

# SCIENTIFIC REPORTS



OPEN

## Transcriptome analysis on the exoskeleton formation in early developmental stages and reconstruction scenario in growth-moulting in *Litopenaeus vannamei*

Yi Gao<sup>1</sup>, Jiankai Wei<sup>3</sup>, Jianbo Yuan<sup>1</sup>, Xiaojun Zhang<sup>1,2</sup>, Fuhua Li<sup>1,2</sup> & Jianhai Xiang<sup>1,2</sup>

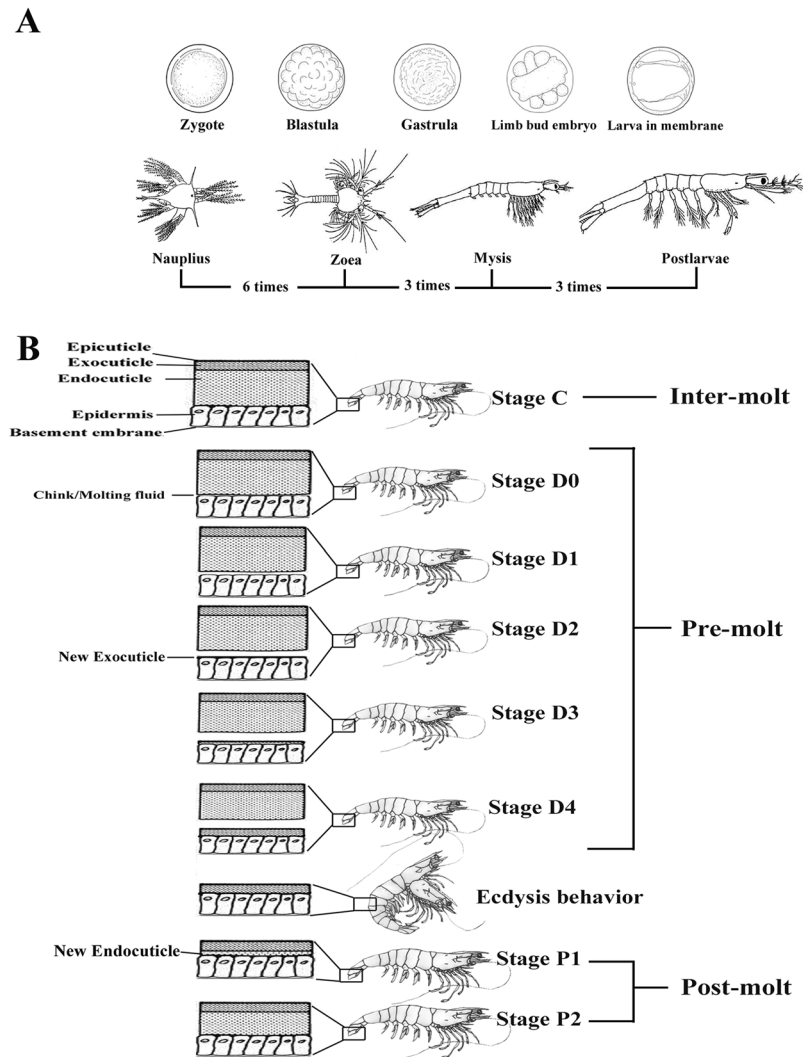
Exoskeleton construction is an important issue in shrimp. To better understand the molecular mechanism of exoskeleton formation, development and reconstruction, the transcriptome of the entire developmental process in *Litopenaeus vannamei*, including nine early developmental stages and eight adult-moulting stages, was sequenced and analysed using Illumina RNA-seq technology. A total of 117,539 unigenes were obtained, with 41.2% unigenes predicting the full-length coding sequence. Gene Ontology, Clusters of Orthologous Group (COG), the Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis and functional annotation of all unigenes gave a better understanding of the exoskeleton developmental process in *L. vannamei*. As a result, more than six hundred unigenes related to exoskeleton development were identified both in the early developmental stages and adult-moulting. A cascade of sequential expression events of exoskeleton-related genes were summarized, including exoskeleton formation, regulation, synthesis, degradation, mineral absorption/reabsorption, calcification and hardening. This new insight on major transcriptional events provide a deep understanding for exoskeleton formation and reconstruction in *L. vannamei*. In conclusion, this is the first study that characterized the integrated transcriptomic profiles cover the entire exoskeleton development from zygote to adult-moulting in a crustacean, and these findings will serve as significant references for exoskeleton developmental biology and aquaculture research.

As the largest phylum of animals with over a million described species, arthropods make up more than 80% of all known living animal species all over the world<sup>1</sup>. This phylum includes trilobites, chelicerates, myriapods, hexapods, and crustaceans, which inhabit the land, sea and air. The common feature of arthropods is a covering of the exoskeleton<sup>2</sup>. The exoskeleton provides an arthropods shape, water-proofing, locomotion, and can also be very important for growth, development and reproduction<sup>3</sup>. However, this structure also confines metamorphosis, body growth and mating. Thus, arthropods periodically discard their old exoskeletons, meanwhile synthesize and form a new one to match the new body, which is a process referred to as moulting<sup>4,5</sup>.

In the past century, the structure, chemical composition and mechanical characteristics of the exoskeleton have been widely studied in arthropods<sup>6–10</sup>. Some bio-mimetic composites with high strength and light weight were built and inspired by the exoskeleton<sup>11–14</sup>. The products extracted from the exoskeleton, such as chitin, chitosan and chitosan oligosaccharides have been widely used in the fields of chemical industry and medicine<sup>15,16</sup>. However, studies on the molecular aspects of exoskeleton development are still scarce, especially analysis of the entire exoskeleton development from zygote to adult. How the exoskeletal regions are constructed and reconstructed to provide their morphological, physiological and mechanical functions is an important issue.

*Litopenaeus vannamei*, one of the most important economic species of decapoda, is the most commonly cultured shrimp species in the world. It is also an ideal model for analysing the exoskeleton development. The

<sup>1</sup>Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, 266071, China. <sup>2</sup>Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao, 266071, China. <sup>3</sup>Ocean University of China, Qingdao, 266071, China. Correspondence and requests for materials should be addressed to X.Z. (email: [xjzhang@qdio.ac.cn](mailto:xjzhang@qdio.ac.cn)) or J.X. (email: [jhxiang@qdio.ac.cn](mailto:jhxiang@qdio.ac.cn))



**Figure 1.** (A) Embryonic and larval stages during early development of *L. vannamei*. Morphologies are modified from Wei, *et al.*, 2014<sup>19</sup>. The moulting times are labelled under the figure. (B) Hierarchical structure of the exoskeleton of *L. vannamei* during the moulting process.

exoskeleton alternate is crucial for metamorphosis and growth throughout the life of shrimp. *L. vannamei* experiences 50 moulting periods during a lifetime on average<sup>17</sup>, which is much more than other arthropods; e.g., silkworms (four times), crabs (~eighteen times), and locusts (five times). In early development, *L. vannamei* has a distinctive metamorphosis consisting of nauplius, zoea, mysis and postlarvae<sup>18,19</sup> (Fig. 1A). The exoskeleton of shrimp is composed of the polysaccharide chitin, cuticle proteins and mineral deposits. It is a four-layered matrix including an epicuticle, exocuticle, endocuticle and epidermis<sup>20</sup>. According to the appearance of the epidermis, pigmentation, the formation of new setae, and the presence of matrix or internal cones in the setal lumen, the moulting cycle of shrimp can be divided into four recurrent stages: inter-moult, pre-moult, the moment of the moulting behaviour/ecdysis and post-moult<sup>21–23</sup> (Fig. 1B). For growth, *L. vannamei* need to shed and replace their old exoskeletons and synthesize a new one, and this process is frequently repeated during the life cycle<sup>24</sup>. Failure of moulting in the metamorphosis and mortality of the moulting shrimps are two important reasons for production reduction in aquaculture, which are all closely related to the formation and reconstruction of the exoskeleton. Furthermore, the decapods are a major source of waste shells in the world. It has been reported that world annual bioproduction of waste shells was estimated to be at 6~8 million tonnes<sup>25</sup>. Therefore, many shells urgently need a natural way to degrade. Studies on the development of the exoskeleton in *L. vannamei*, especially in exoskeleton formation, synthesis, degradation, calcification and hardening processes are very valuable for aquaculture, medicine and industry.

In previous studies, molecular research on the exoskeleton in shrimp were mostly concentrated on cloning genes. Several genes related to chitin synthesis<sup>26</sup>, degradation<sup>27</sup>, and mineral absorption<sup>28</sup> were amplified and sequenced. RNA interference showed that some factors may be associated with exoskeleton reconstruction, such as *CqGAP10*<sup>29</sup>, *CqGAP65*<sup>30</sup>, *LvGPS2*<sup>31</sup>, *CqMIH*<sup>32</sup>. However, the information collected using these methods is incomplete and provided only a fragmented picture. The molecular mechanisms under exoskeleton development

Sample	Raw Reads	Clean reads	Clean bases	Error(%)	Q20(%)	GC(%)
Lv_A_1	28596204	27067376	2.71 G	0.04	96.82	48.94
Lv_A_2	28596204	27067376	2.71 G	0.05	95.34	48.96
Lv_B_1	29925513	28399611	2.84 G	0.04	96.99	49.00
Lv_B_2	29925513	28399611	2.84 G	0.04	95.66	49.05

**Table 1.** Summary of transcriptome sequencing from *L. vannamei*. 1 means left reads, 2 means right reads.

remain poorly understood. The next-generation RNA-sequencing method, which can be used to compare a few genes to the entire transcriptome, could be a vital approach to understand this ontology. This technology has dramatically improved the efficiency and speed of analysis, especially for non-model species, and it is considered to be most powerful tool for dissecting gene networks associated with particular biological and developmental processes at the whole transcriptome level<sup>33–35</sup>.

