Transcriptome Analysis of the *Sepia pharaonis*: Identification of salt stress-related information and Microsatellite Markers

Zhiming Ren

Ningbo University

Yun Zhao

Ningbo University

Weiwei Song (✉ songweiwei@nbu.edu.cn)

Ningbo University <https://orcid.org/0000-0002-5240-9591>

Chunlin Wang

Ningbo University

Changkao Mu

Ningbo University

Ronghua Li

Ningbo University

Research article

Keywords: *Sepia pharaonis*, transcriptome analysis, salt stress-related information, SSR discovered

Posted Date: January 3rd, 2020

DOI: <https://doi.org/10.21203/rs.2.20024/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: *Sepia pharaonis* has great commercial value for aquaculture. However, it is sensitive to salinity fluctuations and lacking in genomic information. The present work utilized throughput transcriptome sequencing to assess the factors associated with salt stress in *Sepia pharaonis*. **Results:** Based on the Illumina paired-end sequencing results, 203,852,818 raw reads were produced, and 130,857 unigenes were assembled having an average of 784.72 bp in length. Transcriptome analysis identified 16013 and 24119 unigenes in the Swiss-Prot protein database and NCBI non-redundant database, respectively. Of note, 12717 unigenes were grouped into 64 Gene Ontology (GO) terms, 5237 unigenes were classified into 332 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, 13808 unigenes were subcategorized into 25 Cluster of orthologous groups for eukaryotic complete genomes (KOG) functional categories based on functional analysis. Besides, 6153 genes were identified as differentially expressed ($p \leq 0.05$), of which 3340 were increased and 2813 were decreased in treatment group relative to the control group. Subsequently, these DEGs were allocated to 226 KEGG pathways and 491 GO terms. Analysis of the transcriptome sequences and DEGs identified several unigenes and pathways involved in salt stress regulation. Moreover, the *Sepia pharaonis* carried 101576 simple sequence repeats (SSRs). **Conclusions:** This is the first time osmoregulation in *Sepia pharaonis* has been explored by transcriptome sequencing. The data presented here reveals key insights into the genetic markers of salt stress in *Sepia pharaonis*.

Background

Sepia pharaonis is commonly located in the [Red Sea](#) and [Persian Gulf](#), from northern Australia to Gulf of Thailand and Japan, and from the Indian Ocean to the Andaman Sea [1]. Being a giant cuttlefish species, it can grow to 42 cm in mantle length and 5 kg in weight. The feeding habits, resistance to disease, tolerance to crowding and handling, short life span, and rapid growth makes this species an ideal commercially viable mariculture [2].

Among the key factors that determine the stable growth and well-being of many aquaculture species in salinity. Thus, the level of salinity should be optimized for effective artificial propagation of *Sepia pharaonis*. During culture, *Sepia pharaonis* often experiences significant changes in salinity caused by heavy rainfalls, and this may lead to high mortality and farm productivity. Osmoregulation is the most direct response to salinity change. Then physiological functions of immune, digestion system will be affected, causing growth and survival threat.

Several types of research about the effect of salinity to *Sepia pharaonis* have been done. LE Ke-xin et al. [3] found the suitable salinity of *Sepia pharaonis* larvae was from 24 to 33‰ and the most appropriate salinity was 27, by researching the impact of salinity on survival and growth similar to the finding of Dai Yuan-tang et al. [4]. HUANG Jian-sheng et al. [5] studied the effects of salinities (15.0, 18.0, 21.0, 24.0, 27.0, 30.0 and 33.0) on fertilized egg hatching and survival activity index (SAI) of the larvae of *Sepia pharaonis* in the laboratory at $22 \pm 1^\circ\text{C}$. WEN Jing et al. investigated the effects of salinity on survival and behavior in juvenile *Sepia pharaonis* in a laboratory [6]. Except for suitable salinity for fertilized egg, larvae and juvenile, the effects of abrupt and gradual changes of salinity on survival rate, specific growth rate, weight gain rate, hepatosomatic index, and enzyme activity (superoxide dismutase, glutamic-pyruvic transaminase, glutamic oxaloacetic transaminase, alkaline phosphatase,) were explored as well [7]. However, current researchers focus on tolerance and general physiology to the salinity of this species, and the molecular mechanism is little to be understood.

Currently, unlike many other species for which genomic sequences are available, such sequences are unavailable for *sepia* and this have substantially limited the exploration of its physiological regulation. To overcome this challenge, researchers have employed next-generation sequencing approach to characterize the molecular process of this species efficiently at low cost [8]. Transcriptome sequencing yields double-stranded cDNA which are annotated in relevant databases [9]. This has expanded our understanding and accelerated the identification of new genes networks, for many organisms for which the genomes are yet to be fully elucidated [10]. Additionally, the widespread application of next-generation sequencing has also led to the discovery and characterization of SSRs for many aquaculture species. This is because its time-saving [11]. SSRs are important factors used in breeding and genetic studies.

