the difference in base structure and number of hydrogen bonds between different base [23,35].

To verify the potential SNPs, 63 primer pairs were designed according to 45 contigs containing SNP with high coverage, of which 21 pairs amplified the exclusive products through 2% agarose gel electrophoresis detection. The distinctly different genotypes on the twenty-one SNPs were validated by HRM genotyping of 96 samples (Fig 9). The polymorphism evaluation indicated that most SNPs (19 of 21) are moderate polymorphic sites and matched to Hardy-Weinberg equilibrium (S4 Table). Moreover, the effective annotations of genes with the identified SNP locations showed that these genes are involved in regulation of cell cycle, cytoskeleton, energy transformation and RNA processing [104].

Taken together, the results obtained in this study indicated that these potential molecular markers identified within the ESTs will enable more detailed studies on evolutionary genomics, comparative mapping, and QTL analysis of *Scophthalmus maximus*.

**Conclusion**

Here we report the comprehensive transcriptome of major tissues in turbot *Scophthalmus maximus*, a commercially important flatfish in China. The large amount of generated sequences (71,107 putative transcripts) will enrich enable genomic resources in turbot and therefore to improve available sequence databases for gene discovery. A significant number of putative genes related to economic traits were identified to facilitate genomics approaches for
controlling sex ratio, improving growth performance and resistance to pathogens in domesticated stocks used for aquaculture. A large amount of genetic markers was detected, providing new tools for genomic studies and management of molecular assistant selection in cultured populations.

Fig 9. Genotyping result using HRM with small amplicon. (A) S10 genotyping; (B) S11 genotyping; (C) S14 genotyping.

doi:10.1371/journal.pone.0149414.g009
Supporting Information

S1 File. All assembled sequences of contigs.
(RAR)

S1 Table. The detailed annotation information of genes.
(XLS)

S2 Table. All identified genes related to immune response.
(XLS)

S3 Table. Summary of EST-SSRs with 2–6 bp motif repeats and primers.
(XLS)

S4 Table. Genetic variability at 21 SNPs in Scophthalmus maximus.
(DOC)

Acknowledgments
We appreciate Tian-yuan Fisheries Co. Ltd. (Yantai, China) for sample offer and the Chinese National Human Genome Center in Shanghai (CHGCS) for transcript sequencing.

Author Contributions
Conceived and designed the experiments: DM AM ZH. Performed the experiments: ZH GW TW. Analyzed the data: DM ZH. Contributed reagents/materials/analysis tools: DX BM. Wrote the paper: DM.

References
1. Qi ZZ, Zhang XH, Boon N, Bossier P. Probiotics in aquaculture of China—current state, problems and prospect. Aquaculture. 2009; 290: 15–21.
2. Rodriguez-Ramilo ST, De La Herran R, Ruiz-Rejon C, Hermida M, Fernandez C, Pereiro P, et al. Identification of quantitative trait loci associated with resistance to viral haemorrhagic septicemia (VHS) in turbot (Scophthalmus maximus): a comparison between bacterium, parasite and virus diseases. Mar Biotechnol. 2014; 16: 265–276. PMID: 24078233
3. Cnaani A, Levavi-Sivan B. Sexual development in fish, practical applications for aquaculture. Sex Dev. 2009; 3: 164–175. doi:10.1159/000223080 PMID: 19684460
4. Devauchelle N, Alexandre JC, Corre NL, Letty Y. Spawning of turbot (Scophthalmus maximus) in captivity. Aquaculture. 1988; 69: 159–184.
5. Piferrer F, Guiguen Y. Fish Gonadogenesis. Part II: Molecular Biology and Genomics of Sex Differentiation. Rev Fish Sci. 2008; 16: 35–55.
6. Devlin RH, Nagahama Y, Devlin RH, Nagahama Y. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. Aquaculture. 2002; 208: 191–364.
7. Cal RM, Vidal S, Gómez C, Álvarez-Biázquez B, Martínez P, Piferrer F. Growth and gonadal development in diploid and triploid turbot (Scophthalmus maximus). Aquaculture. 2006; 251: 99–108.
8. Haffray P, Lebègue E, Jea S, Guennoc M, Guiguen Y, Baroller JF, et al. Genetic determination and temperature effects on turbot Scophthalmus maximus sex differentiation: An investigation using steroid sex-inverted males and females. Aquaculture. 2009; 294: 30–36.
9. De-Santis C, Jerry DR. Candidate growth genes in finfish—Where should we be looking? Aquaculture. 2007; 272: 22–38.
10. Alves E, Faustino MAF, Tomé JPC, Neves MGP, Tomé AC, Cavaleiro JAS, et al. Photodynamic antimicrobial chemotherapy in aquaculture: photoinactivation studies of vibrio fischeri. PLoS One. 2011; 6: e20970. doi: 10.1371/journal.pone.0020970 PMID: 21698119
11. Magnadottir B. Immunological control of fish diseases. Mar Biotechnol. 2010; 12: 361–379. doi: 10.1007/s10126-010-9279-x PMID: 20352271
12. Pardo B, Fernández C, Milán A, Bouza C, Vázquez-López A, Vera M, et al. Expressed sequence tags (ESTs) from immune tissues of turbot (Scophthalmus maximus) challenged with pathogens. BMC Vet Res. 2008; 4: 1–12.