In this study, we analysed the transcriptomic characterization of *L. vannamei* during seventeen dynamic developmental stages from zygote to adult growth-moult by RNA-sequencing method, including nine sequential early developmental stages (zygote, blastula, gastrula, limb bud embryo, larva in membrane, nauplius, zoea, mysis and post-larvae) and eight adult-moulting stages, including inter-moult (C), pre-moult (D0, D1, D2, D3 and D4) and post-moult (P1 and P2). Our analysis underscored the gene expression profiles in dynamic exoskeleton development and deduced transcriptional events involved in the exoskeleton formation, regulation, synthesis, degradation, mineral absorption/reabsorption, calcification and hardening of the exoskeleton. The characterization of the exoskeleton-related genes and pathways will enhanced our understanding of the molecular mechanisms underlying exoskeleton development.

## Result and Discussion

**Sequencing and assembly of the reference transcriptome.** As no genome of *L. vannamei* was available, a reference transcriptome was created first to perform digital gene expression profiling analysis. Using the Illumina RNA-seq technology, two library transcriptomes from adult *L. vannamei* were sequenced, and a total of 11.1 Gb clean bases were obtained (Table 1). Combining our results with the original data previously sequenced by our laboratory (SRR1460493, SRR1460494, SRR1460495, SRR1460504 and SRR1460505) provided a more comprehensive analysis for depicting the transcriptome profiles. As a result, a total of 117,539 unigenes were assembled, with half of the total assembly length (N50) of 1,327 bp and total nucleotides of 83,002,524 bp.

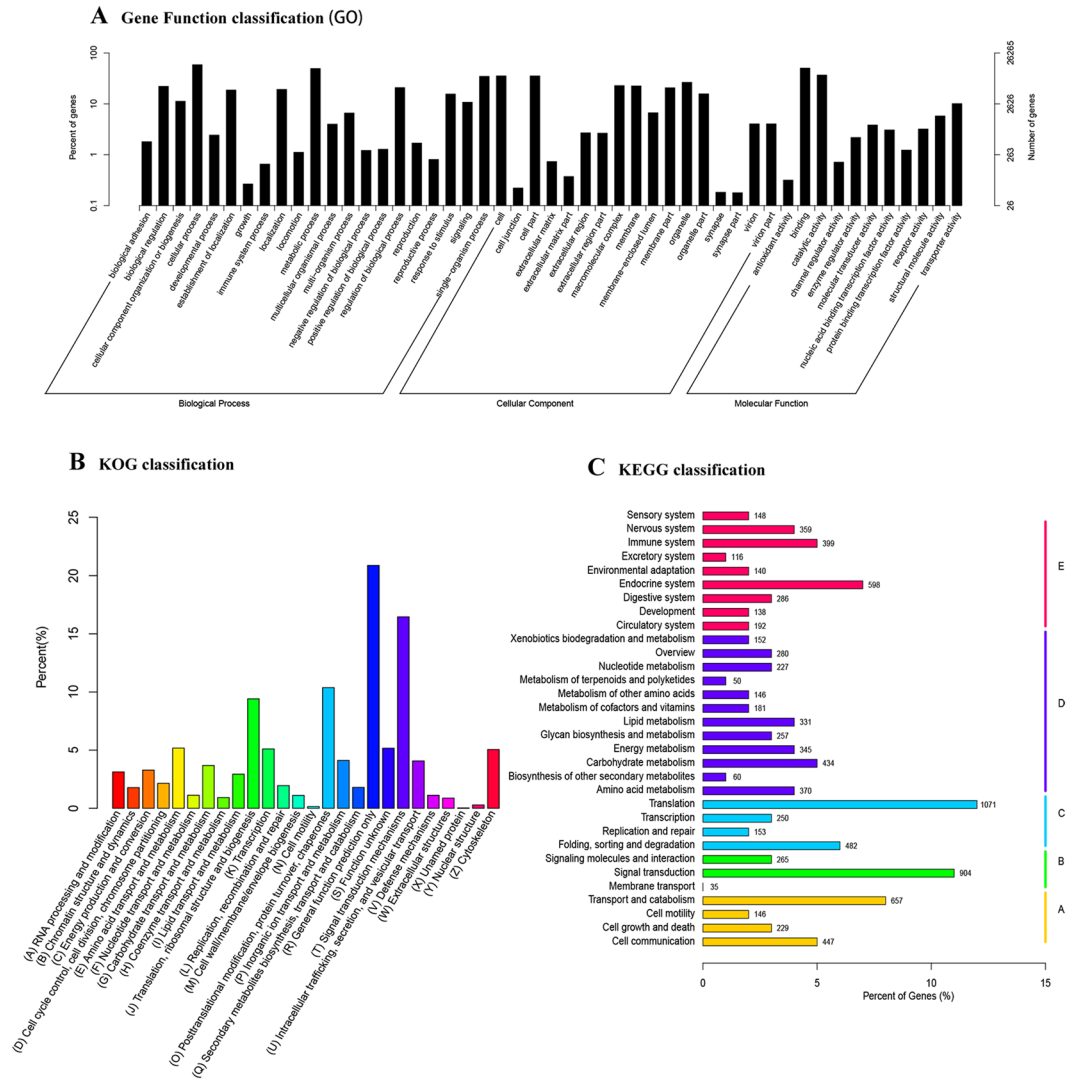
**Functional annotation and classification analysis.** For functional annotation of all unigenes, blast alignment against the National Center for Biotechnology Information (NCBI) non-redundant nucleotide/protein sequence database (NR and NT) and Swiss-Prot databases were performed (Supplementary Table S1). By blast searching with a cutoff E-value  $< 1e^{-5}$ , 32,398 (48.5%) unigenes were found as putative homologues in the NR database, 19,363 (29.0%) unigenes were found as putative homologues in the NT database, and 29,022 (43.4%) were found as putative homologues in the Swiss-Prot database. The E-value and similarity distribution of the best blast hits for unigenes are shown in Supplementary Figure S1A and S1B. The distribution of best hits over species for NR annotation was also analysed (Supplementary Figure S1C). *Daphnia pulex* (10.0%), *Tribolium castaneum* (6.3%) and *Pediculus humanus corporis* (3.9%) possessed the top three maximum unigenes numbers with NR annotation.

To acquire complete functional information and pathway-based analysis, Gene Ontology (GO), Clusters of Orthologous Group (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases were annotated for gene classification and pathway analysis. Through the GO analysis, a total of 158,366 GO terms were associated with all unigenes (Fig. 2A, Supplementary Table S2A). According to the secondary classification of the GO terms, all unigenes were sorted into 49 functional groups that belonged to three main GO categories: biological processes, molecular functions, and cellular components, comprised of 19,445 (34.1%), 21,510 (37.7%), and 16,122 (28.2%) unigenes, respectively (Fig. 2A). In the three categories, the seven subcategories, “cellular process” (15,425 unigenes), “binding” (13,264 unigenes), “metabolic process” (13,030 unigenes), “catalytic activity” (9,673 unigenes), “cell” (9,333 unigenes), “cell part” (9,328 unigenes), “single-organism process” (9,152 unigenes) were included in the top seven most-abundant sub-groups with more than 9000 unigenes.

The COG database is also an important classification system for functional annotation<sup>36</sup>. As for COG classification, 12,368 unigenes (10.52% of all) were classified into 26 functional categories (Fig. 2B, Supplementary Table S2B). The largest group was “general function predicted only” (2,583 unigenes), followed by “Signal transduction mechanisms” (2,036 unigenes), “Posttranslational modification, protein turnover, chaperones” (1,283 unigenes) and “Translation, ribosomal structure and biogenesis” (1,164 unigenes).

Pathway-based analysis helps to further understand the biological pathways of the genes. In the KEGG assignment, 8,498 unigenes were mapped to 259 different pathways (Supplementary Table S2C); a summary of the unigenes assigned to these pathways is shown in Fig. 2C. Of these annotated sequences in KEGG, 28.6% pathways were classified into the pathway related to metabolism, with most of them involved in oxidative phosphorylation (199 unigenes), purine metabolism (192 unigenes), carbon metabolism (169 unigenes), biosynthesis of amino acids (145 unigenes) and pyrimidine metabolism (102 unigenes).

**CDS prediction of the *L. vannamei* transcriptome.** As a species lacks genome information, full-length coding sequence (CDS) analysis is very important for molecular research in *L. vannamei*<sup>37–39</sup>. To better assign full-length gene, the CDSs of all unigenes were analysed by BLASTx and ESTScan. As a result, the CDS of 48,425



**Figure 2.** Gene ontology (GO), Clusters of Orthologous Group (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG) classification of all unigenes. (A) The GO classification results summarized in three main GO categories (cellular component, molecular function, and biological process). The x-axis represents the GO ontology. The left y-axis indicates the percentage of unigenes. The right y-axis indicates the number of unigenes. (B) COG classification of all unigenes. The columns represent the number of unigenes in each subcategory. (C) KEGG classification of all unigenes. The columns represent the number of unigenes in each subcategory.

unigenes (41.2% of all unigenes) were predicted (Supplementary Table S3). Among them, 19,336 CDSs were extracted from BLASTx against the NR and Swiss-Prot databases. The remaining 29,089 CDSs, with no specific matches to the above databases, were predicted by ESTScan. The distribution of predicted CDS lengths is shown in Table 2, with the full-length coding sequence of nearly half of the unigenes obtained. These results will be of great interest for further analysis in *L. vannamei* as a prelude to functional investigations.

**Digital gene expression profiling analysis of dynamical exoskeleton development.** For analysis of the entire exoskeleton developmental process in *L. vannamei*, we sequenced seventeen sequential developmental stages from zygote to adult moult in *L. vannamei*, including nine successive early stages (zygote, blastula, gastrula, limb bud embryo, larva in membrane, nauplius, zoea, mysis and post-larvae) and eight moulting stages, including inter-moult (C), pre-moult (D0, D1, D2, D3 and D4) and post-moult (P1 and P2) by digital gene expression profiling technique.