In the present work, Illumina Hiseq2500 tool was employed to map the gill transcriptome of *Sepia pharaonis*. Genes related to salinity stress were identified. Furthermore, we uncovered and characterized SSR elements in the transcriptome for the first time in *Sepia pharaonis* which helps to understand its genetic information and osmoregulation.

Results

Sequencing and evaluation of the assembly

The sequencing yielded 203,852,818 raw reads. These data were filtered and screened, and 200,000,000 clean reads were harvested (Table 1). Q30 bases rate, reflecting the quality of next-generation sequencing results, was 94.53%. All clean reads generated here were submitted to the NCBI Sequence Read Archive (SRA) database (accession number: SRR2241952, SRR2241953, SRR2241954).

There are 130857 unigenes, among which >500 bp were 62569 unigenes and >1000 bp were 23723 unigenes (Table 2). The unigene (N50) has a median size of 970 bp. the lengths of the unigene were in the range of 301 bp to 39756 bp, and the average length of the unigene is 784.72 bp.

Annotation of Unigenes

We screened for assembled unigenes from the KEGG, COG, Swiss-Prot, NCBI Nr with the use of BLASTx algorithm (E-value, 1.0×10^{-5}) and by matching sequences. As a result, 24119 (18.43%) and 16013 (12.24%) unigenes can be matched to Nr and Swissprot database (Table 3), while 13808 (10.55%), 5237 (4.00%) and 12717 (9.72%) unigenes were subgrouped by KOG, KEGG and GO databases (Table 4).

By data comparison analysis, we found similarity between *Sepia pharaonis* and octopus genomic data published is not high, suggesting it is not suitable for reference analysis. Therefore, alignment was done as no reference genome. Top-hit species distribution for sequences from *sepia* was shown in Figure 1. The distribution of four major species was searched on the Nr database, 5244 (21.74%) unigenes displayed similar features with *Crassostrea gigas*, and *Lottia gigantea* (3304, 13.70%), *Aplysia californica* (2069, 8.58%), *Hydra vulgaris* (1096, 4.54%), *Capitella teleta* (861, 3.57%), *Trichuris suis* (811, 3.36%), *Strongylocentrotus purpuratus* (783, 3.25%), *Stegodyphus mimosarum* (703, 2.91%), *Saccoglossus kowalevskii* (629, 2.61%), others (8619, 35.74%).

KOG, GO and KEGG characterization of Unigenes

The unigenes were matched on the KOG database for classification and functional prediction. Collectively, 13808 unigenes were subcategorized into 25 KOG classes (Figure 2). The most enriched categories included: chromosome partitioning, cell division, Cell cycle control (829), RNA processing and ribosomal biogenesis and structure, translation, (958), vesicular transport secretion, and intracellular trafficking, and (1247), chaperones, protein turnover, posttranslational modification, (2465), general function prediction only (6181), signal transduction mechanisms (4017), Transcription (1424), Function unknown (1376), modification (1185), Cytoskeleton (952).

Gene are often classified into molecular functions, cellular components and biological process using the Gene Ontology (GO) system. In this study, GO analysis identified 64 Go terms for the 12717 unigenes. Of these, 31198 belonged to the biological processes (BP), 16213 belonged to the cellular component (CC), and 17719 belonged to the molecular function (MF) (Figure 3). In the MF subgroup, majority of the unigenes were linked to catalytic activity (6763, 53.18%), binding (8381, 65.9%). In the BP subgroup, majority of the unigenes were associated with metabolic processes (7609, 59.83%), cellular process (7603, 59.78%). In the CC subgroup, a high proportion of the unigenes were linked to cell (3372, 26.51%), cell part (3372, 26.51%).

To comprehensively characterize the biological pathways in *Sepia pharaonis*, the assembled sequences were analyzed against the KEGG database. In total, we mapped the 5237 unigenes corresponding to 332 KEGG pathways (TableS1). Among these unigenes, 3777 were assigned to organismal systems, primarily in the endocrine system (1140), immune system (679), nervous system (617). Metabolism cluster ranked second with (3719) and the lowest cluster was membrane transport with (50). The rest of the genes were as follows: involving carbohydrate metabolism (733), glycan metabolism and biosynthesis (303), lipid metabolism (440), amino acid metabolism (671), energy metabolism (284) and so on. Signaling molecules and interaction (241), signal transduction (2216) and environmental information processing (2507) were also identified. There were 4555, 1668 and 1638 unigenes that were classified into human disease, genetic information processing and cellular processes, respectively.

