A novel gene rearrangement in the mitochondrial genome of Coenobita brevimanus (Anomura: Coenobitidae) and phylogenetic implications for Anomura

Li Gong (gongli1027@163.com)

Xinting Lu
Zhejiang ocean university

Zhifu Wang
Second institude of oceanograph

Wei Shi
south china sea institute of oceanology

Kehua Zhu
Zhejiang Ocean University

Liqin Liu
Zhejiang Ocean University

Lihua Jiang
Zhejiang ocean university

Zhenming Lü
zhejiang ocean university

Bingjian Liu
zhejiang ocean university

Research article

Keywords: Terrestrial hermit crab; Mitogenome; Gene rearrangement; Phylogenetic analysis

Posted Date: June 8th, 2019

DOI: https://doi.org/10.21203/rs.2.10147/v1

Version of Record: A version of this preprint was published at Genomics on March 1st, 2020. See the published version at https://doi.org/10.1016/j.ygeno.2019.10.012.

Abstract

Background: Gene arrangement in vertebrate mitochondrial genomes (mitogenomes) is relatively conserved and fewer gene arrangement is discovered. In contrast, that in invertebrate mitogenomes is relatively common. Although a gradually growing number of gene rearrangement in hermit crabs (Paguridae) has been discovered, it is surprising that gene rearrangement in its close relatives, the terrestrial hermit crab (Coenobitidae), was overlooked until 2018. So far, only few studies focused on the phylogenetic studies of Anomura based on molecular evidences. Results: In the present study, the complete mitogenome of a terrestrial hermit crab, Coenobita brevimanus, was sequenced, and large-scale gene rearrangements were observed. The genomic features of this terrestrial hermit crab were different from those of any other studied crabs. Five gene clusters (or genes) including eleven tRNAs and two PCGs were found to be rearranged with respect to the pancrustacean ground pattern gene order, which was characterized by multiple translocations and inversions. Two phylogenetic trees (ML and BI tree) arrived at a similar topology based on the nucleotide sequences of the 13 concatenated PCGs. Conclusions: We propose tandem duplication-random loss and recombination model to explain the large-scale gene rearrangements in C. brevimanus mitogenome. The phylogenetic trees showed that all Coenobitidae species clustered into one clade. The polyphyly of Paguroidea was well supported, whereas the non-monophyly of Galatheoidea was not in consistence with previous findings. The phylogenetic relationships of Pylochelidae, Lomidae, and Albuneidae were controversial.

Results And Discussion

Genome structure and composition

Although the newly determined complete mitogenome of *C. brevimanus* is almost identical with (98.6% similarity) the published one (GenBank accession number KY352233), the authors mainly focused on the phylogenetic analyses, while hardly described the mitogenome features [5]. Hence, we described the complete mitogenome (MK310257) in detail and focused on the gene rearrangements and possible rearrangement mechanisms. The sequence is 16,393 bp in length, almost the same length with that of the published one (16,390 bp). It comprises 13 PCGs, 22 tRNAs, two rRNAs and one CR (Fig. 1, Table 2), which is identical with that in most crabs [2 , 5 , 44]. The size of *C. brevimanus* mitogenome presented in this study falls within the range of other Anomura mitogenomes from 14,632 bp in *Pagurus lanuginosus* (LC222527) to 17,910 bp in *Munida isos* (NC_039112). The overall nucleotide composition is 27.7% A, 37.3% T, 20.7% G, and 14.3% C, respectively (Table 3). The AT-skew and GC-skew are -0.148 and 0.183, respectively (Table 3), suggesting an obvious bias toward the use of Ts and Gs.

Fig. 1. Gene map of the *Coenobita brevimanus* mitogenome.

The mitogenome of *C. brevimanus* contains 13 PCGs, with a total length of 11,159 bp. Eight PCGs (*COI*, *COII*, *ND2*, *ATP8*, *ATP6*, *COIII*, *ND6* and *Cyt b*) are encoded on the heavy strand (H-strand), while the rest (*ND5*, *ND4*, *ND4L*, *ND1*, and *ND3*) are encoded on the light strand (L-strand). Typically, the *ND3* gene is

encoded on the H-strand. Interestingly, it is inverted to the L-strand, which to our best knowledge, is a quite rare phenomenon only occurring in Coenobitidae mitogenomes [5]. Totally, it encodes 3,708 amino acids. The most frequently used amino acids are *Leu* (15.6%), *Phe* (9.0%), *Ile* (8.2%) and *Val* (7.6%), while the least common amino acids are *Cys* (1.1%), *Arg* (1.6%), *Gln* (1.9%), and *Asp* (1.9%) (Fig. 2A). Relative synonymous codon usage (RSCU) values for the third positions of the 13 PCGs is shown in Fig. 2B. The usage of both two- and four-fold degenerate codons is biased toward the use of codons abundant in T or A, in accord with other crabs. The AT content of the 13 PCGs is 63.7%. The AT-skew and GC-skew are -0.221 and 0.041, respectively (Table 3).

Like most crab mitogenomes, the *C. brevimanus* mitogenome contains a set of 22 tRNA genes [2 , 45 , 46]. In most crab mitogenomes, eight tRNAs are encoded on the L-strand and the other 14 tRNAs are encoded on the H-strand [2 , 45 , 46]. However, the number of tRNAs encoding on the two strands is equal (Fig. 1, Table 2). The tRNA genes range in size from 61 bp (*Arg*) to 70 bp (*Gln*) and the total length of them is 1,457 bp (Tables 2, 3). It shows a moderate AT bias (67.4%), a slight skew of T versus A (AT-skew = -0.009), and strong skew of G versus C (GC-skew = 0.145) (Table 3). The *16S rRNA* is 1,410 bp between *ND1* and *Val* while *12S rRNA* is 797 bp between *Val* and CR (Fig. 1, Table 2). The AT-skew (0.049 and 0.076, respectively) and GC-skew (0.052 and 0.036, respectively) of the two rRNA genes were both positive (Table 3), indicating clearly that more As and Gs than Ts and Cs in rRNAs. The CR is located between *12S rRNA* and Ser_1, with a slight AT bias (62.0%). The AT-skew and GC-skew is -0.031 and 0.042, respectively (Table 3), indicating an obvious bias toward the use of Ts and Gs.

Fig. 2. Amino acid composition in*C. brevimanus* mitogenome (A); Relative synonymous codon usage in *C. brevimanus* mitogenome (B).