13. Ekblom R, Galindo J. Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity. 2011; 107: 1–15. doi: 10.1038/hdy.2010.152 PMID: 21139633

14. Fan ZF, You F, Wang LJ, Weng SD, Wu ZH, Hu JW, et al. Gonadal transcriptome analysis of male and female olive flounder (Paralichthys olivaceus). Biomed Res Int. 2014; 2014: 291067. doi: 10.1155/2014/291067 PMID: 25121093

15. Ribas L, Pardo BG, Fernandez C, Alvarez-Dios JA, Gomez-Tato A, Quiroga MI, et al. A combined strategy involving Sanger and 454 pyrosequencing increases genomic resources to aid in the management of reproduction, disease control and genetic selection in the turbot (Scophthalmus maximus). BMC Genomics. 2013; 14: 180. doi: 10.1186/1471-2164-14-180 PMID: 23497389

16. Poompuang S, Hallerman EM. Toward detection of quantitative trait loci and marker-assisted selection in fish. Rev Fish Sci. 1997; 5: 253–277.

17. Navajas-Pérez R, Robles F, Molina-Luzón M, De La Herran R, Álvarez-Dios J, Pardo BG, et al. Exploitation of a turbot (Scophthalmus maximus L.) immune-related expressed sequence tag (EST) database for microsatellite screening and validation. Mol Ecol Resour. 2012; 12: 706–716. doi: 10.1111/j.1755-0998.2012.03126.x PMID: 22385869

18. Bouza C, Hermida M, Pardo BG, Fernández C, Fortes GG, Castro J, et al. A microsatellite genetic map of the turbot (Scophthalmus maximus). Genetics. 2007; 177: 2457–2467. PMID:18073440

19. Bouza C, Hermida M, Pardo BG, Vera M, Fernández C, De La Herrán R, et al. An Expressed Sequence Tag (EST)-enriched genetic map of turbot (Scophthalmus maximus): a useful framework for comparative genomics across model and farmed teleosts. BMC Genet. 2012; 13: 54. doi:10.1186/1471-2156-13-54 PMID: 22747677

20. Martinez P, Bouza C, Hermida M, Fernandez J, Toro A, Vera M, et al. Identification of the major sex-determining region of turbot (Scophthalmus maximus). Genetics. 2009; 183: 1443–1452. doi: 10.1534/genetics.109.107979 PMID: 19786621

21. Sánchez-Molano E, Cerna A, Toro MA, Bouza C, Hermida M, Pardo BG, et al. Detection of growth-related QTL in turbot (Scophthalmus maximus). BMC Genomics. 2011; 12: 473. doi:10.1186/1471-2164-12-473 PMID: 21958071

22. Rodríguez-Ramilo ST, Toro MA, Bouza C, Hermida M, Pardo BG, Cabaleiro S, et al. QTL detection for Aeromonas salmonicida resistance related traits in turbot (Scophthalmus maximus). BMC Genomics. 2011; 12: 541. doi: 10.1186/1471-2164-12-541 PMID: 22047500