DEGs between Treatment Group and Control Group

Totally, the number of genes that showed marked differential expression were 6153 ($p \leq 0.05$), of which 3340 were up regulated and 2813 were down regulated genes in treatment group relative to the control group. We subsequently conducted KEGG and GO analyses to uncover the functions of the DEGs. We allocated the DEGs to 226 pathways and 491 GO terms on the KEGG platform (Figure 4) (Table S2). The DEGs were preferentially enriched in oxygen transport and neuropeptide signaling pathway in the BP subgroup, extracellular region and endoplasmic reticulum lumen in cellular component, and oxygen transporter activity and dolichyl-diphosphooligosaccharide-protein glycotransferase activity in molecular function (Table S3). In KEGG database, DEGs are greatly enriched in tyrosine metabolism, protein processing in endoplasmic reticulum, betalain biosynthesis, isoquinoline alkaloid biosynthesis and riboflavin metabolism (Figure 5) (Table S4).

qRT-PCR of DEGs Related to Salinity Stress in Two Groups

In the subsequently catheterization, qRT-PCR was performed for validation of the DEGs. This was done for 9 genes using from the control and treatment groups. Figure 6 illustrated that the differential expression of the genes correlated with sequencing results.

SSRs Discovery

Summary of SSR identified from the transcriptome was shown in Table4. In total, 101576 SSRs were recognized in 130857sequences, and 24737sequences harbored > one SSR. The compound formation contained 10725SSRs. Table S5A shows the numbers of tandem repeats and the frequency of SSRs.

Discussion

Sepia pharaonis is of great commercial value for aquaculture. However, *Sepia pharaonis* is sensitive to salinity fluctuations. After heavy rain falls, salinity decreased sharply, making a threat to survival of this species. Understanding the molecular mechanisms is an important step in the discovery of signaling pathways that participate in osmoregulation. Given that little is known about *Sepia pharaonis*, we investigated it with transcriptome sequencing. In our study, a total of 203,852,818 raw reads were produced, and 130,857 unigenes were assembled. It is expected that this will yield key information concerning *Sepia pharaonis* in salinity stress.

Genes and Processes involved in salinity stress of *Sepia pharaonis*

Herein, we performed transcriptome sequencing to reveal the genes regulating salinity stress. The deep sequencing technique was employed for denovo assembly of the transcriptome of *Sepia pharaonis* gill. The development of *Sepia pharaonis* osmoregulation includes a highly regulated interplay between receptors and effectors.

In *Sepia pharaonis*, perception of ionic and osmotic imbalances is drive by receptors on the cell membrane. Ion channels (Cl^- , Ca^{2+} , Na^+ , and K^+ channels) and AQP(Aquaporin TIP4-1) play key roles in the reception of such signals. Studies have recognized that phospholipase A2 (PLA2) modulates the production of osmolytes by finetuning the mobilization of arachidonic acid, which activates cell swelling [12]. Arachidonic acidmetabolism was enriched in KEGG pathway and PLAa was also annotated.

This led to the identification of transduction pathways for salt stress signals. In GO term, response to stimulus (GO:0050896) was identified. In KEGG , calcium signaling pathway(ko04020) , MAPK signaling pathway (ko04010).This results implies that salt stress changes membrane fluidity, inducing cytosolic Ca^{2+} cycling [13].

The mitogen-activated protein kinase (MAPK) cascade modulates many cellular processes such as migration, differentiation and cell proliferation. It is reported that MAPK cascade acts when plant in salt stress status[14].

We uncovered many stress factors at low osmotic stress. (A) water channels, osmolyte transporters, and ion channels (including Na^+ , K^+ , Ca^{2+} and Cl^- channels), were activated altering the osmotic balances. For example, electron carrier activity (GO:0009055), enzyme regulator activity (GO:0030234) were observed. (B) Degradation or denovo synthesis of free amino acids (FAA) which may play a role of compensating or alleviating the effects of ion influx during salinity stress conditions [15]. For example, arginine and proline metabolism ko00330, alanine, aspartate and glutamate metabolism ko00250, glycine, serine and threonine metabolism ko00260, tryptophan metabolism ko00380, cysteine valine, leucine and isoleucine degradation ko00280, and methionine metabolism ko00270, histidine metabolism ko00340, Tyrosine metabolism ko00350, phenylalanine metabolism ko00360, soleucine, leucine, valine, and biosynthesis ko00290, phenylalanine, tyrosine and tryptophan biosynthesis ko00400,Lysine biosynthesis ko00300 were enriched. As shown in Table S1, FAAs metabolic

pathways were activated, which changed osmotic status in the *Sepia pharaonis* and this allowed them to adjust to osmotic stress environment. (C) ROS-related pathways and genes were enriched. Antioxidant activity (GO:0016209), peroxisome (ko04146), glutathione metabolism (ko00480), glutathione peroxidase 3, superoxide dismutase [Mn], mitochondrial, superoxide dismutase [Cu-Zn], peroxisomal membrane protein11A, peroxisomal biogenesis factor3 were observed. (D) Moreover, many other types of metabolism, including immune responses and endocrine system, were shown to be enriched in the KEGG pathway analysis. We found platelet activation (ko04611), leukocyte transendothelial migration (ko04670), B cell receptor signaling pathway (ko04662), Toll-like receptor signaling pathway (ko04620) T cell receptor signaling pathway (ko04660), and complement and coagulation cascades (ko04610). For endocrine system, thyroid hormone signaling pathway (ko04919), Thyroid hormone synthesis (ko04918), Prolactin signaling pathway (ko04917) were observed. These various regulatory pathways may exist in oysters to allow them to cope with salt stress.