Gene rearrangement

The gene arrangement in the complete mitogenome of *C. brevimanus* is shown in Fig. 3. Compared with the gene order in ancestral crustaceans (the pancrutacean ground pattern) mitogenomes [47], the gene order in *C. brevimanus* mitogenome undergoes a large-scale rearrangement. Totally, at least five gene clusters (or genes) dramatically alter the typical order, involving eleven tRNA genes (*G*, *A*, S_1, *P*, L_1, *I*, *Q*, *M*, *W*, *C*, *Y*), and two PCGs (*ND3* and *ND2*). If not considering these gene arrangements, the gene order *COI- L2- COII- K- A- ATP8- ATP6- COIII- R- N- E- F- ND5- H- ND4- ND4L- T- ND6- Cyt b-* S_2*- ND1-16S -V- 12S-* CR remains the same arrangement as that in ancestral crustaceans. Of these five gene rearrangements, *G- ND3- A-* S_1clusteris inverted from the downstream of *COIII* in the H-strand to downstream of the CR in the L-strand (Fig. 3). A single *P* moves from the downstream of *T* to downstream of the S_2 (Fig. 3). A single L_1 moves to the position between S_1*- A- ND3- G*cluster and *Y- W- Q- C* cluster, which is located downstream the CR and forms a large-scale rearranged area (Fig. 3). *I- Q- M- ND2* cluster is divided into two sections, one (*I*, *M* and *ND2*) is shifted to downstream of *K*. The other (*Q*) is shifted to the end of linear mitogenome (Fig. 3).The *W- C- Y* cluster order is changed into *Y- W- C* order, accompanied with *W* and *Y* inversion (Fig. 3).

How did this particular order of mitogenome emerge? Compared the four major common used mechanisms mentioned above, here, we propose that TDRL and recombination model result in the generation of the *C. brevimanus* mitogenome. Firstly, three gene clusters undergo a complete copy, forming three dimeric blocks, (*G- ND3- A- R- N- S_1- E*)- (*G- ND3- A- R- N- S_1- E*) (Fig. 4A), (*I- Q- M- ND2*)- (*I- Q- M- ND2*) (Fig. 4A), and (W- C- Y)- (W- C- Y) (Fig. 4A). Consecutive copies are then followed by a random loss of the duplicated genes. *G- ND3- A- R- N- S_1- E- G- ND3- A- R- N- S_1- E*, *I- Q- M- ND2- I- Q- M- ND2*, and *W- C- Y- W- C- Y* (underline denotes the deleted gene). Then three new gene blocks are formed, *G- ND3- R- N- E- A- S_1*, *Q- M- I- ND2*, and *Y- W- C* (Fig. 4A). Tandem duplication followed by random loss has been widely used to explain this type of translocation of mitochondrial genes [2, 48, 49], hence, we adopt TDRL model to explain these three gene block rearrangements. Subsequently, the two new gene blocks undergo a translocation. *G- ND3- R- N- E- A- S_1* block is translocated downstream to the CR, leaving *R- N- E* in the original position. *Q- M- I- ND2* block is translocated to the *K* and *D* junction (Fig. 4A). According to the reported rearrangements [2, 45, 50], two independent recombination events seem to be the most plausible explanation for these translocations. In the second step, four genes or gene clusters are translocated (Fig. 4B). *Q* is translocated to the position of *W* and *C* junction (Fig. 4B), *P* is translocated to the downstream of S_2 (Fig. 4B), L_1 is translocated to the downstream of S_1 (Fig. 4B), *G- ND3- A- S_1* order is reversed to *S_1- A- ND3- G* in the original position (Fig. 4B). Also, recombination events appear to account for these translocations. Finally, the ultimate gene arrangement of the *C. brevimanus* mitogenome is shown in Fig. 4C.

Fig. 3. Gene rearrangements in *C. brevimanus* mitogenome. PCGs and CR are indicated with boxes, and tRNAs are indicated with columns. Genes labeled above the diagram are encoded on the H-strand and those below the diagram on the L-strand. The gene rearrangement steps are labeled with Figs. (A) The ancestral gene arrangement of crustaceans; (B) The gene order in the *C. brevimanus* mitogenome.

Fig. 4. Inferred intermediate steps between the ancestral gene arrangement of crustaceans and *C. brevimanus* mitogenome. (A) Duplication-loss and translocation in the ancestral mitogenome of crustaceans. The duplicated gene block is boxed in dash and the lost genes are labeled with gray. (B) Translocation. (C) The final gene order in the *C. brevimanus* mitogenome.

Phylogenetic analysis

To further investigate the phylogenetic relationships of Anomura and the position of *C. brevimanus*, two phylogenetic trees (ML tree and BI tree) were constructed based on the nucleotide sequences of the 13 concatenated PCGs. In this study, both trees are largely congruent with each other; consequently, only the BI topology is shown, but both the ML bootstrap values and BI posterior probabilities are shown (Fig. 5). It is obvious that two *C. brevimanus* species cluster together and four *Coenobita* species form a clade. The largest terrestrial crab, *Birgus latro*, has the closest relationship with *Coenobita*, and form a Coenobitidae clade with high support value.

The current phylogenetic analysis of Anomura recovers a polyphyletic Paguroidea similar to previous studies [51-53], with the Coenobitidae + Diogenidae clade dissociates from the other paguroids (Lithodidae + Paguridae + Pylochelidae). The Coenobitidae + Diogenidae clade (*Coenobita* +*Birgus* + *Clibanarius*) is similar to what was reported by McLaughlin et al. [51] based on morphological characters and by Tan et al. [5] based on the amino acid dataset of 13 PCGs. While the other paguroids clade (Lithodidae + Paguridae + Pylochelidae) differs from most morphological results [51 , 54 , 55] and Tan et al.'s [5] molecular result. In these studies, Lithodidae is excluded from Paguroidea and belongs to a new superfamily Lithodoidea. The phylogenetic tree also recovers polyphyletic groups for Galatheoidea. In this study, Galatheoidea consists of two clades: (Porcellanidae + Munidopsidae + Munididae) forms a clade and dissociates from the single Chirostylidae clade. This result is consistent with Tan et al.'s [5] phylogenetic relationship. However, in Tan et al.'s [5] opinion, they treat Chirostylidae as a new superfamily (Chirostyloidea), while not the previous Galatheoidea [51, 56, 57]. Hence, Galatheoidea form a monophyletic group in their findings [5].

In contrast to most other studies that places the Hippoidea in a basal position of Anomura [52, 58, 59], both Tan et al. [5] and our result show Hippoidea at a non-basal position, in which Hippoidea is represented by only a single *Stemonopa insignis* species and the nodal support is low (76% bootstrap value), hence, its novel position is possibly driven by incomplete taxon sample, and should be treated with some level of caution. Similar situations occur in the placement of families Pylochelidae, Lomidae, and Albuneidae, in which the ML bootstrap values of these nodes are relatively low (74%, 77%, and 76%, respectively). Single representative of these families possibly cause the relatively low supporting values. As a result, further taxonomic sampling is needed to confirm the validity of these phylogenetic placement in future studies.

Fig. 5. Phylogenetic tree of Anomura species inferred from the 13 PCGsbased on Bayesian inference (BI) and maximum likelihood(ML) analysis. * at each node indicates 100% supporting value and the number indicates the maximum likelihood bootstrap value. The number after the species name is the GenBank accession number. Superfamilies as recognized by McLaughlin et al. [51]

Declarations

Acknowledgements

This work was supported by the Scientific Research Foundation for the Introduction of Talent of Zhejiang Ocean University.