23. Lv JJ, Liu P, Gao BO, Wang Y, Wang Z, Chen P, et al. Transcriptome analysis of the Portunus trituberculatus: de novo assembly, growth-related gene identification and marker discovery. PLoS One. 2014; 9: e94055. doi: 10.1371/journal.pone.0094055 PMID: 24722690

24. Cui ZX, Li XH, Liu Y, Song CW, Hui M, Shi GH, et al. Transcriptome profiling analysis on whole bodies of microbial challenged Eriocheir sinensis larvae for immune gene identification and SNP development. PLoS One. 2013; 8: e82156. doi: 10.1371/journal.pone.0082156 PMID: 24324760

25. Vera M, Alvarez-Dios JA, Fernandez C, Bouza C, Vilas R, Martinez P, et al. Development and validation of single nucleotide polymorphisms (SNPs) markers from two transcriptome 454-runs of turbot (Scophthalmus maximus) using high-throughput genotyping. Int J Mol Sci. 2013; 14: 5694–5711. doi: 10.3390/ijms14035694 PMID: 23481633

26. Ng P, Wei CL, Sung WK, Chiu KP, Lipovich L, Ang CC, et al. Gene identification signature (GIS) analysis for transcriptome characterization and genome annotation. Nat Methods. 2005; 2: 105–111. PMID: 15782207

27. Rice P, Longden I, Bleasby A. EMBOSS: The European molecular biology open software suite. Trends Genet. 2000; 16: 276–277. PMID: 10827456

28. Chen ZZ, Xue CH, Zhu S, Zhou FF, Ling XFB, Liu GP, et al. GOpipe: Streamlined Gene Ontology annotation for batch anonymous sequences with statistics. Prog Biochem Biophys. 2005; 32: 187–190.

29. Wang LK, Feng ZX, Wang X, Wang XW, Zhang XG. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics. 2010; 26: 136–138. doi: 10.1093/bioinformatics/btp612 PMID: 19855105

30. Seipp MT, Durttschi JD, Liew MA, Williams J, Damjanovich K, Pont-Kingdon G, et al. Unlabeled oligonucleotides as internal temperature controls for genotyping by amplicon melting. J Mol Diagn. 2007; 9: 284–289. PMID: 17591926
31. Li SH, Zhang XJ, Sun Z, Li FH, Xiang JH. Transcriptome analysis on Chinese shrimp *Fenneropenaeus chinensis* during WSSV acute Infection. PLoS One. 2013; 8: e58627. doi: 10.1371/journal.pone.0058627 PMID: 23527000

32. Li CZ, Weng SP, Chen YG, Yu XQ, Lü L, Zhang HQ, et al. Analysis of Li SH, Zhang XJ, Sun Z, Li FH, Xiang JH. Transcriptome analysis on Chinese shrimp *Fenneropenaeus chinensis* during WSSV acute Infection. PLoS One. 2013; 8: e58627. doi: 10.1371/journal.pone.0058627 PMID: 23527000

33. Robledo D, Ronza P, Harrison PW, Losada AP, Bermudez R, Pardo BG, et al. RNA-seq analysis reveals significant transcriptome changes in turbot (*Scophthalmus maximus*) suffering severe entero- myxosis. BMC Genomics. 2014; 15: 1149. doi: 10.1186/1471-2164-15-1149 PMID: 25526753

34. Jung H, Lyons RE, Dinh H, Hurwood DA, McWilliam S, Mather PB. Transcriptomics of a giant freshwater prawn (*Macrobrachium rosenbergii*): de novo assembly, annotation and marker discovery. PLoS One. 2011; 6: e27938. doi: 10.1371/journal.pone.0027938 PMID: 22174756

35. Ma KY, Qiu GF, Feng JB, Li JL. Transcriptome analysis of the oriental river prawn, *Macrobrachium nipponense* using 454 pyrosequencing for discovery of genes and markers. PLoS One. 2012; 7: e39727. doi: 10.1371/journal.pone.0039727 PMID: 22106843

36. Wang JPZ, Lindsay BG, Leebens-Mack J, Cui LY, Wall K, Miller WC, et al. EST clustering error evaluation and correction. Bioinformatics. 2004; 20: 2973–2984. PMID: 15189818

