Fish and Shellfish Immunology 94 (2019) 634–642

Contents lists available at ScienceDirect

Fish and Shellfish Immunology

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

Full length article

Immunity-associated long non-coding RNA and expression in response to bacterial infection in large yellow croaker (*Larimichthys crocea*)

Xiaoxu Liu[a], Weiye Li[b,c], Lihua Jiang[a,**], Zhenming Lü[a,*], Minhai Liu[b], Li Gong[a], Bingjian Liu[a], Liqin Liu[a], Xiaolong Yin[b]

[a] National Engineering Research Center of Marine Facilities Aquaculture, College of Marine Science, Zhejiang Ocean University, No. 1 Haida South Road, Dinghai District, Zhoushan, Zhejiang Province, 316022, China
[b] Administration of Ocean and Fisheries of Zhoushan, No 21,Chenghe xi Road, Dinghai District, Zhoushan, Zhejiang Province, 316021, China
[c] School of Marine Sciences Ningbo University, No 818 Fenghua Road, Jiangbai District, Ningbo City, Zhejiang Province, 315211, China

A B S T R A C T

Long non-coding RNA refers to an RNA transcript of a non-coding protein with a sequence length greater than 200 bp. More and more reports indicated that lncRNA was involved in the regulation of gene expression as a signalling molecule, an inducing molecule, a leader molecule and a scaffold molecule. Previous studies have sequenced the draft genome and several transcriptome data sets for protein-coding genes of the large yellow croaker (*Larimichthys crocea*), but little is known about the expression and function of lncRNAs in this species. In order to obtain a catalogue of lncRNAs for this croaker, *Vibrio parahaemolyticus* infection challenge experiment was conducted and long non-coding RNA sequences were obtained. Using high-throughput sequencing of lncRNA, a total of 73,233 high-confidence transcripts were reconstructed in 32,726 loci, recovering most of the expressed reference transcripts, and 6473 novel expressed loci were identified. The tissue expression profile revealed that most lacunas were specifically enriched in distinct tissues. A set of 163 lncRNAs were identified as being specifically expressed in the spleen and may be involved in the immune response. It is the first time to identify specific lncRNAs in the *L. crocea* systematically in this croaker, aiming to benefit the future genomic study of this species.

1. Introduction

Large-scale transcriptomic studies have led to surprising discoveries, including the fact that less than 10% of the mammalian genome is dedicated to protein coding and that the genome contains a vast amount of non-protein coding transcripts, which has resulted in debate about the role of non-coding RNAs (ncRNAs) in cell biology [1,2]. In contrast to protein-coding genes, it is possible that the non-coding portion of the genome is related to organism complexity and crucial regulatory processes [3,4]. Traditionally, the regulatory functions of RNA were thought to be limited to ribosomal, messenger and transfer RNA roles, however, the FANTOM and ENCODE projects have annotated thousands of non-coding RNAs including tRNA, rRNA, microRNA and other non-coding RNA genes [5,6], and reported on the function of ncRNAs in various life activities.

In this study, we mainly explore long non-coding RNA (lncRNA). LncRNAs are an important class of ubiquitous nucleotides involved in a variety of biological functions [7]. LncRNAs regulate the expression of other genes by the identity of the signal molecule [8]. For example, the lncRNA-Xist gene and Xist gene function are known to cause the X chromosome to lose activity. An X chromosome in a mammalian female self-regulates silencing by the X inactivation centre (Xic) in order to maintain a dose-compensating effect. Binding of polycomb repressive complex 2 (PRC2) to the repeat-A element at the 5′ end of Xist RNA activates the gene. With Xist gene expression, its transcript is transcribed from the inactive X chromosome and encapsulated on the X chromosome, inhibiting gene expression at the overall level of the chromosome [9,10].

LncRNAs will indirectly regulate gene transcription by inducing target protein or miRNA to science the functional sequence, that is, lncRNAs bind to target proteins or miRNAs to dilute their levels within the cell, affecting their function [11–13]. LncRNAs serve as central platforms for the assembly of multiple related molecules [14]. Although the complex scaffold molecule is traditionally considered to be a protein [15], an increasing number of lncRNAs have been found to be involved in gene-regulation in the form of scaffolds. LncRNAs have multiple domains that can bind to different protein complexes, and some domains can bind to multiple elements, leading to time- and space-dependent transcriptional activation or transcriptional repression. For example, lncRNA-TERC (telomerase RNA) acts as a scaffold to activate telomerase catalytic activity. Telomerase catalytic activity requires the use of two universal telomerase subunits: a complete RNA subunit

* Corresponding author.
** Corresponding author.
E-mail addresses: jianglihua@zjou.edu.cn (L. Jiang), nblzmnb@zjou.edu.cn (Z. Lü).

https://doi.org/10.1016/j.fsi.2019.09.015
Received 9 May 2019; Received in revised form 2 September 2019; Accepted 5 September 2019
Available online 15 September 2019

Table 1A
The data filter statistics.

Library	J_P_1	J_P_3	P_P_1	J_P_2	P_P_2	P_P_3
Raw Reads	65,806,016	66,108,182	63,805,758	70,788,006	74,150,380	70,551,778
Clean Reads	58,757,374	59,065,410	56,529,328	62,666,684	65,945,296	63,505,576
Clean Reads Rate	89.29%	89.35%	88.60%	88.53%	88.93%	90.01%
Total Q30	95.62%	95.63%	95.46%	95.43%	95.61%	95.84%

Table 1B
The mapping rate.

Library	J_P_1	J_P_3	P_P_1	J_P_2	P_P_2	P_P_3
Total Reads	58,757,374	59,065,410	56,529,328	62,666,684	65,945,296	63,505,576
Mapped Reads	44,345,720	44,737,208	42,310,604	47,532,529	49,754,920	46,869,514
Mapping Rate	75%	76%	75%	76%	75%	74%
UnMapped Reads	14,411,654	14,328,202	14,218,724	15,134,155	16,190,376	16,636,062