Authors' contributions

LG, XTL, KHZ and LQL conceived and designed the research; ZFW, LHJ and ZML collected and analyzed the datasets; LG and BJL wrote the manuscript; XTL and KHZ performed the experiments. All authors have read and approved the manuscript.

Ethics approval and consent to participate

This study was carried out in strict accordance with the recommendations and guidelines of the National Institutes of Health. All experimental protocols were approved by the Research Ethics Committee of Chinese Academy of Sciences. No specific permits were required because the specimen used in this study was dead before being collected.

Competing interests

The authors declare that they have no conflict of interests.

Corresponding author. E-mail address: gongli1027@163.com (L. Gong) bjetbj@163.com (B. Liu)
These authors contributed equally to this paper.

References

1. Boore JL: Animal mitochondrial genomes. *Nucleic Acids Res* 1999, 27(8):1767-1780.

2. Gong L, Jiang H, Zhu K, Lu X, Liu L, Liu B, Jiang L, Ye Y, Lü Z: Large-scale mitochondrial gene rearrangements in the hermit crab *Pagurus nigrofascia* and phylogenetic analysis of the Anomura. *Gene* 2019, 695:75-83.

3. Ma H, Ma C, Li C, Lu J, Zou X, Gong Y, Wang W, Chen W, Ma L, Xia L: First mitochondrial genome for the red crab (*Charybdis feriata*) with implication of phylogenomics and population genetics. *Sci Rep* 2015, 5:11524.

4. Wang Z, Wang Z, Shi X, Wu Q, Tao Y, Guo H, Ji C, Bai Y: Complete mitochondrial genome of *Parasesarma affine* (Brachyura: Sesarmidae): Gene rearrangements in Sesarmidae and phylogenetic analysis of the Brachyura. *Int J Biol Macromol* 2018, 118:31-40.

5. Tan MH, Gan HM, Lee YP, Linton S, Grandjean F, Bartholomei-Santos ML, Miller AD, Austin CM: ORDER within the chaos: Insights into phylogenetic relationships within the Anomura (Crustacea: Decapoda) from mitochondrial sequences and gene order rearrangements. *Mol Phylogenet Evol* 2018, 127:320-331.

6. Gong L, Shi W, Yang M, Li D, Kong X: Novel gene arrangement in the mitochondrial genome of *Bothus myriaster* (Pleuronectiformes: Bothidae): evidence for the Dimer-Mitogenome and Non-random Loss model. *Mitochondrial DNA Part A* 2015, 27(05):3089-3092.

7. Zhang J-Y, Zhang L-P, Yu D-N, Storey KB, Zheng R-Q: Complete mitochondrial genomes of *Nanorana taihangnica* and *N. yunnanensis* (Anura: Dicroglossidae) with novel gene arrangements and phylogenetic relationship of Dicroglossidae. *BMC Evol Biol* 2018, 18(1):26.

8. Liu J, Yu J, Zhou M, Yang J: Complete mitochondrial genome of *Japalura flaviceps:* Deep insights into the phylogeny and gene rearrangements of Agamidae species. *Int J Biol Macromol* 2019, 125:423-431.

9. Eberhard JR, Wright TF: Rearrangement and evolution of mitochondrial genomes in parrots. *Mol Phylogenet Evol* 2016, 94(Pt A):34-46.

10. Arndt A, Smith MJ: Mitochondrial gene rearrangement in the sea cucumber genus *Cucumaria*. *Mol Biol Evol* 1998, 15(8):1009-1016.

11. Jiang L, Kang L, Wu C, Chen M, Lü Z: A comprehensive description and evolutionary analysis of 9 Loliginidae mitochondrial genomes. *Hydrobiologia* 2018, 808(1):115-124.

12. Wu X, Li X, Li L, Xu X, Xia J, Yu Z: New features of Asian *Crassostrea* oyster mitochondrial genomes: A novel alloacceptor tRNA gene recruitment and two novel ORFs. *Gene* 2012.

13. Liu Q-N, Xin Z-Z, Zhu X-Y, Chai X-Y, Zhao X-M, Zhou C-L, Tang B-P: A transfer RNA gene rearrangement in the lepidopteran mitochondrial genome. *Biochem Bioph Res Co* 2017, 489(2):149-154.

14. Kim S, Choi H-G, Park J-K, Min G-S: The complete mitochondrial genome of the subarctic red king crab, *Paralithodes camtschaticus* (Decapoda, Anomura). *Mitochondrial DNA* 2013, 24(4):350-352.

15. Sultana Z, Asakura A, Kinjo S, Nozawa M, Nakano T, Ikeo K: Molecular phylogeny of ten intertidal hermit crabs of the genus *Pagurus* inferred from multiple mitochondrial genes, with special emphasis on the evolutionary relationship of *Pagurus lanuginosus* and *Pagurus maculosus*. *Genetica* 2018, 146(4-5):369-381.

16. Hickerson M, Cunningham C: Dramatic mitochondrial gene rearrangements in the hermit crab *Pagurus longicarpus* (Crustacea, Anomura). *Mol Biol Evol* 2000, 17(4):639-644.

17. Cantatore P, Gadaleta MN, Roberti M, Saccone C, Wilson AC: Duplication and remoulding of tRNA genes during the evolutionary rearrangement of mitochondrial genomes. *Nature* 1987, 329(6142):853-855.

18. C Moritz, T E Dowling a, Brown WM: Evolution of Animal Mitochondrial DNA: Relevance for Population Biology and Systematics. *Annu Rev Ecol Syst* 1987, 18(1):269-292.

19. Lunt DH, BC H: Animal mitochondrial DNA recombination. *Nature* 1997, 387(6630):247.

20. Lavrov DV, Boore JL, Brown WM: Complete mtDNA sequences of two millipedes suggest a new model for mitochondrial gene rearrangements: duplication and nonrandom loss. *Mol Biol Evol* 2002, 19(2):163-169.

21. McLaughlin PA: Annotated checklist of anomuran decapod crustaceans of the world (exclusive of the Kiwaoidea and families Chirostylidae and Gal: atheidae of the Galatheoidea) Part I?Lithodoidea, Lomisoide and Paguroidea. *The Raffles B Zool* 2010, 23:5-107.

22. Hamasaki K, Iizuka C, Sanda T, Imai H, Kitada S: Phylogeny and phylogeography of the land hermit crab *Coenobita purpureus* (Decapoda: Anomura: Coenobitidae) in the Northwestern Pacific region. *Marine Ecology* 2017, 38(1):e12369.

23. Poupin J: Crustacea Decapoda of French Polynesia (Astacidea, Palinuridea, Anomura, Brachyura). Atoll Research Bulletin, vol. 442: National Museum of Natural History, Smithsonian Institution, Washington D.C.; 1996.

