Genomics Data 7 (2016) 155–158

Contents lists available at ScienceDirect

Genomics Data

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

Data in Brief

De novo assembly and comparison of the ovarian transcriptomes of the common Chinese cuttlefish (*Sepiella japonica*) with different gonadal development

Zhenming Lü [a], Wan Liu [a], Liqin Liu [a], Huilai Shi [b], Hongling Ping [b], Tianming Wang [a], Changfeng Chi [a], Changwen Wu [a], Ching-Hung Chen [c], Kang-Ning Shen [d,*], Chung-Der Hsiao [e,**]

[a] National Engineering Research Center for Facilitated Marine Aquaculture, Marine Science College, Zhejiang Ocean University, Zhoushan 316004, China
[b] Marine Fisheries Research Institute of Zhejiang, Zhoushan 316021, China
[c] WeThink Biotech Inc., Taoyuan, Taiwan
[d] Center of Excellence for the Oceans, National Taiwan Ocean University, Keelung 20224, Taiwan
[e] Department of Bioscience Technology, Chung Yuan Christian University, Chung-Li 32023, Taiwan

ARTICLE INFO

Article history:
Received 30 November 2015
Accepted 17 December 2015
Available online 21 December 2015

Keywords:
Cuttlefish
Gonad development
Transcriptome

ABSTRACT

The common Chinese cuttlefish (*Sepiella japonica*) has been considered one of the most economically important marine Cephalopod species in East Asia and seed breeding technology has been established for massive aquaculture and stock enhancement. In the present study, we used Illumina HiSeq2000 to sequence, assemble and annotate the transcriptome of the ovary tissues of *S. japonica* for the first time. A total of 53,116,650 and 53,446,640 reads were obtained from the immature and matured ovaries, respectively (NCBI SRA database SRX1409472 and SRX1409473), and 70,039 contigs (N50 = 1443 bp) were obtained after de novo assembling with Trinity software. Digital gene expression analysis reveals 47,288 contigs show differential expression profile and 793 contigs are highly expressed in the immature ovary, while 38 contigs are highly expressed in the mature ovary with FPKM >100. We hope that the ovarian transcriptome and those stage-enriched transcripts of *S. japonica* can provide some insight into the understanding of genome-wide transcriptome profile of cuttlefish gonad tissue and give useful information in cuttlefish gonad development.

Specifications [standardized info for the reader]	
Organism/cell line/tissue	*Sepiella japonica*/ovaries
Sex	Female
Sequencer or array type	Illumina Hiseq2000
Data format	Raw and processed
Experimental factors	Transcriptome profiling of ovaries at two different maturation stages
Experimental features	Ovarian tissues at two developmental stages were dissected for total RNA extraction. Prepared cDNA libraries were paired-end sequenced by HiSeq2000 system. The obtained data was subjected for de novo transcriptome assembly using Trinity, and coding regions were predicted by BLAST. We performed BLASTP against the NR, NT, Swiss-Prot, KEGG, COG and GO database using BLAST with an e-value cut off of 1e-5 to annotate identified proteins.
Consent	N/A
Sample source location	National engineering research center for facilitated marine aquaculture, Marine science college, Zhejiang Ocean University, Zhoushan, China

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/sra/SRX1409472 for immature ovary.
http://www.ncbi.nlm.nih.gov/sra/SRX1409473 for mature ovary.

2. Introduction

The common Chinese cuttlefish (*Sepiella japonica*), also known as Japanese spineless cuttlefish, is an economically important Cephalopod species in Chinese coastal regions, south from Hong Kong and north to Korea and west Japan. It was once one of four major cephalopod species harvested in these regions. However, since the 1980s, wild-caught *S. japonica* fisheries have suffered a severe decline due to over-fishing [1,2]. To enhance production, artificial breeding methods have been

* Correspondence to: K.-N. Shen, Center of Excellence for the Oceans, National Taiwan Ocean University, Keelung, 20224, Taiwan.
** Correspondence to: C.-D. Hsiao, 200 Chung-Pei Rd., Department of Bioscience Technology, Chung Yuan Christian University, 32023 Chung-Li, Taiwan.
E-mail addresses: knshen@email.ntou.edu.tw (K.-N. Shen), cdhsiao@cycu.edu.tw (C.-D. Hsiao).