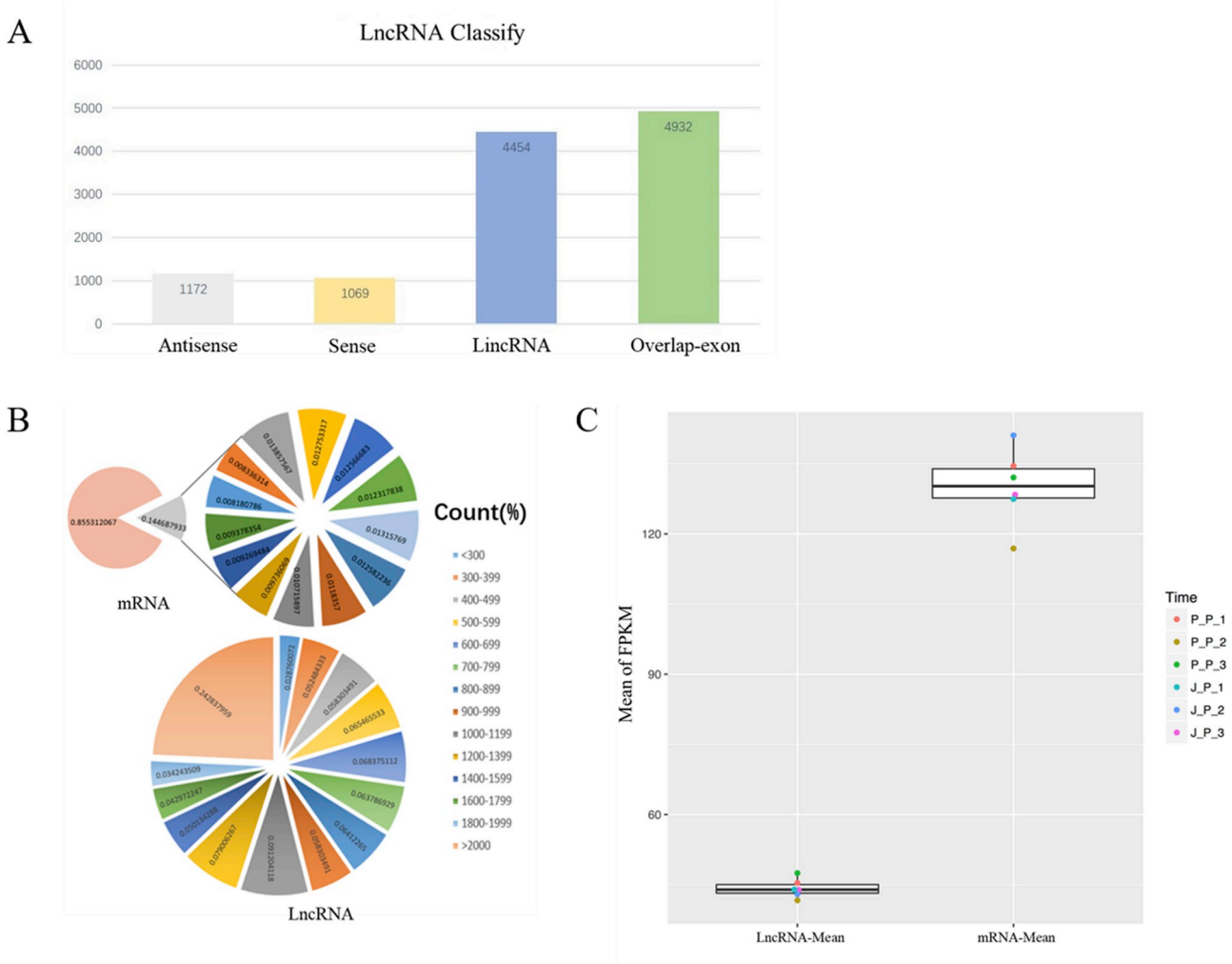

Fig. 1. A. The classification of lncRNA. B. Above is the pie chart of the length of the coding gene, and below is the length of the lncRNA sequence. C. Average expression of lncRNA and coding genes.

Table 2
Number of Long non-coding RNAs shared with other species.

Species	Human	Cow	Zebra fish	Mouse	Chicken
L.crocea lncRNA	34	7	35	44	3
Alignment of lncRNA	111	26	65	83	5

(TERC) that provides repeat template synthesis and support, and a catalytic protein subunit (TERT), as well as several species-specific accessory proteins. TERC has structural and catalytic activities that contribute to TERT-binding and provide stability [16].

An increasing number of reports show that lncRNAs play an important role in the process of life [17–20]. In addition, the continuous development of sequencing technology means more and more fish

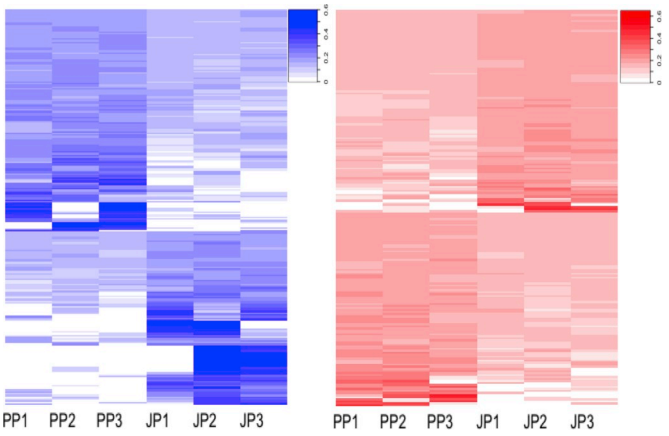

Fig. 2. The hierarchical clustering of gene and lncRNA. Blue indicates the clustering expression pattern of the coding gene, and red indicates the lncRNA. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

genomes have been published, and there are more and more accumulated transcriptome data. However, Studies on lncRNA in fishes immunity are still rare, research of lncRNA has been reported in zebrafish [21,22], salmon [23] and rainbow trout [24,25]. Pauli A et al. (2012) systematically identified long non-coding RNAs expressed during zebrafish embryogenesis. They performed time series of RNA-seq experiments in eight stages during early development in zebrafish, found that zebrafish lncRNA has many of mammalian counterparts: relatively short length, low exon number, low expression and a conservative level comparable to introns. And the temporal expression profile of lncRNAs reveals two novel properties: lncRNAs are expressed in a narrower window of time than protein-encoding genes and lncRNAs are particularly enriched in early embryos [22]. Heena et al. (2015) integrated and provided all relevant dataset resources for zebrafish lncRNA to provide analysis of spatiotemporal expression patterns of lncRNA in the context of various regulatory markers including histone modifications and transcription factors [21]. Sebastian Boltaña et al. (2016) reported the transcriptomic regulation of lncRNA in *Atlantic salmon* during ISA virus infection [23]. At the same time, the long non-coding RNA regulatory mechanism of *Rainbow Trout* was integrated and reported [24,25].

Larimichthys crocea is one of the most economically important marine fishes farmed in China. However, the population is suffering from various infectious diseases due to the increasing density and scale of the mariculture. *Vibrio parahaemolyticus* is a common pathogen causing Vibrio, a disease with high morbidity and mortality that can result in serious economic loss [26]. The whole-genome sequence of the large yellow croaker demonstrated the existence of a well-developed innate immune system and laid the foundation for genome-wide studies in this species [27]. However, few studies have focused on lncRNAs and their biological function in the croaker. In this study, we report on lncRNA related to the immune mechanisms of this croaker in order to better understand the complex immune system and provide basic information about the innate immunity of the species. Furthermore, deciphering the expression pattern of lncRNAs in the large yellow croaker would enable a better understanding of how genes are regulated in teleost.

2. Materials and methods

2.1. Sample preparation

Healthy large yellow croakers (weight 150 ± 15g) were obtained from Zhejiang Dahaiyang Technology Co., Ltd. (Zhoushan, Zhejiang Province, China). Fish were maintained at 25 °C in an aerated seawater tank and fed a commercial diet for two weeks prior to the beginning of the experiment. Water in the tank was changed daily. After acclimation, two groups of 10 individuals were randomly chosen for challenge experiments. *L. crocea* were then intraperitoneally infected with *V. parahaemolyticus* (1×10^8 CFU/mL, resuspended in PBS, pH 7.4) or PBS (as control; 300 mL/200 g). Animals from all groups were anaesthetised by immersion in MS222 before tissue sampling, as required. The spleen tissues were harvested from three fish per group at 72 h after injection. All procedures were in accordance with the guidelines of the Regulations for the Administration of Laboratory Animals (Decree No. 2 of the State Science and Technology Commission of the People's Republic of China, November 14, 1988) and were approved by the Animal Ethics Committee of Zhejiang Ocean University (Zhoushan, China). Total RNA was extracted from the spleen samples using TRIzol Reagent (Invitrogen, Carlsbad, CA, USA). First strand cDNA was synthesised from 3 mg of total RNA using a Tr\anscriptor First Strand cDNA Synthesis Kit (Roche, Indianapolis, IN, USA).