37. Mittapalli O, Bai X, Mamidala P, Rajarapu SP, Bonello P, Herms DA. Tissue-specific transcriptomics of the exotic invasive insect pest emerald ash borer (*Agrilus planipennis*). PLoS One. 2010; 5: e13708. doi: 10.1371/journal.pone.0013708 PMID: 21060843

38. Harris MA, Deegan JI, Lomax J, Ashburner M, Tweedie S, Carbon S, et al. The Gene Ontology project: a knowledgebase of functions and classes. Nucleic Acids Res. 2004; 32: D258–261. PMID: 14697332

39. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000; 28: 29–34. PMID: 10592173

40. Piferrer F, Cal RM, Gómez C, Álvarez-Blázquez B, Castro J, Martínez P. Induction of gynogenesis in *Xenopus laevis* using *454* pyrosequencing for discovery of genes and markers. PLoS One. 2012; 7: e9338598.

41. Schepers GE, Teasdale RD, Koopman P. Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families. Dev Cell. 2002; 3: 167–170. PMID: 12198448

42. Sekido R, Lovell-Badge R. Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. Nature. 2008; 453: 930–934. doi: 10.1038/nature06944 PMID: 18454134

43. Kent J, Wheatley SC, Andrews JE, Sinclair AH, Koopman P. A male-specific role for SOX9 in vertebrate sex determination. Development. 1996; 122: 2813–2819. PMID: 8787755

44. Herpin A, Schartl M. Sox11 genes at the crossroads: a widespread and central class of sexual development factors in fish. Cell Tissue Res. 2009; 336: 271–282. PMID: 19281449

45. Chen SL, Zhang GJ, Shao CW, Huang QF, Liu G, Zhang P, et al. Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. Nat Genet. 2014; 46: 253–260. doi: 10.1038/ng.2890 PMID: 24487278

46. Guiguen Y, Fostier AF, Chang CF. Ovarian aromatase and estrogens: a pivotal role for gonadal sex differentiation and sex change in fish. Gen Comp Endocr. 2010; 165: 352–366. doi: 10.1016/j.ygcen.2009.03.002 PMID: 19289125

47. Navarro L. Characterisation and expression during sex differentiation of Sox19 from the sea bass *Dicentrarchus labrax*. Comp Biochem Phys B. 2012; 163: 316–323.

48. Navarro-Martín L, Galay-Burgos M, Piferrer F. Different *sox17* transcripts during sex differentiation in sea bass, *Dicentrarchus labrax*. Mol Cell Endocrinol. 2009; 299: 240–251. doi: 10.1016/j.mce.2008.11.013 PMID: 19071190

49. Ikenishi K, Tanaka TS. Involvement of the protein of *Xenopus vasa* homolog (*Xenopus* vasa-like gene 1, *XVLG1*) in the differentiation of primordial germ cells. Dev Growth Differ. 1997; 39: 625–633. PMID: 9338598
53. Winkler C, Hornung U, Kondo M, Neuner C, Duschl J, Shima A, et al. Developmentally regulated and non-sex-specific expression of autosomal dmrt genes in embryos of the medaka fish (Oryzias latipes). Mech Develop. 2004; 121: 997–1005.

54. Li Q, Zhou X, Guo YQ, Shang X, Chen H, Lu H, et al. Nuclear localization, DNA binding and restricted expression in neural and germ cells of zebrafish Dmrt3. Biol Cell. 2008; 100: 453–463. doi: 10.1042/BC20070114 PMID: 18282142

55. Sheng Y, Chen B, Zhang L, Luo MJ, Cheng HH, Zhou RJ. Identification of Dmrt genes and their up-regulation during gonad transformation in the swamp eel (Monopterus albus). Mol Biol Rep. 2014; 41: 1237–1245. doi: 10.1007/s11033-013-2988-6 PMID: 24390316

56. Xia XH, Zhao J, Du QY, Chang ZJ. cDNA cloning and expression analysis of two distinct Sox8 genes in Paramisgurnus dabananus (Cypriniformes). J Genet. 2010; 89: 183–192(110). PMID: 20861569

57. Liu QY, Lu HJ, Zhang LH, Xie J, Shen WY, Zhang WM. Homologues of sox8 and sox10 in the orange-spotted grouper Epinephelus coioides: sequences, expression patterns, and their effects on cyp19a1a promoter activities in vitro. Comp Biochem Phys B. 2012; 163: 86–95.