http://dx.doi.org/10.1016/j.gdata.2015.12.011

developed, and successful aquaculture techniques have been established in China [3]. Meanwhile, the restoration programs have been initiated in Chinese coastal waters to address the decline of wild populations of this species [1,3]. However, previous studies has revealed that *S. japonica* had undergone a sophisticated precocious puberty when bred under artificial captivated conditions, which has undercut its possible potential for breeding along the coast of China. In some circumstance, captivated *S. japonica* can even get puberty and sex matured at age of 3–4 months when the average wet body weight reached 100 g during aquaculture [4]. Therefore, how to prevent the early puberty of this species in artificial environment has become quite crucial for the development of *S. japonica* aquaculture in the future, but the first step is necessary to understand the neuro-endocrine and molecular basis for gonad development and maturation of this species.

Oocyte growth is a period of intense RNA synthesis, replication and redistribution of cytoplasmic organelles, and nutrient incorporation in oviparous animals, which may influence timing of puberty, fecundity, egg quality, and early embryogenesis [5]. The knowledge of how and what endocrine and/or intra-ovarian factors regulate oocyte growth is crucial for understanding the physiological basis for animal's puberty. In marine animals, with the aid of recent large-scale transcriptome studies which mainly conducted on teleosts, such as, trout [10], cod [6], and striped bass [7], a number of ovarian genes have been sequenced and deposited in databases making it possible to identify many mRNAs, profile their expression, and determine their functions. Genes involved in early gametogenesis [10], vitellogenesis [16], and final oocyte maturation [9] have been revealed and received considerable research attention. Possible precocious puberty related genes have been investigated in certain teleosts [8]. However, relatively few ovarian genes have been profiled in marine invertebrate Cephalopod and little is known about temporal gene expression during their oocyte growth. The objective of this study is to identify differentially expressed genes during oocyte growth and maturation process with the ultimate goal of establishing what genes regulate these stages of oogenesis and identifying the intra-ovarian factors that drive the puberty onset of the marine Cephalopod species, *S. japonica.*

3. Experimental design, materials and methods

3.1. RNA extraction

The immature (GSI = 0.49) and mature (GSI = 1.13) ovaries were dissected from artificial cultured *S. japonica* and immediately stored in RNA later (Qiagen, Hilden, Germany) and then stored at −80 °C prior to RNA extraction. Total RNAs were extracted by using the TRIZOL Kit (Invitorgen, Carlsbad, CA, USA) following manufacturer's instructions. We also performed histological section to validate the developmental stage of immature (Fig. S1A) and mature (Fig. S1B) ovaries. Total RNA samples were then digested by DNase I to remove potential genomic DNA contamination. Integrity and size distribution were checked with Bioanalyzer 2100 (Agilent technologies, Santa Clara, CA, USA). Equal amounts of the high quality RNA samples from each tissue were then proceed to perform cDNA synthesis and next generation sequencing.

3.2. RNA isolation, library construction and Illumina sequencing

Initially, about 2.5 μg of starting total RNAs were used to synthesize the cDNA libraries by following the standard protocols of the Illumina TruSeq RNA Sample Preparation Kit (Illumina, San Diego, CA, USA). The final library had an average fragment size of 180 bp and final yields of ~400 ng. After KAPA quantitation and dilution, the library was sequenced on an Illumina HiSeq 2000 with 101 bp paired-end reads. A total of 53,116,650 and 53,446,640 paired-end clean reads were generated with a read length of 101 bp for immature and mature gonads, respectively. Adaptor sequences were trimmed and reads with low quality or length less than 70 were further removed by SolexaQA software. After the removal of ambiguous nucleotides, duplicates and low-quality sequences (Phred quality scores <20), a total of 106,563,290 cleaned reads were obtained. The raw transcriptomic sequences for immature and mature ovaries were deposited in the NCBI SRA database with accession number of SRX1409472 and SRX1409473.