2.2. Transcriptome assembly

We used LncRNA high-throughput sequencing and filtered the original downstream data sequence to obtain high-quality clean readings. The RNA sequencing reads from the spleen tissues were aligned independently with the large yellow croaker genome (GenBank accession no. JPYK00000000.1) using TopHat (v2.0.12) software (http://tophat.cbcb.umd.edu/) [28–30]. Prior to the alignment, we used Bowtie 2 to pre-process the clean reads, which can improve the accuracy and speed of the alignment [29]. To gain enough sequencing data for lncRNAs assembly, all alignment results were combined using Samtools v1.3.1 [31]. The mapped reads were then assembled into transcripts by Cufflinks v2.2.1 (http://cufflinks.cbcb.umd.edu/) [32,33], guided by a reference annotation of the croaker.

2.3. Identification of lncRNAs

All transcripts were evaluated for protein coding potential by the Coding Potential Calculator (CPC, http://cpc.cbi.pku.edu.cn/) [34], and Coding-Non-Coding Index (CNCI, www-bioinfo-org/CNCI) [35], which distinguishes coding and non-coding transcripts with high accuracy. CPC applies sequence-based features to predict the protein coding potential of transcripts and has been widely used to discover lncRNAs. CNCI effectively distinguishes protein-coding and non-coding sequences independent of known annotations by profiling adjoining nucleotide triplets. A transcript with a negative score is considered to be a non-coding transcript [1,36,37]. The non-coding transcripts were filtered to remove sequences less than 200 bp in length. All non-coding transcripts were then aligned to the Rfam database in order to eliminate all non-coding transcripts that had been previously annotated as rRNA, miRNA, or other small non-coding RNA transcripts [38]. All remaining transcripts were identified as lncRNAs in large yellow croaker and used for further functional analysis.

2.4. Classification of lncRNAs

The resulting set of lncRNAs was subdivided into four groups: (1) lncRNAs without any overlap with protein-coding gene loci were classified as large intergenic non-coding RNA (lincRNA); (2) the remaining lncRNAs that overlapped with protein-coding gene regions were classified as overlapping-lncRNA; (3) Sense lncRNAs were those transcribed from the sense strand of protein-coding genes, containing exons from protein-coding genes; (4) Antisense lncRNAs were those transcribed from the antisense strand of protein-coding genes.

2.5. Annotation of protein-coding transcripts

Gene ontology (GO: http://www.geneontology.org/) is used to

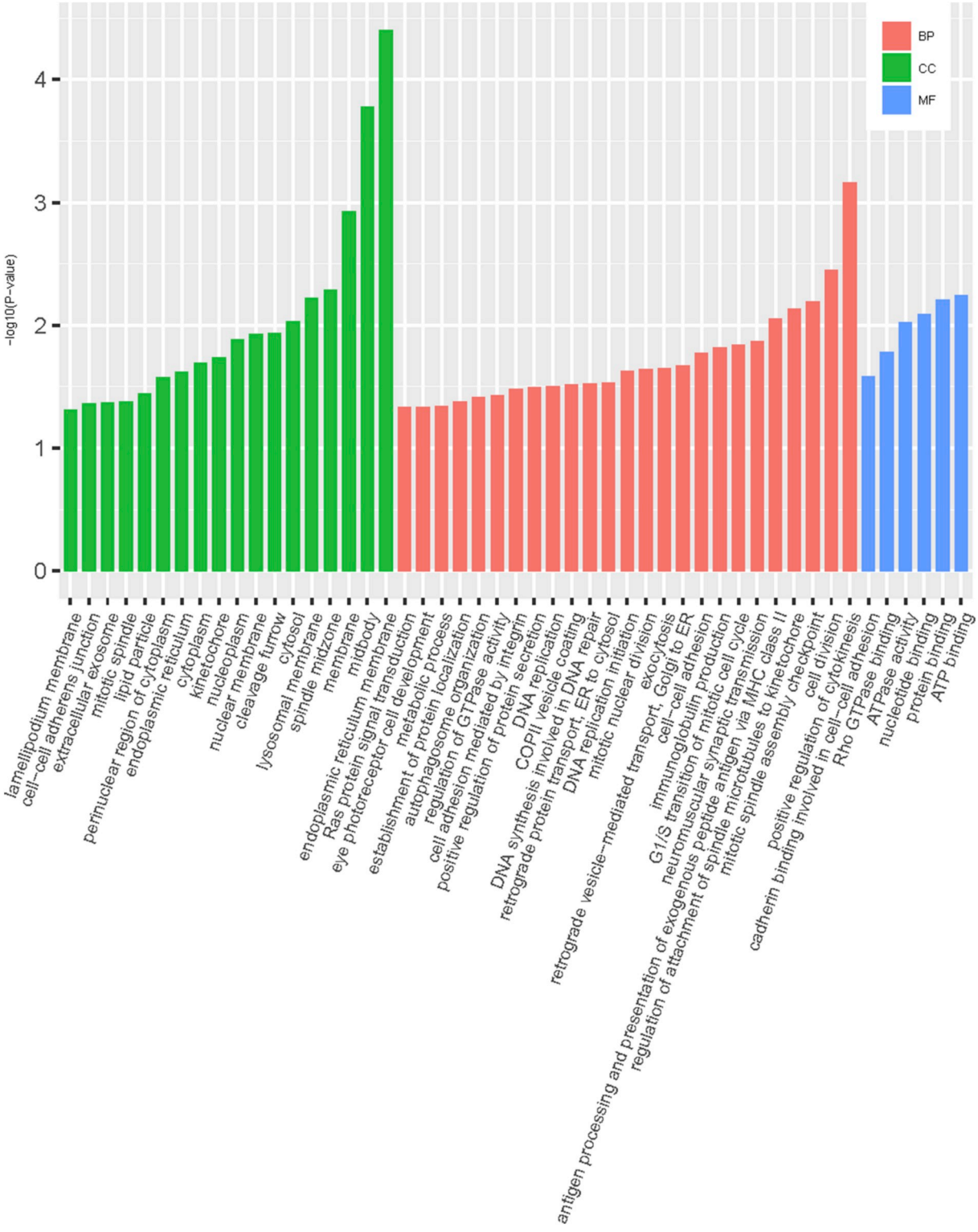

Fig. 3. The function annotation of mRNA.

describe genes and gene product attributes, including cellular components, molecular functions and biological processes [39,40]. GO not only organizes genes into hierarchical categories, but also uncovers the gene regulatory network based on biological processes and molecular functions. The Database for Annotation, Visualization, and Integrated Discovery v6.8 (DAVID) method was used to select the main pathway, with the significance threshold set at $P < 0.05$ [41].