As a result, seventeen stages produced a total of 233.78 million raw reads with 98.5% Q20 values using the Illumina HiSeq 2500 platform. After removal of low-quality reads, adaptor sequences, duplicate sequences and ambiguous reads, 231.44 million high-quality clean reads (15.73 Gb) remained. Then we mapped the filtered high-quality reads to the reference transcriptome, and 83.11~93% filtered reads were mapped. An overview of the sequencing and assembly results is summarized in Table 3.

Gene length(bp)	Number	Percent (%)
51–200	16,677	34.4
201–500	15,278	31.6
501–1000	6,632	13.7
1001–1500	3,921	8.1
1501–2000	2,285	4.7
>2000	3,632	7.5
ALL	48,425	100

**Table 2.** The length distribution of coding sequence (CDS) predictions. For 48,425 of the predicted full-length coding sequences (CDSs), the size distribution can be divided into six species. The total number and the percentage of all genes are presented in the table.

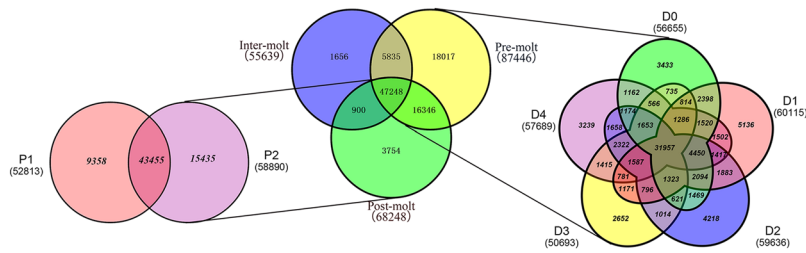
Sample	Raw Reads	Q20(%)	Clean reads	Total mapped	Unigene number	Unigene average length	Unigene N50
Zygo	9074040	98.61	8938449	8249875(92.30%)	27982	1359	2523
blas	11607799	98.45	11464762	10501144(91.59%)	32982	1229	2402
gast	10995688	98.29	10856202	9669177(89.07%)	35979	1166	2323
Lbe	9754060	98.29	9597397	8673620(90.37%)	35526	1174	2298
Lim	8429023	97.83	8335436	7591593(91.08%)	37455	1286	2364
Nauplius	8926413	98.1	8789630	8174251(93.00%)	37531	1337	2348
Zoea	9997945	98.27	9823628	9135914(93.00%)	41711	1299	2264
Mysis	7537908	98.1	7420229	6841843(92.21%)	42655	1319	2239
Postlarvae	8090871	98.1	7961868	7267765(91.28%)	46853	1246	2161
C_1	8462163	98.64	8378139	7205223(86.00%)	50205	1136	2152
C_2	8356671	98.68	8281105	7228674(87.29%)	52156	1112	2109
D0_1	9729632	98.63	9658189	8372908(86.69%)	56217	1063	2059
D0_2	9001963	98.74	8949633	7698406(86.02%)	50493	1133	2147
D1_1	7069133	98.52	7010089	5949698(84.87%)	53411	1102	2085
D1_2	8289817	98.58	8237292	9226254(86.42%)	54246	1085	2097
D2_1	10758027	98.66	10676498	9226254(86.42%)	55917	1060	2059
D2_2	9878904	98.74	9801263	8596987(87.71%)	56264	1068	2070
D3_1	8732484	98.65	8665896	7228584(83.41%)	43427	1189	2256
D3_2	8414766	98.68	8350596	6939910(83.11%)	48107	1139	2160
D4_1	10154005	98.79	10078307	8667790(86.00%)	55468	1064	2066
D4_2	9650580	98.81	9598685	8199145(85.42%)	54727	1082	2077
P1_1	10570861	98.71	10500814	9255363(88.14%)	51858	1117	2126
P1_2	10017393	98.69	9937550	8650085(87.04%)	51350	1122	2133
P2_1	10205898	98.78	10135988	8787490(86.70%)	57581	1047	2030
P2_2	10073944	98.70	9993119	8832160(88.38%)	55285	1071	2068

**Table 3.** Summary of sequencing and assembly of the transcriptome.

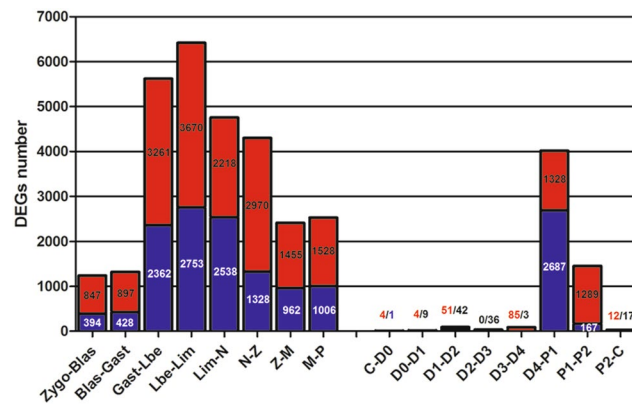
**Expression profile and differentially expressed gene (DEG) analysis.** The expression value of each unigene was estimated by the fragments per kilobase of exon per million fragments mapped (FPKM) method. Among the seventeen developmental stages, the number of expressed genes (FPKM  $\geq$  0.3) were 27,982 (zygote), 32,982 (blastula), 35,979 (gastrula), 35,526 (limb bud embryo), 37,455 (larva in membrane), 37,531 (nauplius), 41,711 (zoea), 42,655 (mysis), 46,853 (post-larvae), 55,639 (C), 56,655 (D0), 60,115 (D1), 59,636 (D2), 50,693 (D3), 57,689 (D4), 52,813 (P1) and 58,890 (P2). Gene expression during seventeen stages is summarized in Supplementary Table S4 for early developmental stages and Supplementary Table S5 for adult moulting. It showed that the expressed genes were increasing gradually in early development. Also, Venn diagrams showed the number of common and unique genes in different moulting stages (Fig. 3). In this study, differentially expressed gene (DEG) was defined as the fold change of the normalized FPKM values of at least two-fold as well as a false discovery rate (FDR)  $\leq$  0.001. As a result, a set of 19,809 genes exhibited significant differential expression between two consecutive developmental stages (Fig. 4, Supplementary Table S6).

To validate the sequencing data, eight DEGs were selected for real-time qPCR both in the early development and moulting stages. Expression patterns of the selected genes determined by real-time qPCR (Supplementary Figure S2) had good consistency with the RNA-seq results, which corroborated our results.





**Figure 3.** A venn diagram showing the genes expressed in three major stages (inter-moult, pre-moult and post-moult) of the moulting process. The details of expressed gene among eight moulting stages are shown.

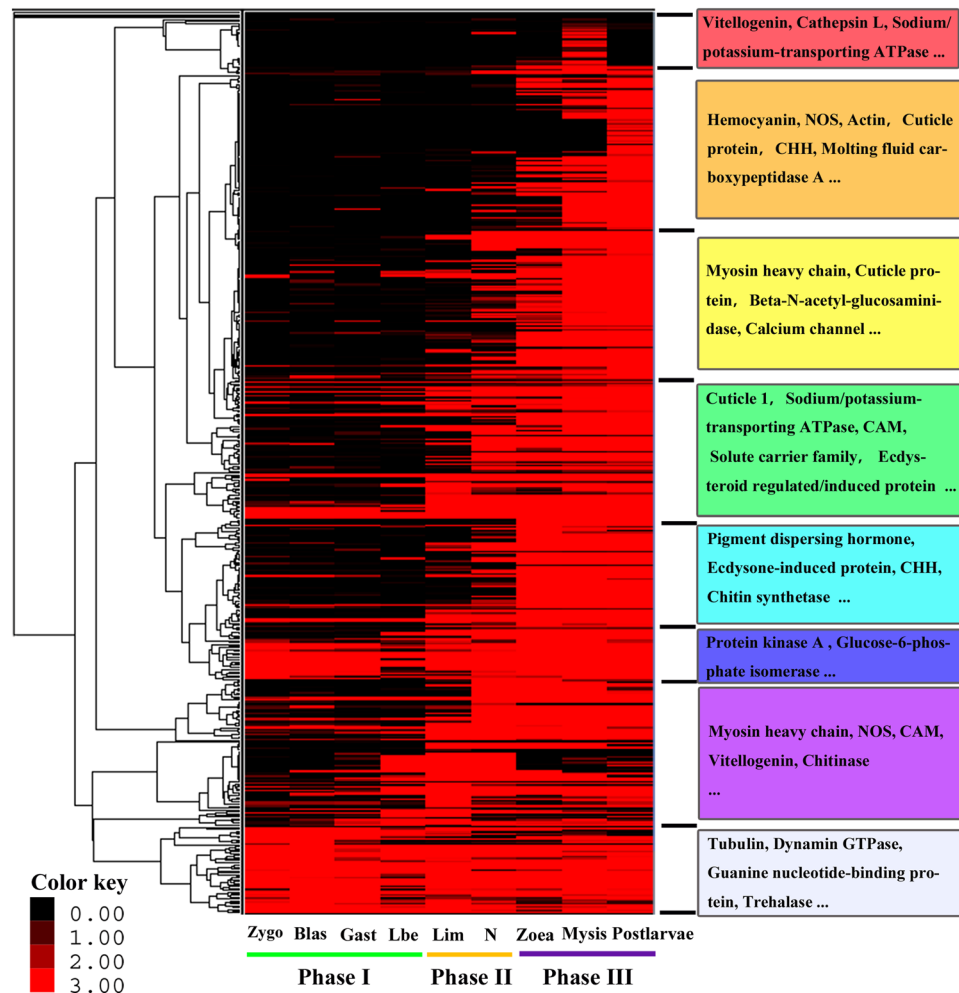


**Figure 4.** Differentially expressed genes (DEGs) detected between adjacent stages. The numbers of up-regulated and down-regulated genes are revealed by red and blue, respectively.