24. Smith MJ, Arndt A, Gorski S, Fajber E: The phylogeny of echinoderm classes based on mitochondrial gene arrangements. *J Mol Evol* 1993, 36(6):545-554.

25. Yuan M-L, Zhang Q-L, Zhang L, Guo Z-L, Liu Y-J, Shen Y-Y, Shao R: High-level phylogeny of the Coleoptera inferred with mitochondrial genome sequences. *Mol Phylogenet Evol* 2016, 104:99-111.

26. Kayal E, Bentlage B, Cartwright P, Yanagihara AA, Lindsay DJ, Hopcroft RR, Collins AG: Phylogenetic analysis of higher-level relationships within Hydroidolina (Cnidaria: Hydrozoa) using mitochondrial genome data and insight into their mitochondrial transcription. *PeerJ* 2015, 3:e1403.

27. Yang J-S, Yang W-J: The complete mitochondrial genome sequence of the hydrothermal vent galatheid crab *Shinkaia crosnieri* (Crustacea: Decapoda: Anomura): A novel arrangement and incomplete tRNA suite. *BMC Genomics* 2008, 9(1):257.

28. Dierckxsens N, Mardulyn P, Smits G: NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res* 2016, 45(4):e18-e18.

29. Stothard P, Wishart DS: Circular genome visualization and exploration using CGView. *Bioinformatics* 2005, 21(4):537-539.

30. Lowe TM, Chan PP: tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res* 2016, 44(W1):W54-W57.

31. Kumar S, Stecher G, Tamura K: MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016, 33(7):1870-1874.

32. Perna NT, Kocher TD: Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J Mol Evol* 1995, 41(3):353-358.

33. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R *et al*: Clustal W and Clustal X version 2.0. *Bioinformatics* 2007, 23(21):2947-2948.

34. Hall TA: BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 1999, 41(41):95-98.

35. Stamatakis A, Hoover P, Rougemont J: A Rapid Bootstrap Algorithm for the RAxML Web Servers. *Syst Biol* 2008, 57(5):758-771.

36. Huelsenbeck JP, Ronquist F: MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 2001, 17(8):754-755.

37. Posada D, Crandall KA: Modeltest: testing the model of DNA substitution. *Bioinformatics* 1998, 14(9):817-818.

38. Nylander JA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey J: Bayesian phylogenetic analysis of combined data. *Syst Biol* 2004, 53(1):47-67.

39. Zhang D, Zhou Y, Cheng H, Wang C: The complete mitochondrial genome of a yeti crab *Kiwa tyleri* Thatje, 2015 (Crustacea: Decapod: Anomura: Kiwaidae) from deep-sea hydrothermal vent. *Mitochondrial DNA Part B* 2017, 2(1):141-142.

40. Gan HY, Gan HM, Tan MH, Lee YP, Austin CM: The complete mitogenome of the hermit crab *Clibanarius infraspinatus* (Hilgendorf, 1869), (Crustacea; Decapoda; Diogenidae) – a new gene order for the Decapoda. *Mitochondrial DNA Part A* 2016, 27(6):4099-4100.

41. Tan MH, Gan HM, Lee YP, Austin CM: The complete mitogenome of the porcelain crab *Petrolisthes haswelli* Miers, 1884 (Crustacea: Decapoda: Anomura). *Mitochondrial DNA Part A* 2016, 27(6):3983-3984.

42. Shen H, Braband A, Scholtz G: Mitogenomic analysis of decapod crustacean phylogeny corroborates traditional views on their relationships. *Mol Phylogenet Evol* 2013, 66(3):776-789.

43. Lee CW, Song J-H, Min G-S, Kim S: The complete mitochondrial genome of squat lobster, *Munida gregaria* (Anomura, Galatheoidea, Munididae). *Mitochondrial DNA Part B* 2016, 1(1):204-206.

44. Tang B-P, Xin Z-Z, Liu Y, Zhang D-Z, Wang Z-F, Zhang H-B, Chai X-Y, Zhou C-L, Liu Q-N: The complete mitochondrial genome of *Sesarmops sinensis* reveals gene rearrangements and phylogenetic relationships in Brachyura. *PLOS ONE* 2017, 12(6):e0179800.

45. Tang B-P, Liu Y, Xin Z-Z, Zhang D-Z, Wang Z-F, Zhu X-Y, Wang Y, Zhang H-B, Zhou C-L, Chai X-Y *et al*: Characterisation of the complete mitochondrial genome of *Helice wuana* (Grapsoidea: Varunidae) and comparison with other Brachyuran crabs. *Genomics* 2017.

46. Shi G, Cui Z, Hui M, Liu Y, Chan T-Y, Song C: Unusual sequence features and gene rearrangements of primitive crabs revealed by three complete mitochondrial genomes of Dromiacea. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 2016, 20:65-73.

47. Boore JL, Lavrov DV, Brown WM: Gene translocation links insects and crustaceans. *Nature* 1998, 392:667.

48. Xin Z-Z, Zhang D-Z, Wang Z-F, Zhang H-B, Tang B-P, Zhou C-L, Chai X-Y, Liu Q-N: Mitochondrial genome of *Helice tientsinensis* (Brachyura: Grapsoidea: Varunidae): Gene rearrangements and higher-level phylogeny of the Brachyura. *Gene* 2017, 627:307-314.

49. Shi W, Gong L, Wang S-Y, Miao X-G, Kong X-Y: Tandem duplication and random loss for mitogenome rearrangement in *Symphurus* (Teleost: Pleuronectiformes). *BMC genomics* 2015, 16(1):355.

50. Kong X, Dong X, Zhang Y, Shi W, Wang Z, Yu Z: A novel rearrangement in the mitochondrial genome of tongue sole, *Cynoglossus semilaevis*: control region translocation and a tRNA gene inversion. *Genome* 2009, 52(12):975-984.

51. McLaughlin PA, Lemaitre R, Sorhannus U: Hermit Crab Phylogeny: A Reappraisal and Its “Fall-Out”. *J Crustacean Biol* 2007, 27(1):97-115.

52. Ahyong S, Schnabel K, W Maas E: Anomuran Phylogeny: New Insights from Molecular Data. Boca Raton: CRC Press; 2009.

53. Ahyong ST, O'Meally D: Phylogeny of the Decapoda reptantia: Resolution using three molecular loci and morphology. *Raffles B Zool* 2004, 52(2):673-693.

54. McLaughlin PA: Hermit Crabs—are They Really Polyphyletic? *J Crustacean Biol* 1983, 3(4):608-621.

55. Tomoyuki K: A check list of Thalassinidea and Anomura (Crustacea: Decapoda) from the South China Sea. *The Raffles B Zool* 2001, 48(8):343-376.

56. Ahyong S, Baba K, Macpherson E, Poore G: A new classification of the Galatheoidea (Crustacea: Decapoda: Anomura). *Zootaxa* 2010, 2676:57-68.