2.6. Co-expression network analysis of lncRNA and function annotation

A correlation analysis was performed using the FPKM values of all significantly differentially expressed genes and lncRNAs at each time point. A co-expression network was constructed between the differentially expressed genes and lncRNAs using R version 3.3.0. The ENViz module of Cytoscape 3.2.0 [42] was used to annotate and visualise the functional annotation network. The FPKM data for the lncRNAs and

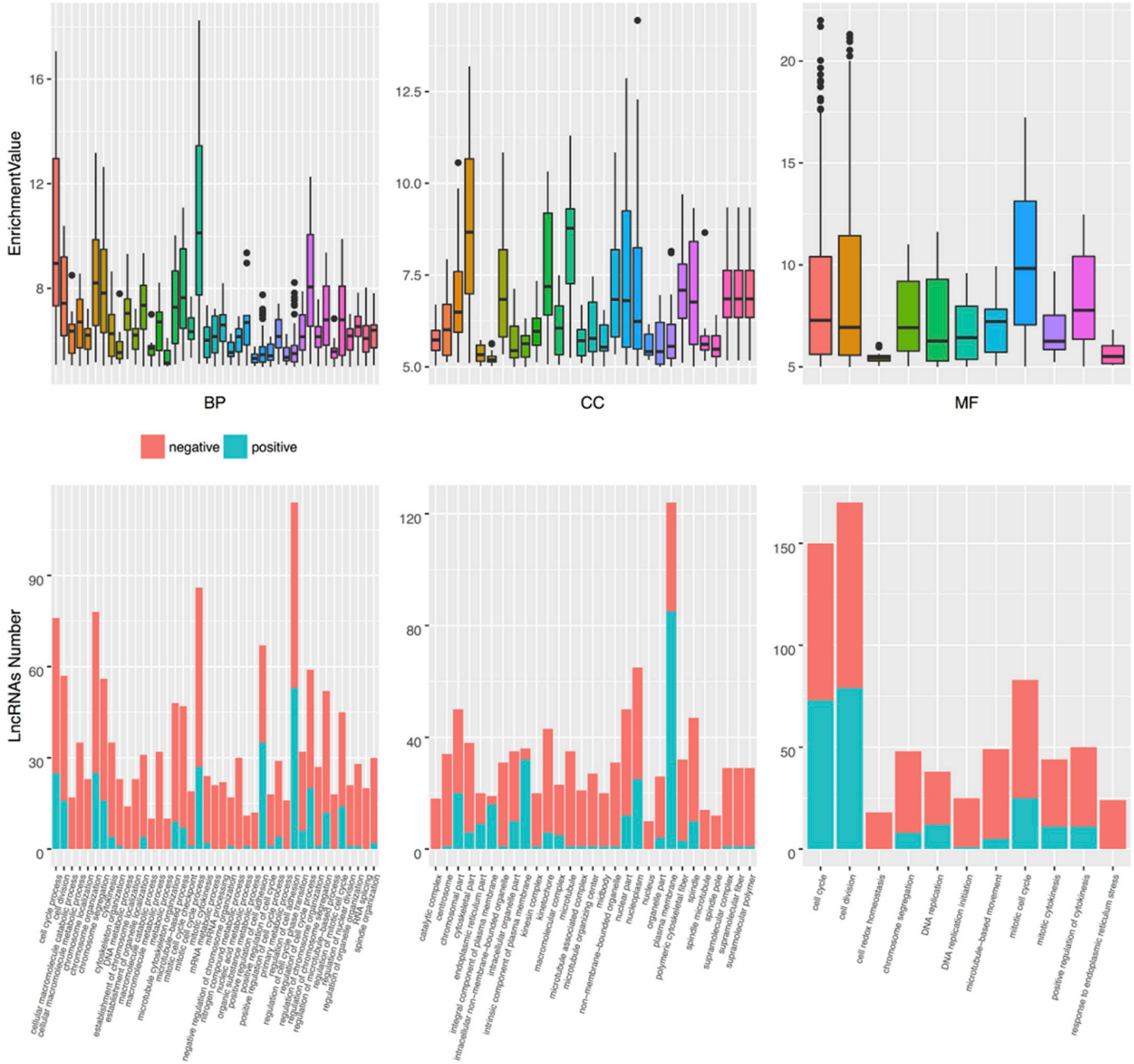

Fig. 4. The function annotation of lncRNA.

genes were analysed using Spearman's correlation coefficient. The GO term nodes were displayed with a cumulative enrichment cut-off value of five and were coloured from yellow to red based on their cumulative enrichment values.

2.7. Quantitative real-time polymerase chain reaction (PCR) analysis

Real-time PCR was performed using SuperReal PreMix Plus (Tiangen, Shanghai, China) with SYBR® Green fluorescent dye, in a total volume of 20 μl on an ABI 3700 Thermal Cycler (Life Technologies, Carlsbad, CA, USA). The primers mRNA-CCL20, CD28, eIF2α, PKD2, lncRNA-XLOC_001904, XLOC_006086, XLOC_008164, XLOC_028673, and β-actin were designed based on gene sequences identified in the current study (Table S1). The cycling conditions were as follows: 2 min at 50 °C, 10 min at 95 °C, 40 cycles of 15 s at 95 °C and 1 min at 60 °C, followed by 15 s at 95 °C, 15 s at 60 °C, and 15 s at 95 °C as the standard dissociation stage. Every sample was run in triplicate and concordance was checked between quintuplicate mean values. The gene expression levels were evaluated relative to the expression level of β-actin using the 2-ΔΔCT method [43]. Statistically significant differences between the control and treatment groups were investigated by one-way ANOVA analysis.

3. Results

3.1. RNA-sequence data generation and processing

In the study, L. crocea was challenged with *Vibrio parahaemolyticus* and *V. parahaemolyticus*. The spleen tissue of L. crocea was extracted, then the spleen RNA was purified and converted into cDNA libraries that were subjected to paired-end 150 bp sequencing on Illumina's HiSeq platform. A high-throughput sequencing technique was used to

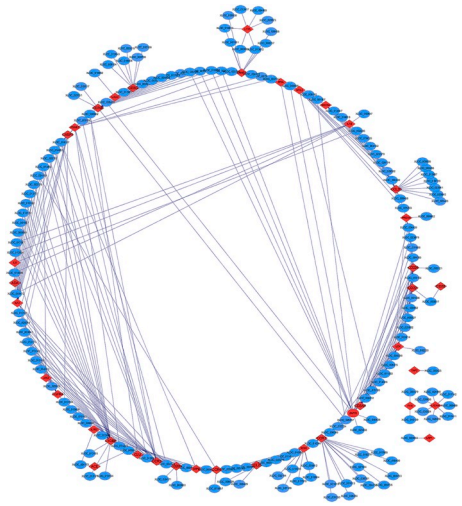

Fig. 5. Co-expression network. Red indicates the immune gene, blue indicates the immune-related lncRNAs, and the relationship between them is represented by a line connection. More details are shown in Supplementary material Table S2. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

carry out the long chain non-coding sequencing of the library, and about 10 G of the raw data were obtained in each of the six libraries. The Q30 (Phred Quality Score) of the base mass level of the sample was over 95%. Using Trimmomatic v0.33 [44] to filter out the linker sequence in Reads, the N bases > 5% of Reads and low quality Reads (< Q30). The filtering ratio ranged from 88.53% to 90.01% and averaged 89.11% (Table 1A). The mapping ratio ranged from 74% to 76% and averaged 75%, which could be aligned to the latest assembly of the large yellow croaker genome sequence (Table 1B). Transcripts were assembled using a step-wise protocol. Briefly, TopHat [28,30] was used to align all reads to the reference genome, including those mapped to span splice junctions. Transcripts were then reconstructed and expression calculated using Cufflinks [32,33], resulting in the assembly of a total number of 73,233 expressed transcripts across all samples. Henceforth, this set is referred to as the standard transcriptome for large yellow croaker and is used in all subsequent analyses.

