Supplementary Materials for

“Omics” data unveil early molecular response underlying limb regeneration in the Chinese mitten crab, *Eriocheir sinensis*

Jun Wang *et al.*

Corresponding author: Michael Hofreiter, hofreiter.michael@googlemail.com; Chenhong Li, chli@shou.edu.cn; Chenghui Wang, wangch@shou.edu.cn

*Sci. Adv.* 8, eabl4642 (2022)
DOI: 10.1126/sciadv.abl4642

The PDF file includes:

Fig. S1 to S34
Tables S1 to S8, S22
Legends for tables S9 to S21
Legend for movie S1

Other Supplementary Material for this manuscript includes the following:

Tables S9 to S21
Movie S1
Fig. S1. Information on Hi-C interactions among 70 pseudochromosomes. The figure shows Hi-C interactions among the 70 chromosomes. Strong and weak interactions were labeled with dark red and light yellow, respectively. The stronger interactions indicate the more accurate scaffold clustering by Hi-C in the final assembly. LG01 to LG70 indicate the 70 pseudochromosomes.
Fig. S2. De novo genome assembly pipeline for *E. sinensis*. We conducted *de novo* assembly using Nanopore long sequencing reads, Illumina short sequencing reads, BioNano optical maps, and Hi-C data. Our assembly pipeline was the following: First, we used Nanopore long reads to get primary contigs. Second, we use Illumina short sequencing reads and Nanopore long sequencing reads to polish the primary contigs. Third, with the assistance of BioNano optical maps, we obtained the preliminary longer scaffolds. Finally, we used Hi-C data to cluster the scaffolds into our final assembly.
Fig. S3. Summary statistics of BUSCO prediction on the assembled *E. sinensis* genome. The number of complete BUSCOs in the *E. sinensis* genome is 1,009 (complete and single-copy BUSCOs: 942, complete and duplicated BUSCOs: 67), that of fragmented BUSCOs is 7, and that of missing BUSCOs is 50. The total number of BUSCOs in the Arthropod lineage of the BUSCO database is 1066 (Table S4).
Fig. S4. Venn diagram of gene annotation information of the *E. sinensis* genome using different databases. Numbers in the figure indicate the numbers of genes annotated with different databases.
Fig. S5. Phylogeny of 18 metazoan species based on 62 one-to-one orthologous genes with regenerative ability labeled on each clade.
**Fig. S6. Potential gene loss events that may be associated with limb regeneration of *E. sinensis*.** Pink and cyan box indicates absence and presence of genes in the indicated species. Detailed information of the potential gene loss events in *E. sinensis* can be accessed in Table S14. The gene loss events identified in this study maybe overestimated and should be interpreted with caution as further improvements of our genome assembly are still needed (gaps are still existed).
Fig. S7. Cluster dendrogram showing modules based on co-expression topological overlap of genes during the limb regeneration process, identified with WGCNA. Genes in the same module share highly similar expression patterns (co-expression). Different horizontal color bars represent different modules of co-expressed genes. A total of 22 modules were identified in “Merged dynamic” after merging modules with gene expression correlation above 0.75 based on “Dynamic tree cut” by WGCNA. Co-expressed genes in the same module are of biological interest since these genes may be controlled by the same transcriptional regulatory program, be functionally associated, or be in the same signaling pathway.
Fig. S8. Relationships of consensus module eigengenes and limb regeneration stage identified by WGCNA analysis. Through WGCNA analysis, we identified 22 gene co-expression modules in response to temporal changes during the limb regeneration process of *E. sinensis* with specific modules showing strong correlation with different limb regeneration stages (1 dpa, 13 dpa, and 30 dpa). Numbers in the box indicate the correlation coefficient between module and limb regeneration stages and numbers in the brackets indicate *p*-values. Highest correlation coefficients and small *p*-values indicate that the co-expressed genes in the module were strongly associated with the regeneration stages. dpa: days post autotomy.