**Expression profile of genes related to exoskeleton development.** Based on the functional gene profiling analysis, we summarized the genes related to exoskeleton formation, development and reconstruction from the following perspectives, including: (i) Cuticle protein (cuticle protein, chitin binding protein, structural constituent of cuticle), (ii) Moulting-hormone (moulting-inhibiting hormone, crustacean hyperglycemic hormone *etc.*), (iii) Exoskeleton degradation (chitinase, beta-N-acetyl-glucosaminidase, acetyl-hexosaminidase), (iv) Mineral absorption (solute carrier family, ferritin heavy chain *etc.*), (v) Mineral reabsorption (sodium/potassium-transporting ATPase, clathrin heavy chain *etc.*), (vi) Exoskeleton synthesis (chitin synthetase, Glutamine:Fructose-6-phosphate aminotransferase *etc.*), (vii) Ion channel (calcium channel, potassium channel, chloride channel *etc.*), (viii) Moulting signaling pathway (calmodulin, nitric oxide synthase, moulting defective family member *etc.*) and (ix) Late gene of moulting (actin/tubulin/myosin, hemocyanin *etc.*). As a result, 603 unigenes related to exoskeleton formation, development and reconstruction were identified (Supplementary Table S7). We performed hierarchical clustering of all exoskeleton related genes to examine the similarity and diversity of expression profiles (Fig. 5). The hierarchical clustering generated a global view of the gene expression pattern related to exoskeleton development, indicating that the majority of genes were up-regulated during the middle (lim/nauplius) and late of early developmental stages (zoeta/mysis).

**Exoskeleton formation and development during early developmental stages in *L. vannamei*.** During the early developmental stages of *L. vannamei*, both morphological and physiological features change dramatically<sup>19</sup> (Fig. 1A), and this developmental mode provides a unique opportunity to examine exoskeleton formation and development<sup>40, 41</sup>. Based on exoskeleton development, the early developmental stages could be divided into three phases: Phase I, where there is no exoskeleton cover around the body, including zygote, blastula, gastrula, and limb bud embryo; Phase II, the initiation stage of exoskeleton formation, including larva in membrane and nauplius; Phase III, the metamorphosis process, in which there is rapid growth of the exoskeleton. The developmental stages in this phase include zoea, mysis and postlarvae. The transcriptome profile showed a rather different expression pattern among different early stages:

(i) Phase I was significantly enriched in histones, Zinc finger proteins, methyltransferase, and transcription initiative factors, which is related to maternal genes or embryo development; few exoskeleton-related genes were expressed in this phase. Additionally, it appeared that transcriptional repressors were enriched in this stage, and the expression levels of them reduce gradually during the early development. The decrease of the expression levels of repressors meant an active transcription level of metabolism<sup>42</sup>, which was in accordance with our results that gene expression was increasing gradually in early development. (ii) Phase II is the triggering stage of exoskeleton formation, and the specific expressed genes in this phase might be important for exoskeleton formation. Genes of Hox, hormone, transposon-encoded related proteins from larva in membrane to nauplius were significantly enriched, which were related to body plan and hormone regulation<sup>43, 44</sup>; additionally, it appeared that



**Figure 5.** Hierarchical clustering of all exoskeleton-related genes. The clustering indicates similar expression patterns among the developmental stages (x-axis) and among the genes (y-axis). Colours represent the gene expression levels from black (low) to red (high). For each cluster, the major enriched genes are listed.

some important exoskeleton-related genes emerged in this phase. For example, the cuticle proteins, which are an important composition of the exoskeleton, were found to be sharply increased to thousands of times in this phase (Table 4). Moreover, moulting-inhibiting hormone (MIH) gene and crustacean hyperglycemic hormone (CHH) gene were triggered their transcription during Phase II. It is now well established that MIH and CHH peptide families serve as key regulators of hormones controlling the moulting process<sup>45, 46</sup>. Their initial expression represents the beginning of moulting-regulation<sup>46</sup>. Therefore, Phase II was the important time node for exoskeleton formation and hormone-regulation, and transcription of the exoskeleton-related genes were triggered in this phase.

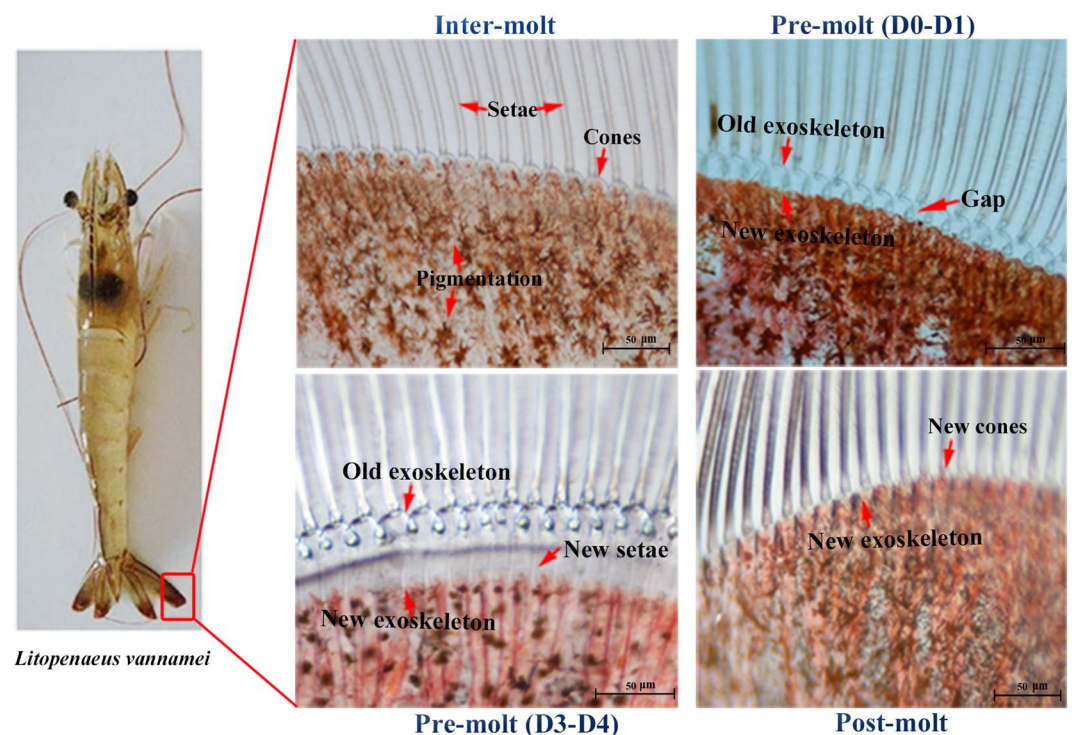
(iii) In the metamorphosis stages (Phase III), a large number of genes related to trypsin, transcription factors, ion channels and receptors, transmembrane receptors, and actin/tubulin/myosin were found to be expressed, which was consistent with some characteristics when converting from larvae to adult, such as larval feeding habits<sup>47</sup> and formation of the eye<sup>48</sup>. We found that the genes related to chitinase, calcification of the exoskeleton and chitin metabolism occurred at Phase III. Moreover, the downstream hormone-factors under the regulation of MIH and CHH, including ecdysteroid regulated-like protein, ecdysone receptor (ECR), moulting fluid carboxypeptidase, ecdysone-induced proteins *etc.* appeared in these stages. These factors are essential for exoskeleton development and are especially important for synthesis, degradation, regulation, calcification and hardening in the moulting cycle.

In summary, the transcriptome results showed a distinct expression time and profile of developmental genes in the early development of *L. vannamei*. Also, it suggested that the exoskeleton formation was mainly initiate in lim/nauplius stages (phase II). Since factors related to exoskeleton synthesis, degradation, regulation, calcification and hardening were mainly expressed in the phase III, it indicated that an advanced system for exoskeleton development was gradually perfected in late phase.

**Analysis of the exoskeleton degradation in the moulting.** In shrimp, the exoskeleton reconstruction is a stepwise process that has several landmark events; i.e., apolysis (the separation of the cuticle from the

Gene ID	NR Description	zygo	blas	gast	lbe	lim	Nauplius	Zoea	Mysis	Postlarvae
c56881_g2	arthrodial cuticle protein AMP8.1	1.59	3.10	0.00	0.00	<b>24.42</b>	<b>3695.25</b>	1016.29	1194.86	2757.89
c76513_g1	cuticle protein	4.86	2.27	0.83	0.15	<b>2.89</b>	<b>3096.47</b>	143.43	568.54	525.30
c56881_g1	arthrodial cuticle protein AMP8.1	2.19	1.64	0	0.66	<b>35.78</b>	<b>3009.29</b>	575.29	923.33	1855.53
c62024_g1	cuticle protein CUT8	0.00	1.24	0.35	0.00	<b>0.33</b>	<b>1479.42</b>	337.21	1396.92	2408.11
c39382_g1	cuticular protein RR-2 motif 78 precursor	0.65	0.90	0.35	0.00	<b>1.35</b>	<b>948.43</b>	61.46	86.86	401.31
c65344_g1	cuticle protein	0.66	0.26	0	0.00	<b>0.00</b>	<b>841.99</b>	166.13	605.69	1011.85
c77367_g1	calcified cuticle protein CP14.1	0.41	0.77	0.35	0.00	<b>7.17</b>	<b>754.50</b>	457.04	479.15	993.26
c76821_g2	putative cuticle protein (Cpr64Ac)	0.44	1.35	0.00	0.00	<b>8.84</b>	<b>504.28</b>	0.41	0.51	0.00
c66289_g1	early cuticle protein 6	0.00	0.55	1.89	0.45	<b>0.00</b>	<b>247.32</b>	139.12	174.46	382.73
c73025_g1	early cuticle protein 6	0.21	0.49	0.00	0.00	<b>0.00</b>	<b>241.66</b>	29.55	137.76	772.74
c74284_g1	structural constituent of cuticle	0.00	0	0.73	0.20	<b>0.00</b>	<b>233.06</b>	16.61	25.36	315.62
c66706_g1	structural constituent of cuticle	3.54	113.39	264.43	148.89	<b>496.93</b>	<b>596.02</b>	21.09	5.46	1.86
c67645_g1	calcified cuticle protein CP8.5	0.00	0.33	0.00	0.00	<b>0.00</b>	<b>97.619</b>	20.40	10.42	18.36