Finally, amino acid metabolism, ion currents and water currents were observed to be balanced to maintain isoosmotic conditions.

DEGs between Treatment Group and Control Group 500

Differential analysis of gene expression between salinity22 and salinity28 provide an opportunity to understand the critical genes in the osmoregulation, including the regulation to low salty stress. As a result, 6153 genes were found to be significantly differentially expressed ($p < 0.05$), including 3340 up regulated and 2813 down regulated genes. 665 (10.81%) of DEGs were annotated successfully to at least one database. To understand the functions of these DEGs, GO and KEGG analyses were performed. All DEGs were assigned to 491 GO terms and 226 pathways through the KEGG database (Table S4).

Of these genes, there were 67 with little to no expression in salinity28 but high expression in salinity22, while 49 with little to no expression in salinity22 but high expression in salinity28.

Zinc finger protein 92 homolog and Zinc finger protein 300 were down-regulated. Zinc finger protein 41, Zinc finger protein 333, AN1-type zinc finger protein 4 and Zinc finger protein 26 were up-regulated. That means zinc finger protein affect each other in salinity stress. Zinc fingers were first identified in a study of transcription in the [African clawed frog, *Xenopus laevis*](#) [16], widely distributed in animals, plants and microorganisms. Proteins that contain zinc fingers (zinc finger proteins) is mostly as interaction modules that bind DNA, RNA, proteins, or other small, useful molecules, and variations in structure serve primarily to alter the binding specificity of a particular protein. In plants, like *Oryza sativa* [5, 17], *Arabidopsis thaliana* [18], zinc finger proteins were found to play an important role when plants faced with drought, high salt and low temperature. Histone-lysine N-methyltransferase SETMAR is another one. SETMAR (alternative name: Metnase) is a domesticated primate transposase that enhances DNA repair, replication, and decatenation [19, 20]. It has a histone methyltransferase activity and methylates 'Lys-4' and 'Lys-36' of histone H3, Specifically, it mediates dimethylation of H3 'Lys-36' at sites of DNA double-strand break and may recruit proteins required for efficient DSB repair through non-homologous end-joining [21, 22]. Histones modifications, part of epigenetic processes, influence the efficiency of stress-induced gene expression in plant [23]. All of these implies the possible function of SETMAR in sepia to salinity stress.

To understand the functions of these DEGs, GO and KEGG enrichment were done. For up-regulated genes, adipocytokine signaling pathway, fatty acid metabolism, PPAR signaling pathway, fatty acid degradation, and

insulin resistance are the top ones. Peroxisome proliferator-activated receptors (PPARs) are nuclear hormone receptors, activated by fatty acids and their derivatives. PPAR has three subtypes (PPA Ralpha, beta/delta, and gamma), showing different expression patterns in vertebrates. Each of them is encoded in a separate gene and binds fatty acids and eicosanoids. Additionally, it is reported that during long term salinity exposure, fatty acid composition of gill membrane in *Gammarus duebeni* changes, being a highly effective method to reduce ion diffusion and water inox [32]. Genes (CPT-1,CPT-2) are both up-regulated in PPAR signaling pathway and fatty acid degradation. Further study should be done to help understand the function of CPT-1,CPT-2 in the pathway. For down-regulated genes, protein processing in endoplasmic reticulum, tyrosine metabolism, betalain biosynthesis, isoquinoline alkaloid biosynthesis, riboflavin metabolism are the top ones. During culture, *Sepia pharaonis* gives out ink when salinity suddenly changes. Tyrosine metabolism is the biological processes behind this phenomenon. 8 genes down-regulated in tyrosine metabolism may play an important role in osmoregulation in *Sepia pharaonis*. Down-regulated genes also enriched in Genes oxygen transport, oxygen transporter activity, response to stress, and oxidation–reduction process (Go term). We found related genes: hemocyanin G-type, units Oda to Odg (ODHCY), hemocyanin A-type, units Ode to Odg (Fragment), hemocyanin, units G and H (Fragments), and superoxide dismutase [Cu-Zn]. [Osmoregulation influences respiratory gas exchange and transport](#) [24]. The answer to gas exchange and transport, involves mechanisms of salt and water balance. Potassium channel subfamily K member 16 (KCNK16), voltage-gated hydrogen channel 1 (HVCN1) and sodium channel protein 1 brain were identified. Besides, sodium- and chloride-dependent GABA transporter 2 (SLC6A13), sodium-and chloride-dependent glycine transporter 2 (SLC6A5), solute carrier family 35 member B1 (slc35b1), solute carrier family 22 member 21 (Slc22a21) and electroneutral sodium bicarbonate exchanger 1 (SLC4A8) were also observed. Especially, SLC6A5 has little to no expression in salinity22 but high expression in salinity 28. GlyT2, a protein that in humans is encoded by the SLC6A5 gene, is a specific marker of glycinergic neurons and a member of the Na⁺ and Cl⁻-coupled transporter family SLC6. These ion channel and SLC provides an view to study salt and water balance in *Sepia pharaonis*.

