1 **Transcriptional profiles of early stage red sea urchins (Mesocentrotus franciscanus) reveal**

2 **differential regulation of gene expression across development**

3 Juliet M. Wong [a,*], Juan D. Gaitán-Espitia [b], and Gretchen E. Hofmann [a]

4

5 [a] Department of Ecology, Evolution and Marine Biology, University of California Santa Barbara,

6 Santa Barbara, CA 93106, USA

7 [b] The Swire Institute of Marine Science, School of Biological Sciences, The University of Hong

8 Kong, Pokfulam Road, Hong Kong, SAR, People's Republic of China

9

10 [*] Corresponding author

11 *Email addresses*: julietmwong@ucsb.edu (J. Wong), jdgaitan@hku.hk (J. Gaitán-Espitia),

12 hofmann@ucsb.edu (G. Hofmann)

13

14   **Abstract**

15   The red sea urchin, Mesocentrotus franciscanus, is an ecologically important kelp forest species

16   that also serves as a valuable fisheries resource. In this study, we have assembled and annotated a

17   developmental transcriptome for M. franciscanus that represents eggs and six stages of early

18   development (8- to 16-cell, morula, hatched blastula, early gastrula, prism and early pluteus).

19   Characterization of the transcriptome revealed distinct patterns of gene expression that

20   corresponded to major developmental and morphological processes. In addition, the period

21   during which maternally-controlled transcription was terminated and the zygotic genome was

22   activated, the maternal-to-zygotic transition (MZT), was found to begin during early cleavage

23   and persist through the hatched blastula stage, an observation that is similar to the timing of the

24   MZT in other sea urchin species. The presented developmental transcriptome will serve as a

25   useful resource for investigating, in both an ecological and fisheries context, how the early

26   developmental stages of this species respond to environmental stressors.

27   **Keywords [4-6 words]**

28   Red sea urchin, *Mesocentrotus franciscanus*, RNA-seq, *De novo* assembly, Early development

29

30   **1. Introduction**

31   The red sea urchin *Mesocentrotus franciscanus* is found along the West Coast of North

32   America, ranging from Baja California, Mexico to Kodiak, Alaska (Ebert, et al. 1999). *M.*

33   *franciscanus* is harvested for its gonads (i.e., roe) and has suffered historical overfishing and

34   exploitation as a high-demand fishery species (Andrew, et al. 2002, Keesing and Hall 1998). The

35   export value of roe in the United States was estimated to be approximately $28.7 million in 2011

36   (Rogers-Bennett 2013), and the red sea urchin fishery remains as one of the top five fisheries in

37   the state of California (Kalvass 2000). Ecologically, *M. franciscanus* acts as an important

38   ecosystem engineer by controlling algae populations, particularly in kelp forest ecosystems, and

39   are capable of transforming algal communities into urchin barrens (Leighton, et al. 1966, Rogers-

40   Bennett 2007). Given the high economic and ecological importance of *M. franciscanus*, genomic

41   resources are very useful for studying and monitoring this species.

42   The introduction of next-generation sequencing and increasing affordability of various

43   technologies has expanded the molecular resources and knowledge available for non-model

44   species, particularly those of high value to fisheries and aquaculture. Annotated and assembled

45   *de novo* transcriptomes have been published across a variety of valuable fisheries and

46   aquaculture species, including mollusks (Coppe, et al. 2012, De Wit and Palumbi 2012, Tian, et

47   al. 2018, Zhao, et al. 2012), crustaceans (Ghaffari, et al. 2014, Lv, et al. 2014, Souza, et al.

48   2018), echinoderms (Gaitán-Espitia, et al. 2016, Gillard, et al. 2014, Jo, et al. 2016), and fishes

49   (Carruthers, et al. 2018, Ji, et al. 2012, Liao, et al. 2013). These transcriptomes are useful for

50   investigating subjects important to fisheries health and management, such as population

51   dynamics, evolutionary processes, the effects of abiotic stress, disease susceptibility and

52   resilience, and stock assessments (Valenzuela-Quiñonez 2016, Wenne, et al. 2007). For example,

53   assembled transcriptomes have been used to investigate salinity stress in the Pacific oyster

54   *Crassostrea gigas* (Zhao, et al. 2012), and viral infection in the Pacific whiteleg shrimp

55   *Litopenaeus vannamei* (Chen, et al. 2013). Transcriptomic data have also been used to

56   investigate patterns of gene flow and local adaptation in the red abalone *Haliotis rufescens* (De

57   Wit and Palumbi 2012). Here, we used RNA sequencing (RNA-seq) to assess the fine regulatory

58   control of gene expression during early development of the economically and ecologically

59   important sea urchin, *M. franciscanus*.

60   *Mesocentrotus franciscanus* is an important organism to study within a climate change

61   context because it is likely to be impacted by altered ocean conditions in the future. Numerous

62   studies have used transcriptomics to investigate how marine organisms respond to changes in

63   their environment that are related to climate change, such as elevated temperatures and lowered

64   pH and oxygen concentrations (Ekblom and Galindo 2011, Franks and Hoffmann 2012, Reusch

65   and Wood 2007, Strader, et al. In revision). Describing the transcriptional dynamics of *M.*

66   *franciscanus* during its early development is particularly pertinent as the early stages of

67   development are believed to be the most vulnerable times during the life history of many marine

68   organisms (Byrne 2011, Dupont and Thorndyke 2009, Gosselin and Qian 1997, Kurihara 2008).

69   A reduction in fitness at the embryological and larval stages leading to poor recruitment could

70   have devastating impacts on marine population dynamics. As rapid environmental change

71   continues, the early life stages may act as a bottleneck that dictates whether a species will be

72   successful in the future (Byrne 2012, Byrne and Przeslawski 2013, Kurihara 2008).

73   A developmental transcriptome is a useful tool for understanding the potential

74   susceptibility of *M. franciscanus* to environmental stress during its early life stages. Several

75   recent studies have reported developmental transcriptomes for marine organisms (Brekhman, et

76   al. 2015, Gildor, et al. 2016, Heyland, et al. 2011, Lenz, et al. 2014, Zeng, et al. 2011), and have

77   described stage-specific expression of many transcription factors. Developmental transcriptomes

78   help to identify which genes are important for development as well as the timing of expression of

79   these genes. These genomic resources offer the opportunity to unveil mechanisms underlying

80   developmental plasticity and its role buffering different abiotic stressors across ontogeny.

81   We collected eggs from *M. franciscanus* as well as embryos and larvae from six stages of

82   development: 8- to 16-cell, morula (composed of approximately 64 cells), hatched blastula, early

83   gastrula, prism, and early pluteus. One group of processes that occur during development

84   involves the change of control from the maternal to zygotic genome, identified as the maternal-

85   to-zygotic transition (MZT), during which time there is a shift in expression from maternal to

86   zygotic transcripts (Shier 2007, Tadros and Lipshitz 2009). ~~The timing of the MZT has not yet~~

87   ~~been determined for *M. franciscanus*.~~ The developmental transcriptome presented here will

88   provide insight into the timing of the MZT and will help identify genes and regulatory pathways

89   that are important for successful development at each stage. Overall, this will be a useful

90   resource for transcriptomic analyses of this species, including for studies that use gene

91   expression to assess the response of early developmental stages to changing environmental

92   conditions.

93

94   **2. Materials and Methods**

95   *2.1. Animal collection and culturing*

96   Adult sea urchins were collected in May 2016 at Mohawk Reef, CA, USA (34° 23.606'

97   N, 199° 43.807' W) using CA Scientific Collection permit XX. Urchins were immediately

98   transported to the Marine Science Institute and the University of California, Santa Barbara

99   (UCSB) (Santa Barbara, CA), and maintained flow-through seawater tanks for approximately

100   one week prior to spawning. Spawning was induced via intracoelomic injection of 1.0 M KCl.

101   Egg samples (EG) were collected by gently transferring ~5,000 eggs into a 1.5 mL Eppendorf

102   tube, quickly pelleting the sample by centrifugation, removing the excess seawater, and flash

103   freezing the sample using liquid nitrogen. All samples were stored at -80 °C. Test fertilizations

104   were performed to verify egg-sperm compatibility. The eggs from 2 females were gently pooled

105   together and were fertilized from sperm from a single male. To avoid polyspermy, dilute sperm

106 was slowly added to the eggs until at least 95% fertilization success was reached. The newly

107 fertilized embryos were then placed into each of three replicate culture vessels at a concentration

108 of ~9 embryos per mL of seawater.

109   All *M. franciscanus* EDS cultures were raised in 0.35 μm filtered, UV-sterilized seawater

110 (FSW). The EDS cultures were raised at ~15 °C and ~425 μatm pCO_2. These conditions were

111 chosen to represent average, "normal" conditions that populations of *M. fransciscanus* have been

112 observed to experience *in situ* near Mohawk reef (Hofmann and Washburn 2015). Water

113 temperature was controlled using a Delta Star® heat pump with a Nema 4x digital temperature

114 controller (AquaLogic, San Diego, CA, USA), which maintained culturing temperatures at ~15

115 °C. A flow-through CO_2-mixing system modified from Fangue, et al. (2010) was used to ensure

116 stable carbonate chemistry conditions throughout development. The CO_2 system was used to

117 establish a 5-gallon reservoir tank, in which water was treated to the target pCO_2 level prior to

118 transporting the treated water to each culture vessel.

119   Each culture vessel was composed of two, nested 5-gallon buckets (12 L capacity). The

120 inner bucket has a dozen holes 5.5 cm in diameter, each fitted with 64-micron mesh to prevent

121 the loss of embryos or larvae while allowing for a flow-through of seawater. Seawater flow to

122 each vessel was controlled using irrigation button drippers (DIG Corporation), which regulated

123 the flow to a rate of 4 L/hr. Each vessel contained a 15 cm x 15 cm plastic paddle driven by a 12-

124 V motor to allow for continuous, gentle mixing and to prevent embryos from settling to the

125 bottom of the bucket.

126   Embryos and larvae were sampled at six developmental stages: 8- to 16- cell (CL; ~4

127 hours post-fertilization (hpf)), morula (MA; ~7 hpf), hatched blastula (BL; ~16 hpf), early

128 gastrula (GA; ~29 hpf), prism (PR; ~44 hpf), and early pluteus (PL; ~64 hpf). While the

129 development of *M. franciscanus* in culture is generally synchronous, during early cellular

130 divisions, it is unlikely to collect a large batch of embryos that are exhibiting identical timing. As

131 such, the samples collected at the 8- to 16-cell stage (CL) were composed of a mixture of

132 embryos undergoing their third and fourth cleavage divisions. At the morula stage (MA), the

133 embryos were composed of masses approximately 64 or more cells. The blastula stage (BL) was

134 designated by the enzymatic digestion of the fertilization envelope and emergence of swimming

135 blastula. The early gastrula (GA) stage was designated by the formation of mesenchyme cells

136 and an archenteron extended to approximately one-half the body length. The prism stage (PR)

137 was identified by the formation of the pyramid-like prism shape, the archenteron becoming

138 tripartite, and the early development of skeletal rods. Lastly, the early pluteus state (PL) was

139 defined as having internal structures, including the mouth, esophagus, stomach and anus, as well

140 as anterolateral and postoral skeletal body rods and the early formation of feeding arms. Two

141 replicate samples from each of the three culture vessels were taken at each developmental stage.

