Digging deeper: new gene order rearrangements and distinct patterns of codons usage in mitochondrial genomes among shrimps from the Axiidea, Gebiidea and Caridea (Crustacea: Decapoda)

Mun Hua Tan[1,2], Han Ming Gan[1,2], Yin Peng Lee[1,2], Gary C.B. Poore[3] and Christopher M. Austin[1,2,4]

[1] School of Science, Monash University Malaysia, Bandar Sunway, Selangor, Malaysia
[2] Genomics Facility, Tropical Medicine and Biology Platform, Monash University Malaysia, Bandar Sunway, Selangor, Malaysia
[3] Museums Victoria, Melbourne, VIC, Australia
[4] School of Life and Environmental Sciences, Deakin University, Burwood, VIC, Australia

ABSTRACT

Background: Whole mitochondrial DNA is being increasingly utilized for comparative genomic and phylogenetic studies at deep and shallow evolutionary levels for a range of taxonomic groups. Although mitogenome sequences are deposited at an increasing rate into public databases, their taxonomic representation is unequal across major taxonomic groups. In the case of decapod crustaceans, several infraorders, including Axiidea (ghost shrimps, sponge shrimps, and mud lobsters) and Caridea (true shrimps) are still under-represented, limiting comprehensive phylogenetic studies that utilize mitogenomic information.
Methods: Sequence reads from partial genome scans were generated using the Illumina MiSeq platform and mitogenome sequences were assembled from these low coverage reads. In addition to examining phylogenetic relationships within the three infraorders, Axiidea, Gebiidea, and Caridea, we also investigated the diversity and frequency of codon usage bias and mitogenome gene order rearrangements.
Results: We present new mitogenome sequences for five shrimp species from Australia that includes two ghost shrimps, *Callianassa ceramica* and *Trypaea australiensis*, along with three caridean shrimps, *Macrobrachium bullatum*, *Alpheus lobidens,* and *Caridina* cf. *nilotica*. Strong differences in codon usage were discovered among the three infraorders and significant gene order rearrangements were observed. While the gene order rearrangements are congruent with the inferred phylogenetic relationships and consistent with taxonomic classification, they are unevenly distributed within and among the three infraorders.
Discussion: Our findings suggest potential for mitogenome rearrangements to be useful phylogenetic markers for decapod crustaceans and at the same time raise important questions concerning the drivers of mitogenome evolution in different decapod crustacean lineages.

Subjects Aquaculture, Fisheries and Fish Science, Biodiversity, Evolutionary Studies, Genomics, Marine Biology

INTRODUCTION

Sequencing of animal mitochondrial genomes has exploded in recent years. Over 40,000 animal mitogenomes are currently lodged on the NCBI database compared to fewer than 5,000 sequences a decade ago. As mitogenomic data have accumulated, it has become apparent that mitochondrial gene order is not as conserved as first thought, and that interesting and phylogenetically useful patterns that invite further research are emerging (*Boore, 2006*; *Dowton, Castro & Austin, 2002*; *Gan et al., 2016*; *Gissi, Iannelli & Pesole, 2008*; *Lin et al., 2012*; *Poulsen et al., 2013*; *Tan et al., 2015*).

In addition to gene order evolution, properties of mitogenome sequences that are of phylogenetic significance have emerged that are of interest to comparative mitogenomics such as gene loss and duplications, AT bias, strand asymmetry in nucleotide composition, length, and structure of the control region, features of intergenic non-coding regions, codon usage, variation in gene length, variation in start and stop codons, gene diversity levels, mutation rates, and signals of selection and secondary structures of ribosomal genes (*Bauzà-Ribot et al., 2009*; *Gissi, Iannelli & Pesole, 2008*; *Jia & Higgs, 2008*; *Li, Huang & Lei, 2015*; *Oliveira et al., 2008*; *Poulsen et al., 2013*; *Qian et al., 2011*; *Shoemaker et al., 2004*). However, despite the rapidly accumulating mitogenomic resources, there are gaps in taxonomic representation and more data are required to fully evaluate the usefulness of mitogenome gene rearrangements as phylogenetic markers in specific groups (*Mao, Gibson & Dowton, 2014*; *Tan et al., 2015*) and for broadly based comparative studies to detect patterns and investigate evolutionary hypotheses (*Boore, 2006*; *Castellana, Vicario & Saccone, 2011*; *Gissi, Iannelli & Pesole, 2008*; *Jiang et al., 2007*).

In general, the overall structure and function of animal mitochondrial genomes are remarkably stable. It is a circular, double-stranded DNA molecule of usually 15–20 kb in length, generally containing a consistent set of 37 genes made up of two ribosomal RNA genes (rRNA), 13 protein-coding genes (PCG) and 22 transfer RNA genes (tRNA) (*Bernt et al., 2013*; *Castellana, Vicario & Saccone, 2011*; *Gissi, Iannelli & Pesole, 2008*). Intergenic regions are usually minimal, although all species contain at least one large AT-rich region associated with strand replication e.g., the control region. Mutations in the mitochondrial DNA can range from point mutations and infrequent insertions/deletions to gene order rearrangements. However, the type, scale and distribution of mutations can vary widely across various taxonomic levels (*Gissi, Iannelli & Pesole, 2008*; *Tan et al., 2015*).