3.2. Identification and initial classification of lncRNAs

To retrieve putative non-coding transcripts in the large yellow croaker, all transcripts were analysed using Coding Potential Calculator (CPC) and Coding-Non-Coding Index (CNCI). A total of 8936 (12.2%) large yellow croaker transcripts were classified as non-coding transcripts, while 64,297 (87.8%) were identified as traditional protein-coding transcripts. Secondly, to eliminate small RNA transcripts, a three-step filter was applied to the non-coding transcripts: (1) non-coding transcripts that matched entries in the Rfam database were deleted; (2) transcripts shorter than 200 bp were eliminated; (3) non-coding transcripts were mapped to the rRNA, tRNA and miRNA databases. After filtering the transcripts, 8463 lncRNAs in the large yellow croaker were identified. Of these, 4454 lncRNAs were in intergenic regions and were defined as lincRNAs. The remaining 4932 transcripts overlapped protein-coding genes; 1172 were antisense lncRNAs and just 1069 sense lncRNAs (Fig. 1A).

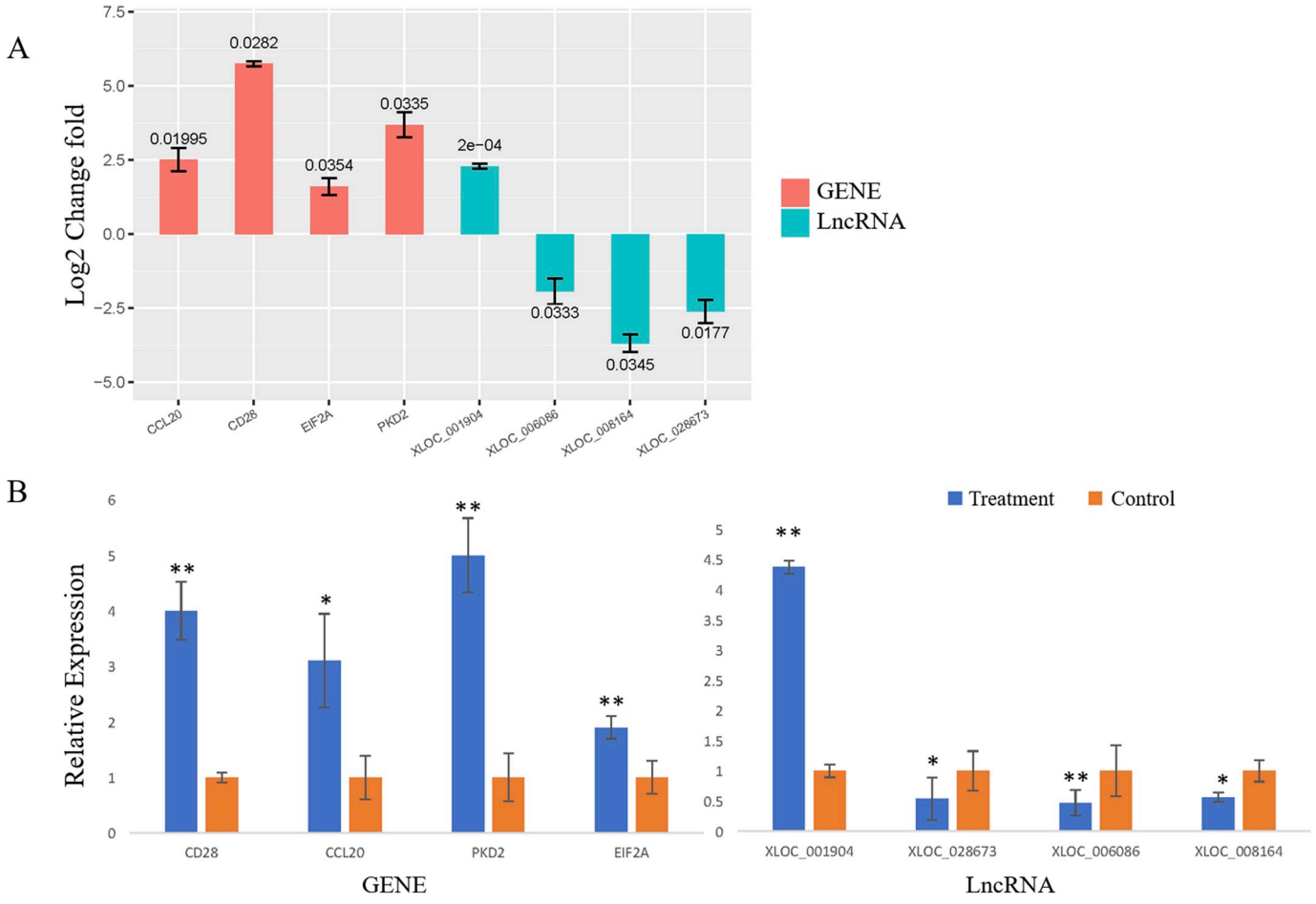

Fig. 6. A. The expression of mRNA-lncRNA in Transcriptome Analysis. B. The figure of verification data result.

3.3. Characterization of lncRNAs

We analysed the structure and expression levels of RNAs of the large yellow croaker challenged with *V. parahaemolyticus* infection and we found that 0.85 (85%) of protein-coding transcripts were over 2000 bp in length, whereas less than 0.24 (24%) of lncRNAs were over 2000 bp in length (Fig. 1B). Moreover, the lncRNAs had fewer exons per transcript than an average protein-coding gene. Notably, the average lncRNA expression level in the spleen tissue was 10-fold lower than the protein-coding genes (Fig. 1C). To investigate the sequence conservation features of the lncRNAs identified in the spleen of this croaker, candidate lncRNAs were used as queries to search for corresponding homologs in other species. Of the 8463 lncRNAs in the spleen of the large yellow croaker, only 34 lncRNAs had hits with 111 human lncRNAs in the NONCODE V4 database, with E-value < 0.00001. In addition, 7 lncRNAs had hits with 26 lncRNAs from the cow, 35 lncRNAs had hits with 65 lncRNAs from the zebrafish (*Danio rerio*), 44 lncRNAs had hits with 83 lncRNAs from the mouse, and 3 lncRNAs had hits with 5 lncRNAs from the chicken (Table 2).

3.4. Expression profiles and function annotation of coding genes and lncRNAs

To ascertain the potential biological functions of coding genes and lncRNAs, we examined coding genes and lncRNA expression profiles in the spleen of the large yellow croaker using RNA-sequence (RNA-seq) analysis. We used authoritative data sources that contained over 64,297 coding genes and 8463 lncRNAs to identify the most significant candidates. The coding genes and lncRNAs with at least a 1.5-fold change in expression were selected. Using this criterion, coding genes and lncRNAs expression profiles at the different time points were evaluated by hierarchical clustering (Fig. 2), which indicated there is a huge difference in lncRNAs with regulated expression between *V. parahaemolyticus* infected and non-infected fish. The functional annotation was divided into three parts, corresponding to the three Gene Ontology (GO) categories: biological process, molecular function and cellular component. Fig. 3 is the functional annotation map of the encoded gene, and Fig. 4 is the functional annotation map of the long-chain non-coding RNA.