Fig. S9. GO enrichment analysis for co-expressed genes in Module 1 (A) and Module 2 (B) from Fig. 3 by WGCNA analysis.
Fig. S10. GO enrichment analysis for co-expressed genes in Module 3 (A) and Module 4 (B) from Fig. 3 by WGCNA analysis.
Fig. S11. GO enrichment analysis for co-expressed genes in Module 5 (A) and Module 7 (B) from Fig. 3 by WGCNA analysis.
Fig. S12. Differentially expressed genes in *E. sinensis* during the limb regeneration process. 
(A) Expression heatmap of differentially expressed genes during the limb regeneration process. 
(B) GO enrichment analysis on representative biological process categories of upregulated genes at 1 dpa, 13 dpa and 30 dpa during the limb regeneration process. dpa: days post autotomy. Rich Factor indicates the ratio of the number of genes in the differential expressed gene list to the number of genes in the whole gene set (background) annotated in specific GO terms. Larger rich factors indicate higher enrichment level. The differential expressed genes with detailed information can be accessed in Table S16.
Fig. S13. Volcano map of differentially expressed genes between 0 dpa and 1 dpa (fold change >2, \(p < 0.05\)). Red dots indicate significantly upregulated genes, green dots indicate significantly downregulated genes at 1 dpa, and blue dots indicate genes that were not significantly differentially expressed. Upregulated *Innexin* gene IDs at 1 dpa were labeled with red color. The *Innexin* gene IDs that were not differential expressed were labeled with blue color. One *Innexin* gene (Esin.LG37.0153) that was not expressed (TPM = 0) was not labeled in the figure. dpa: days post autotomy. Detailed information of other differentially expressed genes can be accessed in Table S16.
Fig. S14. Volcano map of differentially expressed genes between 13 dpa and 1 dpa (fold change >2, p <0.05). Red dots indicate significantly upregulated genes, green dots indicate significantly downregulated genes at 1 dpa, and blue dots indicate genes that were not significantly differentially expressed. Upregulated gene IDs at 13 dpa with log2FC >3 associated with cell cycle process were labeled with red color. dpa: days post autotomy. Detailed information of other differentially expressed cell cycle related genes can be accessed in Table S16.
**Fig. S15. Volcano map of differentially expressed genes between 30 dpa and 13 dpa** (fold change $>2$, $p < 0.05$). Red dots indicate significantly upregulated genes, green dots indicate significantly downregulated genes at 1 dpa, and blue dots indicate genes that were not significantly differentially expressed. Upregulated and downregulated gene IDs at 30 dpa annotated to be associated with cuticle development were labeled with red and green color, respectively. dpa: days post autotomy. Detailed information of other differentially expressed genes can be accessed in Table S16.
Fig. S16. Expression of arthropod-specific genes in *E. sinensis* during the limb regeneration process. The proportion of arthropod-specific genes in DEGs was significantly higher than that of arthropod-specific genes in non-DEGs (*p* < 0.001, Chi-squared test).
Fig. S17. GO enrichment analysis on differentially expressed arthropod-specific (A) and crustacean-specific (B) genes at biological process level during the limb regeneration process.
Fig. S18. Information on 16 Innexin genes identified in the E. sinensis genome. (A) Gene structure of the 16 identified Innexin genes. (B) Chromosomal distribution of Innexin genes in the E. sinensis genome.
Fig. S19. Conserved motifs of the *Innexin* genes identified by the MEME software.
*indicates the motif sequences were annotated to be *Innexin* domain by PFAM search.
Fig. S20. Phylogenetic analysis of *Innexin* genes in the arthropod lineage. Dpu: *D. pulex*; Dme: *D. melanogaster*; Dma: *D. magna*; Esin: *E. sinensis*; Ham: *H. americanus*; Haz: *H. azteca*; Lva: *L. vannamei*; Pha: *P. hawaiensis*; Spa: *S. paramamosain*. *Innexin* genes of *E. sinensis* and *D. melanogaster* are labeled with red color and yellow color, respectively.
A