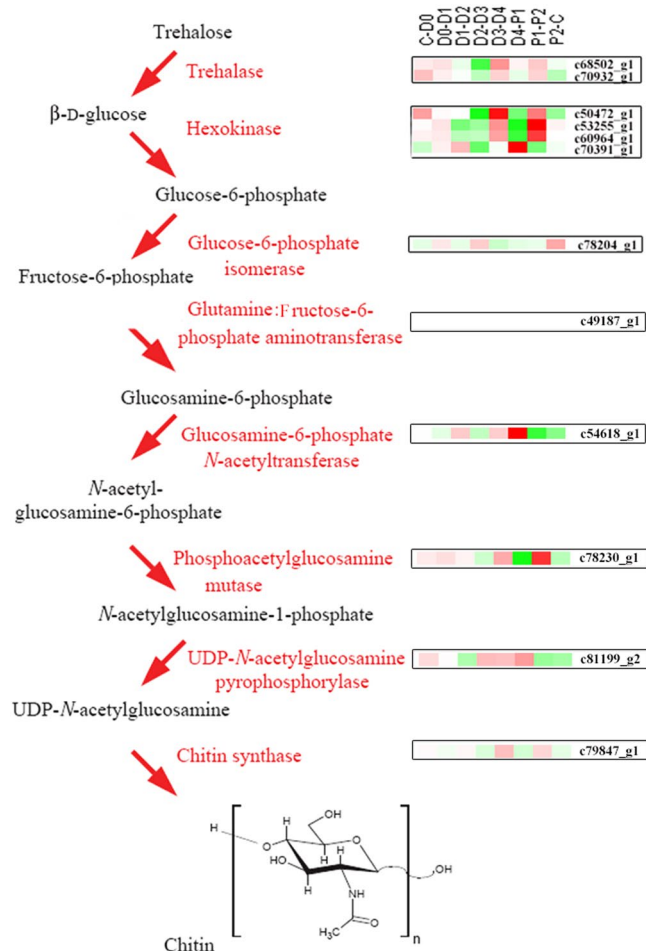
**Table 4.** A heatmap showing the expression profile of cuticle proteins (part) during the early development of *L. vannamei*. The gene ID, expression profile (FPKM value) and NR description is depicted. Bold text denotes up-regulation at phase II (larva in membrane and nauplius).



**Figure 6.** Morphological changes of shrimp uropods during moulting stages under a light microscope.

epidermis), new exoskeleton synthesis and shedding of the old exoskeleton<sup>49–51</sup>. As the major component that breaks down the glycosidic bonds in chitin and partially digested old exoskeleton, chitinase was the major downstream factors of ecdysone to break off the old exoskeleton<sup>50</sup>. Forty-four unigenes related to chitinase were identified in the moulting cycle of *L. vannamei* (Supplementary Table S7). It revealed that these genes corresponding to chitin degradation were preferentially expressed during the D0-D1 and period D4. The activity of chitinase in the early pre-moult phase might be associated with the separation of the old and new exoskeleton (Fig. 6).





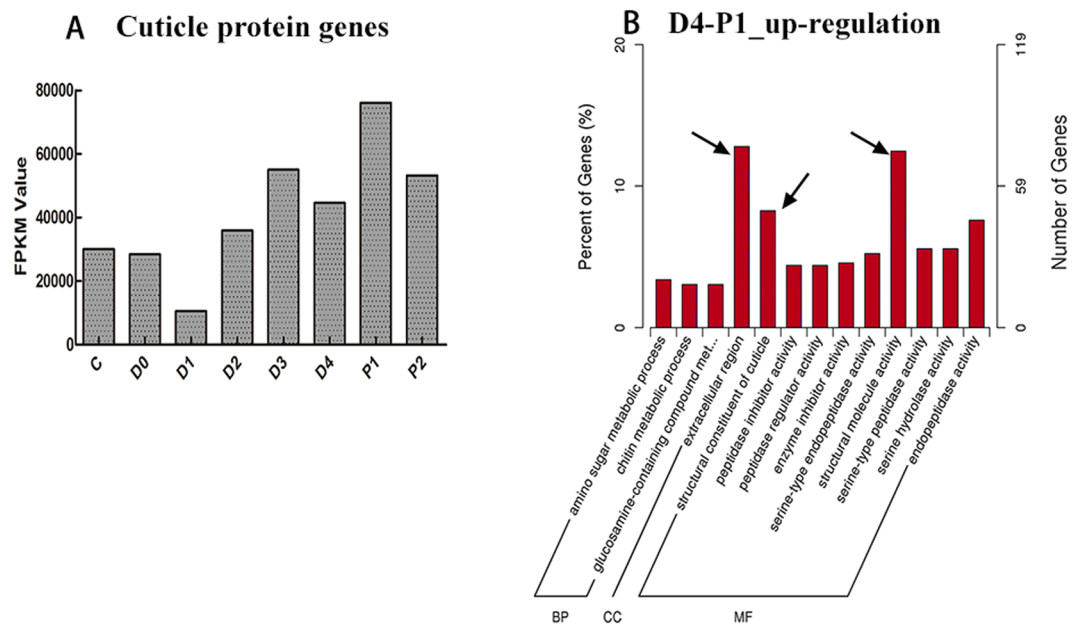
**Figure 7.** Expression profile of chitin synthesis pathways using the insect analogy. It displays expression patterns for all chitin synthesis factors in each comparison. The colour key represents FPKM normalized log<sub>2</sub> transformed counts. Red represents up-regulated expression and green represents down-regulated expression. The diagrammatic representation is based on previous studies in insects<sup>57</sup>.

Also, during ecdysis behaviour, the shrimp needed to digest the old exoskeleton for exfoliation. This might be the reason for obvious up-regulation of chitin degradation related genes in period D4. After moulting (P1), the expression of chitin degradation genes reaches the minimum. Similar results regarding the activity of chitinase during the post-moult phase were found in the Antarctic krill *Euphausia superb*<sup>52</sup>, the fiddler crab *Uca pugilator*<sup>53</sup> and the whiteleg shrimp<sup>26</sup>, which have up-regulated expression of chitinase in the pre-moult phase.

These results provided significant clues for a better understanding of the regulatory mechanisms under exoskeleton degradation in *L. vannamei*, and revealed that the D0-D1 and D4 periods were the important time nodes for exoskeleton degradation and indicated some key genes that have potential value for scientific research and applications.

It is known that world annual bioproduction of waste chitin shells is enormous, and the traditional method for extracting chemicals from waste shells is destructive, wasteful and expensive<sup>25</sup>. For example, to produce 1 kilogram of chitosan from shrimp shells, more than 1 tonne of water is need and as well as a large amount of 40% sodium hydroxide solution. Thus, developing a natural way to degrade the waste shells will be promising and necessary<sup>54,55</sup>. From our results, several chitin degradation-related genes with high expression in the moulting process could be targeted candidates. For instance, the FPKM values of the c78849\_g2 (*chitinase*) increased more than 2,500 at the D3-D4 period, which might be an important factor for chitin degradation. In future practical applications, by a proper combination of practical manufacturing technology, such as nano-material and film-material<sup>56</sup>, a novel method for degradation of waste shells could have great potential.

**Analysis of chitin synthesis in *L. vannamei*.** Chitin, a polymer of N-acetyl-β-D-glucosamine, is the major component of the shrimp exoskeleton<sup>26</sup>. As a framework material, chitin supports the exoskeleton of shrimp. Moreover, the moulting of shrimp is strictly dependent on the capability to synthesize chitin<sup>57</sup>. However, knowledge of its biosynthesis is still fragmentary in crustaceans, and is far behind this analysis in insects<sup>58-60</sup>. Based on the transcriptome results, we searched for chitin synthesis pathways by referring to the insect analysis<sup>57</sup>. Identifying the pathways of chitin synthesis in the *L. vannamei* analogy of the insect seemed to be considerable conserved (Fig. 7). Expression profile analysis showed that an up-regulated expression trend was found in the



**Figure 8.** Expression profiles of cuticle proteins and GO enrichment of up-regulated differentially expressed genes (DEGs) in D4-P1. **(A)** Expression profiles of cuticle protein genes. Row axis represents the developmental stages. Columns represent the total FPKM value of genes. **(B)** The GO terms of up-regulated DEGs in D4-P1. The arrows indicate the terms related to exoskeleton. Extracellular region (GO:0005576) with 76 DEGs are dominant in all categories. Structural molecule activity (GO:0005198) and Structural constituent of cuticle (GO:0042302) consisted of 74 and 49 DEGs that were dominant in categories of molecular function. The x-axis represents the name of GO subcategories. The left y-axis indicates the percent of genes. The right y-axis indicates the number of DEGs expressed in a given sub-category. The English in this document has been checked by native English-language editors. And the logical presentation of ideas and the structure of the paper were also checked during the editing process. The edit was performed by professional editors at American Journal Experts ([www.aje.com](http://www.aje.com)), which is recommended by editorial staffs of Scientific Reports.

chitin pathway during the D4 and P2 period. Among the pathways, the expression of a chitin synthetase gene (c79847\_g1), which encodes the key enzyme in chitin synthesis, significantly increased expression to the peak in the D4 and P2 periods. It is well known that partial digestion of the exoskeleton in the pre-moult phase is accompanied by the synthesis of new underlying exocuticle and epicuticle<sup>61,62</sup>. Therefore, this up-regulation of the chitin synthesis pathway, together with the functional annotation of the physiological process, suggested that the synthesis of the new exoskeleton was mainly occurring in the D4 and P2 periods. These results provide a vital resource for further research into chitin synthesis in shrimp. Moreover, the temporal stage-related scheduling of the expression of the factors may serve as markers in future studies for monitoring the chitin synthesis process in *L. vannamei*.