142 All samples were preserved using the same methods for preserving the eggs.

143   Temperature, salinity, pH, and total alkalinity (TA) was recorded daily to monitor the

144 culturing conditions throughout development. Temperature was measured using a wire

145 thermocouple (Thermolyne PM 20700 / Series 1218), and salinity was measured using a

146 conductivity meter (YSI 3100). Daily pH measurements were conducted by following the

147 standard operating procedure (SOP) 6b (Dickson, et al. 2007), using a spectrophotometer (Bio

148 Spec-1601, Shimadzu) and *m*-cresol purple (Sigma-Aldrich) indicator dye. Water samples for

149 TA were poisoned with saturated 0.02% mercuric chloride. TA was estimated using SOP 3b

150 (Dickson, et al. 2007). Using the carbonic acid dissociation constants from Mehrbach, et al.

151   (1973) refit by Dickson and Millero (1987), parameters of pCO$_2$, Ω_{ara}, and Ω_{cal} were calculated

152   using CO$_2$calc (Robbins, et al. 2010).

153

154   *2.2. RNA extractions and sequencing*

155   Total RNA was extracted using 500 µL of Trizol® reagent, following the manufacturer's

156   instructions (Invitrogen). Briefly, each sample was homogenized in Trizol® reagent by passing

157   the sample three times through decreasing sizes of needles (21-guage, 23-gauge, and then 25-

158   gauge). A chloroform addition and centrifugation were used to isolate the RNA-containing upper

159   aqueous phase. The RNA was precipitated in isopropyl alcohol, washed using ethanol, and

160   resuspended in DEPC-treated water. RNA purity, quantity, and quality were verified using a

161   NanoDrop® ND100, a Qubit® fluorometer, and a Tapestation 2200 system (Agilent technologies).

162   Samples of the highest quality were selected for library preparation and sequencing.

163   Three libraries were generated from triplicate samples of eggs. For each developmental

164   stage, one library was generated for each of the three replicate culture vessels. This resulted in a

165   total of 21 libraries. Libraries were generated using high quality total RNA (RIN values > 9.1)

166   using a TruSeq Stranded mRNA Library Preparation Kit (Illumina) following the manufacturer's

167   instructions. The quantity and quality of each library was verified using a Qubit® fluorometer and

168   a Tapestation 2200 system (Agilent). The libraries were submitted to the Genome Center at the

169   University of California, Davis for sequencing on an Illumina HiSeq 4000 sequencer on two

170   lanes with 150 base-pair (bp) paired-end reads.

171

172   *2.3.* De novo *transcriptome assembly*

173   Additional *M. franciscanus* raw sequence data from Gaitán-Espitia and Hofmann (2017)

174   were included with sequence data from our 21 libraries to generate the *de novo* transcriptome.

175   These data represented gastrula stage embryos (GenBank accession numbers SRS823202 and

176   SRS823216) and pluteus larvae (accession numbers SRS823218 and SRS82322) of bioproject

177   PRJNA272924. Any potential adapter sequence contamination as well as any base pairs with

178   quality scores below 30 were removed from all raw sequence data using Trim Galore! (version

179   0.4.1) (Krueger 2015). Sequence quality was verified using FastQC (version 0.11.5) (Andrews

180   2010).

181   The transcriptome was assembled following a pipeline available from the National Center

182   for Genome Analysis Support (NCGAS) at Indiana University (https://github.com/NCGAS/de-

183   novo-transcriptome-assembly-pipeline). This workflow generates a combined *de novo* assembly

184   that uses multiple assemblers with multiple parameters. Prior to assembly, the data was

185   normalized using the in silico read normalization function in Trinity (version 2.6.6) (Grabherr, et

186   al. 2011). Multiple *de novo* assemblies were created using Trinity (version 2.6.6) (kmer = 25),

187   SOAPdenovo-Trans (version 1.03) (Xie, et al. 2014) (kmers = 35, 45, 55, 65, 75, and 85), Velvet

188   (version 1.2.10) (Zerbino and Birney 2008) and Oases (version 0.2.09) (Schulz, et al. 2012)

189   (kmers = 35, 45, 55, 65, 75, and 85), and Trans-ABySS (version 2.0.1) (Robertson, et al. 2010)

190   (kmers = 35, 45, 55, 65, 75, and 85). These 19 transcriptomes were then combined using

191   EvidentialGene (version 2013.07.27) (Gilbert 2013), which removes perfect redundancy and

192   fragments to reduce false transcripts while predicting unique transcripts within the final

193   assembly. Quast (version 5.0.0) (Gurevich, et al. 2013) was used to generate basic quality

194   metrics of the final assembly. BUSCO (version 3.0.2) (Simão, et al. 2015) was used to assess

195   completeness of the final assembly using the single-copy ortholog reference for metazoa.

196

197 *2.4 Gene prediction and functional annotation*

198   Gene models from the do novo assembled transcriptome were inferred and annotated

199 using the BLASTP (against the nr database), BLASTN (against he eukaryotic nt database) and

200 BLASTX (against the Uniprot database, Swiss-Prot and TrEMBL) algorithms with an e-value

201 cutoff of 1e-5. Annotated sequences were further searched for Gene Ontology (GO) terms using

202 Blast2GO software (www.blast2go.com; version 5.2.5)(Conesa et al., 2005) according to the

203 main categories of Gene Ontology (GO; molecular functions, biological processes and cellular

204 components) (Ashburner et al., 2000). Complementary annotations were done with the

205 InterProScan v.5 software (Jones et al., 2014. Finally, the annotation results were further fine-

206 tuned with the Annex and GO slim functions and the enzyme code annotation tool of the Kyoto

207 Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto 2000) implemented in

208 Blast2GO.

209

210 *2.5 Expression quantitation and differential expression analyses*

211   Trimmed sequence data from the 21 libraries were mapped onto the *de novo* reference

212 transcriptome and expression values were calculated using RSEM (version 1.3.0) (Li and Dewey

213 2011) and bowtie2 (version 2.3.2) (Langmead and Salzberg 2012). Using the LIMMA package

214 (Ritchie, et al. 2015) in R (version 3.4.4), the data were filtered to sequences that have more than

215 0.5 counts per million mapped reads across at least three of the 21 samples. A trimmed mean of

216 M-values (TMM) normalization method (Robinson and Oshlack 2010) was used to apply scale

217 normalization to the read counts. The data were voom-transformed using LIMMA to convert the

218 read counts to long-counts per million while accounting for sample-specific quality weights and

219   blocking design (i.e., biological replicates). The filtered, normalized and voom-transformed data

220   were used to perform a principal component analysis (PCA) using the prcomp function in R.

221   Hierarchical clustering on the principal components was performed using the HCPC function of

222   the FactoMineR package (Le, et al. 2008) in R.

223   Using the WGCNA package (Langfelder and Horvath 2008) in R, a Weighted Gene Co-

224   Expression Network Analysis (WGCNA) was performed on the same filtered, normalized and

225   voom-transformed data to identify clusters of similarly expressed genes into modules, in which

226   each module contained a minimum of 30 genes. Modules with highly correlated eigengenes were

227   merged using a threshold of 0.27 (i.e., a height cut-off of 0.27 and a correlation of 0.73 for

228   merging). Eigengene expression was correlated with each developmental stage (i.e., EG, CL,

229   MA, BL, GA, PR, and PL), and a heatmap was generated to visualize significant correlations

230   between each stage and module.

231   Functional enrichment analyses were performed on lists of genes within modules with

232   significant correlations to a developmental stage (r^2 correlation $\geq$ 0.50 and p-value $\leq$ 0.05).

233   Enrichment analyses were performed in Blast2GO (version 5.2.5) using a Fisher's Exact Text

234   with an FDR filter value of 0.05 to identify gene ontology (GO) terms within the GO categories:

235   biological process, molecular function, and cellular component.

236

237   *2.6. Survey of the maternal-to-zygotic transition*

238   To investigate the timing of the maternal-to-zygotic transition (MZT), two aspects of the

239   data were probed: 1) the loss of maternally-derived transcripts, and 2) the initiation of zygotic

240   transcription. To examine the loss of maternally-derived transcripts, putative genes related to the

241   removal of maternal RNAs were identified. These included putative genes for the

Table 1. Transcripts that play a functional role during the MZT.

242   *microprocessor complex subunit DGCR8* (*dgcr8*), *endoribonuclease dicer* (*dicer*), *smaug*

243   (*smaug1*), and *nonsense mediated mRNA decay* (*smg7*, *smg8*, and *smg9*) (Table 1). The

244   expression levels of these transcripts were examined across the eggs and early development.

245   Additionally, a heatmap was constructed to visualize the expression of maternal transcripts

246   throughout development. Transcripts expressed in the eggs were considered to be maternal. The

247   heatmap was constructed using the top 500 transcripts expressed during the egg stage (EG).

Transcript ID	Name	Description
058424	*dgcr8*	microprocessor complex subunit DGCR8-like
103747	*dicer*	endoribonuclease Dicer
035070	*smaug*	protein Smaug homolog 1 isoform X1
089348	*smg7*	protein SMG7 isoform X2
052093	*smg8*	protein smg8
035665	*smg9*	protein SMG9-like
089934	*alx*	aristaless-like homeobox protein
082528	*bra*	transcription factor Brachyury
015822	*dri*	protein dead ringer homolog
036581	*gcm*	glial cells missing transcription factor
035045	*gsc*	homeobox protein goosecoid-like
047087	*hox11/13b*	transcription factor Hox11/13b
104463	*lefty2*	left-right determination factor 2-like
047266	*nodal*	nodal homolog 2-A-like
076942	*wnt8*	Wnt8

248

249   To examine the initiation of zygotic transcription, transcripts associated with zygotic

250   development were targeted. Putative genes for *aristaless-like homeobox* (*alx*), *brachyury* (*bra*),

251   *dead ringer* (*dri*), *glial cells missing* (*gcm*), *goosecoid* (*gsc*), *homeobox 11/13b (hox11/13b)*, *left-*

252   *right determination factor 2* (*lefty2*), *nodal*, and *wnt8* were targeted within *M. franciscanus*

253   (Table 1). The expression levels of these transcripts were examined across the eggs and early

254   development.