Metamorphosis

- Megalopa
- Larvae

B

Aerial respiration

- Control
- Day 5

C

Molting

- Intermolt
- Premolt
Fig. S21. Expression of the *Innexin* gene family in metamorphosis development (A), aerial respiration (B), and molting process (C) of *E. sinensis*. Illumina raw sequencing reads from different molting stages of hepatopancreas (Intermolt and Premolt), gill from the aerial respiration phase (control versus five days out of water), and whole individual from different developmental stages (megalopa and larvae I stage of *E. sinensis*) were downloaded from the NCBI SRA database (PRJNA271233, PRJNA480555) and National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation (GSA: CRA003690). * indicates fold change >2 and \( p < 0.05 \), ** indicates fold change >2 and \( p < 0.01 \).
Fig. S22. Differentially expressed *Innexin* genes in *Litopenaeus vannamei* (A), *Macrobrachium rosenbergii* (B), and *Macrobrachium nipponensis* (C) after autotomy at 1 dpa compared with 0 dpa. * indicates fold change >2 and $p < 0.05$, ** indicates fold change >2 and $p < 0.01$. All expression values of other *Innexin* genes are presented in Table S19.
Fig. S23. qRT-PCR analysis on related genes and Ca\textsuperscript{2+} measurement during the limb regeneration process. (A) qRT-PCR analysis of Inx2 gene during the limb regeneration process. Groups labeled with different lowercase letters indicate significant difference at \( p < 0.05 \), using one-way analysis of variance (ANOVA). (B) qRT-PCR analysis of Slc7a5 gene during the limb regeneration process. Groups labeled with different lowercase letters indicate significant difference at \( p < 0.05 \), using one-way analysis of variance (ANOVA). (C) qRT-PCR analysis of Calmodulin (CaM) gene during the limb regeneration process. (D) Measurement of Ca\textsuperscript{2+} at 0 dpa and 1 dpa. Ca\textsuperscript{2+} was measured using a Ca\textsuperscript{2+} detection kit (Nanjing Jiancheng Bioengineering Institute, NanJing, China). hpa: hours post autotomy, dpa: days post autotomy.
Fig. S24. Co-expression network showing interaction of Inx2 with genes annotated as transmembrane transport. The network was visualized with the software Cytoscape with the results of WGCNA as input. The network indicates that Inx2 may interact with the LAT1 protein, which is encoded by the Slc7a5 gene.
Fig. S25. The expression of genes in the mTORC1 signaling pathway after autotomy in *E. sinensis*. (A) Activated mTORC1 signaling pathway at 1 dpa. Red arrow indicates upregulated genes by RNA-seq analysis. (B) qRT-PCR analysis of *Vatb, Rraga, Rragb, Lamtor1, Lamtor2, Lamtor3, Lamtor4*, and *Lamtor5* genes involved in the mTORC1 signaling pathway during the limb regeneration process. hpa: hours post autotomy, dpa: days post autotomy. Groups labeled with different lowercase letters indicate significant difference at $p < 0.05$, using one-way analysis of variance (ANOVA).
Fig. S26. Rapamycin injection experiment after autotomy (n=9). The length of regenerated papillae was measured at 9 dpa. The length of the regenerated papillae in the experimental group (injected with rapamycin) was significantly smaller than that in the control group (injected with DMSO) at 9 dpa. dpa: days post autotomy.
Fig. S27. Hemocyanin gene family in the arthropod lineage and expression pattern during limb regeneration of *E. sinensis*. (A) Number of genes in the hemocyanin, cryptocyanin, and phenoloxidases families in 12 species that show that hemocyanin is present only in the arthropod lineage. (B) Expression patterns of 22 hemocyanin genes during the limb regeneration process of *E. sinensis* with most of the genes showing upregulation trend at 1 dpa. dpa: days post autotomy.
Fig. S28. Expression patterns of LGBP (A), ALPS (B), CFB (C), PCE (D) genes during the limb regeneration process. dpa: days post autotomy. * indicates fold change >2 and $p < 0.05$, ** indicates fold change >2 and $p < 0.01$ between 0 dpa and 1 dpa.
Fig. S29. Enhanced and effective crustacean specific pro-phenoloxidase system (ProPo-AS) during the limb regeneration process of *E. sinensis*. Red arrows indicate related gene families/genes that are upregulated at 1 dpa compared to 0 dpa during the limb regeneration process.
Fig. S30. Expression patterns of all the identified SMYDA genes in the regeneration process of regenerated limb bud tissue of *E. sinensis*. dpa: days post autotomy. * indicates fold change >2 and *p* < 0.05, ** indicates fold change >2 and *p* < 0.01 between 0 dpa and 1 dpa.
Fig. S31. Expression patterns of all the identified SMYDA genes in *Litopenaeus vannamei* (A), *Macrobrachium rosenbergii* (B), and *Macrobrachium nipponensis* (C) at 1 dpa compared with 0 dpa. dpa: days post autotomy. * indicates $p < 0.05$, ** indicates $p < 0.01$. Detailed information on the annotation and expression values of the Smyda genes are presented in Table S19.
Fig. S32. Fluorescence in situ hybridization of the SMYDA1 (Esin.LG14.0065) gene at 0 dpa and 1 dpa. Autotomy membrane (AM), muscle cell (MC), epidermal cell (Ep), melanized scab (S). The white arrow indicates the positive signal of the SMYDA1 gene. The yellow box indicates the magnified region of the white box.
Fig. S33. Maximum-likelihood phylogenetic tree reconstructed using SMYD4 protein sequences identified from 18 metazoan species. EsinLG52.0105 and Esin.LG18.0008 labeled with red color are two SMYD4 genes of *E. sinensis* that showed dynamic expression in the limb regeneration process and clustered in the arthropod lineage. Hvu: *H. vulgaris*; Cel: *C. elegans*; Cbr: *C. briggsae*; Lva: *L. vannamei*; Dpu: *D. pulex*; Dml: *D. melanogaster*; Cte: *C. teleta*; Hro: *H. robusta*; Cvi: *C. virginica*; Osi: *O. sinensis*; Aja: *A. japonicus*; Aru: *A. rubens*; Dre: *D. rerio*; Xla: *X. laevis*; Gga: *G. gallus*; Aca: *A. carolinensis*; Hsa: *H. sapiens*
Fig. S34. Overview of the limb regeneration process of *E. sinensis*. 
Table S1. Genome sequencing information for *E. sinensis*.