3.3. De novo transcriptome assembly and functional annotation of gonad expressed genes

Cleaned reads were de novo assembled into unigenes by Trinity software [12] with default parameter settings. The transcriptome was assembled, combining 106,563,290 cleaned reads into 70,039 unigenes, ranging from 200 to 15,925 bp in length. The average length was 800 bp, the N50 length was 1443 bp (Fig. 1). The assembled unigenes were subjected to similarity search against NR, NT, Swiss-Prot, KEGG, COG and GO database using BLAST [13] with an e-value cut off of 1e-5. Gene name, descriptions and Gene ontology (GO) were assigned to each unigene based on the BLASTx results and we totally obtained 28,492 unigenes (41% annotation rate) has a significant hit. The E-value, similarity and species distribution for BLAST search were summarized in Fig. 2. KEGG pathways were assigned to assembled

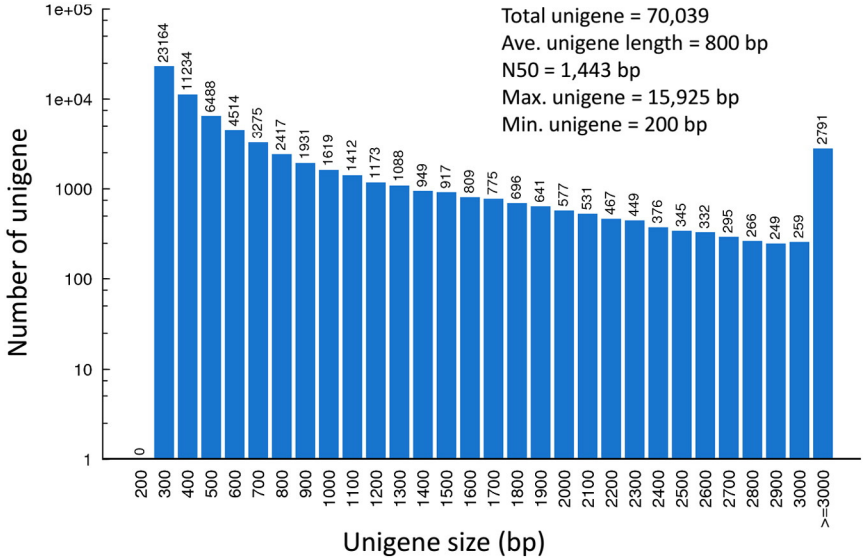

Fig. 1. Length distribution of the assembled unigene of *Sepiella japonica* ovary transcriptome.