SSR discovery

SSRs are useful molecular markers for genetic and breeding studies[25]. Next-generation sequencing has recently been used to discover SSRs in aquaculture species, and is considered a time-saving, highly efficient approach. In total, 101576 SSRs were recognized in 130857sequences, and 24737sequences contained more than one SSR. These microsatellites are expected to be useful for genetic linkage mapping and other genetic studies.

Conclusion

This study is the first time to investigate osmoregulation in *Sepia pharaonis* by transcriptome sequencing. Our transcriptome sequencing produced 203,852,818 raw reads, and 130,857 unigenes were assembled. DEGs analysis revealed 6153 significant genes between two different groups. Based on insightful information from unigenes and DEG, we analyzed related pathways and genes in salinity stress of *Sepia pharaonis*. Moreover, we discover 101576 SSRs in 130857sequences . These microsatellites will be useful for genetic linkage mapping and other genetic studies. Overall, the current study provided information to understand the molecular process of *Sepia pharaonis* osmoregulation, which will provide more useful insight for aquaculture practice.

Methods

Ethics statement

All experiments were conducted in compliance with guidelines and approval of the Animal Care and Use Committee of Ningbo University.

RNA extraction, cDNA library construction, and Illumina sequencing

Total RNA was extracted using Trizol reagent (Invitrogen, USA) following the manufacturer's instructions. RNA quality and quantity were analyzed using a Bio Analyzer 2100 (Agilent Technology, Santa Clara, CA) and Nano Drop 2000 spectrophotometer (Infinigen Biotechnology Inc., City of Industry, CA), respectively. They were fragmented by treating with heat and divalent cations before cDNA synthesis. The cDNA was reverse transcribed with random hexamer primers, end repaired by DNA polymerase and adapter ligated with T4 DNA ligase, according to Illumina manufacturer's protocol. This was followed by second-strand cDNA synthesis, end repair, and adaptor ligation. Finally, libraries with insert lengths of ~280 bp were created by PCR amplification and purification. Each library was sequenced on an Illumina HiSeq2500 in 125PE mode (Illumina Inc., San Diego, CA, USA). Short reads were deposited in the NCBI Sequence Read Archive (SRA) under Accession numbers SRR2241952, SRR2241953, and SRR2241954.

Transcriptome De novo Assembly and Annotation

In order to ensure reliable assembly results, the quality of raw reads was checked by Fast QC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and filtered by the high-throughput quality control (HTQC) toolkit [26]. The following quality-filtering criteria were applied: 1) the 5-bp window was omitted if the average quality was lower than 20; 2) reads were removed if the percentage of unknown bases was higher than 10%; 3) read pairs were filtered if any one read end was shorter than 50 bp. The resulting cleaned reads were used in the following bioinformatics pipelines. The Trinity package was used for transcript assembly using default parameters [27]. The assembled transcripts were then processed through the Evigene package (http://arthropods.eugenes.org/Evidential_Gene/) to eliminate sequence redundancy with default parameters [28]. The resulting transcripts that showed significant similarities (>90%) were then clustered and the longest transcripts for each group were selected as representative unigenes, which were then used for functional annotation. Sequence-length statistics of the assembled transcriptome were performed using our own Python scripts.

For gene annotation, BLAST package (with an E-value threshold of 1×10^{-5}) of all unigenes were performed in the National Center for Biotechnology Information (NCBI) non-redundant protein (nr) database. Transcripts were further annotated using the Swiss-Prot, the Cluster of Orthologous Groups (COG), Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases with Blast2GO.

Differently expressed genes (DEGs) Identification and Enrichment Analysis

Single-end RNA sequencing results from two groups were mapped to the assembled transcriptome using bowtie2 (<http://bowtiebio.sourceforge.net/bowtie2/manual.shtml>) [27]. The mapped read counts for each transcript were normalized by RPKM [29], which is able to eliminate the influence of different gene lengths and

sequencing levels using RSEM [30]. Therefore, the calculated gene expression can be directly used for comparing differences in gene expression between samples. Then, we calculated differences in the abundance of expression of each gene on each transcript between two samples using DESeq [31] with a $p < 0.05$. Biological repetitive analysis was carried out to evaluate the quality of sequence data.