| Types    | Method    | Library size | Raw data  | Clean data | Average Length |
|----------|-----------|--------------|-----------|------------|----------------|
| Genome   | Illumina  | 400 bp       | 91.0 Gb   | 81.2 Gb    | 150 bp         |
|          | Hiseq     |              |           |            |                |
| Genome   | Oxford    | >20kb        | 99.1 Gb   | 81.7 Gb    | 20.0 kb        |
|          | Nanopore  |              |           |            |                |
| Genome   | Hi-C      | 300-600 bp   | 321 Gb    | 300.5 Gb   | 150 bp         |
| Genome   | BioNano   | /            | 480 Gb    | 442 Gb     | /              |
Table S2. Genome assembly statistics by different sequencing methods.

| Assembly               | Contigs | Scaffolds | Contig N50 | Scaffold N50 | Assembly size (contigs) |
|------------------------|---------|-----------|------------|--------------|-------------------------|
| Oxford Nanopore (ONT)  | 1712    | NA        | 1,664,518  | NA           | 1,669,425,693           |
| BioNano Optical Map    | NA      | 515       | NA         | 9,062,000    | 1,644,499,000           |
| ONT+BioNano            | 4796    | 2752      | 739,031    | 8,569,554    | 1,669,425,693           |
| ONT+Bionano+HiC       | 4808    | 2160      | 717,335    | 16,975,517   | 1,667,381,268           |

NA: Not Applicable
Table S3. Scaffold cluster results of 70 pseudochromosomes by Hi-C scaffolding.