57. Schnabel K, Ahyong S: A new classification of the Chirostyloidea (Crustacea: Decapoda: Anomura). *Zootaxa* 2010, 2687:56-64.

58. Porter ML, Pérez-Losada M, Crandall KA: Model-based multi-locus estimation of decapod phylogeny and divergence times. *Mol Phylogenet Evol* 2005, 37(2):355-369.

59. Jones W, Macpherson E, Segonzac M: A new squat lobster family of Galatheoidea (Crustacea, Decapoda: Anomura) from the hydrothermal vents of the Pacific-Antarctic Ridge. *Zoosystema* 2005, 723:709-723.

Tables

Table 1. List of 38 Anomura species used in this paper

Species	Family	Superfamily	Length (bp)	Accession No.	Reference
Lithodes nintokuae	Lithodidae	Paguroidea	15731	NC_024202	unpublished
Paralithodes camtschaticus	Lithodidae	Paguroidea	16720	NC_020029	[12]
Paralithodes brevipes	Lithodidae	Paguroidea	16303	NC_021458	unpublished
Pagurus japonicus	Paguridae	Paguroidea	16401	LC222532	[13]
Pagurus filholi	Paguridae	Paguroidea	15674	LC222528	[13]
Pagurus minutus	Paguridae	Paguroidea	14939	LC222533	[13]
Pagurus gracilipes	Paguridae	Paguroidea	16051	LC222534	[13]
Pagurus nigrofascia	Paguridae	Paguroidea	15423	MH756635	unpublished
Pagurus lanuginosus	Paguridae	Paguroidea	14632	LC222527	[13]
Pagurus maculosus	Paguridae	Paguroidea	15420	LC222524	[13]
Pagurus sp.	Paguridae	Paguroidea	14648	LC222535	[13]
Pagurus longicarpus	Paguridae	Paguroidea	15630	NC_003058	[14]
Pylocheles mortensenii	Pylochelidae	Paguroidea	15093	KY352242	[15]
Lomis hirta	Lomidae	Lomoidea	17239	KY352239	[15]
Aegla aff. longirostri	Aeglidae	Aegloidea	15387	MF457407	[15]
Kiwa tyleri	Kiwaidae	Kiwaoidea	16865	NC_034927	[33]
Gastroptychus roger	Chirostylidae	Galatheoidea	16504	KY352238	[15]
Gastroptychus investigatoris	Chirostylidae	Galatheoidea	16423	KY352237	[15]
Stemonopa insignis	Albuneidae	Hippoidea	15596	KY352240	[15]
Clibanarius infraspinatus	Diogenidae	Paguroidea	16504	NC_025776	[34]
Birgus latro	Coenobitidae	Paguroidea	16411	KY352241	[15]
Coenobita brevimanus	Coenobitidae	Paguroidea	16390	KY352233	[15]
Coenobita brevimanus	Coenobitidae	Paguroidea	16393	MK310257	This study
Coenobita perlatus	Coenobitidae	Paguroidea	16447	KY352234	[15]
Coenobita variabilis	Coenobitidae	Paguroidea	16421	KY352236	[15]

Coenobita rugosus	Coenobitidae	Paguroidea	16427	KY352235	[15]
Petrolisthes haswelli	Porcellanidae	Galatheoidea	15348	NC_025572	[35]
Neopetrolisthes maculatus	Porcellanidae	Galatheoidea	15324	NC_020024	[36]
Shinkaia crosnieri	Munidopsidae	Galatheoidea	15182	NC_011013	[37]
Munida gregaria	Munididae	Galatheoidea	16326	NC_030255	[38]
Munida isos	Munididae	Galatheoidea	17910	NC_039112	[15]
Tubuca polita	Ocypodidae	Ocypodoidea	15672	MF457400	[15]
Tubuca capricornis	Ocypodidae	Ocypodoidea	15629	MF457401	[15]
Cranuca inversa	Ocypodidae	Ocypodoidea	15677	MF457405	[15]
Pachygrapsus marmoratus	Grapsidae	Grapsoidea	15406	MF457403	[15]
Cardisoma carnifex	Gecarcinidae	Grapsoidea	15597	MF461623	[15]
Epixanthus frontalis	Oziidae	Xanthoidea	15993	MF457404	[15]
Pilumnus vespertilio	Pilumnidae	Pilumnoidea	16222	MF457402	[15]

Table 2. Features of the mitochondrial genome of *Coenobita brevimanus*

Gene	Position		Length (bp)	Amino acid	Start/Stop codon	Anticodon	Intergenic region*	Strand
	From	To						
COI	1	1539	1539	512	ATG/TAA		-5	H
Leu(L2)	1535	1600	66			TAG	7	H
COII	1608	2297	690	229	ATG/TAG		9	H
Lys(K)	2307	2371	65			TTT	7	H
Met(M)	2379	2445	67			CAT	11	H
Ile(I)	2457	2522	66			GAT	16	H
ND2	2539	3574	1036	345	ATT/T		0	H
Asp(D)	3575	3639	65			GTC	0	H
ATP8	3640	3798	159	52	ATC/TAG		-7	H
ATP6	3792	4466	675	224	GTG/TAA		-1	H
COIII	4466	5257	792	263	ATG/TAA		15	H
Arg(R)	5273	5333	61			TCG	0	H
Asn(N)	5334	5399	66			GTT	5	H
Glu(E)	5405	5470	66			TTC	3	H
Phe(F)	5474	5538	65			GAA	13	L
ND5	5552	7249	1698	565	ATA/TAA		18	L
His(H)	7268	7334	67			GTG	48	L
ND4	7383	8723	1341	446	ATG/TAA		-7	L
ND4L	8717	9001	285	94	TTG/TAA		20	L
Thr(T)	9022	9089	68			TGT	10	H
ND6	9100	9633	534	177	GTG/TAA		-17	H
Cyt b	9617	10748	1132	377	ATA/T		0	H
Ser(S2)	10749	10814	66			TGA	-1	H
Pro(P)	10814	10879	66			TGG	1	L
ND1	10881	11807	927	308	ATC/TAA		0	L
16S	11808	13217	1410				0	L

Val(V)	13218	13285	68			TAC	0	L
12S	13286	14082	797				0	L
CR	14083	15459	1377				0	H
Ser(S1)	15460	15525	66			TCT	4	L
Ala(A)	15530	15593	64			TGC	16	L
ND3	15610	15960	351	116	ATG/TAG		0	L
Gly(G)	15961	16026	66			TCC	3	L
Leu(L1)	16030	16095	66			TAA	-1	L
Tyr(Y)	16095	16161	67			GTA	5	H
Trp(W)	16167	16235	69			TCA	14	L
Gln(Q)	16250	16319	70			TTG	3	L
Cys(C)	16323	16389	67			GCA	0	L

*Intergenic region: non-coding bases between the feature on the same line and the line below, with a negative number indicating an overlap.