58. Seo KW, Wang YD, Kubot K, Kettower DA, Johnson RL. Targeted disruption of the DM domain containing transcription factor Dmrt2 reveals an essential role in somite patterning. Dev Biol. 2006; 290: 200–210. PMID: 16387292

59. Gross R, Nilsson J. Restriction fragment length polymorphism at the growth hormone 1 gene in Atlantic salmon (Salmo salar L.) and its association with weight from the offspring of a hatchery stock. Aquaculture. 1999; 173: 73–80.

60. Kang JH, Lee SJ, Park SR, Ryu HY. DNA polymorphism in the growth hormone gene and its association with weight in olive flounder Paralichthys olivaceus. Fisheries Sci. 2002; 68: 494–498.

61. Wargelius A, Fjelldal PG, Benedet S, Hansen T, Björnsson BT, Nordgarden U. A peak in expression of sox8 gene during gonad transformation in the swamp eel (Monopterus albus). Mech Develop. 2004; 121: 997–1005. doi: 10.1016/S0195-6681(04)00001-7

62. Powell-Braxton L. IGF-I is required for normal embryonic growth in mice. Gene Dev. 1993; 7: 2609–2617. PMID: 8276243

63. Moriyama S, Ayson FG, Kawachi H. Growth regulation by insulin-like growth factor-I in fish. Biosci, Biotech Bioch. 2000; 64: 1553–1562.

64. Cheong HS, Yoon D-H, Kim LH, Park BL, Choi YH, Chung ER, et al. Growth hormone-releasing hormone (GHRH) polymorphisms associated with carcass traits of meat in Korean cattle. BMC Genet. 2006; 7: 35. PMID: 16749938

65. Tao W, Boulding E. Associations between single nucleotide polymorphisms in candidate genes and growth rate in Arctic char (Salvelinus alpinus L.). Heredity. 2003; 91: 60–69. PMID: 12815454

66. McPherron AC, Lawler AM, Lee S-J. Regulation of skeletal muscle mass in mice by a new TGF-beta superfamily member. Nature. 1993; 366: 83–89. PMID: 839826

67. Sawatari E, Seki R, Adachi T, Hashimoto H, Uji S, Wakamatsu Y, et al. Overexpression of the dominant-negative form of myostatin results in doubling of muscle-fiber number in transgenic medaka (Oryzias latipes). Comp Biochem Phys A. 2010; 155: 183–189.

68. Hickford J, Forrest R, Zhou H, Fang Q, Han J, Frampton CM, et al. Polymorphisms in the ovine myostatin gene (MSTN) and their association with growth and carcass traits in New Zealand Romney sheep. Anim Genet. 2010; 41: 64–72.

69. Wang XL, Meng XY, Song B, Qiu XM, Liu HY. SNPs in the myostatin gene of the mollusk Chlamys farrii: association with growth traits. Comp Biochem Phys B. 2010; 155: 327–330.

70. Tang YK, Li JL, Yu JH, Chen XF, Li HX. Genetic structure of MSTN and association between its polymorphisms and growth traits in genetically improved farmed tilapia (GIFT). J Fish Sci China. 2010; 17: 44–51.

71. Li C, Basarab J, Snelling W, Benkel B, Murdoch B, Hansen C, et al. Assessment of positional candidate genes myf5 and igf1 for growth on bovine chromosome 5 in commercial lines of Bos taurus. J Anim Sci. 2004; 82: 1–7.

72. Chen YH, Lee WC, Liu CF, Tsai HJ. Molecular structure, dynamic expression, and promoter activity of zebrafish (Danio rerio) myf-5 gene. Genesis. 2001; 29: 22–35. PMID: 11135459

73. Tan X, Zhang Y, Zhang P-J, Xu P, Xu Y. Molecular structure and expression patterns of flounder (Paralichthys olivaceus) Myf-5, a myogenic regulatory factor. Comp Biochem Phys B. 2006; 145: 204–213.