| Chr  | Length       | Scaffold Numbers |
|------|--------------|-----------------|
| LG01 | 45,379,147   | 86              |
| LG02 | 31,954,432   | 29              |
| LG03 | 29,173,192   | 26              |
| LG04 | 29,084,662   | 44              |
| LG05 | 28,108,396   | 30              |
| LG06 | 27,314,786   | 73              |
| LG07 | 27,014,105   | 49              |
| LG08 | 25,820,439   | 17              |
| LG09 | 25,704,940   | 40              |
| LG10 | 24,916,309   | 15              |
| LG11 | 24,776,463   | 48              |
| LG12 | 24,411,283   | 16              |
| LG13 | 23,812,216   | 22              |
| LG14 | 22,909,581   | 16              |
| LG15 | 22,682,036   | 2               |
| LG16 | 22,092,210   | 32              |
| LG17 | 21,810,268   | 7               |
| LG18 | 21,593,893   | 8               |
| LG19 | 21,385,840   | 41              |
| LG20 | 21,245,621   | 52              |
| LG21 | 20,632,807   | 16              |
| LG22 | 20,595,259   | 24              |
| LG23 | 20,502,797   | 50              |
| LG24 | 20,108,279   | 23              |
| LG25 | 19,885,966   | 12              |
| LG26 | 19,743,373   | 27              |
| LG27 | 19,727,531   | 12              |
| LG28 | 19,593,204   | 7               |
| LG29 | 18,884,131   | 29              |
| LG30 | 18,566,567   | 20              |
| LG31 | 18,276,216   | 5               |
| LG32 | 18,177,313   | 19              |
| LG33 | 18,109,958   | 10              |
| LG34 | 18,105,521   | 13              |
| LG35 | 18,098,480   | 9               |
| LG36 | 17,730,161   | 23              |
| LG37 | 17,516,247   | 6               |
| LG38 | 17,374,747   | 10              |
| LG39 | 17,195,986   | 5               |
| LG40 | 16,975,517   | 11              |
| LG41 | 16,949,729   | 17              |
| LG42 | 15,414,004   | 46              |
| LG43 | 15,377,322   | 16              |
| LG44 | 15,200,582 | 12 |
| LG45 | 14,719,648 | 19 |
| LG46 | 14,298,206 | 14 |
| LG47 | 13,895,859 | 19 |
| LG48 | 13,203,802 | 13 |
| LG49 | 13,078,750 | 19 |
| LG50 | 12,993,234 | 13 |
| LG51 | 12,937,974 | 19 |
| LG52 | 12,750,814 | 12 |
| LG53 | 12,567,781 | 32 |
| LG54 | 12,250,771 | 17 |
| LG55 | 12,242,209 | 1 |
| LG56 | 12,201,330 | 10 |
| LG57 | 12,024,539 | 24 |
| LG58 | 11,946,249 | 11 |
| LG59 | 11,757,766 | 11 |
| LG60 | 11,648,666 | 11 |
| LG61 | 11,596,831 | 18 |
| LG62 | 11,439,990 | 16 |
| LG63 | 11,336,473 | 16 |
| LG64 | 10,710,421 | 24 |
| LG65 | 10,428,594 | 18 |
| LG66 | 10,309,680 | 16 |
| LG67 | 10,233,791 | 8 |
| LG68 | 8,892,716 | 13 |
| LG69 | 8,632,099 | 26 |
| LG70 | 7,952,464 | 18 |
**Table S4. Summary statistics of BUSCO prediction on the assembled *E. sinensis* genome.**

| Type                                           | Number | Percent (%) |
|-----------------------------------------------|--------|-------------|
| Complete BUSCOs (C)                           | 1,009  | 94.65       |
| Complete and single-copy BUSCOs (S)           | 942    | 88.37       |
| Complete and duplicated BUSCOs (D)            | 67     | 6.29        |
| Fragmented BUSCOs (F)                         | 7      | 0.66        |
| Missing BUSCOs (M)                            | 50     | 4.69        |
| Total BUSCO groups searched                   | 1,066  | 100         |
Table S5. Summary statistics of repeat sequences in the genome of *E. sinensis* and comparison with three other malacostracans with high-quality genome assemblies.