The most common mitogenome gene order rearrangements involve the translocation of single tRNA genes and occasionally with a change of transcriptional polarity or duplication. Less frequent is the repositioning of multiple tRNAs, duplication of the control region or changes to the order and orientation of protein coding and rRNA genes. The paradox of mitogenome gene rearrangements is that the molecule can be highly conserved among phylogenetically distant species such as some insect and decapod

crustacean species but can, in a restricted number of taxonomic groups, also vary substantially among species in the same family or genus. Examples of taxa with a concentration of reported mitogenome gene rearrangements include ticks (*Fahrein et al., 2007*; *Shao et al., 2005*), hymenopterans (*Dowton et al., 2003*), gulper eels (*Poulsen et al., 2013*), salamanders (*Chong & Mueller, 2013*), tunicates (*Gissi, Iannelli & Pesole, 2004*), and in several crustacean groups (*Kilpert & Podsiadlowski, 2006*; *Kim et al., 2012*; *Miller et al., 2004*; *Stokkan et al., 2016*; *Tan et al., 2015*).

Codon usage is another aspect of comparative mitogenomics that is attracting increasing attention as research moves from describing patterns to understanding them within an evolutionary and molecular genetic context (*Castellana, Vicario & Saccone, 2011*; *Gissi, Iannelli & Pesole, 2008*; *Jiang et al., 2007*). Amino acids can be encoded by two to six codons, but alternative codons for the same amino acid often do not occur at equal frequencies either between species for the same gene or between different genes in the one species. Patterns of differential codon usage have been attributable to selection, variable mutation rates, translational efficiency, and random factors (genetic drift) (*Castellana, Vicario & Saccone, 2011*; *Jia & Higgs, 2008*; *Prat et al., 2009*; *Whittle & Extavour, 2015*). However, it has been rarely addressed among crustacean species (*Cook, Yue & Akam, 2005*; *García-Machado et al., 1999*; *Rota-Stabelli et al., 2013*).

As in other major animal groups, mitogenome sequences are becoming increasingly available for decapod crustacean species, contributing to the understanding of the evolution of this taxonomically challenging group due to its high diversity, deep lineages, and highly flexible body plan (*Shen, Braband & Scholtz, 2013*; *Tan et al., 2015*). In addition, intriguing and taxonomically unevenly distributed gene order rearrangements are emerging, requiring further investigation and raising questions regarding the dynamics and drivers of mitogenome gene order evolution in several groups (*Gan et al., 2016b*; *Kilpert & Podsiadlowski, 2006*; *Miller et al., 2004*). Two of the less well-represented decapod crustacean infraorders in mitochondrial databases are the shrimp infraorders Axiidea (ghost shrimps, sponge shrimps, and mud lobsters) and Caridea (true shrimps) (*Lin et al., 2012*; *Tan et al., 2015*). Further, the phylogenetic relationships within and among these shrimp groups remain largely unresolved and disputed (*Lin et al., 2012*; *Timm & Bracken-Grissom, 2015*). A major limiting factor to the resolution of phylogenetic relationships within and among these shrimp lineages and the determination of the distribution and evolutionary significance of mitogenome gene order rearrangements is inadequate taxon sampling (*Lin et al., 2012*; *Shen, Braband & Scholtz, 2013*; *Tan et al., 2015*).

At the present time, the Axiidea, Gebiidea, and Caridea are represented by only six, five and 17 complete mitogenomes respectively on the NCBI public database. To support ongoing phylogenetic and comparative mitogenomic studies, this paper reports five new mitogenome sequences of shrimp species sampled from Australia. These include two mitogenomes from the ghost shrimp, *Callianassa ceramica* Fulton & Grant, 1906 and *Trypaea australiensis* Dana, 1852 (Callianassidae), and three from caridean shrimps *Macrobrachium bullatum* Fincham, 1987 (Paleamonidae), *Alpheus lobidens* De Haan,

Table 1 List of source and accession number of mitogenome samples.

	Callianassa ceramica	*Callianassa ceramica*	*Trypaea australiensis*	*Alpheus lobidens*	*Caridina* cf. *nilotica*	*Macrobrachium bullatum*
Family	Callianassidae	Callianassidae	Callianassidae	Alpheidae	Atyidae	Palaemonidae
Subfamily	Callianassinae	Callianassinae	Callianassinae	N/A	N/A	Palaemoninae
Location	Anglesea, South west of Geelong, Victoria	South of Port Authority Pier, Queenscliff, Victoria	Stony Point, Western Port Bay, Victoria	Nightcliff, Darwin, Northern Territory	Amy Ward Drive, Darwin, Northern Territory	Rapid Creek, Darwin, Northern Territory
Voucher	N/A	NMV J40715	NMV J40711	MAGNT Cr.18581	N/A	N/A
Verification	99.8% (660 bp *COI*)[1]	N/A	N/A	N/A	100% (326 bp 16S)[2]	100% (447 bp 16S)[3]
Accession #	KU362925.1	KU350630.1	KM501040.2	KP276147.1	KU726823.1	KM978918.1

Notes:
NMV, Museum Victoria; MAGNT, Museum and Art Galleries of the Northern Territory.
[1] NMV J70519, Point Roadknight, Anglesea, Victoria.
[2] *Page, von Rintelen & Hughes (2007)* – DQ478508.1 – MAGNT Cr. 9399.
[3] *Murphy & Austin (2004)* – AY282778.1.