74. Hauerland NH, Spener F. Fatty acid-binding proteins—insights from genetic manipulations. Prog Lipid Res. 2004; 43: 329–349. PMID: 15234551
90. Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like

91. Wu XY, Xiang LX, Huang L, Jin Y, Shao JZ. Characterization, expression, evolution analysis of Toll-

92. Wei YC, Pan TS, Chang MX, Huang B, Xu Z, Luo TR, et al. Cloning and expression of Toll-like recep-

75. Zimmerman A, Veerkamp J. New insights into the structure and function of fatty acid-binding proteins. Cell Mol Life Sci. 2002; 59: 1096–1116. PMID: 12222958

76. Torstensen B, Nanton D, Olsvik P, Sundvold H, Stubhaug I. Gene expression of fatty acid-binding pro-
teins, fatty acid transport proteins (cd36 and FATP) and β-oxidation-related genes in Atlantic salmon (Samo salar L.) fed fish oil or vegetable oil. Aquacult Nutr. 2009; 15: 440–451.

77. Her GM, Chiang CC, Wu JL. Zebrafish intestinal fatty acid binding protein (I-FABP) gene promoter drives gut-specific expression in stable transgenic fish. Genesis. 2004; 38: 26–31. PMID: 14755801

78. Alvarez-Pellitero P. Fish immunity and parasite infections: from innate immunity to immunoprophylac-
tic prospects. Vete Immunol Immunop. 2008; 126: 171–198.

79. Power D. Developmental ontogeny of prolactin and its receptor in fish. Gen Comp Endocr. 2005; 142:

80. Chou MY, Yang CH, Lu FI, Lin HC, Hwang PP. Modulation of calcium balance in tilapia larvae (Oreochromis mossambicus) acclimated to low-calcium environments. J Comp Physiol B. 2002; 172: 109–114. PMID: 11916107

81. Shepherd BS, Sakamoto T, Nishioka RS, Richman NH, Mori I, Madsen SS, et al. Somatotropic actions of the homologous growth hormone and prolactins in the euryhaline teleost, the tilapia, Oreochromis mossambicus. P Natl Acad Sci USA. 1997; 94: 2068–2072.

82. Yang BY, Greene M, Chen TT. Early embryonic expression of the growth hormone family protein genes in the developing rainbow trout, Onchorhynchus mykiss. Mol Reprod Dev. 1999; 53: 127–134. PMID: 10331450

83. Sandra O, Le Rouziec P, Cauty C, Edery M, Prunet P. Expression of the prolactin receptor (tiPRL-R) gene in tilapia Oreochromis niloticus: tissue distribution and cellular localization in osmoregulatory organs. J Mol Endocrinol. 2000; 24: 215–224. PMID: 10750022

84. Yada T, Misumi I, Muto K, Azuma T, Schreck CB. Effects of prolactin and growth hormone on proliferation and survival of cultured trout leukocytes. Gen Comp Endocr. 2004; 136: 298–306. PMID: 15028535

85. Yada T, Uchida K, Kajimura S, Azuma T, Hirano T, Grau EG. Immunomodulatory effects of prolactin and growth hormone in the tilapia, Oreochromis mossambicus. J Endocrinol. 2002; 173: 483–492. PMID: 12065238

86. Watts M, Munday BL, Burke CM. Immune responses of teleost fish. Aust Vet J. 2001; 79: 570–574. PMID: 11599820

87. Pereiro P, Balseiro P, Romero A, Dios S, Forn-Cuni G, Fuste B, et al. High-throughput sequence anal-

ysis of turbot (Scophthalmus maximus) transcriptome using 454-pyrosequencing for the discovery of antiviral immune genes. PLoS One. 2012; 7: 224. doi: 10.1371/journal.pone.0035369 PMID: 22629298

88. Janeway CA. Innate immune recognition. Annu Rev Immunol. 2002; 20: 197–216. PMID: 11861602

89. Akina S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. Cell. 2006; 124: 783–801. PMID: 16497588

90. Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. Nat Immunol. 2010; 11: 373–384. doi: 10.1038/ni.1863 PMID: 20404851