| Type            | *Eriocheir sinensis* | *Litopenaeus vannamei* | *Homarus americanus* | *Scylla paramamosain* |
|-----------------|----------------------|------------------------|----------------------|------------------------|
|                 | Number | Length       | Percent* | Number | Length       | Percent* | Number | Length       | Percent* | Number | Length       | Percent* |
| SINEs           | 120,507 | 16,430,316 | 0.93     | 2,934  | 1,021,571 | 0.06     | 50,455 | 9,820,586 | 0.43     | 13,693 | 1,291,564 | 0.08     |
| LINE            | 1,146,379 | 281,910,139 | 15.95   | 126,578 | 46,932,128 | 2.82    | 398,238 | 234,140,880 | 10.22   | 358,994 | 137,134,826 | 8.87    |
| LTR             | 805,343  | 127,473,128 | 7.21     | 36,163  | 10,259,678 | 0.62     | 198,624 | 103,793,034 | 4.53     | 240,204 | 93,304,374 | 6.04     |
| DNA elements    | 2,694,297 | 361,789,213 | 20.46   | 932,683 | 155,145,405 | 9.33    | 366,009 | 111,029,823 | 4.84     | 303,586 | 40,153,253 | 2.60     |
| RC              | 99,504   | 5,566,745  | 0.31     | /       | /          | /        | 9,534  | 2,333,674  | 0.10     | /       | /          | /        |
| Unclassified    | 289,926  | 97,612,770 | 5.52     | 216,901 | 55,650,866 | 3.35    | 1,460,061 | 597,162,181 | 26.05   | 176,777 | 97,385,157 | 6.3      |
| Satellites      | 855,345  | 104,688,182 | 5.92   | 13,527  | 1,550,312 | 0.09    | 11,142  | 3,840,580  | 0.17     | 6,345  | 724,855  | 0.05     |
| Simple repeats  | 220,406  | 32,340,540 | 1.83     | 5,515,074 | 398,243,493 | 23.93 | 1,709,134 | 143,273,836 | 6.25   | 2,293,823 | 163,133,947 | 10.55 |
| Low complexity  | 36,764   | 7,113,224  | 0.40     | 1,103,699 | 154,532,775 | 9.29  | /       | /          | /        | 256,787 | 19,597,723 | 1.27     |
| Other           | 234,203  | 26,010,794 | 1.47     | /       | /          | /        | 22,442  | 7,826,535  | 0.34     | 1,258  | 412,430  | 0.03     |
| Total           | 6,502,674 | 1,060,935,051 | 60.01 | 7,947,559 | 823,336,228 | 49.48 | 4,225,639 | 1,213,221,129 | 52.93 | 3,651,467 | 55,313,8129 | 35.79 |

*indicates percent of the assembled genome, “/” indicates no data available.
**Table S6. Comparison of statistics of four assembled *E. sinensis* genomes.**

|                               | Assembled genome in this study | Previously published genome by Cui et al* | Previously Published genome by Tang et al* | Previously published genome by Song et al* |
|-------------------------------|-------------------------------|------------------------------------------|--------------------------------------------|--------------------------------------------|
| Total assembled genome size (bp) | 1,767,846,446                 | 1,567,615,418                            | 1,270,960,592                              | 1,118,179,523                              |
| Total number of scaffolds      | 2,160                         | 101,205                                  | 4,311                                      | 1,768,649                                  |
| No. of scaffolds ≥1000 bp      | 2,158                         | 63,883                                   | 4,310                                      | 51,121                                     |
| No. of scaffolds ≥5000 bp      | 2,107                         | 13,440                                   | 4,097                                      | 5,610                                      |
| Longest length (bp)            | 45,379,147                    | 50,864,308                               | 31,480,327                                 | 2,002,076                                  |
| Scaffold N50 (bp) / No. of N50  | 16,975,517 / 40               | 17,127,685 / 30                          | 17,608,299 / 30                            | 111,755 / 2,066                            |
| Contig N50 (bp) / No. of Contig N50 | 717,335 / 434              | 26,045 / 12,717                          | 3,161,423 / 96                             | 2,873 / 67,473                             |
| Scaffold N75 (bp) / No. of N75  | 1,296,936 / 106               | 9,971,905 / 60                           | 11,039,611 / 52                            | 625 / 114,142                              |
| Contig N75 (bp) / No. of N75   | 296,290 / 1383                | 10,845 / 31,987                          | 521,706 / 352                              | 409 / 311,430                              |
| GC (%)                         | 41.21                         | 42.43                                    | 41.96                                      | 41.87                                      |
| Total length anchored on pseudochromosomes (bp) | 1,265,982,173                 | 1,267,002,578                            | 1,131,993,911                              | NA                                         |