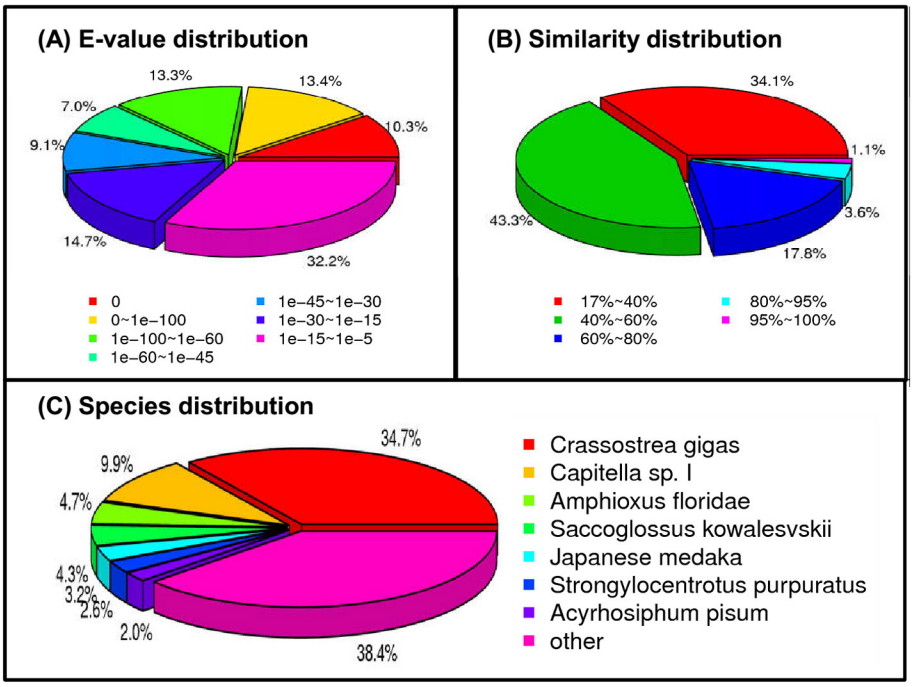

Fig. 2. The E-value (A), nucleotide similarity (B) and species distribution (C) for BLAST search of the *Sepiella japonica* ovarian transcriptome.

unigenes using the online KEGG Automatic Annotation Server (KAAS) (http://www.genome.jp/tools/kaas/). The Bi-directional Best Hit (BBH) method was used to obtain KEGG Orthology (KO) assignment. Gene ontology (GO) analysis was conducted on those 28,492 unigenes unigenes by Blast2Go. A total of 10,969 unique proteins were assigned at least one GO term for describing biological processes, molecular functions and cellular components.

3.4. Gene expression quantification and differential expression analysis

The high-quality cleaned reads of each RNA-seq library were mapped to the assembled transcripts with Bowtie program [14]. The counting of alignments was done using RSEM [15]. The differential expression statistical analysis was done using the statistical method described in the R package [11]. Differentially expressed gene (fold changes >2 and adjusted p-value <0.001) between two samples were identified with the software. Mapping results show 47,288 contigs display significant differential expression between the immature and mature ovaries (Fig. 3). The contigs with FPKM ≥ 100 was defined as abundant expressed genes (AEGs). In total, 793 immature ovary-AEGs and 38 mature ovary-AEGs were identified in the common Chinese cuttlefish gonad transcriptome (Table S1). Apparently, the immature stage-biased AEGs including cyclin-O, piwi-like protein 1-like, vasa-like protein, histone H1, histone H1.6, histone H2A, gonadal-like histone H2B, histone H2B.1/H2B.2-like, histone H3.2-like, histone H4-like, sex-determining protein fem-1, DNA replication licensing factor mcm5, proliferating cell nuclear antigen (PCNA) and proliferation-associated protein 2G4 participate in germ plasm formation, early meiotic division resumption, oocyte division and cell proliferation. On the contrary, the mature stage-biased AEGs like vitellogenin-6, mitochondrial ferritin, transferrin and Cathepsin L are mostly associated with yolk protein synthesis, oocyte quality control, oocyte development/maturation, vitellin degradation and yolk protein utilization. We hope that the ovarian transcriptome and those stage-enriched transcripts of *S. japonica* can provide some insight into the understanding of genome-wide transcriptome profile of cuttlefish gonad tissue and give useful information in cuttlefish gonad development.

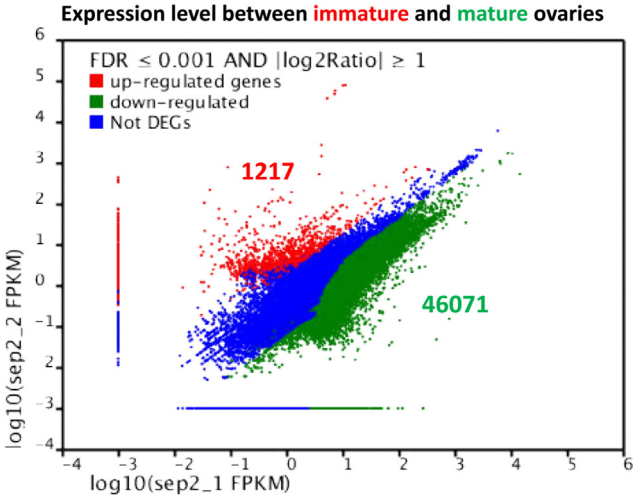

Fig. 3. The scattered plot showing the differential expressed unigenes between the immature (red) and mature (green) ovaries. The unigene highlighted in blue showing no differential expressed pattern.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gdata.2015.12.011.