255

3. Results and Discussion

3.1 Embryological and larval development conditions

The EDS cultures developed normally and were highly synchronistic. Little to no mortality was observed throughout development. Temperature and seawater chemistry conditions were stable throughout the ~64 hour culturing period. Across all replicate culture vessels, the temperature was 15.3 ± 0.1 °C, the salinity was 33.4 ± 0.04, the pH was 8.00 ± 0.03, the pCO_2 level was 438 ± 33.1 µatm, and the TA was 2228.97 ± 3.19 µmol kg^{-1}.

3.2 Summary statistics of the transcriptome assembly and annotation

Sequencing of the 21 libraries yielded a total of 751,578,474 150-bp paired-end reads. After trimming to remove any adapter contamination or low quality reads, an average of 35.4 ± 5.8 million reads remained per library. FastQC reports of the trimmed reads from all libraries showed high sequence quality (scores >30) with limited adapter contamination or presence of overrepresented sequences. The transcriptome generated by the NCGAS pipeline was 96.74 megabases (Mb) with 115,719 contigs, a N50 of 3,292 bp, and a GC content of 42.58% (Table 2). The BUSCO analysis which used metazoan as the single-copy ortholog reference showed high transcriptome completeness with a

Table 2. Statistics of *de novo* transcriptome assembly

Assembly statistic	Value
No. contigs	115,719
No. contigs > 1kb	19,511
Assembly size (Mb)	96.74
Mean contig length (bp)	836
Median contig length (bp)	341
Max contig length (bp)	132,566
GC content	42.58
N_{50} (bp)	3,292
L_{50} (bp)	6,526
BUSCO completeness (%)	92.2
BUSCO fragmented (%)	0.8
BUSCO missing (%)	7.0

279   complete BUSCO score of 92.2% (Table 2). Therefore, this transcriptome should offer a suitable

280   foundation for transcriptomic analyses of *M. franciscanus*.

281   The gene discovery and functional annotation analyses identified 35,632 contigs that

282   blasted to known proteins in the public databases. From these, 24,900 contigs were linked to GO

283   classifications. Hypothetical or predicted proteins in these databases were excluded by discarding

284   matches associated to "hypothetical", "predicted", "unknown" and "putative" categories. Over

285   95% of the annotated contigs hit against the genomes of the purple sea urchin *Strongylocentrotus*

286   *pupuratus*, followed by the sea star Acanthaster planci and the sea cucumber *Apostichopus*

287   *japonicus*. The functional annotation analysis retrieved 48,990 GO terms, with 23,053 linked to

288   molecular function (mainly protein binding), 15,754 biological process (mainly G protein-

289   coupled receptor signaling pathway, oxidation-reduction process and transmembrane transport),

290   and 10,183 to cellular component (mainly integral component of membrane). Finally, the

291   enzyme code annotation with KEGG mapping identified 1,948 transcripts, which represented

292   433 enzymes in 122 unique pathways (Table 3, S1). KEGG pathways included those related to

293   purine metabolism, biosynthesis of antibiotics, T cell receptor signaling pathway, Th1 and Th2

294   cell differentiation, and ether lipid metabolism. The complete list of KEGG pathways are

295   available in Supplementary file XX.

296

Table 3. Top 10 KEGG pathways in the transcriptome.

Pathway	Pathway ID	No. transcripts	No. enzymes
Purine metabolism	map00230	825	47
Thiamine metabolism	map00730	727	6
Drug metabolism - other enzymes	map00983	233	16
Biosynthesis of antibiotics	map01130	185	100
T cell receptor signaling pathway	map04660	132	2
Th1 and Th2 cell differentiation	map04658	129	1
Glutathione metabolism	map00480	76	13

|			
Ether lipid metabolism	map00565	69	7
Cysteine and methionine metabolism	map00270	60	23
Sphingolipid metabolism	map00600	58	14

297

298   *3.3 Gene expression patterns pre-fertilization and throughout early development*

299   After the count matrix data generated from RSEM was filtered to those with more than

300   0.5 counts per million mapped reads across at least three of the 21 samples, 35,126 sequences

301   remained. A PCA was used to examine sample-to-sample distances (Fig. 1). Triplicate samples

302   clustered the closest together within each of their representative stages. The first and second

303   dimensions captured 62.94% and 13.09% of the variation, respectively, and revealed a clear

304   separation between early egg/embryonic stages and later developmental stages. Hierarchical

Fig. 1. PCA of *Mesocentrotus franciscanus* eggs and early developmental stages showing the first two
dimensions and hierarchical clustering of the samples. Sample colors denote the different stages, which include:
egg (EG), 8- to 16-cell (CL), morula (MO), blastula (BL), gastrula (GA), prism (PR), and pluteus (PL).
Hierarchical clustering show two main clusters (1 and 2), which each contain two clusters (a and b).

305   clustering revealed the two primary clusters (Fig. 1). Cluster 1 included cluster 1a, which

306   contained eggs (EG), and cluster 1b, which included the 8- to 16-cell (CL) and morula (MO)

307   stages. Cluster 2 included cluster 2a, which contained blastula (BL) and gastrula (GA) stages,

308   and cluster 2b, which contained prism (PR) and pluteus stages (PL). The clear demarcation

309   between stages is represented by major alterations in development processes and morphology.
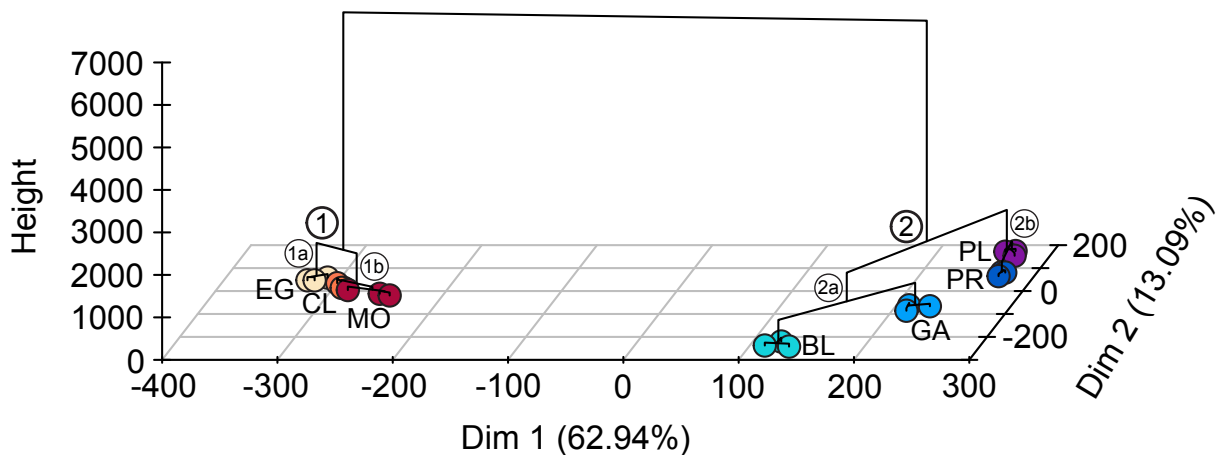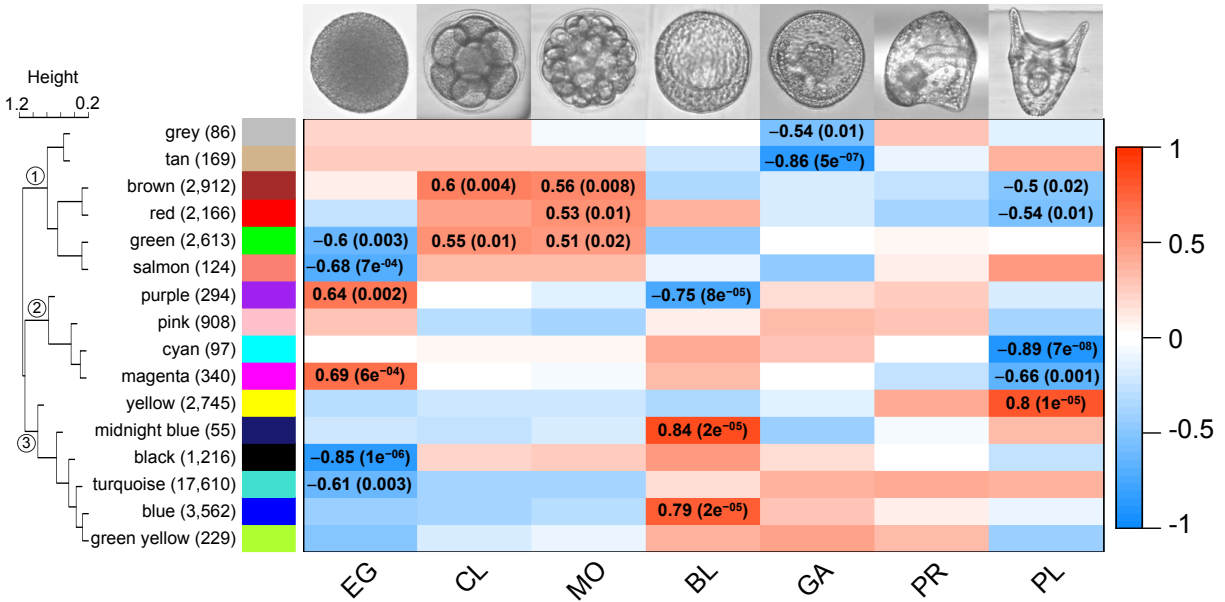

310

311

312   WGCNA was used to highlight groups of genes that were co-expressed within eggs and

313   each developmental stage. After filtering, normalizing and voom-transforming the data, the

314   remaining 35,126 genes were assigned into module eigengenes containing similarly expressed

315   genes. Only 86 genes remained unclustered and unassigned and were grouped into the grey

316   "module" (Fig. 2). All other genes were assigned into 15 different modules that were designated

317   by color. Hierarchical clustering of the module eigengenes revealed three main clusters (Fig. 2).

318   Each module was related to each stage to generate eigengene networks with positive or negative

319   correlation values ranging from 1 to -1 (Fig. 2). Of the 15 module eigengenes, pink (908 genes)

320   and green yellow (229) were not significantly correlated to any stage (r^2 correlation $\leq$ 0.50, *p*-

321   value $\geq$ 0.05). Functional enrichment analyses did not identify any GO terms within the module

322   eigengenes purple (294 genes), cyan (97 genes), magenta (340 genes), or midnight blue (55

323   genes). There were nine remaining module eigengenes that were significantly correlated to at

324   least one stage (r^2 correlation $\geq$ 0.50 and *p*-value $\leq$ 0.05) and in which functional enrichment

325   analyses successfully identified GO terms. These module eigengenes were tan (169 genes),

326   brown (2,912 genes), red (2,166 genes), green (2,613 genes), salmon (124 genes), yellow (2,745

327   genes), black (1,216 genes), turquoise (17,610 genes), and blue (3,562).