* The genome assembly data from the other papers were obtained from http://www.genedatabase.cn/esi_genome.html (17), NCBI database https://www.ncbi.nlm.nih.gov/assembly/GCA_013436485.1 (19), and Gigadb http://gigadb.org/dataset/100186 (18).
NA: Not applicable.
Table S7. Summary statistics of annotated protein-coding genes.

| Gene set   | Total number of genes | Average gene length (bp) | Average CDS length (bp) | Average exon number per gene | Average exon length (bp) | Average intron length (bp) |
|------------|-----------------------|--------------------------|-------------------------|-----------------------------|-------------------------|---------------------------|
| De novo    | 43,791                | 5,163.90                 | 895.25                  | 3.93                        | 227.55                  | 1,454.73                  |
| Homology   | 25,122                | 27,240.96                | 1,117.34                | 4.86                        | 230.13                  | 6,776.23                  |
| RNA-seq    | 15,568                | 28,598.60                | 4,176.12                | 20.99                       | 199.0                   | 1,221.98                  |
| Final set  | 20,286                | 14,504.96                | 1,283.93                | 5.61                        | 229.05                  | 2,870.68                  |
Table S8. Information on gene function annotation of the *E. sinensis* genome.

| Type        | Number | Percent (%) |
|-------------|--------|-------------|
| Annotation  |        |             |
| Swissprot   | 13,276 | 65.44       |
| KEGG        | 7,800  | 38.45       |
| KOG         | 11,529 | 56.83       |
| GO          | 9,716  | 47.90       |
| NR          | 18,119 | 89.32       |
| Total       |        |             |
| Annotated genes | 18,507 | 91.23       |
| Genes       | 20,286 | -           |
Table S9. Annotation information of 20,286 predicted genes in *E. sinensis*.

Table S10. Identification of arthropod-specific, crustacean-specific, and *E. sinensis*-specific gene families in *E. sinensis*.

Table S11. Information on expanded and contracted gene families in *E. sinensis*.

Table S12. List of 104 differential expressed genes annotated to be Zinc finger proteins during the limb regeneration process of *E. sinensis*.

Table S13. Expanded gene families associated with cuticle development and their expression data during limb regeneration in *E. sinensis*.

Table S14. Identification of potential gene loss events in the *E. sinensis* lineage.

Table S15. Annotation and expression of co-expressed genes in module 1 to module 8 through WGCNA analysis.

Table S16. Differential expressed genes identified during the limb regeneration process of *E. sinensis*.

Table S17. Information on *Innexin* genes identified in the *E. sinensis* genome.

Table S18. Expression patterns of *Innexin* genes in *E. sinensis* during the limb regeneration process.

Table S19. Expression patterns of *Innexin* and *Smyda* genes in *Litopenaeus vannamei*, *Macrobrachium nipponensis*, and *Macrobrachium rosenbergii* after autotomy at 1 dpa.

Table S20. Identified genes annotated to be C-lectin and their expression during the limb regeneration process of *E. sinensis*.