3.5. Co-expression network of lncRNAs

A correlation coefficient of > 0.95 or < -0.95, and $P < 0.05$, was selected to construct the co-expression network, using the FPKM values of all the differentially expressed genes and lncRNAs for each time point. A total of 163 pairs of lncRNAs and genes (Table S2) were harvested. The mRNA-lncRNA co-expression network is shown in Fig. 5.

3.6. Validation of predicted functional lncRNAs related to antiviral response

It is not realistic to perform functional experiments for every lncRNA due to limited resources and time. Therefore, four predicted pairs of mRNA-lncRNAs were studied. We selected lncRNA-XLOC_001904 associated with the CCL20 gene, lncRNA-XLOC_006086 associated with the CD28 gene, lncRNA-XLOC_008164 associated with the eIF2α gene, and lncRNA-XLOC_028673 associated with the PKD2 gene. To verify whether these lncRNAs were functional in the large yellow croaker immune response, the changes in their level of expression following *V. anguillarum* treatment were measured using RT-PCR. In this study, RNA-seq analysis found that the expression level of lncRNA associated with the immune gene CCL20 was positively correlated with CCL20, and the expression level of lncRNA associated with the immune genes CD28, eIF2α, PKD2 was inversely correlated with the immune gene (Fig. 6A). The results of RT-PCR experimental validation are consistent with RNA-seq analysis results (Fig. 6B). This shows that the long chain non-coding RNAs screened here are immune-related.

4. Discussions

In the previous studies, most of them were focused on protein-coding gene function [45–48], such as the study of innate immunity genes of teleost fish. Gao et al., 2016, reported the post-innate immune responses in various tissues of the Asian swamp eel after *Aeromonas hydrophila*, *Acanthocephalan pallisentis*, and poly I:C challenges [49]. Under stress by immunostimulants lipopolysaccharide (LPS), poly I:C, Edwards disease and streptococcus, the rainbow trout congenital gene CXCR4 has different expression patterns in the head, kidney and spleen [50]. The mechanism of innate immune gene expression in the spleen and eyes of grouper was studied after LPS and necrosis virus stimulation [51]. Using different stimulation experiments, genomic and transcriptome analysis/identification, phylogenetic trees, multiple sequence alignment, homology alignment, and other analytical methods, the mechanism of action of the genes involved in the immune responses of teleost fish has been reported extensively.

Since lncRNAs can be used as an important regulator of different biological processes, this study screened and identified lncRNAs related to the large yellow croaker, providing a basis for systematic research on fish immune mechanisms.

Non-coding RNAs are functional RNA molecules that are not normally translated into a protein, and play different biological processes such as development, differentiation and epigenetic regulation [52]. More and more ncRNA studies on teleost fish have been reported. For example, reports indicate that teleost fish miRNAs and lncRNAs are involved in a variety of biological processes, including development, differentiation, response to environmental stimuli and so on [25,53–55]. Importantly, there are also a few reports on the function of ncRNAs in fish immune mechanisms [25,56,57]. Sebastian Boltaña et al. (2016) reported the transcriptomic regulation of lncRNA in *Atlantic salmon* during ISA virus infection [23], found 5636 putative lncRNAs, most of which were regulated by ISAV infection and found that most lncRNA is associated with innate immune genes. Diego Valenzuelamiranda et al. (2016) identified 918 lncRNAs from during *Piscirickettsia salmonis* infection in Atlantic salmon [58], then studied in different pathogens of *Atlantic salmon* infection with these lncRNAs [59]. The results indicate that lncRNAs are involved in the immune regulation of teleost fish and are closely related to immune-related genes involved in innate immunity, lncRNAs are regulated by pathogen specificity during pathogen infection. And it has been shown that the expression of lncRNA shows high specificity in different tissues and cell types [6,36,60]. LncRNAs play an important role in complex biological processes and are involved in the regulation of several innate and adaptive immune pathways [60]. In fish (such as *Atlantic salmon, Danio rerio*), lncRNA regulation was observed during pathogen infection [21,23,61,62]. A group of lncRNAs were reported to be upregulated in Atlantic salmon infected with the Infectious Salmon Anaemia (ISA) virus [23]. Pathogen-specific lncRNA regulation was observed in Atlantic salmon infected with bacteria, viruses and ectoparasites, indicating a key role for lncRNA in salmon immunity [62]. Highly expressed immune-related genes in large yellow croaker [63]. Furthermore, we compared the spleen tissue of the infected and control fish. We used the deribosomal chain-specific database construction method and lncRNA high-throughput sequencing, which are more targeted and accurate than traditional transcriptome sequencing methods.

High regulation of lncRNAs neighboured to coding genes has been reported to be associated with biological processes such as stress, the immune response and the inflammatory response [64]. This study, by functional annotation and construction of co-expression networks, reports that 163 immune-related lncRNA were identified and the relationships between these lncRNA and known immune genes was obtained. We also randomly selected eight groups of immune-related lncRNAs to verify that we screened immune-related lncRNAs involved

in the regulation of the immune system of the large yellow croaker. The chemokine CCL20 gene is a CC subfamily of chemokines that is involved in the directional movement of T cells and dendritic cells after interaction with the CC6R receptor. The CCL20 gene plays an important role in immunity and is involved in immune system pathology such as autoimmune diseases and tumour immunity. The CC12 gene has a strong antibacterial effect [65]. The CD28 gene plays an important role in the activation of T cells: the CD28 family binds to the B7 family to produce co-stimulatory signals that activate T cells [66]. The eIF2α gene plays a role in fish immunity by inhibiting virus replication; when cells are infected by the virus, the eIF2α kinase of Protein kinase R (PKR) is activated, activating the eIF2α gene to block the initial synthesis of the protein, thereby suppressing viral replication [67]. The immune gene PKD2 is a renal disease gene. Mutations and deletions of the PKD2 gene cause abnormal secretion, abnormal apoptosis, and abnormal function of renal epithelial cells [68].

Actually, further study of lncRNAs in this croaker is required. The lncRNA transcripts with tissue-specific expression manner may have specific functional roles in different biological pathways. The lncRNAs identified here may provide a better understanding of immune response regulatory network including lncRNA, mRNA and miRNA. The strategy developed in our study for novel lncRNAs identification can be also applied to other newly assembled genomes for understanding the function of lncRNA transcripts.

Conflicts of interest

The authors declare there is no conflict of interest regarding the publication of this paper.

Acknowledgements

This study was supported by the projects from Science and Technology Major Project of the Ministry of Science and Technology of Zhejiang Province (No. 2016C02055-7) and Science and Technology Program of Zhoushan (No 2018C31091).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fsi.2019.09.015.