328

Fig. 2. WGCNA identified significant correlations between module eigengenes (rows) and stages (columns). The number of genes within each module eigengene is noted in parenthesis following each color name. The red-blue color scale represents the strength of the correlation (1 to -1). Each correlation value (r^2) is followed by a *p*-value in parenthesis. Hierarchical clustering of the module eigengenes revealed three primary clusters of gene expression (1-3).

	EG	CL	MO	BL	GA	PR	PL
grey (86)					−0.54 (0.01)		
tan (169)					−0.86 (5e⁻⁰⁷)		
brown (2,912)		0.6 (0.004)	0.56 (0.008)				−0.5 (0.02)
red (2,166)			0.53 (0.01)				−0.54 (0.01)
green (2,613)	−0.6 (0.003)	0.55 (0.01)	0.51 (0.02)				
salmon (124)	−0.68 (7e⁻⁰⁴)						
purple (294)	0.64 (0.002)			−0.75 (8e⁻⁰⁵)			
pink (908)							
cyan (97)							−0.89 (7e⁻⁰⁸)
magenta (340)	0.69 (6e⁻⁰⁴)						−0.66 (0.001)
yellow (2,745)							0.8 (1e⁻⁰⁵)
midnight blue (55)				0.84 (2e⁻⁰⁵)			
black (1,216)	−0.85 (1e⁻⁰⁶)						
turquoise (17,610)	−0.61 (0.003)						
blue (3,562)				0.79 (2e⁻⁰⁵)			
green yellow (229)							

Eggs (EG) were significantly correlated to six module eigengenes, the greatest number of any other stage. Eggs had a significant positive correlation with module eigengenes purple (r^2 = 0.64, *p*-value = 0.002) and magenta (r^2 = 0.69, *p*-value = 0.0006), which were both clustered together (module eigengene cluster 2, Fig. 2), but functional enrichment analyses failed to reveal any GO terms within these modules. Eggs had a significant negative correlation with module eigengenes green (r^2 = -0.6, *p*-value = 0.003), salmon (r^2 = -0.68, *p*-value = 0.0007), black (r^2 = -0.85, *p*-value = 0.000001), and turquoise (r^2 = -0.61, *p*-value = 0.003). Green and salmon clustered together within module eigengene cluster 1 (Fig. 2), and functional enrichment analyses identified 695 and 18 GO terms, respectively, within biological process, molecular function, and cellular component categories (Tables 4, S2). GO terms in green included those related to ATP binding, integral component of membrane, regulation of RNA metabolic process, and oxidation-reduction process (Tables 4, S2). Ontologies in salmon were related to protein binding, metabolic process, methyltransferase activity, and phospholipid binding (Tables 4, S2).

344   Module eigengenes black and turquoise clustered together within module eigengene cluster 3

345   (Fig. 2), and functional enrichment analyses revealed 111 and 1596 GO terms within each

346   module, respectively (Tables 4, S2). The black module included ontologies related to protein

347   binding, DNA binding, and RNA processing, while the turquoise module included those related

348   to ATP binding, DNA integration, calcium ion binding, and RNA-dependent DNA polymerase

349   activity (Tables 4, S2). Overall, the egg transcriptome was characterized by a down-regulation of

350   genes related to metabolic processes and catalytic activity relative to the measured

351   developmental stages post-fertilization.

352   Following fertilization, the embryo undergoes radial, holoblastic cleavage (Strathmann

353   1987). Here, these early cell divisions are represented by the 8- to 16-cell stage (CL) and the

354   morula stage (MO), whose gene expression patterns are highly similar to one another (Fig. 1).

355   CL was positively correlated to two module eigengenes, brown ($r^2 = 0.6$, p-value $= 0.004$) and

356   green ($r^2 = 0.55$, p-value $= 0.01$), both of which were clustered together (module eigengene

357   cluster 1, Fig. 2). Enrichment analyses identified 309 GO terms within module eigengene brown,

358   which included those related to integral component of membrane, ATP binding, protein kinase

359   activity, transmembrane transport, and catalytic activity acting on RNA (Tables 4, S2).

360   Ontologies for module eigengene green was previously described for the egg stage (EG), except

361   green was positively, rather than negatively, correlated with the CL stage. Similar to the 8- to 16-

362   cell stage, the morula stage (MO) was positively correlated with module eigengenes brown ($r^2 =$

363   0.56, p-value $= 0.008$) and green ($r^2 = 0.51$, p-value $= 0.02$). The MO stage, however, also had a

364   significant correlation with module eigengene red ($r^2 = 0.53$, p-value $= 0.01$), which was also

365   within module eigengene cluster 1 (Fig. 2). Module eigengene red contained GO terms related to

366   protein binding, integral component of membrane, ATP binding, mRNA splicing, and regulation

367   of transcription (Tables 4, S2), which were highly similar to those found in module terms brown

368   and green. During these early cell divisions, the embryos are enriched with genes related to

369   metabolic processes, catalytic activity, and organelle and membrane formation, which likely

370   reflect various processes involved in cell proliferation.

371   Processes vital for blastula formation, gastrulation, organ construction, and

372   skeletogenesis all involve cell differentiation. In developing into a blastula, the embryo forms a

373   blastocoel, cilia, and enzymes required to digest the fertilization membrane during the hatching

374   process (Barrett and Edwards 1976, Lepage and Gache 1989). Here, the blastula stage (BL) was

375   negatively correlated with module eigengene purple (r^2 = -0.75, p-value = 0.00008) and

376   positively correlated with module eigengenes midnight blue (r^2 = 0.84, p-value = 0.00005) and

377   blue (r^2 = 0.79, p-value = 0.00005). Module eigengenes midnight blue and blue were clustered

378   together (module eigengene cluster 3, Fig. 2). While functional enrichment analyses failed to

379   reveal any GO terms within module eigengenes purple or midnight blue, module eigengene blue

380   contained 146 identified GO terms. GO terms within blue were related to RNA-directed DNA

381   polymerase activity, protein binding, DNA integration, G protein-coupled receptor activity, and

382   transmembrane transport (Tables 4, S2). The enrichment of these genes are in alignment with

383   other studies in which genes related to DNA replication and energy production are expressed

384   during zygotic development (Gildor, et al. 2016, Tadros and Lipshitz 2009).

385   Gastrulation is a major and fundamental process of metazoan development (Wolpert

386   1992) that begins by invagination at the vegetal plate and the formation of the archenteron (Dan

387   and Okazaki 1956, Ettensohn 1984). Somewhat surprisingly, there were few correlations

388   between the gastrula stage (GA) and module eigengenes identified by WGCNA. GA was not

389   positively correlated with any module eigengenes, and was negatively correlated with only the

390   module eigengene tan (r^2 = -0.86, *p*-value = 5e^{-7}) (module eigengene cluster 1, Fig. 2). GA was

391   also negatively correlated with the grey "module" (r^2 = -0.54, *p*-value = 0.01), which contained

392   the unclustered and unassigned genes. Functional enrichment analysis of module eigengene tan

393   revealed 39 GO terms, including oxidation-reduction process, integral component of membrane,

394   ion transmembrane transport, and ATPase activity Tables 4, S2.

395   The digestive tract and supporting skeletal rods are formed during the prism and early

396   pluteus stages, which are necessary for the planktotrophic feeding strategy of the urchin larvae

397   (Burke 1980, Ettensohn and Malinda 1993). The prism stage (PR) was not significantly

398   correlated to any module eigengene. However, its expression patterns were similar to that of the

399   pluteus stage (PL) (Figs. 1, 2). PL was negatively correlated to module eigengenes brown (r^2 = -

400   0.5, *p*-value = 0.02) and red (r^2 = -0.54, *p*-value = 0.01). As previously mentioned in the

401   description of hatched blastula module results, GO terms within brown and red modules include

402   those related to protein binding, integral component of membrane, and ATP binding, and

403   transmembrane transport (Tables 4, S2). PL was also negatively correlated with module

404   eigengenes cyan (r^2 = -0.89, *p*-value = 7e^{-8}) and magenta (r^2 = -0.66, *p*-value = 0.001), although

405   functional enrichment analyses were unable to identify GO terms within these modules. Lastly,

406   PL was positively correlated with one module eigenegene, yellow (r^2 = 0.8, *p*-value = 0.00001),

407   which was within module eigengene cluster 3 (Fig. 2). Enrichment analysis identified 368 GO

408   terms within yellow, which included G protein-coupled receptor signaling pathway, oxidation-

409   reduction process, calcium ion binding, proteolysis, acetylcholine-gated cation-selective channel

410   activity, and ion transmembrane transport (Tables 4, S2). This expression pattern likely reflects

411   the energy production and biomineralization processes necessary to support gut and skeletal

412   formation in the developing larvae.

Table 4. Select GO term results from functional enrichment analyses of WGCNA module eigengenes

| GO ID	GO term name	GO category	FDR value	No. transcripts (% of ref)
tan: GA (r2 = -0.86)				
GO:0016021	integral component of membrane	Cellular Component	6.67E-06	26 (0.6)
GO:0055114	oxidation-reduction process	Biological Process	2.27E-05	12 (1.2)
GO:0016491	oxidoreductase activity	Molecular Function	2.82E-04	11 (1.1)
GO:0034220	ion transmembrane transport	Biological Process	3.13E-03	6 (2)
GO:0015267	channel activity	Molecular Function	5.97E-03	6 (1.7)
GO:0042623	ATPase activity, coupled	Molecular Function	5.97E-03	5 (2.5)
brown: CL (r2 = 0.6), MO (r2 = 0.56), PL (r2 = -0.5)				
GO:0005515	protein binding	Molecular Function	4.71E-25	253 (5.6)
GO:0016021	integral component of membrane	Cellular Component	4.81E-12	207 (4.6)
GO:0005524	ATP binding	Molecular Function	1.63E-10	79 (6.5)
GO:0004672	protein kinase activity	Molecular Function	4.66E-06	36 (7.5)
GO:0055085	transmembrane transport	Biological Process	1.12E-05	57 (5.7)
GO:0140098	catalytic activity, acting on RNA	Molecular Function	3.75E-03	20 (7.2)
red: MO (r2 = 0.53), PL (r2 = -0.54)				
GO:0005515	protein binding	Molecular Function	2.56E-18	188 (4.1)
GO:0016021	integral component of membrane	Cellular Component	9.33E-14	172 (3.8)
GO:0005524	ATP binding	Molecular Function	1.18E-11	68 (5.6)
GO:0000398	mRNA splicing, via spliceosome	Biological Process	1.29E-07	13 (21.7)
GO:0006355	regulation of transcription, DNA-templated	Biological Process	2.01E-05	34 (5.4)
GO:0055114	oxidation-reduction process	Biological Process	8.93E-04	39 (4.1)
green: EG (r2 = -0.6), CL (r2 = 0.55), MO (r2 = 0.02)				
GO:0005524	ATP binding	Molecular Function	2.79E-34	117 (10)
GO:0005096	GTPase activator activity	Molecular Function	3.59E-11	15 (39.5)
GO:0140098	catalytic activity, acting on RNA	Molecular Function	1.48E-10	31 (11.7)
GO:0016021	integral component of membrane	Cellular Component	4.84E-09	177 (3.9)
GO:0055114	oxidation-reduction process	Biological Process	5.06E-07	53 (5.7)
GO:0016073	snRNA metabolic process	Biological Process	6.98E-07	8 (57.1)
salmon: EG (r2 = -0.68)				
GO:0003824	catalytic activity	Molecular Function	1.81E-07	34 (0.4)
GO:0005515	protein binding	Molecular Function	9.29E-04	20 (0.4)
GO:0008152	metabolic process	Biological Process	5.40E-03	26 (0.3)
GO:0008168	methyltransferase activity	Molecular Function	5.40E-03	6 (1.7)
GO:0005543	phospholipid binding	Molecular Function	5.40E-03	4 (4.7)
GO:0016020	membrane	Cellular Component	9.21E-03	21 (0.3)
yellow: PL (r2 = 0.8)				
GO:0007186	G protein-coupled receptor signaling pathway	Biological Process	3.62E-30	100 (11)
GO:0055114	oxidation-reduction process	Biological Process	8.55E-23	87 (9.7)
GO:0005509	calcium ion binding	Molecular Function	1.44E-16	77 (8.3)
GO:0006508	proteolysis	Biological Process	3.35E-14	70 (7.9)
GO:0022848	acetylcholine-gated cation-selective channel activity	Molecular Function	1.78E-07	9 (60)
GO:0034220	ion transmembrane transport	Biological Process	4.70E-07	27 (9.8)