Acknowledgments

This work was supported by the National Key Technology R&D Program Foundation of China (2011BAD13B08).

References

[1] C.W. Wu, C.F. Chi, G.Y. He, Z.M. Lü, M.Y. Xu, Isolation via enrichment and characterization of ten polymorphic microsatellite loci in the cuttlefish, *Sepiella maindroni* de Rochebruns. Acta Oceanol. Sin. 29 (6) (2010) 121–124.

[2] L.H. Jiang, A.Y. Zhu, C.W. Wu, Y.Q. Su, J.S. Zhang, Z.Y. Dong, Tetracycline Immersion tagging of cuttlefish, Sepiella japonica, larvae. J. World Aquacult. Soc. 45 (3) (2014) 342–349.
[3] F. Yin, P. Sun, S.M. Peng, B.J. Tang, D. Zhang, C.L. Wang, C.K. Mu, Z.H. Shi, The respiration, excretion and biochemical response of the juvenile common Chinese cuttlefish, *Sepiella maindroni* at different temperatures. Aquaculture 402 (15) (2013) 127–132.
[4] X.D. Zheng, X.Z. Lin, Z.K. Wang, R.H. Yu, C.Y. Tian, Q. Li, Studies on life cycle of cultured spineless cuttlefish Sepiella Japonica. Trans. Oceanol. Limnol. 3 (2010) 24–28.
[5] J.A. Luckenbach, D.B. Iliev, F.W. Goetz, P. Swanson, Identification of differentially expressed ovarian genes during primary and early secondary oocyte growth in coho salmon, *Oncorhynchus kisutch*. Reprod. Biol. Endocrinol. 6 (2008) 2.
[6] L. Kleppe, R.B. Edvardsen, T. Furmanek, G.L. Taranger, A. Wargelius, Global transcriptome analysis identifies regulated transcripts and pathways activated during oogenesis and early embryogenesis in Atlantic cod. Mol. Reprod. Dev. 81 (2014) 619–635.
[7] B.J. Reading, R.W. Chapman, J.E. Schaff, E.H. Scholl, C.H. Opperman, C.V. Sullivan, An ovary transcriptome for all maturational stages of the striped bass (*Morone saxatilis*), a highly advanced perciform fish. BMC Res. Notes 5 (2012) 111.
[8] R. Patiño, C.V. Sullivan, Ovarian follicle growth, maturation, and ovulation in teleost fish. Fish Physiol. Biochem. 26 (2002) 57–70.
[9] J. Bobe, T. Nguyen, B. Jalabert, Targeted gene expression profiling in the rainbow trout (*Oncorhynchus mykiss*) ovary during maturational competence acquisition and oocyte maturation. Biol. Reprod. 71 (2004) 73–82.
[10] D. Baron, R. Houlgatte, A. Fostier, Y. Guiguen, Large-scale temporal gene expression profiling during gonadal differentiation and early gametogenesis in rainbow trout. Biol. Reprod. 73 (2005) 959–966.
[11] F. Cordero, M. Beccuti, M. Arigoni, S. Donatelli, R.A. Calogero, Optimizing a massive parallel sequencing workflow for quantitative miRNA expression analysis. PLoS One 7 (2012), e31630.
[12] B.J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P.D. Blood, J. Bowden, M.B. Couger, D. Eccles, B. Li, M. Lieber, et al., De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8 (2013) 1494–1512.
[13] A. Julich, Implementations of BLAST for parallel computers. Comput. Appl. Biosci. 11 (1995) 3–6.
[14] B. Langmead, Aligning short sequencing reads with Bowtie. in: A.D. Baxevanis, et al., (Eds.), Current Protocols in Bioinformatics/Editoral Board, 2010 (Chapter 11, Unit 11 17).
[15] B. Li, C.N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinf. 12 (2011) 323.
[16] R. Patino, C.V. Sullivan, Ovarian follicle growth, maturation, and ovulation in teleost fish. Fish Physiol. Biochem. 26 (2002) 57–70.