References

[1] I. Ulitsky, A. Shkumatava, C.H. Jan, H. Sive, D.P. Bartel, Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution, Cell 147 (7) (2011) 1537–1550.
[2] L. Josset, N. Tchitchek, L.E. Gralinski, M.T. Ferris, A.J. Eisfeld, R. Green, M.J. Thomas, J. Tisoncikgo, G.P. Schroth, Y. Kawaoka, Annotation of long non-coding RNAs expressed in Collaborative Cross founder mice in response to respiratory virus infection reveals a new class of interferon-stimulated transcripts, RNA Biol. 11 (7) (2014) 875–890.
[3] P.P. Amaral, M.E. Dinger, T.R. Mercer, J.S. Mattick, The eukaryotic genome as an RNA machine, Science 319 (5871) (2008) 1787–1789.
[4] J.S. Mattick, RNA regulation: a new genetics? Nat. Rev. Genet. 5 (4) (2004) 316–323.
[5] E. Arner, C.O. Daub, K. Vittingseerup, R. Andersson, B. Lilje, F. Drablos, A. Lennartsson, M. Ronnerblad, O. Hrydziuszko, M. Vitezic, Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells, Science 347 (6225) (2015) 1010–1014.
[6] T. Derrien, R. Johnson, G. Bussotti, A. Tanzer, S. Djebali, H. Tilgner, G. Guernec, D. Martin, A. Merkel, D.G. Knowles, The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression, Genome Res. 22 (9) (2012) 1775–1789.
[7] K.C. Wang, H.Y. Chang, Molecular mechanisms of long noncoding RNAs, Mol. Cell 43 (6) (2011) 904–914.
[8] M. Guttman, I. Amit, M. Garber, C. French, M.F. Lin, D.M. Feldser, M. Huarte, O. Zuk, B.W. Carey, J.P. Cassady, Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals, Nature 458 (7235) (2009) 223–227.
[9] A. Wutz, T.P. Rasmussen, R. Jaenisch, Chromosomal silencing and localization are mediated by different domains of Xist RNA, Nat. Genet. 30 (2) (2002) 167–174.
[10] D.B. Pontier, J. Gribnau, Xist regulation and function eXplored, Hum. Genet. 130 (2) (2011) 223–236.
[11] I. Martianov, A.S. Ramadass, A.S. Barros, N. Chow, A. Akoulitchev, Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript, Nature 445 (7128) (2007) 666–670.
[12] M.G. Guenther, S.S. Levine, L.A. Boyer, R. Jaenisch, R.A. Young, A chromatin landmark and transcription initiation at most promoters in human cells, Cell 130 (1) (2007) 77–88.
[13] T. Kino, D.E. Hurt, T. Ichijo, N. Nader, G.P. Chrousos, Noncoding RNA Gas5 is a growth arrest– and starvation-associated repressor of the glucocorticoid receptor, Sci. Signal. 3 (107) (2010).
[14] R.C. Spitale, M. Tsai, H.Y. Chang, RNA templating the epigenome: long noncoding RNAs as molecular scaffolds, Epigenetics 6 (5) (2011) 539–543.
[15] R. Bonasio, S. Tu, D. Reinberg, Molecular signals of epigenetic states, Science 330 (6004) (2010) 612–616.
[16] K. Collins, Physiological assembly and activity of human telomerase complexes, Mech. Ageing Dev. 129 (1) (2008) 91–98.
[17] J. Cao, The functional role of long non-coding RNAs and epigenetics, Biol. Proced. Online 16 (1) (2014) 42.
[18] J.T. Kung, D. Colognori, J.T. Lee, Long noncoding RNAs: past, present, and future, Genetics 193 (3) (2013) 651–669.
[19] M.K. Iyer, Y.S. Niknafs, R. Malik, U. Singhal, A. Sahu, Y. Hosono, T.R. Barrette, J.R. Prensner, J.R. Evans, S. Zhao, The landscape of long noncoding RNAs in the human transcriptome, Nat. Genet. 47 (3) (2015) 199–208.
[20] P.J. Batista, H.Y. Chang, Long noncoding RNAs: cellular address codes in development and disease, Cell 152 (6) (2013) 1298–1307.
[21] H. Dhiman, S. Kapoor, A. Sivadas, S. Sivasubbu, V. Scaria, zflncRNApedia: a comprehensive online resource for zebrafish long non-coding RNAs, PLoS One 10 (6) (2015).
[22] A. Pauli, E. Valen, M.F. Lin, M. Garber, N.L. Vastenhouw, J.Z. Levin, L. Fan, A. Sandelin, J.L. Rinn, A. Regev, Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis, Genome Res. 22 (3) (2012) 577–591.
[23] S. Boltana, D. Valenzuelamiranda, A. Aguilar, S. Mackenzie, C. Gallardoescarate, Long noncoding RNAs (lncRNAs) dynamics evidence immunomodulation during ISAV-Infected Atlantic salmon (*Salmo salar*), Sci. Rep. 6 (4) (2016) 22698-22698.
[24] R. Al-Tobasei, B. Paneru, M. Salem, Genome-wide Discovery of long non-coding RNAs in rainbow trout, PLoS One 11 (2) (2016) e0148940.
[25] J. Wang, L. Fu, P.P. Koganti, L. Wang, J.M. Hand, H. Ma, J. Yao, Identification and functional prediction of large intergenic noncoding RNAs (lincRNAs) in rainbow trout (*Oncorhynchus mykiss*), Mar. Biotechnol. 18 (2) (2016) 271–282.
[26] M. Liu, Y.S. De Mitcheson, Profile of a fishery collapse: why mariculture failed to save the large yellow croaker, Fish Fish. 9 (3) (2008) 219–242.
[27] C. Wu, D. Zhang, M. Kan, Z. Lv, A. Zhu, Y. Su, D. Zhou, J. Zhang, Z. Zhang, M. Xu, The draft genome of the large yellow croaker reveals well-developed innate immunity, Nat. Commun. 5 (2014) 5227-5227.
[28] D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, S.L. Salzberg, TopHat 2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions, Genome Biol. 14 (4) (2013) 1–13.
[29] B. Langmead, C. Trapnell, M. Pop, S.L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome, Genome Biol. 10 (3) (2009) 0-0.
[30] C. Trapnell, L. Pachter, S.L. Salzberg, TopHat: discovering splice junctions with RNA-Seq, Bioinformatics 25 (9) (2009) 1105–1111.
[31] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G.T. Marth, G.R. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools, Bioinformatics 25 (16) (2009) 2078–2079.
[32] C. Trapnell, B.A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M.J. Van Baren, S.L. Salzberg, B.J. Wold, L. Pachter, Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation, Nat. Biotechnol. 28 (5) (2010) 511–515.
[33] C. Trapnell, A. Roberts, L.A. Goff, G. Pertea, D. Kim, D.R. Kelley, H. Pimentel, S.L. Salzberg, J.L. Rinn, L. Pachter, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks, Nat. Protoc. 7 (3) (2012) 562–578.
[34] L. Kong, Y. Zhang, Z.Q. Ye, X.Q. Liu, S.Q. Zhao, L. Wei, G. Gao, CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine, Nucleic Acids Res. 35 (2007) W345 (Web Server issue).
[35] L. Sun, H. Luo, D. Bu, G. Zhao, K. Yu, C. Zhang, Y. Liu, R. Chen, Y. Zhao, Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts, Nucleic Acids Res. 41 (17) (2013).
[36] M. Guttman, M. Garber, J.Z. Levin, J. Donaghey, J.T. Robinson, X. Adiconis, L. Fan, M.J. Koziol, A. Gnirke, C. Nusbaum, Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs, Nat. Biotechnol. 28 (5) (2010) 503–510.
[37] V.A. Moran, R.J. Perera, A.M. Khalil, Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs, Nucleic Acids Res. 40 (14) (2012) 6391–6400.
[38] E.P. Nawrocki, S.W. Burge, A. Bateman, J. Daub, R.Y. Eberhardt, S.R. Eddy, E.W. Floden, P.P. Gardner, T.A. Jones, J.G. Tate, Rfam 12.0: updates to the RNA families database, Nucleic Acids Res. 43 (2015) 130–137.
[39] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H.L. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium, Nat. Genet. 25 (1) (2000) 25–29.
[40] G.O. Consortium, Creating the gene ontology resource: design and implementation, Genome Res. 11 (8) (2001) 1425–1433.
[41] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of