black: EG (r2 = -0.85)

GO:0005515	protein binding	Molecular Function	1.18E-11	111 (2.4)
GO:0003677	DNA binding	Molecular Function	8.07E-07	39 (3.3)
GO:0006464	cellular protein modification process	Biological Process	3.49E-05	38 (2.8)
GO:0046872	metal ion binding	Molecular Function	4.01E-05	69 (2.1)
GO:0006396	RNA processing	Biological Process	1.46E-04	16 (4.9)
GO:0016021	integral component of membrane	Cellular Component	1.59E-04	86 (1.8)

turquoise: EG (r2 = -0.61)

GO:0005524	ATP binding	Molecular Function	1.50E-92	499 (63.2)
GO:0015074	DNA integration	Biological Process	5.19E-40	271 (53.3)
GO:0005509	calcium ion binding	Molecular Function	3.64E-36	315 (45.5)
GO:0003964	RNA-directed DNA polymerase activity	Molecular Function	3.09E-33	415 (38)
GO:0006278	RNA-dependent DNA biosynthetic process	Biological Process	4.22E-33	415 (37.9)
GO:0005525	GTP binding	Molecular Function	2.54E-19	136 (51.7)

blue: BL (r2 = 0.79)

GO:0016021	integral component of membrane	Cellular Component	7.53E-28	296 (6.7)
GO:0003964	RNA-directed DNA polymerase activity	Molecular Function	5.84E-24	132 (9.6)
GO:0005515	protein binding	Molecular Function	5.01E-21	277 (6.1)
GO:0015074	DNA integration	Biological Process	1.28E-10	65 (9.1)
GO:0004930	G protein-coupled receptor activity	Molecular Function	1.24E-08	70 (7.8)
GO:0055085	transmembrane transport	Biological Process	4.85E-06	67 (6.7)

414

3.4. The maternal-to-zygotic transition

To examine the timing of the MZT, the decline of maternally-derived transcripts and the increase of zygotic transcription were examined across development. Upon targeting genes that play a role in the degradation of maternal RNAs, one *DGCR8*-like gene (*dgcr8*), one *dicer* gene (*dicer*), one *smaug* homolog (*smaug1*) and three putative *smg* genes (*smg7*, *smg8*, and *smg9*) were identified within the *M. franciscanus* developmental transcriptome (Table 1). The expression levels of *dgcr8, dicer, smg7, smg8,* and *smg9* all peaked during the 8- to 16-cell (CL) and morula (MO) stages (Fig. 3A). The *dgcr8* gene plays a role in processing microRNAs that are required for degrading mRNAs in mammals (Marlow 2010, Wang, et al. 2007). The Mediterranean sea urchin, *Paracentrotus lividius*, exhibited a similar pattern of expression of *dgcr8* as reported here, in which there was a peak in expression within 8- and 16-cell embryos (Gildor, et al. 2016). The authors attributed this observation to the role of *dgcr8* in degrading

427   maternal mRNAs (Gildor, et al. 2016). *Dicer* is involved in clearing maternal messages in

428   zebrafish and mice (Giraldez, et al. 2005, Marlow 2010), and mutations in the *dicer* gene are

429   known to alter and arrest embryonic development in some species (Murchison, et al. 2007).

430   Therefore, the peak in expression of both *dgcr8* and *dicer* during the CL and MO stages supports

431   that maternal mRNAs are degraded during this period of embryonic development.

432
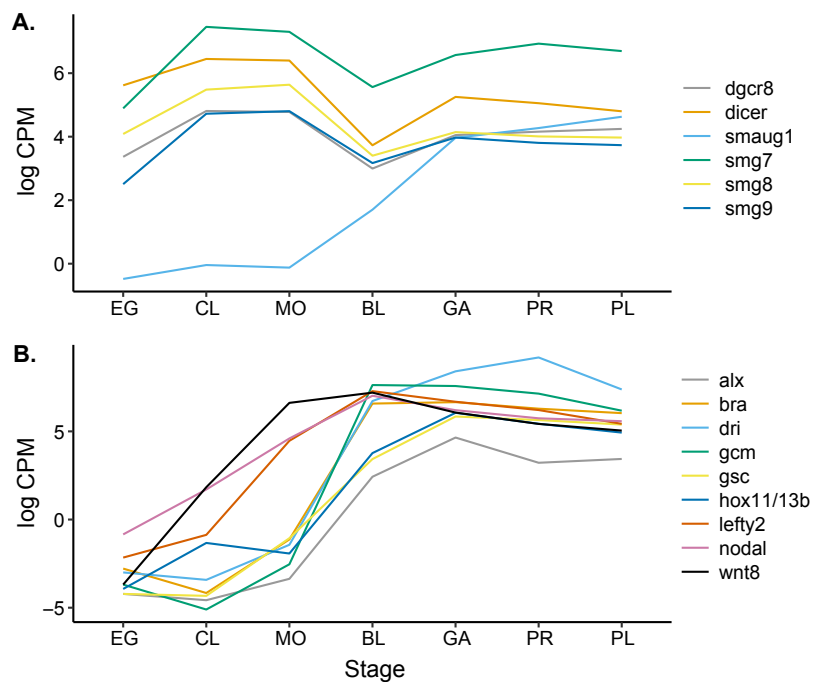
433

Fig. 3. The expression of putative genes that play a functional role during the
MZT. These genes **A.** regulate the removal of mRNA, and **B.** regulate   434
zygotic development. The expression data is in log2 counts per million reads
                                                                        435

436   The *smg* genes code for proteins that function in nonsense-mediated mRNA decay

437   (NMD) in a variety of organisms (Okada-Katsuhata, et al. 2012, Pulak and Anderson 1993,

438   Yamashita, et al. 2009). The NMD pathway detects and degrades mRNAs, and is often described

439   as a surveillance pathway that serves as a quality-control mechanism to remove mRNAs with

440   premature termination codons (Chang, et al. 2007, Hentze and Kulozik 1999). However, the

441   NMD pathway also serves functional roles that shape gene expression and are important for

442   differentiation and development (Lykke-Andersen and Jensen 2015). The NMD pathway has

443   also been shown to selectively degrade mRNA transcripts with longer 3' UTRs, causing a

444   relative enrichment of shorter 3' UTR transcripts (Bao, et al. 2016). In zebrafish embryos, 3'

445   UTR length affects the stability of maternal mRNAs because longer 3' UTRs confer resistance to

446   codon-mediated deadenylation, the first step required for mRNA decay (Mishima and Tomari

447   2016). Therefore, the removal of long 3' UTR transcripts via the NMD pathway may increase

448   the relative proportion of short 3' UTR transcripts available for deadenylation and decay during

449   the CL and MO stages.

450   In contrast to the expression of *dgcr8*, *dicer*, and *smg* genes, *smaug1* was not expressed

451   until the blastula stage (Fig. 3A). The *smaug* gene is a transcriptional regulator known to bind to

452   and target maternal RNAs for degradation in *Drosophila melanogaster*, that is highly conserved

453   across taxa (Tadros, et al. 2007). It is therefore possible, that degradation of maternal mRNAs is

454   still ongoing at the blastula stage. This differs from observations in *S. purpuratus*, in which

455   maternal degradation appears to end prior to the blastula stage (Tadros and Lipshitz 2009, Wei,

456   et al. 2006). With the exception of *smaug1* expression, the degradation of maternal transcripts

457   appears to primarily occur during the 8- to 16-cell (CL) and morula (MO) stages. This is

458   additionally supported by the WGCNA analysis, which revealed genes related to catalytic

459   activity acting on RNA in module eigengenes brown and green, both of which share significant,

460   positive correlations with the CL and MO stages (Table 4).

461   Evidence of maternal transcript degradation is also reflected by a decrease in expression

462   of maternal transcripts, which are represented by those expressed during the egg stage (EG). A

463   heatmap of the top 500 transcripts expressed in eggs revealed that expression of these transcripts

464   began to decline at the 8-to 16-cell and morula stages (Fig. 4). By the blastula stage, the overall

465 expression patterns of the maternal transcripts had completely changed, with some moderate

466 expression of maternal transcripts remaining although the majority have dramatically decreased.