Table 3. Composition and skewness of *Coenobita brevimanus* mitogenome

	A%	T%	G%	C%	A+T%	AT-skew	GC-skew	Length(bp)
Mitogenome	27.7	37.3	20.7	14.3	65.0	-0.148	0.183	16393
PCGs	24.8	38.9	18.9	17.4	63.7	-0.221	0.041	11159
COI	20.7	40.7	24.4	14.2	61.4	-0.326	0.264	1539
COII	22.3	40.6	23.8	13.3	62.9	-0.291	0.283	690
ND2	18.4	47.1	21.6	12.8	65.5	-0.438	0.256	1036
ATP8	29.6	39.6	21.4	9.4	69.2	-0.145	0.390	159
ATP6	20.7	45.9	20.0	13.3	66.6	-0.378	0.201	675
COIII	19.8	42.7	22.9	14.6	62.5	-0.366	0.221	792
ND5	32.1	31.2	13.3	23.4	63.3	0.014	-0.275	1698
ND4	31.4	31.7	14.5	22.4	63.1	-0.005	-0.214	1341
ND4L	25.3	36.5	15.1	23.2	61.8	-0.181	-0.211	285
ND6	23.2	44.9	19.9	12.0	68.1	-0.319	0.248	534
Cyt b	19.7	44.2	21.2	14.9	63.9	-0.383	0.175	1132
ND1	29.3	34.4	14.9	21.4	63.7	-0.080	-0.179	927
ND3	28.8	33.3	14.0	23.9	62.1	-0.072	-0.261	351
tRNAs	33.4	34.1	18.6	13.9	67.4	-0.009	0.145	1457
16S	38.0	34.5	14.5	13.0	72.5	0.049	0.052	1410
12S	37.1	31.9	16.1	14.9	69.0	0.076	0.036	797
CR	30.1	31.9	19.8	18.2	62.0	-0.031	0.042	1377

Figures

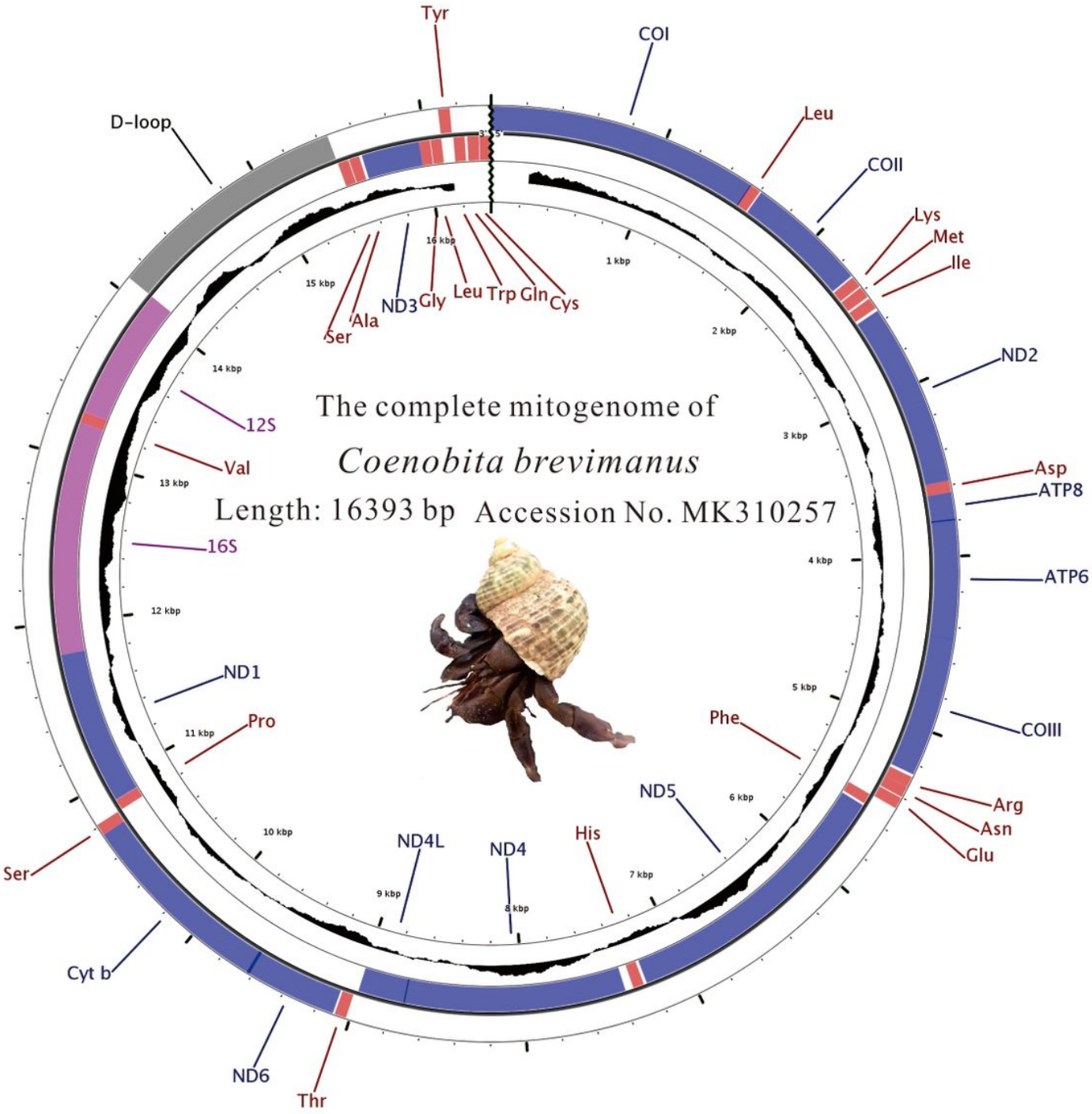

Figure 1

Gene map of the Coenobita brevimanus mitogenome.