large gene lists using DAVID bioinformatics resources, Nat. Protoc. 4 (1) (2009) 44.
[42] I. Steinfeld, R. Navon, M.L. Creech, Z. Yakhini, A. Tsalenko, ENViz: a Cytoscape App for integrated statistical analysis and visualization of sample-matched data with multiple data types, Bioinformatics 31 (10) (2015) 1683–1685.
[43] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, Methods 25 (4) (2001) 402–408.
[44] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30 (15) (2014) 2114–2120.
[45] C. Wu, D. Zhang, M. Kan, Z. Lv, A. Zhu, Y. Su, D. Zhou, J. Zhang, Z. Zhang, M. Xu, The draft genome of the large yellow croaker reveals well-developed innate immunity, Nat. Commun. 5 (1) (2014) 5227-5227.
[46] J. Ao, Y. Mu, L. Xiang, D. Fan, M. Feng, S. Zhang, Q. Shi, L. Zhu, T. Li, Y. Ding, Genome sequencing of the perciform fish *Larimichthys crocea* provides insights into molecular and genetic mechanisms of stress adaptation, PLoS Genet. 11 (4) (2015).
[47] F.C. Jones, M. Grabherr, Y.F. Chan, P. Russell, E. Mauceli, J.A. Johnson, R. Swofford, M. Pirun, M.C. Zody, S.D.M. White, The genomic basis of adaptive evolution in threespine sticklebacks, Nature 484 (7392) (2012) 55–61.
[48] B. Venkatesh, A.P. Lee, V. Ravi, A.K. Maurya, M.M. Lian, J.B. Swann, Y. Ohta, M.F. Flajnik, Y. Sutoh, M. Kasahara, Elephant shark genome provides unique insights into gnathostome evolution, Nature 505 (7482) (2014) 174–179.
[49] W. Gao, L. Fang, D. Yang, K. Ai, K. Luo, G. Tian, J. Zhou, W. Hu, H. Yuan, Q. Xu, Cloning and expression of Asian swamp eel (Monopterus albus) cxcr4 paralogues, and their modulation by pathogen infection, Aquaculture 457 (2016) 50–60.
[50] W.S. Thulasitha, N. Umasuthan, K.S. Revathy, I. Whang, J. Lee, Molecular characterization, genomic structure and expressional profiles of a CXC chemokine receptor 4 (CXCR4) from rock bream Oplegnathus fasciatus, Fish Shellfish Immunol. 44 (2) (2015) 471–477.
[51] C.Y. Lin, Y.M. Chen, H.H. Hsu, C.T. Shiu, H.C. Kuo, T. Chen, Grouper (Epinephelus coioides) CXCR4 is expressed in response to pathogens infection and early stage of development, Dev. Comp. Immunol. 36 (1) (2012) 112–120.
[52] M. Wang, S. Jiang, W. Wu, F. Yu, W. Chang, P. Li, K. Wang, Non-coding RNAs function as immune regulators in teleost fish, Front. Immunol. 9 (2018).
[53] T.T. Bizuayehu, S. Johansen, V. Puvanendran, H. Toften, I. Babiak, Temperature during early development has long-term effects on microRNA expression in Atlantic cod, BMC Genomics 16 (1) (2015) 305-305.
[54] J.A. Mennigen, Micromanaging metabolism-a role for miRNAs in teleost energy metabolism, Comp. Biochem. Physiol. B 199 (2016) 115–125.
[55] S. Basu, Y. Hadzhiev, G. Petrosino, C. Nepal, J. Gehrig, O. Armant, M. Ferg, U. Strahle, R. Sanges, F. Mã¼Ller, The Tetraodon nigroviridis reference transcriptome: developmental transition, length retention and microsynteny of long non-coding RNAs in a compact vertebrate genome, Sci. Rep. 6 (33210) (2016) 33210.
[56] T. Xu, Q. Chu, J. Cui, R. Huo, MicroRNA-216a inhibits NF-κB-mediated inflammatory cytokine production in teleost fish by modulating p65, Infect. Immun. 86 (6) (2018).
[57] B.C. Zhang, J. Zhang, L. Sun, In-depth profiling and analysis of host and viral microRNAs in Japanese flounder (*Paralichthys olivaceus*) infected with megalocytivirus reveal involvement of microRNAs in host-virus interaction in teleost fish, BMC Genomics 15 (1) (2014) 878.
[58] D. Valenzuelamiranda, K. Etebari, S. Asgari, C. Gallardoescarate, Long noncoding RNAs: unexplored players in the drug response of the sea louse Caligus rogercresseyi, Agri Gene 4 (2017) 1–7.
[59] E. Tarifeño-Saldivia, D. Valenzuela-Miranda, C. Gallardo-Escárate, In the shadow: the emerging role of long non-coding RNAs in the immune response of Atlantic salmon, Dev. Comp. Immunol. 73 (2017) 193.
[60] T. Ravasi, H. Suzuki, C.V. Cannistraci, S. Katayama, V.B. Bajic, K. Tan, A. Akalin, S. Schmeier, M. Kanamorikatayama, N. Bertin, An atlas of combinatorial transcriptional regulation in mouse and man, Cell 140 (5) (2010) 744–752.
[61] L. Jiang, W. Liu, A. Zhu, J. Zhang, J. Zhou, C. Wu, Transcriptome analysis demonstrate widespread differential expression of long noncoding RNAs involve in *Larimichthys crocea* immune response, Fish Shellfish Immunol. 51 (2016) 1–8.
[62] E. Tarifenosaldivia, D. Valenzuelamiranda, C. Gallardoescarate, In the shadow: the emerging role of long non-coding RNAs in the immune response of Atlantic salmon, Dev. Comp. Immunol. 73 (2017) 193–205.
[63] X. Liu, L. Kang, W. Liu, B. Lou, C. Wu, L. Jiang, Molecular characterization and expression analysis of the large yellow croaker (*Larimichthys crocea*) chemokine receptors CXCR2, CXCR3, and CXCR4 after bacterial and poly I:C challenge, Fish Shellfish Immunol. 70 (2017) 228.
[64] V. Valenzuelamunoz, D. Valenzuelamiranda, C. Gallardoescarate, Comparative analysis of long non-coding RNAs in Atlantic and Coho salmon reveals divergent transcriptome responses associated with immunity and tissue repair during sea lice infestation, Dev. Comp. Immunol. 87 (2018) 36–50.
[65] H.L. Yue, D.Z. Peng, Structure and function of CCL20, Immunological Journal 20 (3) (2004) 100–103.
[66] X. Hao, F. Zhang, Research advancement of CD28, Journal of Beihua University 5 (2) (2004) 133–137.
[67] T. Liu, Y. Zhang, Y. Liu, F. Sun, J. Gui, Cooperative roles of fish protein kinase containing Z-DNA binding domains and double-stranded RNA-dependent protein kinase in interferon-mediated antiviral response, J. Virol. 85 (23) (2011) 12769–12780.
[68] A. Li, X. Tian, X. Zhang, S. Huang, Y. Ma, D. Wu, G. Moeckel, S. Somlo, G. Wu, Human polycystin-2 transgene dose-dependently rescues ADPKD phenotypes in Pkd2 mutant mice, Am. J. Pathol. 185 (10) (2015) 2843–2860.