467 Most of these maternal transcripts continued to show low levels of expression relative to the eggs

468 during the remaining stages of development (i.e., gastrula through pluteus stages). Taken

469 together, the degradation of maternal RNAs and the resulting reduction in expression of maternal

470 transcrpts begin as early as the 8-cell stage, although it is possible that the process begins even

471 sooner after fertilization at a stage prior to what was examined in this study (e.g., at the 2-cell or

472 4-cell stage). This result is similar to the timing of maternal mRNA degradation in the purple sea

473 urchin, *Strongylocentrotus purpuratus*, in which maternal transcripts are destabilized by the 2-

474 cell embryonic stage (Tadros and Lipshitz 2009, Wei, et al. 2006).
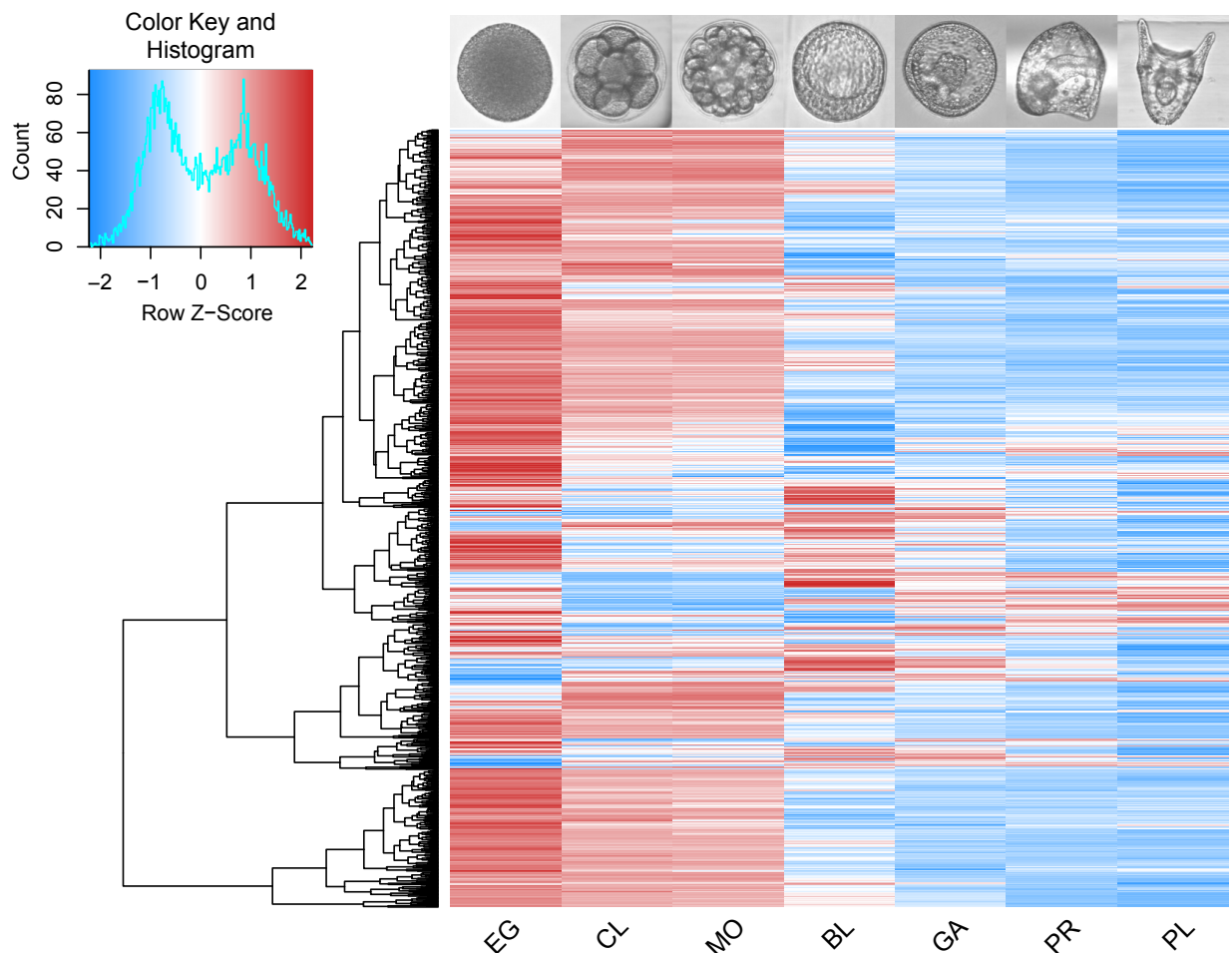
475

Fig. 4. Heatmap of the top 500 maternal transcripts expressed during the egg stage (EG). The rows are transcripts and columns are in order of developmental stage. Transcript expression data are in log2 counts per million reads (log2 CPM), and the data are scaled by row.

476   To examine the timing of zygotic genome activation, nine putative genes important for

477   zygotic development were identified within the *M. franciscanus* transcriptome (Table 1). Of

478   these, *hox11/13b*, *lefty2*, *nodal*, and *wnt8* increased in expression between egg and the earliest

479   measured developmental stage, the 8- to 16-cell stage (Fig. 3B). The expression levels of *wnt8*,

480   *nodal*, and *lefty2* increased further during the morula stage before plateauing and maintained

481   relatively consistent levels of expression from the blastula stage through the remainder of

482   development. This is somewhat similar to the expression patterns observed in *S. purpuratus*, in

483   which many of these zygotic genes reached peak expression levels at the blastula stage (Tadros

484   and Lipshitz 2009, Wei, et al. 2006). The *homeobox 11/13b* (*hox11/13b*) gene is one of the

485   earliest transcription factors necessary for endoderm cell specification in echinoderms (Peter and

486   Davidson 2010). Similar transcriptional mechanisms may underlie both *left-right determination*

487   *factor-2* (*lefty2*) and *nodal* expression, which function together to establish the oral-aboral

488   embryonic axis (Adachi, et al. 1999, Duboc, et al. 2004, Duboc, et al. 2008). Lastly, *wnt8* is

489   required for endomesoderm development in sea urchins, including cell differentiation and

490   gastrulation processes (Minokawa, et al. 2005, Wikramanayake, et al. 2004). In agreement with

491   our results, in *S. purpuratus* expression of a *wnt8* homolog has been observed beginning at the

492   16-cell stage (Wikramanayake, et al. 2004). The expression of these transcripts during the CL

493   and MO stages support that the activation of zygotic transcription in *M. franciscanus* may occur

494   as early as early as fourth cleavage.

495   The remaining five zygotic transcripts, *alx*, *bra*, *dri*, *gcm*, and *gsc* remained at low levels

496   of expression until the blastula stage (BL), at which time most reached their peak levels of

497   expression (Fig. 3B). The expression of these transcripts remained fairly consistent for the

498   remainder of development. *Aristaless-like homeobox* (*alx*) expression has a role in primary

499 mesenchyme cell formation, and acts as an early regulatory gene for skeletogenesis (Ettensohn,

500 et al. 2003, Ettensohn 2009). *Brachyury* (*bra*) functions in gastrulation and endoderm

501 development (Peterson, et al. 1999, Rast, et al. 2002). The *dead ringer* (*dri*) gene is required for

502 normal embryological development and is highly conserved across taxa (Shandala, et al. 1999),

503 functioning in skeletogenesis and oral ectoderm formation in sea urchins (Amore, et al. 2003).

504 The *glial cells missing* (*gsm*) gene functions in endomesoderm specification, particularly those of

505 pigment cells (Ransick, et al. 2002, Ransick and Davidson 2006). Lastly, in sea urchin embryos,

506 *goosecoid* (*gsc*) plays a role in regulating cell specification along the animal-vegetal and oral-

507 aboral axes (Angerer, et al. 2001). The expression patterns of *alx*, *bra*, *dri*, *gcm*, and *gsc* may

508 represent a second wave of zygotic genome activation that occurs at the BL stage. This pattern of

509 zygotic genome activation is very similar to that of *S. purpuratus*, in which there is a minor wave

510 of zygotic transcription during early cell divisions of the embryo, followed by a major wave of

511 zygotic transcription at the blastula stage (Tadros and Lipshitz 2009, Wei, et al. 2006). Overall,

512 the timing of the MZT in *M. franciscanus* appears to span from early cleavage through the

513 blastula stage, in which 1) maternal degradation begins at or before the 8- to 16-cell stage and

514 persists to the blastula stage, and 2) zygotic activation occurs as a minor wave at the 8- to 16-cell

515 and morula stages and as a major wave by the blastula stage.

516

517 **4. Conclusions**

518   The transcriptome presented here is a useful molecular resource for studying *M.*

519 *franciscanus*, a non-model organism and important fisheries species. This reference will support

520 future investigations into the early development of *M. franciscanus*, and its response to

521 environmental stress. These studies will facilitate our understanding of a species that possesses a

522 significant ecological role in kelp forest ecosystems and is economically valuable as a fisheries

523 organism. Finally, the examination of the timing of the MZT will inform future gene expression

524 studies that aim to target stages in which the zygotic transcriptome is fully activated.

525

526 **5. Data deposition**

527 The raw sequence data and the *M. franciscanus* developmental transcriptome are

528 available under NCBI Bioproject # and Sequence Read Archive accession #s. The transcriptome

529 annotation can be found in the NCBI Transcriptome Shotgun Assembly (TSA) database under

530 accession #.

531

547

548

References

550

551 Adachi, H., Y. Saijoh, K. Mochida, S. Ohishi, H. Hashiguchi, A. Hirao, H. Hamada,
552   Determination of left/right asymmetric expression of *nodal* by a left side-specific
553   enhancer with sequence similarity to a *lefty-2* enhancer, Genes & Development, 13
554   (1999) 1589-1600.
555 Amore, G., R.G. Yavrouian, K.J. Peterson, A. Ransick, D.R. McClay, E.H. Davidson,
556   *Spdeadringer*, a sea urchin embryo gene required separately in skeletogenic and oral
557   ectoderm gene regulatory networks, Developmental Biology, 261 (2003) 55-81.
558 Andrew, N., Y. Agatsuma, E. Ballesteros, A. Bazhin, E. Creaser, D. Barnes, L. Botsford, A.
559   Bradbury, A.K. Campbell, J. Dixon, S. Einarsson, P. Gerring, K. Hebert, M. Hunter, S.
560   Hur, C. Johnson, Juinio-Meñez, P.E. Kalvass, R. Miller, C. Moreno, J. Palleiro, D. Rivas,
561   S. Robinson, S.C. Schroeter, R. Steneck, R. Vadas, D. Woodby, Z. Xiaoqi, Status and
562   management of world sea urchin fisheries, Oceanogr Mar Biol, 40 (2002) 343-425.
563 Andrews, S., FASTQC: a quality control tool for high throughput sequence data, Available
564   online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc, 2010.
565 Angerer, L.M., D.W. Oleksyn, A.M. Levine, X. Li, W.H. Klein, R.C. Angerer, Sea urchin
566   goosecoid function links fate specification along the animal- vegetal and oral-aboral
567   embryonic axes, Development, 128 (2001) 4393-4404.
568 Bao, J., K. Vitting-Seerup, J. Waage, C. Tang, Y. Ge, B.T. Porse, W. Yan, UPF2-Dependent
569   Nonsense-Mediated mRNA Decay Pathway Is Essential for Spermatogenesis by
570   Selectively Eliminating Longer 3'UTR Transcripts, PLoS Genetics, 12 (2016) e1005863.
571 Barrett, D., B.F. Edwards, Hatching Enzyme of the Sea Urchin *Strongylocentrotus purpuratus*
572   barrett, Methods in enzymology, Academic Press, 1976.
573 Brekhman, V., A. Malik, B. Haas, N. Sher, T. Lotan, Transcriptome profiling of the dynamic life
574   cycle of the scypohozoan jellyfish *Aurelia aurita*, Bmc Genomics, 16 (2015).
575 Burke, R.D., Morphogenesis of the digestive tract of the pluteus larva of *Strongylocentrotus*
576   *purpuratus*: shaping and bending, International Journal of Invertebrate Reproduction, 2
577   (1980) 13-21.
578 Byrne, M., Impact of ocean warming and ocean acidification on marine invertebrate life history
579   stages: Vulnerabilities and potential for persistence in a changing ocean, Oceanogr Mar
580   Biol, 49 (2011) 1-42.
581 Byrne, M., Global change ecotoxicology: Identification of early life history bottlenecks in
582   marine invertebrates, variable species responses and variable experimental approaches,
583   Mar Environ Res, 76 (2012) 3-15.
584 Byrne, M., R. Przeslawski, Multistressor impacts of warming and acidification of the ocean on
585   marine invertebrates' life histories, Integrative and Comparative Biology, 53 (2013) 582-
586   596.