Using the BLASTx tool, which was developed to evaluate the similarities between two sequences, the DEGs were searched against National Center for Biotechnology Information (NCBI) non-redundant protein(Nr) database, Swiss-Prot protein database, euKaryotic Ortholog Groups(KOG), and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database.

Quantitative real-time PCR of DEGs related to salinity stress in two groups

Total RNA was extracted from each example using TRIzol Reagent (Invitrogen), and converted to single stranded cDNA using Superscrip III (Invitrogen). Quantitative real-time PCR analysis was performed using SYBR green PCR system (Applied Biosystems) on ABI prism 7900HT system (Applied Biosystems). Primer sequences are listed in Table 5. Relative gene expression level, normalized to 18s rRNA expression, was calculated by $2^{-(\Delta\Delta C_t)}$. All the experiments were repeated three times. The results were presented as fold change to the level expressed in samples.

Simple Sequence Repeat (SSR) Detection and Primer Design

Potential SSR markers were detected among the 130857 unigenes using the MicroSAteellite identification tool (MISA)(<http://pgrc.ipk-gatersleben.de/misa/>) [32]. We searched for SSRs with motifs ranging from di- to hexa-nucleotides in size. Given difficulties in distinguishing genuine mono-nucleotide repeats from polyadenylation products and some mono-nucleotide repeats that were generated by base mismatches or sequencing errors, mono-nucleotide repeats were discarded in this study. According to the MISA results, primer pairs flanking each SSR locus were designed using the Primer3 program (http://www.broadinstitute.org/genome_software/other/primer3.html) [33].

Abbreviations

GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; DEGs: Differently expressed genes; NCBI: National Center for Biotechnology Information; Nr: non-redundant protein; HTQC: high-through-put quality control; SSRs: simple sequence repeats

Declarations

Ethics approval and consent to participate

All aspects of the experiment were conducted according to the guidelines provided by the Ethical Committee of Experimental Animal Care at Ningbo University (Ningbo, Zhejiang, China).

Consent for publication

Not applicable.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Competing Interests

The authors have declared that no competing interests exist.

Funding

The design of the study, experimentation, and interpretation of the data was funded by Ningbo agricultural major projects (201401C1111001), the Open Fund of Zhejiang Provincial Top Key Discipline of Aquaculture in Ningbo University. All the authors were sponsored by K.C.Wong Magna Fund in Ningbo University.

Authors' contributions

Conceived and designed the experiments: WWS, ZMR, CLW. Performed the experiments: ZMR, YZ. Analyzed the data: WWS, ZMR, CLW, YZ, RHL, CKM. Contributed reagents/materials/analysis tools: WWS, ZMR, YZ, CLW, CKM. Wrote the paper: ZMR, WWS. All authors read and approved the final manuscript.

Acknowledgements

The authors are grateful to Yi-Meng Yuan, Meng Han and Hui Dong (Key Laboratory of the Ministry of Education for Applied Marine Biotechnology, Ningbo University) for continuous technical advice and helpful discussion. This manuscript has been thoroughly edited by a native English speaker from an editing company.

Author details

¹Key Laboratory of Applied Marine Biotechnology, Ministry of Education, Ningbo University, Ningbo City, Zhejiang Province, 315832, P.R. China

²Collaborative Innovation Center for Zhejiang Marine High-efficiency and Healthy Aquaculture, Ningbo City, Zhejiang Province, 315832, P.R. China

References

1. Roper CFE, Sweeney MJ, Nauen CE: **Cephalopods of the world**. 2005.
2. Minton JW, Walsh LS, Lee PG, Forsythe JW: **First multi-generation culture of the tropical cuttlefish *Sepia pharaonis* Ehrenberg, 1831**. *Aquaculture International* 2001, **9**(5):379-392.
3. Le KX, Jiang XM, Peng RB, Luo J, Tang F, Wang CL: **Effects of four ecological factors on the growth and survival of *Sepia pharaonis* larvae**. 2014.
4. Dai YT, Xie XH, Huang GG, Liang WF, Shen NN, Zhao HH: **Tolerance of the Cuttlefish *Sepia pharaonis* Larvae to Water Salinities and pH**. *Hebei Fisheries* 2012.