Table S21. Genome assembly information of species used in this study.
| Primers       | Sequences (5'-3')                       | Usage                |
|--------------|----------------------------------------|----------------------|
| Inx2-F       | CTGTTCCCGCGGATGACCAA                   | qPCR for Inx2        |
| Inx2-R       | GCTGTCGGCAACAAACACA                    | qPCR for Inx2        |
| S1c7a5-F     | GTCCCCGCTGGCACCGTTGTTG                | qPCR for S1c7a5      |
| S1c7a5-R     | GGGCCGTGTCACCGTTGACA                  | qPCR for S1c7a5      |
| S27-F        | GGTCGATGACAAATTGCAAGA                 | qPCR for S27         |
| S27-R        | CCACAGTACTGCCTGACTCAA                 | qPCR for S27         |
| CaM-F        | GGCAACATCAACACCAAGGA                  | qPCR for CaM         |
| CaM-R        | ACCGCCATCATCGTAAGGA                   | qPCR for CaM         |
| Vatb-F       | TCTTCCCTGAACCTGGCAAT                  | qPCR for Vatb        |
| Vatb-R       | GGACGTCCTTCCACACTGG                   | qPCR for Vatb        |
| Lamtor1-F    | AGCACACAGCCAGTAGTATCAT                | qPCR for Lamtor1     |
| Lamtor1-R    | TCTACAGCTGCTGCTGATCTA                | qPCR for Lamtor1     |
| Lamtor2-F    | CAATCTGCTGCTGATCTCT                  | qPCR for Lamtor2     |
| Lamtor2-R    | GCAGAGGTCCATCAAGGTAG                  | qPCR for Lamtor2     |
| Lamtor3-F    | TGGCCGAGGAATGAAGAAG                   | qPCR for Lamtor3     |
| Lamtor3-R    | CCTCCTTCTCCGAGACTTTC                 | qPCR for Lamtor3     |
| Lamtor4-F    | CCAAGTGAGCCATCAAGATAC                 | qPCR for Lamtor4     |
| Lamtor4-R    | ATGGGCTCTGCTGAGAACAAA                | qPCR for Lamtor4     |
| Lamtor5-F    | CAGCCTTGGACCGATCATCT                 | qPCR for Lamtor5     |
| Lamtor5-R    | TGGAGACTTGTAGACGGCAA                  | qPCR for Lamtor5     |
| Rragd-F      | TCCGACGACCAACAGATTGA                 | qPCR for Rragd       |
| Rragd-R      | GTTCGTAGATGGAGGTAGG                   | qPCR for Rragd       |
| RragA-F      | ATCTCCCCACTACTGACATCCG               | qPCR for RragA       |
| RragA-R      | TGCAAGAGAGCTTGAACCTGCG               | qPCR for RragA       |
| Inx2-dsRNA-F | taatacgactcataagggAGAAACGCGAGACAGCC  | dsRNA for Inx2       |
| Inx2-dsRNA-R | taatacgactcataagggTCCTGCGATTCCTCCTT  | dsRNA for Inx2       |
| GFP-dsRNA-F  | taatacgactcataagggCAGTGCTTCAGCCGCTACC | dsRNA for GFP        |
| Probe          | Sequence                                                                 | Description                           |
|----------------|--------------------------------------------------------------------------|---------------------------------------|
| GFP-dsRNA-R    | taatacgactcatactagggAGTTCACTTTAGCCGTTCTT                                 | dsRNA for GFP                          |
| Probe-Inx2     | FAM-CAGTGTGGGATGAAGTTGAGAATCTCGCAGAAGCAGAATCTTCATGG                      | In situ hybridization probe for Inx2, 5’-FAM labeled |
| Probe-Slc7a5   | Cy3-CGTATGATTACCACGTTGACCCATAGCGTGAGGAAGGGCAGGAAGG                       | In situ hybridization probe for Slc7a5, 5’-Cy3 labeled |
| Probe-SMYDA1   | FAM-CGTCATCGTGTGGTGCTGTGTTGGTCAGAAGTGAGGAGAAGTGGGCCAAGAAGTGGAAGG         | In situ hybridization probe for SMYDA1, 5’-FAM labeled |
| Probe-Control  | FAM-TTGACTACAAAAAGTACTG                                                 | In situ hybridization probe for Scrambled control, 5’-FAM labeled |
Movie S1. The limb regeneration process of *E. sinensis* during molting. The video recorded that the new limb was fully regenerated after molting of *E. sinensis*. 