Carruthers, M., A.A. Yurchenko, J.J. Augley, C.E. Adams, P. Herzyk, K.R. Elmer, *De novo* transcriptome assembly, annotation and comparison of four ecological and evolutionary model salmonid fish species, Bmc Genomics, 19 (2018).

Chang, Y.-F., J.S. Imam, M.F. Wilkinson, The Nonsense-Mediated Decay RNA Surveillance Pathway, Annual Review of Biochemistry, 76 (2007) 51-74.

Chen, X., D. Zeng, X. Chen, D. Xie, Y. Zhao, C. Yang, Y. Li, N. Ma, M. Li, Q. Yang, Z. Liao, H. Wang, Transcriptome Analysis of *Litopenaeus vannamei* in Response to White Spot Syndrome Virus Infection, PLoS One, 8 (2013) e73218.

Coppe, A., S. Bortoluzzi, G. Murari, I.A.M. Marino, L. Zane, C. Papetti, Sequencing and Characterization of Striped Venus Transcriptome Expand Resources for Clam Fishery Genetics, PLoS One, 7 (2012) e44185.

Dan, K., K. Okazaki, Cyto-Embryological Studies of Sea Urchins. III. Role of the Secondary Mesenchyme Cells in the Formation of the Primitive Gut in Sea Urchin Larvae, Bio Bull, 110 (1956) 29-42.

De Wit, P., S. Palumbi, Transcriptome-wide polymorphisms of red abalone (Haliotis rufescens) reveal patterns of gene flow and local adaptation, Mol Ecol, (2012).

Dickson, A., F. Millero, A comparison of the equilibrium constants for the dissociation of carbonic acid in seawater media, Deep-Sea Res Pt I, 34 (1987) 1773-1743.

Dickson, A., C. Sabine, J. Christian, SOP 6b. Determination of the pH of seawater using the indicator dye *m*-cresol purple. Ver. 3.01. Jan 28, 2009, 2007.

Dickson, A., C. Sabine, J. Christian, SOP 3b. Determination of total alkalinity in seawater using an open-cell titration, Ver. 3.01 2008, 2007.

Duboc, V., E. Röttinger, L. Besnardeau, T. Lepage, Nodal and BMP2/4 Signaling Organizes the Oral-Aboral Axis of the Sea Urchin Embryo, Developmental Cell, 6 (2004) 397-410.

Duboc, V., F. Lapraz, L. Besnardeau, T. Lepage, Lefty acts as an essential modulator of Nodal activity during sea urchin oral–aboral axis formation, Developmental Biology, 320 (2008) 49-59.

Dupont, S., M. Thorndyke, Impact of CO_2-driven ocean acidification on invertebrates early life-history – What we know, what we need to know and what we can do, Biogeosciences, 6 (2009) 3109-3131.

Ebert, T.A., J.D. Dixon, S.C. Schroeter, P.E. Kalvass, N.T. Richmond, WA, D.A. Woodby, Growth and mortality of red sea urchins *Strongylocentrotus franciscanus* across a latitudinal gradient, Marine Ecology Progress Series, 190 (1999) 189-209.

Ekblom, R., J. Galindo, Applications of next generation sequencing in molecular ecology of non-model organisms, Heredity, 107 (2011) 1-15.

Ettensohn, C.A., Primary Invagination of the Vegetal Plate During Sea Urchin Gastrulation, American Zoologist, 24 (1984) 571-588.

Ettensohn, C.A., K.M. Malinda, Size regulation and morphogenesis: a cellular analysis of skeletogenesis in the sea urchin embryo, Development, 119 (1993) 155-167.

Ettensohn, C.A., M.R. Illies, P. Oliveri, D.L. De Jong, Alx1, a member of the Cart1/Alx3/Alx4 subfamily of Paired-class homeodomain proteins, is an essential component of the gene network controlling skeletogenic fate specification in the sea urchin embryo, Development, 130 (2003) 2917-2928.

Ettensohn, C.A., Lessons from a gene regulatory network: echinoderm skeletogenesis provides insights into evolution, plasticity and morphogenesis, Development, 136 (2009) 11-21.

Fangue, N.A., M.J. O'Donnell, M.A. Sewell, P.G. Matson, A.C. MacPherson, G.E. Hofmann, A laboratory-based, experimental system for the study of ocean acidification effects on marine invertebrate larvae, Limnology and Oceanography: Methods, 8 (2010) 441-452.

Franks, S.J., A.A. Hoffmann, Genetics of climate change adaptation, Annual Review of Genetics, 46 (2012) 185-208.

Gaitán-Espitia, J.D., R. Sánchez, P. Bruning, L. Cárdenas, Functional insights into the testis transcriptome of the edible sea urchin *Loxechinus albus*, Sci Rep, 6 (2016).

Gaitán-Espitia, J.D., G.E. Hofmann, Gene expression profiling during the embryo-to-larva transition in the giant red sea urchin *Mesocentrotus franciscanus*, Ecol Evol, 7 (2017) 2798-2811.

Ghaffari, N., A. Sanchez-Flores, R. Doan, K.D. Garcia-Orozco, P.L. Chen, A. Ochoa-Leyva, A.A. Lopez-Zavala, J.S. Carrasco, C. Hong, L.G. Brieba, E. Rudiño-Piñera, P.D. Blood, J.E. Sawyer, C.D. Johson, S.V. Dindot, R.R. Sotelo-Mundo, M.F. Criscitiello, Novel transcriptome assembly and improved annotation of the whiteleg shrimp (*Litopenaeus vannamei*), a dominant crustacean in global seafood mariculture, Sci Rep, 4 (2014).

Gilbert, D., Gene-omes built from mRNA seq not genome DNA, 7th annual arthropod genomics symposium, Notre Dame, 2013.

Gildor, T., A. Malik, N. Sher, L. Avraham, S. Ben-Tabou de-Leon, Quantitative developmental transcriptomes of the Mediterranean sea urchin *Paracentrotus lividus*, Marine Genomics, 25 (2016) 89-94.

Gillard, G.B., D.J. Garama, C.M. Brown, The transcriptome of the NZ endemic sea urchin Kina (*Evechinus chloroticus*), Bmc Genomics, 15 (2014).

Giraldez, A.J., R.M. Cinalli, M.E. Glasner, A.J. Enright, J.M. Thomson, S. Baskerville, S.M. Hammond, D.P. Bartel, A.F. Schier, MicroRNAs Regulate Brain Morphogenesis in Zebrafish, Science, 308 (2005) 833-838.

Gosselin, L., P.-Y. Qian, Juvenile mortality in benthic marine invertebrates., Marine Ecology Progress Series, 146 (1997) 265-282.

Grabherr, M.G., B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B.W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, A. Regev, Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data, Nature Biotechnology, 29 (2011) 644-652.

Gurevich, A., V. Saveliev, N. Vyahhi, G. Tesler, QUAST: quality assessment tool for genome assemblies, Bioinformatics, 29 (2013) 1072-1075.

Hentze, M.W., A.E. Kulozik, A Perfect Message: RNA Surveillance and Nonsense-Mediated Decay, Cell, 96 (1999) 307-310.

Heyland, A., Z. Vue, C.R. Voolstra, M. Medina, L.L. Moroz, Developmental transcriptome of *Aplysia californica*, Journal of Experimental Zoology, Part B: Molecular and Developmental Evolution, 0 (2011) 113-134.

Hofmann, G., L. Washburn, SBC LTER: Ocean: Time-series: Mid-water SeaFET and CO2 system chemistry at Mohawk Reef (MKO), ongoing since 2012-01-11, Santa Barbara Coastal LTER, 2015.

Ji, P., G. Liu, J. Xu, X. Wang, J. Li, Z. Zhao, X. Zhang, Y. Zhang, P. Xu, X. Sun, Characterization of Common Carp Transcriptome: Sequencing, *De Novo* Assembly, Annotation and Comparative Genomics, PLoS One, 7 (2012) e35152.

Jo, J., J. Park, H.-G. Lee, E.M.A. Kern, S. Cheon, S. Jin, J.-K. Park, S.-J. Cho, C. Park, Comparative transcriptome analysis of three color variants of the sea cucumber *Apostichopus japonicus*, Marine Genomics, 28 (2016) 21-24.

Kalvass, P.E., Riding the rollercoaster: boom and decline in the California red sea urchin fishery, Journal of Shellfish Research, 19 (2000) 621-622.

Kanehisa, M., S. Goto, KEGG: Kyoto Encyclopedia of Genes and Genomes, Nucleic Acids Res, 28 (2000) 27-30.

Keesing, J., K. Hall, Review of harvests and status of world sea urchin fisheries points to opportunities for aquaculture, Journal of Shellfish Research, 17 (1998) 1597-1604.

Krueger, F., Trim Galore!: A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, Available at: www.bioinformatics.babraham.ac.uk/projects/trim_galore/, 2015.

Kurihara, H., Effects of CO_2-driven ocean acidification on the early developmental stages of invertebrates, Marine Ecology Progress Series, 373 (2008).

Langfelder, P., S. Horvath, WGCNA: an R package for weighted correlation network analysis, Bmc Bioinformatics, 9 (2008).

Langmead, B., S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, Nature Methods, 9 (2012) 357-359.

Le, S., J. Josse, F. Husson, FactoMineR: An R Package for Multivariate Analysis, Journal of Statistical Software, 25 (2008) 1-18.

Leighton, D., L. Jones, W. North, Ecological relationships between giant kelp and sea urchins in southern California, in: E. Young, J. Maclachlan (Eds.) Proceedings of the 5th international seaweed symposium, Pergamon Press, 1966, pp. 141-153.

Lenz, P.H., V. Roncalli, R.P. Hassett, L.-S. Wu, M.C. Cieslack, D.K. Hartline, A.E. Christie, De Novo Assembly of a Transcriptome for *Calanus finmarchicus* (Crustacea, Copepoda) – The Dominant Zooplankter of the North Atlantic Ocean, PLoS One, 9 (2014) e88589.