5. Huang JS, Chen G, Zhang JD, Feng LH: **Effects of Salinity on Fertilized Egg Hatch and Survival Activity Index of the Larvae of *Sepia pharaonis***. *Journal of Guangdong Ocean University* 2012.
6. Wen J, Cao GR, Shi-Ying LI, Chen QX, Chen HD, Chen MF, Chen DH: **Effects of Environmental Factors on Survival and Behavior in Juvenile Cuttlefish *Sepia pharaonis***. *Fisheries Science* 2011.
7. Kexin LE, Jiang X, Wang Y, Peng R, Jiang M, Han Q: **Salinity effects on growth and enzyme activity of juvenile**. 2015.
8. De Vos S, Van Stappen G, Sorgeloos P, Vuylsteke M, Rombauts S, Bossier P: **Identification of salt stress response genes using the *Artemia* transcriptome**. *Aquaculture* 2019, **500**:305-314.
9. Ma B, Ran Z, Xu X, Xu J, Liao K, Cao J, Yan X: **Comparative transcriptome analyses provide insights into the adaptation mechanisms to acute salt stresses in juvenile *Sinonovacula constricta***. *Genes & Genomics* 2019, **41**(5):599-612.
10. Mardis ER: **The impact of next-generation sequencing technology on genetics**. *Trends in Genetics* 2008, **24**(3):133-141.
11. Jia Y, Bai J-Q, Liu M-L, Jiang Z-F, Wu Y, Fang M-F, Li Z-H: **Transcriptome analysis of the endangered *Notopterygium incisum*: Cold-tolerance gene discovery and identification of EST-SSR and SNP markers**. *Plant Diversity* 2019, **41**(1):1-6.
12. Lambert IH: **Regulation of the cellular content of the organic osmolyte taurine in mammalian cells**. *Neurochemical Research* 2004, **29**(1):27-63.
13. Xiong L, Schumaker KS, Zhu JK: **Cell Signaling during Cold, Drought, and Salt Stress**. *The Plant cell* 2002, **14**(Suppl):S165.
14. Zhu JK: **Salt and drought stress signal transduction in plants**. *Annu Rev Plant Biol* 2002, **53**:247-273.
15. Huong DT, Yang WJ, Okuno A, Wilder MN: **Changes in free amino acids in the hemolymph of giant freshwater prawn *Macrobrachium rosenbergii* exposed to varying salinities: relationship to osmoregulatory ability**. *Comparative biochemistry and physiology Part A, Molecular & integrative physiology* 2001, **128**(2):317-326.
16. Miller J, McLachlan AD, Klug A: **Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus oocytes***. *The EMBO journal* 1985, **4**(6):1609-1614.
17. Ciftciyilmaz S, Morsy MR, Song L, Coutu A, Krizek BA, Lewis MW, Warren D, Cushman J, Connolly EL, Mittler R: **The EAR-motif of the Cys2/His2-type zinc finger protein Zat7 plays a key role in the defense response of *Arabidopsis* to salinity stress**. *Journal of Biological Chemistry* 2007, **282**(12):9260-9268.
18. Barth O, Vogt S, Uhlemann R, Zschiesche W, Humbeck K: **Stress induced and nuclear localized HIP26 from *Arabidopsis thaliana* interacts via its heavy metal associated domain with the drought stress related zinc finger transcription factor ATHB29**. *Plant molecular biology* 2009, **69**(1-2):213-226.
19. Shaheen M, Williamson E, Nickoloff J, Lee SH, Hromas R: **Metnase/SETMAR: a domesticated primate transposase that enhances DNA repair, replication, and decatenation**. *Genetica* 2010, **138**(5):559.
20. Lee SH, Oshige M, Durant ST, Rasila KK, Williamson EA, Ramsey H, Kwan L, Nickoloff JA, Hromas R: **The SET domain protein Metnase mediates foreign DNA integration and links integration to nonhomologous end-joining repair**. *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102**(50):18075-18080.

21. Hromas R, Williamson EA, Fnu S, Lee YJ, Park SJ, Beck BD, You JS, Leita A, Nickoloff JA, Lee SH: **Chk1 phosphorylation of Metnase enhances DNA repair but inhibits replication fork restart.** *Oncogene* 2012, **31**(38):4245-4254.
22. Fnu S, Williamson EA, De Haro LP, Brennenman M, Wray J, Shaheen M, Radhakrishnan K, Lee SH, Nickoloff JA, Hromas R: **Methylation of histone H3 lysine 36 enhances DNA repair by nonhomologous end-joining.** *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(2):540-545.
23. Golldack D, Luking I, Yang O: **Plant tolerance to drought and salinity: stress regulating transcription factors and their functional significance in the cellular transcriptional network.** *Plant cell reports* 2011, **30**(8):1383-1391.
24. Mangum CP: **Osmoregulation in marine and estuarine animals: Its influence on respiratory gas exchange and transport.** *Bolletino Di Zoologia* 1986, **53**(1):1-7.
25. Tulsani NJ, Hamid R, Jacob F, Umretiya NG, Nandha AK, Tomar RS, Golakiya BA: **Transcriptome landscaping for gene mining and SSR marker development in Coriander (Coriandrum sativum L.).** *Genomics* 2019.
26. Yang X, Liu D, Liu F, Wu J, Zou J, Xiao X, Zhao F, Zhu B: **HTQC: a fast quality control toolkit for Illumina sequencing data.** *BMC bioinformatics* 2013, **14**(1):1-4.
27. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nature methods* 2012, **9**(4):357.
28. Nakasugi K, Crowhurst R, Bally J, Waterhouse P: **Combining Transcriptome Assemblies from Multiple De Novo Assemblers in the Allo-Tetraploid Plant Nicotiana benthamiana.** *PloS one* 2014, **9**(3):e91776.
29. Mortazavi A, Williams BA, Mccue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq.** *Nature methods* 2008, **5**(7):621.
30. Li B, Dewey CN: **RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome.** *BMC bioinformatics* 2011, **12**(1):323.
31. Anders S, Huber W: **Differential expression analysis for sequence count data.** *Genome biology* 2010, **11**(10):R106.
32. Thiel T, Michalek W, Varshney RK, Graner A: **Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (Hordeum vulgare L.).** *Theoretical & Applied Genetics* 2003, **106**(3):411-422.
33. Rozen S, Skaletsky H: **Primer3 on the WWW for general users and for biologist programmers.** *Methods in molecular biology* 2000, **132**(3):365.