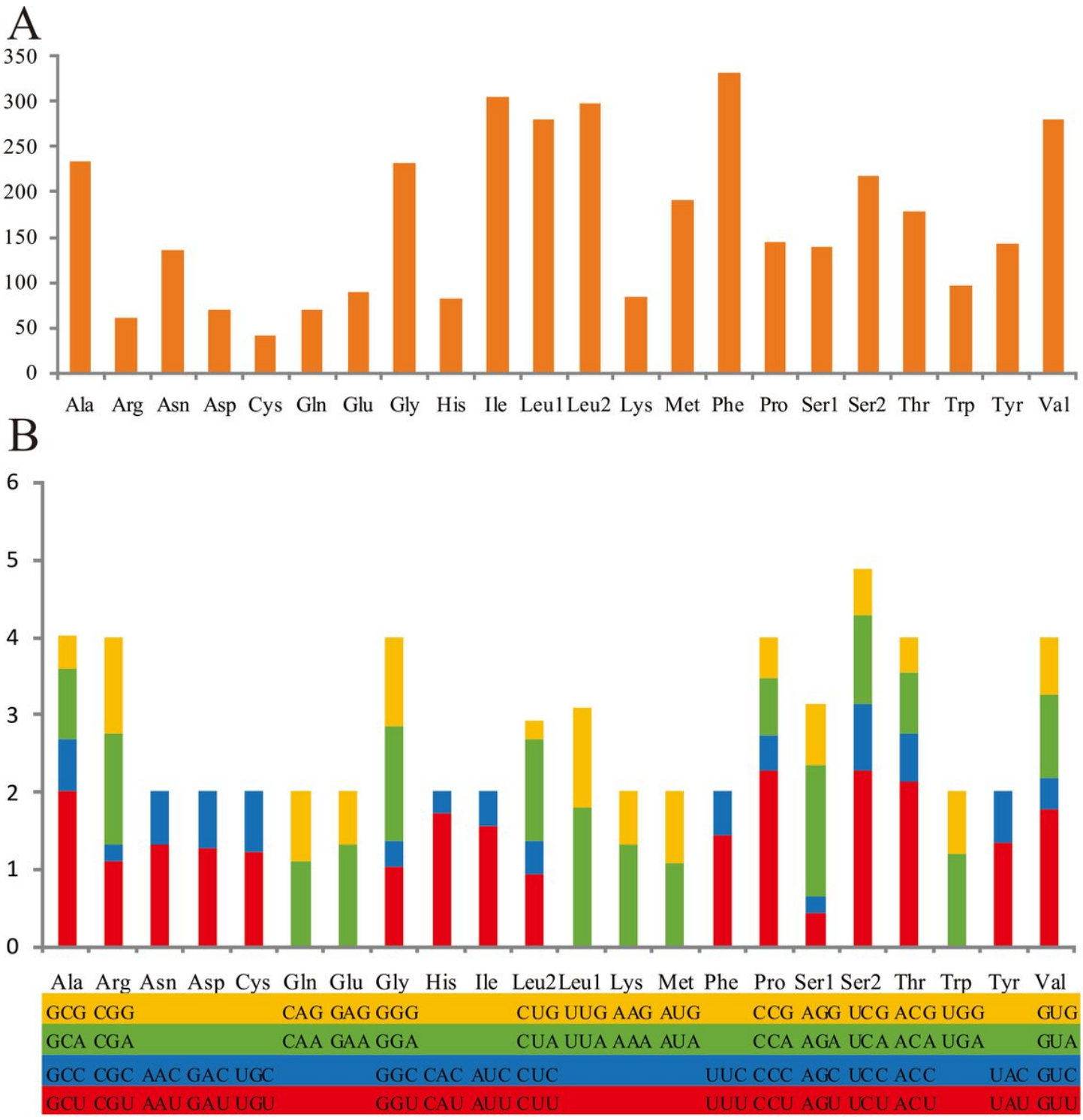

Figure 2

Amino acid composition in C. brevimanus mitogenome (A); Relative synonymous codon usage in C. brevimanus mitogenome (B).

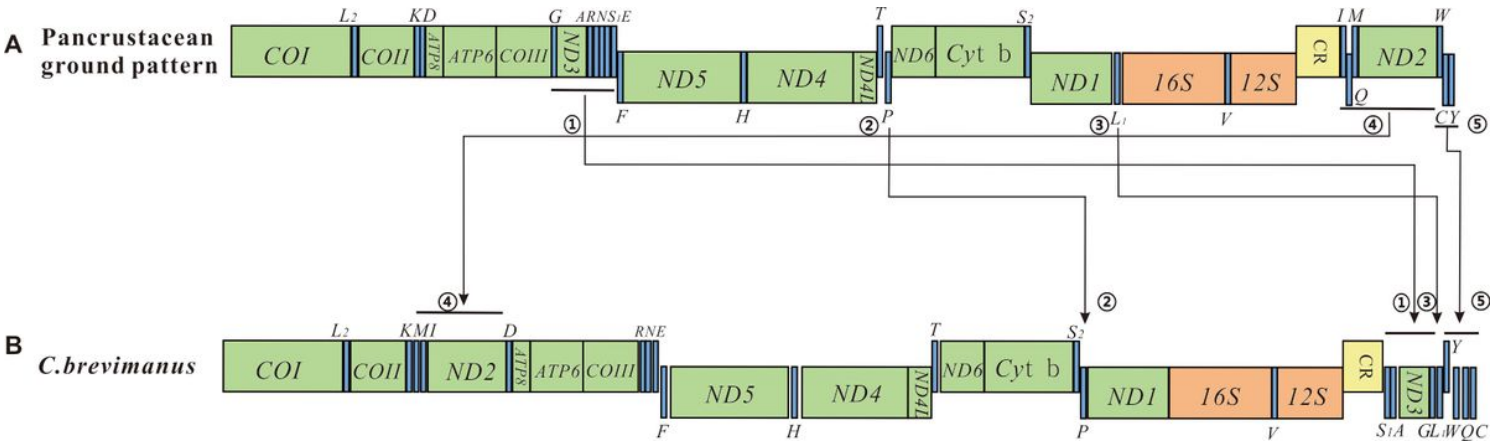

Figure 3

Gene rearrangements in C. brevimanus mitogenome. PCGs and CR are indicated with boxes, and tRNAs are indicated with columns. Genes labeled above the diagram are encoded on the H-strand and those below the diagram on the L-strand. The gene rearrangement steps are labeled with Figs. (A) The ancestral gene arrangement of crustaceans; (B) The gene order in the C. brevimanus mitogenome.