**Mineral absorption/reabsorption and the exoskeleton hardening process during moulting.** There is mineralization in shrimp exoskeleton, typically calcium carbonate, which gives the exoskeleton mechanical rigidity<sup>2</sup>. It is well known that there is a structure called gastrolith in crayfish that can collapse into the stomach and dissolve to promote minerals for the new exoskeleton<sup>63,64</sup>. However, there is no such structure that has been reported in shrimp, thus the mineral for moulting was supposed to be absorbed from seawater or reabsorbed in *L. vannamei*<sup>65</sup>. The annotated transcriptome was searched for mineral absorption and reabsorption encoding genes using the KEGG database. As a result, ko04978 “Mineral absorption” and ko04961 “Endocrine and other factor-regulated calcium reabsorption” were analysed, and the expression levels of the members are summarized in Supplementary Table S8. The results showed that mineral absorption was mainly occurring in the D1 and D4 period, and reabsorption was mainly occurring in the D2 and D4 period. During the early stage of pre-moult, minerals derived from the degradation of the old exoskeleton are not available<sup>66,67</sup>, hence the mineral absorption (D1) was earlier than the reabsorption of minerals (D2). In addition, they all followed the initial degradation in D0-D1, which was thought to be the onset of moulting.

As the protein embedded in the chitin, the physical properties of the exoskeleton are determined largely by the cuticle proteins<sup>51</sup>, especially the calcified cuticle proteins, which could support the hardness of the exoskeleton<sup>68</sup>. It was reported that proteins with the cuticle\_1 domain were associated with exoskeleton hardening in decapod crustaceans<sup>69,70</sup>. In terms of RNA-seq results, a total of seventy-three cuticle protein related genes were identified from our transcriptome data (Supplementary Table S7), and the expression pattern on the whole is summarized in Fig. 8A. We found that the expression of cuticle proteins accumulated at a peak after ecdysis (P1). GO analysis also showed an obvious up-regulation of GO terms related to the structure (“GO:0005576” extracellular region, “GO:0042302” structural constituent of cuticle, “GO:0005198” structural molecule activity) between D4 and P1

Stage	Transcriptional events
D0-D1	Initiative transcription for exoskeleton degradation
D1	Mineral absorption
D2	Mineral reabsorption
D4	Important period for old exoskeleton shed, new chitin synthesis; mineral absorption/reabsorption
P1	Period for new cuticle protein synthesis
P2	Period for new chitin synthesis

**Table 5.** A summary of transcriptional events applicable to the early development and moulting cycle in *L. vannamei*

(Fig. 8B). The abundant expression of cuticle proteins and structure related terms in P1 was considered to be the method for supporting and hardening the soft exoskeleton after moulting.

In summary, a coordination of exoskeleton degradation, chitin synthesis and exoskeleton hardening combined with mineral absorption and reabsorption were collectively occurring during moulting with a programmed procedure and strict control that beginning with the initial degradation in D0-D1, then mineral absorption in D1, followed by mineral reabsorption in D2 for mineral reserve. Then the synthesis of new chitin (D4) occurred, which was the main component of the outer layer-epicuticle<sup>71</sup>. The cuticle proteins, the important parts of the inner three layers<sup>66</sup>, were produced in P1. Finally, the new chitin continuously synthesized (P2) for a complete exoskeleton. The transcriptome analysis on the exoskeleton reconstruction scenario in growth-moulting was summarised in Table 5. This high degree of order is the key element in the many occurrences of moulting repeated during the life cycle of shrimp.

## Conclusion

This paper presents the first transcriptome analysis to cover the entire and dynamic exoskeleton formation, development and reconstruction in shrimp, including continuous early development stages and the entire moulting cycle. It establishes a transcriptional scenario for exoskeleton development and deduces the transcriptional events involved from the several perspective, including exoskeleton formation, regulation, synthesis, degradation, mineral absorption/reabsorption, calcification and hardening, *etc.* Moreover, it provides useful insights into the molecular processes underlying these transcriptional events. It indicates that a complex regulatory network of genes is involved in exoskeleton formation, development and construction in *L. vannamei*. Moreover, these genes show a strong coordination with a high degree of order during development. These results will provide molecular evidence for enhancing our understanding of exoskeleton development in shrimp and may possibly serve as a blueprint for other arthropods in future research.

## Methods

**Sample collection and RNA isolation.** Healthy adult *Litopenaeus vannamei* with an average body length of 14–16 cm were collected from culture pond of the laboratory for *de novo* sequencing. They were from the same generation and had cultured more than half year to minimize the smallest effects by the environment. Two *de novo* libraries were constructed with mixed samples.

The early developmental samples were collected from a farm in Wenchang, Hainan, China. A total of nine samples, including zygote, cell, blastula, gastrula, limb bud embryo, larva in membrane, nauplius, zoea, mysis and postlarvae, were collected as the pool based on both morphological classification and physiological characters (Fig. 1A). They were reared in a 25 m<sup>3</sup> indoor ponds with seawater at 31 °C, salinity of 2.5%. Embryos and larvae were collected with screen mesh when 90% of the population had reached the objective stage.

The samples at different moulting stages were collected from the whole individual shrimp with an average body length of 10–12 cm, which were cultured in the pond of our laboratory (seawater at 25 °C, salinity of 3%). The moulting stages including eight periods, inter-moult (C), pre-moult (D0, D1, D2, D3 and D4) and post-moult (P1 and P2) were separately collected, identified by the formation of new setae, the appearance of new epidermis, and the presence of matrix or internal cones in the setal lumen<sup>21</sup>. Each stage used one individual in order to get the accurate moulting-stage, and each stage was analysed with biological replicates to ensure the accuracy. These materials were immediately frozen in liquid nitrogen and kept at –80 °C. All procedures involving animals were performed per the institutional ethic committee guidelines of Institute of Oceanology Chinese Academy of Sciences (IOCAS).

Total RNA was extracted from samples using the Trans-up RNA reagent according to the manufacturer's protocol. DNase digestion was completed during RNA purification using the RNase Free DNase Set (QIAGEN). The yield, purity and integrity of total RNA were assessed by electrophoresis in 1% agarose gels and quantified with a NanoDrop 2000 spectrophotometer (Thermo Scientific, USA) and Agilent 2100 bioanalyser.

**Library construction, sequencing and assembly.** The RNA-sequencing libraries were prepared according to Illumina's RNA sample preparation protocol. Magnetic beads with oligo dT were used to isolate poly (A) mRNA from total RNA. Then, fragmentation buffer was added for interrupting mRNA to short fragments. These short fragments were used as templates for random hexamer-primed synthesis of first-strand cDNA. The second-strand cDNA was synthesized using Buffer, RNase H, dNTPs and DNA polymerase I. A paired-end library was synthesized using the Genomic Sample Preparation Kit (Illumina). Short fragments were purified

with a QiaQuick PCR extraction kit (Qiagen, Germany). For *de novo* sequencing, the constructed sequencing library was sequenced using Illumina HiSeq 2500 under the  $2 \times 100$  bp paired-end model. For digital gene expression profiling, agarose gel electrophoresis was used to select the fragments with about 50 bp in size. Then the short fragments were connected with sequencing adapters. Finally, the libraries were sequenced using Illumina HiSeq™ 2500 (Novogene, Beijing).

Raw image data was transformed into raw reads by base calling, and stored in fastq format. After removing empty reads, adaptor sequences and low quality sequences, the raw reads were transformed into clean reads. The quality control of the reads was checked with FASTQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). A reference transcriptome was assembled using a subset of reads from seven libraries, which were combined with the original data previously sequenced by our laboratory (SRR1460493, SRR1460494, SRR1460495, SRR1460504 and SRR1460505). The *de novo* transcriptome was assembled using Trinity (version trinity-naseq\_r2013\_08\_14) to obtain transcripts with min\_kmer\_cov set to 2 and all other parameters set to default<sup>72</sup>. Unigenes were obtained after redundant transcripts were removed by clustering in Tgicl. Then, reads from each of the development stages were mapped back to the reference transcriptome using the Bowtie2 aligner (parameters: mismatch 0) and converted into fragments per kilobase of transcript per million base pairs sequenced (FPKM)<sup>73</sup>.

**Gene annotation, classification and CDS prediction.** Gene function and classification were analysed based on searches against the following databases: National Center for Biotechnology Information (NCBI) non-redundant nucleotide/protein sequence database (NR and NT), SwissProt database, Gene Ontology (GO), Clusters of orthologous groups of proteins (COG), Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The software and methods for annotation and classification were: NCBI blast 2.2.28 + for NR, NT and SwissProt annotation with the E-value threshold was  $< 1e^{-5}$ . Blast2GO (version 3.0) (<https://www.blast2go.com/>) was used for GO annotation with E-value  $< 1e^{-6}$ <sup>74</sup>; WEGO was used to categorize GO term to view the distribution<sup>75</sup>. KAAS (<http://www.genome.jp/tools/kaas/>) and KEGG Automatic Annotation Server (<http://www.genome.jp/kegg/>) were carried out for KEGG annotation and classification<sup>76</sup>. Quantification of gene expression levels was estimated by RSEM (version v1.2.15)<sup>77</sup>. Differential expression analysis between adjacent developmental stages was implemented using the DESeq R package<sup>78</sup>. The absolute value of  $\log_2$ Ratio  $\geq 1$  and false discovery rate (FDR)  $\leq 0.001$  were used as the threshold to judge the significance of each gene expression difference. The heat maps were grouped together according to their FPKM values by Cluster 3.0<sup>79</sup> and visualized by TreeView 1.6<sup>80</sup>. Moreover, in the priority order of NR and SwissProt, the best-matched fragments of the annotated unigenes were predicted for coding regions (CDS) using blastx with E value  $< 0.00001$ . The CDS of the remaining un-annotated unigenes were predicted using ESTScan.

**Real-time qPCR amplification for RNA-Seq verification.** To verify the accuracy of RNA-seq data, eight DEGs were selected for verification by real-time qPCR amplification in the early development and moulting stages respectively, with gene-specific primers designed by Primer 5.0 (Supplementary Table S9). Myosin light chain (c26953\_g1) was used as the control reference gene, and relative gene expression levels were calculated using the comparative Ct method with the formula  $2^{-\Delta\Delta Ct}$ <sup>81</sup>. All samples were run in triplicates in separate tubes and each cDNA sample was run in duplicate (Supplementary Figure S2).