Tables

Table 1. Statistics of assembled transcripts.

mple	Raw reads	Raw bases	Clean reads	Clean bases	Valid ratio(base)	Q30(%)	GC content(%)
ntrol	203852818	30577922740	200000000	25000000000	98.11%	94.20%	40.00%
w inity	203852818	30577922740	200000000	25000000000	98.12%	94.53%	41.00%

Table2. Statistics of assembled unigenes.

	All	>=500 bp	>=1000 bp	N50	Total Length	Max Length	Min Length	Average Length
Unigene	130857	62569	23723	970	102686741	39756	301	784.72

Table3. Unigenes annotation against Nr, Swiss Prot, KOG, KEGG and GO databases.

Database	NR	SWISSPROT	KOG	KEGG	GO
Annotation_numbers	24119	16013	13808	5237	12717
Annotation_ratio	18.43%	12.24%	10.55%	4.00%	9.72%

Table 4. Summary of SSR identified from the transcriptome.

MICROSATELLITE SEARCH	Size
Total number of sequences examined	130857
Total size of examined sequences (bp)	102686741
Total number of identified SSRs	101576
Number of SSR containing sequences	58626
Number of sequences containing more than 1 SSR	24737
Number of SSRs present in compound formation	10725

Table 5. Primer sequence

Prime name	Sequence	Product size(bp)
RPL18-F	AAGAGTGGAGGCGAGATTC	157
RPL18-R	CGTAGGGTTTGGTATGACTG	
CTSL-F	GCCATTGATGCCAGCGATTT	113
CTSL-R	GTGCCGTATCCAACAACCAGA	
Orct2-F	GTGCCATTCTGGTTTGA	162
Orct2-R	GCTGCGTTTCATTACCCT	
chac1-F	TATCGTCGTGATGCCAAAG	84
chac1-R	AGCCGTGAAGCAAAGAAC	
CD163L1-F	CTGGGTACAGGCAAGG	197
CD163L1-R	GCATCAACTGGATAGGGA	
trpa-1-F	TTGGCAGCTATGTCAGGT	106
trpa-1-R	AAGCTTTATGCAGTGGTGT	
CYP2U1-F	TACTATCCTTTGGGCACTC	196
CYP2U1-R	CTGACGACCGAGCAATA	
MMP19-F	CACTGGGACGATATTGCA	200
MMP19-R	ATGCTTGCGTAGGGATGA	
SIL1-F	TGCAAAGCAACCCAAAG	186
SIL1-R	AAACGGACAGTCCACCC	
Slc26a5-F	ACACGGTGATAGAGGTCC	115
Slc26a5-R	GACAAGTGGCTGCGTAAT	
ism1-F	GTTGCCCTTGCATTTATC	132
ism1-R	ATTTGGCTCCTGGTTTGT	

Figures

Figure 1

Top-hit species distribution for sequences from *Sepia pharaonis* submitted BLASTX against the NCBI-Nr database.

Figure 4

GO terms for differentially expressed genes in *Sepia pharaonis* due to low salinity. A stand for down-regulated genes and B stand for up-regulated genes.

A

B

Figure 5

KEGG annotation for differentially expressed genes in *Sepia pharaonis* due to low salinity. A stand for down-regulated genes and B stand for up-regulated genes.

Figure 6

qRT-PCR results of differentially expressed genes due to low salinity.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Thearriveguidelinesandchecklist.docx](#)
- [TableS5SummariesofSSR.xls](#)
- [TableS3DifferentiallyexpressedgenesenrichmentbyGO.xlsx](#)
- [TableS1Importantunigenesinthetranscriptome.xlsm](#)
- [TableS2DEGsbetweenTreatmentGroupandControlGroup.xlsx](#)
- [TableS4Differentiallyexpressedgenesenrichmentbykegg.xlsx](#)