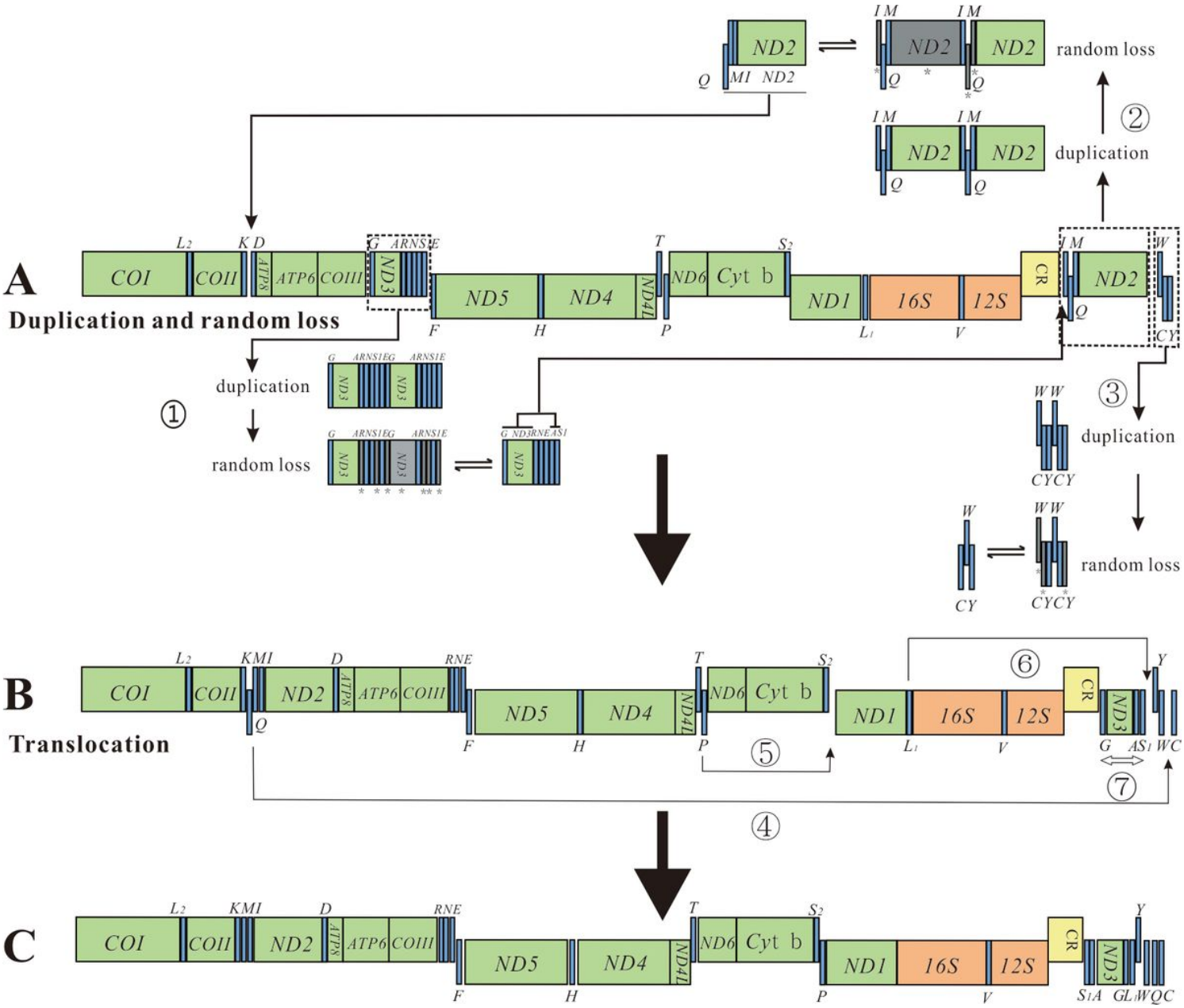

Figure 4

Inferred intermediate steps between the ancestral gene arrangement of crustaceans and C. brevimanus mitogenome. (A) Duplication-loss and translocation in the ancestral mitogenome of crustaceans. The duplicated gene block is boxed in dash and the lost genes are labeled with gray. (B) Translocation. (C) The final gene order in the C. brevimanus mitogenome.

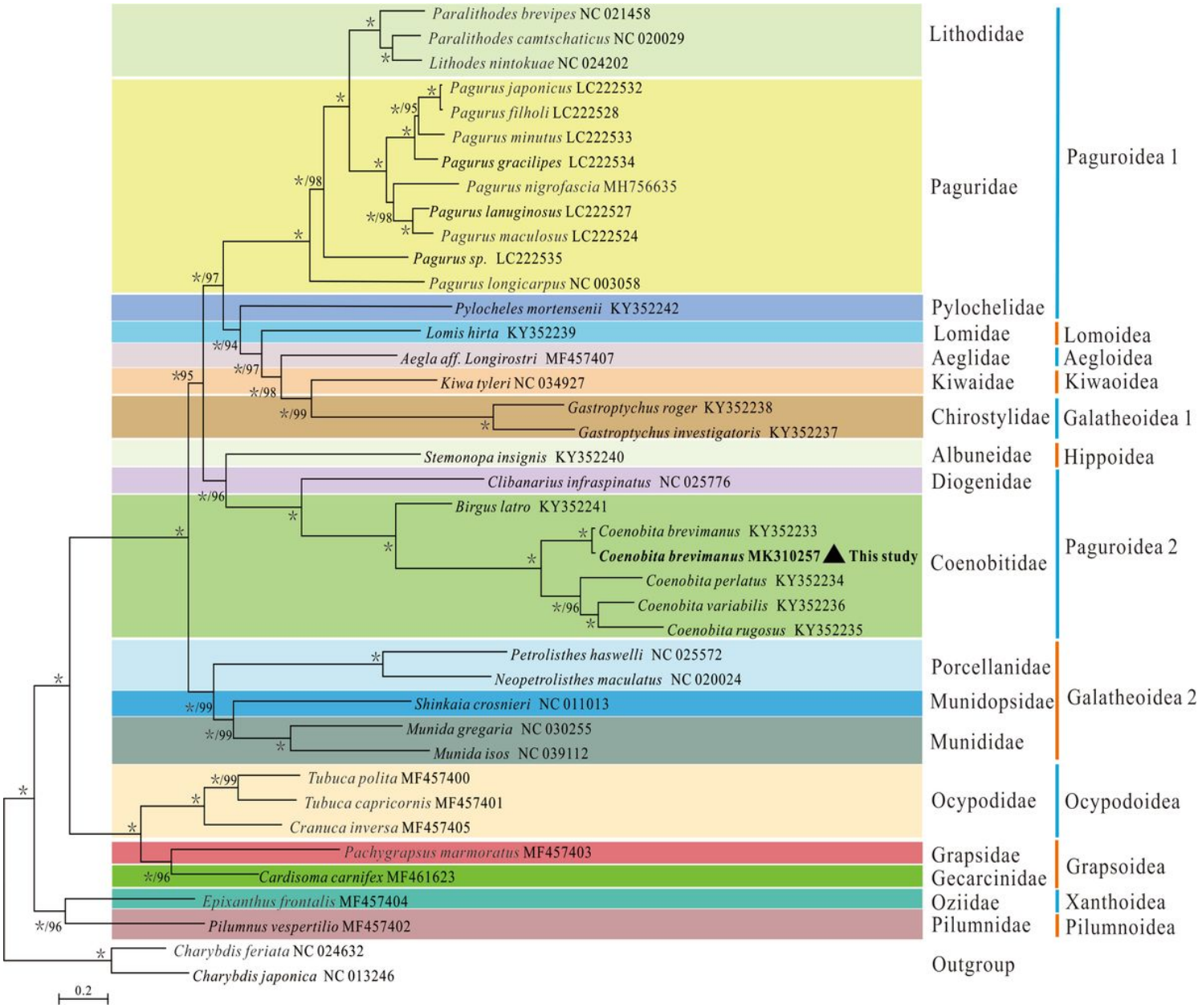

Figure 5

Phylogenetic tree of Anomura species inferred from the 13 PCGsbased on Bayesian inference (BI) and maximum likelihood(ML) analysis. * at each node indicates 100% supporting value and the number indicates the maximum likelihood bootstrap value. The number after the species name is the GenBank accession number. Superfamilies as recognized by McLaughlin et al. [51]