**Data availability.** The sequence data in the present study have been released to the NCBI Sequence Read Archive database (<http://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=studies>) and the accession numbers of samples are SRS1145819 and SRS1828174 for two reference transcriptome, SRP094135 for nine early developmental stages and SRP061180 for eight adult moulting stages.

## References

- Regier, J. C. *et al.* Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* **463**, 1079–1083, doi:10.1038/nature08742 (2010).
- Chen, P. Y., Lin, A. Y., McKittrick, J. & Meyers, M. A. Structure and mechanical properties of crab exoskeletons. *Acta biomaterialia* **4**, 587–596, doi:10.1016/j.actbio.2007.12.010 (2008).
- Moussian, B. Recent advances in understanding mechanisms of insect cuticle differentiation. *Insect biochemistry and molecular biology* **40**, 363–375, doi:10.1016/j.ibmb.2010.03.003 (2010).
- Promwikorn, W., Boonyoung, P. & Kirirat, P. Histological characterization of cuticular depositions throughout the molting cycle of the black tiger shrimp (*Penaeus monodon*). *Songklanakarinn Journal of Science & Technology* **27**, 499–509 (2005).
- Skinner, D. M. Molting and Regeneration. *Integument Pigments & Hormonal Processes*. 43–146 (1985).
- Melnick, C. A., Chen, Z. & Mecholsky, J. J. Hardness and toughness of exoskeleton material in the stone crab, *Menippe mercenaria*. *Journal of Materials Research* **11**, 2903–2907, doi:10.1557/JMR.1996.0367 (1996).
- Michels, J. & Gorb, S. N. Detailed three-dimensional visualization of resilin in the exoskeleton of arthropods using confocal laser scanning microscopy. *Journal of microscopy* **245**, 1–16, doi:10.1111/j.1365-2818.2011.03523.x (2012).
- Tamanis, E., Mihailova, I., Valainis, U. & Gerbreder, V. Specific characteristics of chitin exoskeleton of leaf weevil *phyllobius maculicornis* (germar, 1824). *Acta Biologica Universitatis Daugavpiliensis* **12**, 58–64 (2012).
- Michels, J., Vogt, J. & Gorb, S. N. Tools for crushing diatoms—opal teeth in copepods feature a rubber-like bearing composed of resilin. *Scientific reports* **2**, 465, doi:10.1038/srep00465 (2012).
- Bentov, S., Aflalo, E. D., Tynyakov, J., Glazer, L. & Sagi, A. Calcium phosphate mineralization is widely applied in crustacean mandibles. *Scientific reports* **6**, 22118, doi:10.1038/srep22118 (2016).
- Karlsson, A. M. High Strength and Light-weight Materials Inspired by the Exoskeleton of Arthropods. *Aro Exoskeleton Materials* (2010).
- Shavandi, A., Bekhit, E. D. A., Ali, M. A. & Sun, Z. Bio-mimetic composite scaffold from mussel shells, squid pen and crab chitosan for bone tissue engineering. *International journal of biological macromolecules* **80**, 445–454, doi:10.1016/j.ijbiomac.2015.07.012 (2015).
- Ifuku, S., Ura-kami, T., Izawa, H., Morimoto, M. & Saimoto, H. Preparation of a protein–chitin nanofiber complex from crab shells and its application as a reinforcement filler or substrate for biomineralization. *Rsc Advances* **5**, 1588–1597, doi:10.1039/C5RA12761K (2015).



14. Chen, Y. & Wang, L. Bio-inspired heterogeneous composites for broadband vibration mitigation. *Scientific reports* **5**, 17865, doi:10.1038/srep17865 (2015).
15. Li, K. *et al.* Separation of chito-oligomers with several degrees of polymerization and study of their antioxidant activity. *Carbohydrate Polymers* **88**, 896–903, doi:10.1016/j.carbpol.2012.01.033 (2012).
16. Senel, S. & McClure, S. J. Potential applications of chitosan in veterinary medicine. *Adv. Drug Deliv. Rev.* **56**, 1467–1480, doi:10.1016/j.addr.2004.02.007 (2004).
17. Liao, I. C. & Chien, Y. H. *The Pacific White Shrimp, Litopenaeus vannamei, in Asia: The World's Most Widely Cultured Alien Crustacean.* (2011).
18. Dall, W., Hill, B. J., Rothlisberg, P. C. & Sharples, D. J. *The biology of the Penaeidae.* (Bailliere, Tindall & Cox, 1933).
19. Wei, J. *et al.* Comparative transcriptomic characterization of the early development in Pacific white shrimp *Litopenaeus vannamei*. *PLoS one* **9**, e106201–e106201, doi:10.1371/journal.pone.0106201 (2014).
20. Roer, R., Abehsera, S. & Sagi, A. Exoskeletons across the Pancrustacea: Comparative Morphology, Physiology, Biochemistry and Genetics. *Integrative and comparative biology* **55**, 771–791, doi:10.1093/icb/icv080 (2015).
21. Chang, E. S. Physiological and biochemical changes during the molt cycle in decapod crustaceans: an overview. *Journal of Experimental Marine Biology & Ecology* **193**, 1–14, doi:10.1016/0022-0981(95)00106-9 (1995).
22. Corteel, M. *et al.* Molt cycle of laboratory-raised Penaeus (*Litopenaeus*) *vannamei* and *P. monodon*. *Aquaculture International* **20**, 13–18, doi:10.1007/s10499-011-9437-9 (2012).
23. Kuballa, A. V. & Elizur, A. Differential expression profiling of components associated with exoskeletal hardening in crustaceans. *BMC genomics* **9**, 575, doi:10.1186/1471-2164-9-575 (2008).
24. Aiken, D. E. *Molting and Growth.* (1980).
25. Yan, N. & Chen, X. Don't waste seafood waste. *Nature* **524**, 155–157, doi:10.1038/524155a (2015).
26. Rocha, J. *et al.* Cuticular chitin synthase and chitinase mRNA of whiteleg shrimp *Litopenaeus vannamei* during the molting cycle. *Aquaculture* **330**, 111–115, doi:10.1016/j.aquaculture.2011.12.024 (2012).
27. Zhang, S. Y. *et al.* Six chitinases from oriental river prawn *Macrobrachium nipponense*: cDNA characterization, classification and mRNA expression during post-embryonic development and molting cycle. *Comp. Biochem. Physiol. B-Biochem. Mol. Biol.* **167**, 30–40, doi:10.1016/j.cbpb.2013.09.009 (2014).
28. Tynyakov, J. *et al.* A crayfish molar tooth protein with putative mineralized exoskeletal chitinous matrix properties. *Journal of Experimental Biology* **218**, 3487–3498, doi:10.1242/jeb.123539 (2015).
29. Glazer, L. *et al.* A protein involved in the assembly of an extracellular calcium storage matrix. *The Journal of biological chemistry* **285**, 12831–12839, doi:10.1074/jbc.M109.071068 (2010).
30. Shechter, A. *et al.* A gastrolith protein serving a dual role in the formation of an amorphous mineral containing extracellular matrix. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 7129–7134, doi:10.1073/pnas.0800193105 (2008).
31. Sangsuriya, P. *et al.* Knockdown of a novel G-protein pathway suppressor 2 (GPS2) leads to shrimp mortality by exuvial entrapment during ecdysis. *Fish & shellfish immunology* **37**, 46–52, doi:10.1016/j.fsi.2014.01.007 (2014).
32. Pamuru, R. R. *et al.* Stimulation of molt by RNA interference of the molt-inhibiting hormone in the crayfish *Cherax quadricarinatus*. *General and comparative endocrinology* **178**, 227–236, doi:10.1016/j.ygcen.2012.05.007 (2012).
33. Gao, Y., Xu, H., Shen, Y. & Wang, J. Transcriptomic analysis of rice (*Oryza sativa*) endosperm using the RNA-Seq technique. *Plant molecular biology* **81**, 363–378, doi:10.1007/s11103-013-0009-4 (2013).
34. Eswaran, J. *et al.* RNA sequencing of cancer reveals novel splicing alterations. *Scientific reports* **3**, doi:10.1038/srep01689 (2013).
35. Tom, M. *et al.* Expression of cytoskeletal and molt-related genes is temporally scheduled in the hypodermis of the crayfish *Procambarus clarkii* during premolt. *The Journal of experimental biology* **217**, 4193–4202, doi:10.1242/jeb.109009 (2014).
36. Tatusov, R. L., Koonin, E. V. & Lipman, D. J. A genomic perspective on protein families. *Science* **278**, 631–637, doi:10.1126/science.278.5338.631 (1997).
37. Yang, H., Li, S., Li, F., Wen, R. & Xiang, J. Analysis on the expression and function of syndecan in the Pacific white shrimp *Litopenaeus vannamei*. *Developmental and comparative immunology* **51**, 278–286, doi:10.1016/j.dci.2015.03.013 (2015).
38. Lin, Y. C. *et al.* Characterization of white shrimp *Litopenaeus vannamei* integrin beta and its role in immunomodulation by dsRNA-mediated gene silencing. *Developmental and comparative immunology* **40**, 167–179, doi:10.1016/j.dci.2013.01.001 (2013).
39. Wang, Z. *et al.* Characterization of two types of vascular endothelial growth factor from *Litopenaeus vannamei* and their involvements during WSSV infection. *Fish & shellfish immunology* **47**, 824–832, doi:10.1016/j.fsi.2015.10.026 (2015).
40. Medina, M. Functional genomics opens doors to understanding metamorphosis in nonmodel invertebrate organisms. *Molecular ecology* **18**, 763–764, doi:10.1111/j.1365-294X.2008.04079.x (2009).
41. Wei, J., Zhang, X., Yu, Y., Li, F. & Xiang, J. RNA-Seq reveals the dynamic and diverse features of digestive enzymes during early development of Pacific white shrimp *Litopenaeus vannamei*. *Comparative biochemistry and physiology. Part D, Genomics & proteomics* **11**, 37–44, doi:10.1016/j.cbd.2014.07.001 (2014).
42. Reynolds, N., O'Shaughnessy, A. & Hendrich, B. Transcriptional repressors: multifaceted regulators of gene expression. *Development* **140**, 505–512, doi:10.1242/dev.083105 (2013).
43. Yang, P., Chen, X. M., Liu, W. W., Feng, Y. & Sun, T. Transcriptome analysis of sexually dimorphic Chinese white wax scale insects reveals key differences in developmental programs and transcription factor expression. *Scientific reports* **5**, 8141, doi:10.1038/srep08141 (2015).
44. Di, Z. *et al.* Genome-wide analysis of homeobox genes from *Mesobuthus martensii* reveals Hox gene duplication in scorpions. *Insect biochemistry and molecular biology* **61**, 25–33, doi:10.1016/j.ibmb.2015.04.002 (2015).
45. Chung, J. S. & Webster, S. G. Molt cycle-related changes in biological activity of molt-inhibiting hormone (MIH) and crustacean hyperglycaemic hormone (CHH) in the crab, *Carcinus maenas*. From target to transcript. *European Journal of Biochemistry* **270**, 3280–3288, doi:10.1046/j.1432-1033.2003.03720.x (2003).
46. Gao, Y. *et al.* Whole Transcriptome Analysis Provides Insights into Molecular Mechanisms for Molting in *Litopenaeus vannamei*. *PLoS one* **10**, e0144350, doi:10.1371/journal.pone.0144350 (2015).
47. Muhammad, F., Zhang, Z.-F., Shao, M.-Y., Dong, Y.-P. & Muhammad, S. Ontogenesis of digestive system in *Litopenaeus vannamei* (Boone, 1931) (Crustacea: Decapoda). *Italian Journal of Zoology* **79**, 1–9, doi:10.1080/11250003.2011.590534 (2011).
48. Little, G. The larval development of the shrimp, *Palaemon macrrodactylus* Rathbun, reared in the laboratory, and the effect of eyestalk extirpation on development. *Crustaceana* **17**, 69–87, doi:10.1163/156854069X00079 (1969).
49. Tom, M., Manfrin, C., Giulianini, P. G. & Pallavicini, A. Crustacean oxo-reductases protein sequences derived from a functional genomic project potentially involved in ecdysteroid hormones metabolism - a starting point for function examination. *General and comparative endocrinology* **194**, 71–80, doi:10.1016/j.ygcen.2013.09.003 (2013).
50. Kuballa, A. V., Holton, T. A., Paterson, B. & Elizur, A. Molt cycle specific differential gene expression profiling of the crab *Portunus pelagicus*. *BMC genomics* **12**, 147, doi:10.1186/1471-2164-12-147 (2011).
51. Soares, M. P. *et al.* Genes involved in thoracic exoskeleton formation during the pupal-to-adult molt in a social insect model, *Apis mellifera*. *BMC genomics* **14**, 576, doi:10.1186/1471-2164-14-576 (2013).
52. Buchholz, F. Molt Cycle and Seasonal Activities of Chitinolytic Enzymes in the Integument and Digestive-Tract of the Antarctic Krill. **9**, 311–317, doi:10.1007/BF00287429 (1989).