Lepage, T., C. Gache, Purification and Characterization of the Sea Urchin Embryo Hatching Enzyme, The Journal of Biological Chemistry, 264 (1989) 4787-4793.

Li, B., C.N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome, Bmc Bioinformatics, 12 (2011).

Liao, X., L. Cheng, P. Xu, G. Lu, M. Wachholtz, X. Sun, S. Chen, Transcriptome Analysis of Crucian Carp (*Carassius auratus*), an Important Aquaculture and Hypoxia- Tolerant Species, PLoS One, 8 (2013) e62308.

Lv, J., P. Liu, B. Gao, Y. Wang, Z. Wang, P. Chen, J. Li, Transcriptome Analysis of the *Portunus trituberculatus*: De Novo Assembly, Growth-Related Gene Identification and Marker Discovery, PLoS One, 9 (2014) e94055.

Lykke-Andersen, S., T.H. Jensen, Nonsense-mediated mRNA-decay: an intricate machinery that shapes transcriptomes, Nature Reviews Molecular Cell Biology, 16 (2015) 665-677.

Marlow, F.L., Maternal control of development in vertebrates, Morgan & Claypool Life Sciences, 2010.

Mehrbach, C., C. Culberson, J. Hawley, R. Pytkowicz, Measurement of the apparent dissociation constants of carbonic acid in seawater at atmospheric pressure, Limnology and Oceanography, 18 (1973) 897-907.

Minokawa, T., A.H. Wikramanayake, E.H. Davidson, *cis*-Regulatory inputs of the *wnt8* gene in the sea urchin endomesoderm network, Developmental Biology, 288 (2005) 545-558.

722 Mishima, Y., Y. Tomari, Codon Usage and 30 UTR Length Determine Maternal mRNA Stability
723   in Zebrafish, Molecular Cell, 61 (2016) 874-885.
724 Murchison, E.P., P. Stein, Z. Xuan, H. Pan, M.Q. Zhang, R.M. Schultz, G.J. Hannon, Genes &
725   Development, 21 (2007) 682-693.
726 Okada-Katsuhata, Y., A. Yamashita, K. Kutsuzawa, N. Izumi, F. Hirahara, S. Ohno, N- and C-
727   terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7
728   during NMD, Nucleic Acid Research, 40 (2012) 1251-1266.
729 Peter, I.S., E.H. Davidson, The endoderm gene regulatory network in sea urchin embryos up to
730   mid-blastula stage, Developmental Biology, 340 (2010) 188-199.
731 Peterson, K.J., R.A. Cameron, K. Tagawa, N. Satoh, E.H. Davidson, A comparative molecular
732   approach to mesodermal patterning in basal deuterostomes: the expression pattern of
733   *Brachyury* in the enteropneust hemichordate *Ptychodera flava*, Development, 126 (1999)
734   85-95.
735 Pulak, R., P. Anderson, mRNA surveillance by the *Caenorhabditis elegans smg* genes, Genes &
736   Development, 7 (1993) 1885-1897.
737 Ransick, A., J.P. Rast, T. Minokawa, C. Calestani, E.H. Davidson, New Early Zygotic
738   Regulators Expressed in Endomesoderm of Sea Urchin Embryos Discovered by
739   Differential Array Hybridization, Developmental Biology, 246 (2002) 132-147.
740 Ransick, A., E.H. Davidson, *cis*-regulatory processing of Notch signaling input to the sea urchin
741   *glial cells missing* gene during mesoderm specification, Developmental Biology, 297
742   (2006) 587-602.
743 Rast, J.P., R.A. Cameron, A.J. Poustka, E.H. Davidson, *brachyury* Target Genes in the Early Sea
744   Urchin Embryo Isolated by Differential Macroarray Screening, Developmental Biology,
745   246 (2002) 191-208.
746 Reusch, T.B., T.E. Wood, Molecular ecology of global change, Mol Ecol, 16 (2007) 3973-3992.
747 Ritchie, M., B. Phipson, D. Wu, Y. Hu, C. Law, W. Shi, G. Smyth, *limma* powers differential
748   expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res, 43
749   (2015) e47.
750 Robbins, L., M. Hansen, J. Kleypas, S. Meylan, CO2calc—A user-friendly seawater carbon
751   calculator for Windows, Max OS X, and iOS (iPhone), U.S. Geological Survey Open-
752   File Report, 2010, pp. 17.
753 Robertson, G., J. Schein, R. Chiu, R. Corbett, M. Field, S.D. Jackman, K. Mungall, S. Lee, H.M.
754   Okada, J.Q. Qian, M. Griffith, A. Raymond, N. Thiessen, T. Cezard, Y.S. Butterfield, R.
755   Newsome, S.K. Chan, R. She, R. Varhol, B. Kamon, A.-L. Prabhu, A. Tam, Y. Zhao,
756   R.A. Moore, M. Hirst, M.A. Marra, S.J.M. Jones, P.A. Hoodless, I. Birol, *De novo*
757   assembly and analysis of RNA-seq data, Nature Methods, 7 (2010) 909-912.
758 Robinson, M.D., A. Oshlack, A scaling normalization method for differential expression analysis
759   of RNA-seq data, Genome Biol, 11 (2010) R25.
760 Rogers-Bennett, L., The ecology of *Strongylocentrotus franciscanus* and *Strongylocentrotus*
761   *purpuratus*, Developments in Aquaculture and Fisheries Science, 37 (2007) 393-425.
762 Rogers-Bennett, L., *Strongylocentrotus franciscanus* and *Strongylocentrotus purpuratus*, in: J.M.
763   Lawrence (Ed.) Sea Urchins: Biology and Ecology, 2013, pp. 413-435.
764 Schulz, M.H., D.R. Zerbino, M. Vingron, E. Birney, Oases: Robust de novo RNA-seq assembly
765   across the dynamic range of expression levels, Bioinformatics, 28 (2012) 1086-1092.

766   Shandala, T., R.D. Kortschak, S. Gregory, R. Saint, The *Drosophila dead ringer* gene is required
767   for early embryonic patterning through regulation of *argos* and *buttonhead* expression,
768   Development, 126 (1999) 4341-4349.
769   Shier, A.F., The Maternal-Zygotic Transition: Death and Birth of RNAs, Science, 316 (2007)
770   406-407.
771   Simão, F.A., R.M. Waterhouse, P. Ioannidis, E.V. Kriventseva, E.M. Zdobnov, BUSCO:
772   assessing genome assembly and annotation completeness with single-copy orthologs,
773   Bioinformatics, 31 (2015) 3210–3212.
774   Souza, C.A., N. Murphy, J.M. Strugnell, De novo transcriptome assembly and functional
775   annotation of the southern rock lobster (*Jasus edwardsii*), Marine Genomics, 42 (2018)
776   58-62.
777   Strader, M., J. Wong, L. Kozal, T. Leach, G. Hofmann, Parental environments alter DNA
778   methylation in offspring of the purple sea urchin, *Strongylocentrotus purpuratus*, J Exp
779   Mar Biol Ecol, (In revision).
780   Strathmann, M.F., Reproduction and Development of Marine Invertebrates of the Northern
781   Pacific Coast, University of Washington Press, USA, 1987.
782   Tadros, W., A.L. Goldman, T. Babak, F. Menzies, T. Orr-Weaver, T.R. Hughes, J.T. Westwood,
783   C.A. Smibert, H.D. Lipshitz, SMAUG Is a Major Regulator of Maternal mRNA
784   Destabilization in *Drosophila* and Its Translation Is Activated by the PAN GU Kinase,
785   Developmental Cell, 12 (2007) 143-155.
786   Tadros, W., H.D. Lipshitz, The maternal-to-zygotic transition: a play in two acts, Development,
787   136 (2009) 3033-3042.
788   Tian, K., F. Lou, T. Gao, Y. Zhou, Z. Miao, Z. Han, *De novo* assembly and annotation of the
789   whole transcriptome of *Sepiella maindroni*, Marine Genomics, 38 (2018) 13-16.
790   Valenzuela-Quiñonez, F., How fisheries management can benefit from genomics?, Briefings in
791   Functional Genomics, 15 (2016) 352-357.
792   Wang, Y., R. Medvid, C. Melton, R. Jaenisch, R. Blelloch, DGCR8 is essential for microRNA
793   biogenesis and silencing of embryonic stem cell self-renewal, Nature Genetics, 39 (2007)
794   380-385.
795   Wei, Z., R. Angerer, L. Angerer, A database of mRNA expression patterns for the sea urchin
796   embryo, Developmental Biology, 300 (2006) 476-484.
797   Wenne, R., P. Boudry, J. Hemmer-Hansen, K.P. Lubieniecki, A. Was, A. Kause, What role for
798   genomics in fisheries management and aquaculture?, Aquatic Living Resources, 20
799   (2007) 241-255.
800   Wikramanayake, A.H., R. Peterson, J. Chen, L. Huang, J.M. Bince, D.R. McClay, W.H. Klein,
801   Nuclear β -Catenin-Dependent Wnt8 Signaling in Vegetal Cells of the Early Sea Urchin
802   Embryo Regulates Gastrulation and Differentiation of Endoderm and Mesodermal Cell
803   Lineages, Genesis, 39 (2004) 194-205.
804   Wolpert, L., Gastrulation and the evolution of development, Development, 116 (1992) 7-13.
805   Xie, Y., G. Wu, J. Tang, R. Luo, J. Patterson, S. Liu, W. Huang, G. He, S. Gu, S. Li, X. Zhou,
806   T.-W. Lam, Y. Li, X. Xu, G.K.-S. Wong, J. Wang, SOAPdenovo-Trans: *de novo*
807   transcriptome assembly with short RNA-Seq reads, Bioinformatics, 30 (2014) 1660-
808   1666.
809   Yamashita, A., N. Izumi, I. Kashima, T. Ohnishi, B. Saari, Y. Katsuhata, R. Muramatsu, T.
810   Morita, A. Iwamatsu, T. Hachiya, R. Kurata, H. Hirano, P. Anderson, S. Ohno, SMG-8
811   and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the

812   mRNA surveillance complex during nonsense-mediated mRNA decay, Genes &
813   Development, 23 (2009) 1091-1105.
814 Zeng, V., K.E. Villanueva, B.S. Ewen-Campen, F. Alwes, W.E. Browne, C.G. Extavour, De
815   novo assembly and characterization of a maternal and developmental transcriptome for
816   the emerging model crustacean Parhyale hawaiensis, Bmc Genomics, 12 (2011).
817 Zerbino, D.R., E. Birney, Velvet: algorithms for de novo short read assembly using de Bruijn
818   graphs, Genome Res, 18 (2008) 821-829.
819 Zhao, X., H. Yu, L. Kong, Q. Li, Transcriptomic Responses to Salinity Stress in the Pacific
820   Oyster *Crassostrea gigas*, PLoS One, 7 (2012) e46244.
821