91. Wu XY, Xiang LX, Huang L, Jin Y, Shao JZ. Characterization, expression, evolution analysis of Toll-
like receptor 1 gene in pufferfish (Tetraodon nigroviridis). Int J Immunogenet. 2008; 35: 215–225. doi:
10.1111/j.1744-313X.2008.00759.x PMID: 18312594

92. Wei YC, Pan TS, Chang MX, Huang B, Xu Z, Luo TR, et al. Cloning and expression of Toll-like recep-
tors 1 and 2 from a teleost fish, the orange-spotted grouper Epinephelus coioides. Vet Immunol Immunop. 2011; 141: 173–182.

93. Wu L, Sun JS, Geng XY, Pan BP, Wei JL, Wang XH. Molecular cloning and expression analysis of Toll-like receptor 1 cDNA in Japanese flounder, Paralichthys olivaceus. J Ag Sci Tech. 2012; 13: 2464–2470.

94. Matsuo A, Oshiumi H, Tsujita T, Mitani H, Kasai H, Yoshimizu M, et al. Teleost TLR22 recognizes RNA duplex to induce IFN and protect cells from birnaviruses. J Immunol. 2008; 181: 3474–3485. PMID: 18714020

95. Sundaram AYM, Consuegra S, Kiron V, Fernandes JMO. Positive selection pressure within teleost

TLR subfamilies and their response to temperature stress and microbial components in zebrafish. Mol Biol Rep. 2012; 39: 8965–8975. doi: 10.1007/s11033-012-1765-y PMID: 22722906

96. Yi ZA, Lin WW, Stunz LL, Bishop GA. Roles for TNF-receptor associated factor 3 (TRAF3) in lympho-
cyte functions. Cytokine Growth F R. 2014; 25: 147–156.
97. Lin HS, Lee E, Hestir K, Leo C, Huang M, Bosch E, et al. Discovery of a cytokine and its receptor by functional screening of the extracellular proteome. Science. 2008; 320: 807-. doi:10.1126/science.1154370 PMID:18467591

98. Bei JX, Suetake H, Araki K, Kikuchi K, Yoshiura Y, Lin HR, et al. Two interleukin (IL)-15 homologues in fish from two distinct origins. Mol Immunol. 2006; 43: 860–869. PMID:16055191

99. Fang W, Shao JZ, Xiang LX. Molecular cloning and characterization of IL-15R alpha gene in rainbow trout (Oncorhynchus mykiss). Fish Shellfish Immun. 2007; 23: 119–127.

100. Kohli G, Hu SQ, Clelland E, Di Muccio T, Rothenstein J, Peng C. Cloning of transforming growth factor-beta 1 (TGF-beta 1) and its type II receptor from zebrafish ovary and role of TGF-beta 1 in oocyte maturation. Endocrinology. 2003; 144: 1931–1941. PMID:12697700

101. Hardie LJ, Laing KJ, Daniels GD, Grabowski PS, Cunningham C, Secombes CJ. Isolation of the first piscine transforming growth factor beta gene: analysis reveals tissue specific expression and a potential regulatory sequence in rainbow trout (Oncorhynchus mykiss). Cytokine. 1998; 10: 555–563. PMID:9722928

102. Harms CA, Kennedy-Stoskopf S, Horne WA, Fuller FJ, Tompkins WAF. Cloning and sequencing hybrid striped bass (Morone saxatilis x M. chrysops) transforming growth factor-β (TGF-β), and development of a reverse transcription quantitative competitive polymerase chain reaction (RT-qPCR) assay to measure TGF-β mRNA of teleost fish. Fish Shellfish Immun. 2000; 10: 61–85.

103. Liu ZJ, Cordes JF. DNA marker technologies and their applications in aquaculture genetics. Aquaculture. 2004; 238: 1–37.

104. Wang T, Huang ZH, Ma AJ, Ma DY, Wang XA, Xia DD, et al. Development and polymorphic analysis of SNP markers in Scophthalmus maximus based on transcriptome database. Oceanologia et Limnologia Sinica. 2014; 45: 1300–1307.