53. Zou, E. & Fingerman, M. Chitinase activity in the epidermis and hepatopancreas of the fiddler crab *Uca pugilator* during the molting cycle. *Marine Biology* **133**, 97–101, doi:10.1007/s002270050447 (1999).
54. Saleem, F. *et al.* Molecular characterisation of *Bacillus* chitinase for bioconversion of chitin waste. *Nat. Prod. Res.* **30**, 720–723, doi:10.1080/14786419.2015.1040789 (2016).
55. Bajaj, M., Freiberg, A., Winter, J., Xu, Y. M. & Gallert, C. Pilot-scale chitin extraction from shrimp shell waste by deproteination and decalcification with bacterial enrichment cultures. *Applied microbiology and biotechnology* **99**, 9835–9846, doi:10.1007/s00253-015-6841-5 (2015).
56. Qin, X., Xiang, X. M., Sun, X. W., Ni, H. & Li, L. Preparation of nanoscale *Bacillus thuringiensis* chitinases using silica nanoparticles for nematicide delivery. *International journal of biological macromolecules* **82**, 13–21, doi:10.1016/j.ijbiomac.2015.10.030 (2016).
57. Merzendorfer, H. & Zimoch, L. Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. *J. Exp. Biol.* **206**, 4393–4412, doi:10.1242/jeb.00709 (2003).
58. A-Banat, B. M. A., Kameyama, Y., Yoshioka, T. & Koga, D. Purification and characterization of a 54 kDa chitinase from *Bombyx mori*. *Insect Biochemistry & Molecular Biology* **29**, 537–547, doi:10.1016/S0965-1748(99)00028-4 (1999).
59. Abdelbanat, B. M. A., Zhou, W., Karasuda, S. & Koga, D. Analysis of Hydrolytic Activity of a 65-kDa Chitinase from the Silkworm. *Bioscience Biotechnology & Biochemistry* **66**, 1119–1122, doi:10.1271/bbb.66.1119 (2002).
60. Andersen, S. O., Hojrup, P. & Roepstorff, P. Insect cuticular proteins. *Insect Biochemistry & Molecular Biology* **25**, 153–176, doi:10.1016/0965-1748(94)00052-J (1995).
61. Shechter, A. *et al.* Reciprocal Changes in Calcification of the Gastrolith and Cuticle during the Molt Cycle of the Red Claw Crayfish *Cherax quadricarinatus*. *The Biological Bulletin* **214**, 122–134, doi:10.2307/25066669 (2008).
62. Abehsera, S. *et al.* Binary gene expression patterning of the molt cycle: the case of chitin metabolism. *PLoS one* **10**, e0122602, doi:10.1371/journal.pone.0122602 (2014).
63. Yudkovski, Y. *et al.* Hepatopancreatic multi-transcript expression patterns in the crayfish *Cherax quadricarinatus* during the moult cycle. *Insect molecular biology* **16**, 661–674, doi:10.1111/j.1365-2583.2007.00762.x (2007).
64. Yudkovski, Y. *et al.* Multi-transcript expression patterns in the gastrolith disk and the hypodermis of the crayfish *Cherax quadricarinatus* at premolt. *Comparative biochemistry and physiology. Part D, Genomics & proteomics* **5**, 171–177, doi:10.1016/j.cbd.2010.03.010 (2010).
65. Greenaway, P. Calcium balance and molting in the Crustacea. *Biological Reviews* **60**, 425–454, doi:10.1111/j.1469-185X.1985.tb00424.x (1985).
66. Ueno, M. Calcium transport in crayfish gastrolith disc: Morphology of gastrolith disc and ultrahistochemical demonstration of calcium. *Journal of Experimental Zoology* **213**, 161–171, doi:10.1002/jez.1402130202 (1980).
67. Dillaman, R., Hequembourg, S. & Gay, M. Early pattern of calcification in the dorsal carapace of the blue crab, *Callinectes sapidus*. *Journal of Morphology* **263**, 356–374, doi:10.1002/jmor.10311 (2005).
68. Liang, J., Zhang, L., Xiang, Z. & He, N. Expression profile of cuticular genes of silkworm, *Bombyx mori*. *Bmc Genomics* **11**, 1702–1716, doi:10.1186/1471-2164-11-173 (2010).
69. Andersen, S. O. Exoskeletal proteins from the crab, *Cancer pagurus*. *Comparative Biochemistry & Physiology Part A Molecular & Integrative Physiology* **123**, 203–211, doi:10.1016/S1095-6433(99)00051-3 (1999).
70. Sass, M., Kiss, A. & Locke, M. Classes of integument peptides. *Insect Biochemistry & Molecular Biology* **23**, 845–857, doi:10.1016/0965-1748(93)90073-2 (1993).
71. O'Brien, J. J. & Skinner, D. M. Proteins of Crustacean Exoskeletons: I. Similarities and Differences among Proteins of the Four Exoskeletal Layers of Four Brachyurans. *Biological Bulletin* **181**, 427–441, doi:10.2307/1542363 (1991).
72. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**, 644–652, doi:10.1038/nbt.1883 (2011).
73. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* **5**, 621–628, doi:10.1038/nmeth.1226 (2008).
74. Conesa, A. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676, doi:10.1093/bioinformatics/bti610 (2005).
75. Ye, J. *et al.* WEGO: a web tool for plotting GO annotations. *Nucleic acids research* **34**, 293–297, doi:10.1093/nar/gkl1031 (2006).
76. Bindea, G. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093, doi:10.1093/bioinformatics/btp101 (2009).
77. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *Bmc Bioinformatics* **12**, 93–99, doi:10.1186/1471-2105-12-323 (2011).
78. Wang, L., Feng, Z., Wang, X., Wang, X. & Zhang, X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26**, 136–138, doi:10.1093/bioinformatics/btp612 (2010).
79. de Hoon, M. J., Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. *Bioinformatics* **20**, 1453–1454, doi:10.1093/bioinformatics/bth078 (2004).
80. Saldanha, A. J. Java Treeview—extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248, doi:10.1093/bioinformatics/bth349 (2004).
81. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* **25**, 402–408, doi:10.1006/meth.2001.1262 (2001).

## Acknowledgements

This work was financially supported by National Natural Science Foundation of China (31672632, 41376165, 31272397). And the Scientific and Technological Innovation Project of Qingdao National Laboratory for Marine Science and Technology (No. 2015ASKJ02).

## Author Contributions

Conceived and designed the experiments: X.Z., F.L., J.X. Performed the experiments: Y.G., J.W., X.Z. Analyzed the data: Y.G., X.Z. Contributed reagents/materials/analysis tools: J.W., J.Y. Contributed to the writing of the manuscript: Y.G., X.Z., J.X.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-01220-6

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017