1849 (Alpheidae) and *Caridina* cf. *nilotica* Roux, 1833 (Atyidae), each of which represents highly diverse superfamilies, families and genera within Caridea. This study compares the mitogenomic features of these five species together with additional representatives of their infraorders and Gebiidea available from the NCBI database. In addition to exploring evolutionary relationships within each infraorder, we uncover distinctive signatures and patterns with respect to sequence composition, codon usage bias, and gene rearrangements that can possibly act as synapomorphies for specific shrimp taxonomic groups, suggesting the potential of these features for phylogenetic inferences at different evolutionary scales.

MATERIALS AND METHODS

Sample collection

Two species belonging to the infraorder Axiidea (*C. ceramica*, *T. australiensis*) and three from Caridea (*Macrobrachium bullatum*, *A. lobidens*, *Caridina* cf. *nilotica*) were collected from different locations in Australia (Table 1). *C. ceramica* is represented by two individuals: one from a vouchered specimen (Museum Victoria J40715; GenBank accession number KU350630.1); and the other a previously sequenced sample published incorrectly as *T. australiensis* (KM501040.1) (*Gan et al., 2016*). The mitogenome sequence for this latter sample is now registered as *C. ceramica* under the accession number KU726823.1 and has a genetic similarity of 99.8% to the COI region of a vouchered *C. ceramica* specimen (Museum Victoria J70519) collected from the same general locality.

To maintain continuity with the original NCBI accession number and species name, the original mitogenome sequence lodged on NCBI for *T. australiensis* was updated with the newly sequenced mitogenome from a vouchered *T. australiensis* specimen (Museum Victoria J40711; GenBank accession number: KM501040.2). The accession numbers for each species and associated collecting and identification-related information are detailed in Table 1 and Data S3, including voucher numbers for specimens lodged in Museum

Table 2 List of samples and their corresponding accession numbers included in phylogenetic and comparative analyses.

Infraorder	Family	Species	Accession	Reference
Axiidea	Callianassidae	*Callianassa ceramica*	KU350630.1	This study
		Callianassa ceramica	KU362925.1	*Gan et al. (2016)*
		Corallianassa coutierei	NC_020025.1	*Shen, Braband & Scholtz (2013)*
		Nihonotrypaea japonica	NC_020351.1	*Kim et al. (2013)*
		Nihonotrypaea thermophila	NC_019610.1	*Lin et al. (2012)*
		Paraglypturus tonganus	NC_024651.1	*Kim et al. (2016b)*
		Trypaea australiensis[1]	KM501040.2	This study
	Strahlaxiidae	*Neaxius acanthus*[2]	NC_019609.1	*Lin et al. (2012)*
Gebiidea	Thalassinidae	*Thalassina kelanang*	NC_019608.1	*Lin et al. (2012)*
	Upogebiidae	*Austinogebia edulis*	NC_019606.1	*Lin et al. (2012)*
		Upogebia major	NC_019607.1	*Lin et al. (2012)*
		Upogebia pusilla	NC_020023.1	*Shen, Braband & Scholtz (2013)*
		Upogebia yokoyai	NC_025943.1	*Yang et al. (2016)*
Caridea	Alvinocarididae	*Alvinocaris chelys*	NC_018778.1	*Yang et al. (2012)*
		Alvinocaris longirostris	NC_020313.1	*Yang et al. (2013)*
		Nautilocaris saintlaurentae	NC_021971.1	*Kim, Pak & Ju (2015a)*
		Opaepele loihi	NC_020311.1	*Yang et al. (2013)*
		Rimicaris exoculata	NC_027116.1	*Yu et al. (2015)*
		Rimicaris kairei	NC_020310.1	*Yang et al. (2013)*
	Alpheidae	*Alpheus distinguendus*	NC_014883.1	*Qian et al. (2011)*
		Alpheus lobidens	KP276147.1	This study
	Atyidae	*Caridina gracilipes*	NC_024751.1	*Xu et al. (2016)*
		Caridina cf. *nilotica*	KU726823.1	This study
		Halocaridina rubra	NC_008413.1	*Ivey & Santos (2007)*
		Neocaridina denticulata	NC_023823.1	*Yu, Yang & Yang (2014)*
		Paratya australiensis	NC_027603.1	*Gan et al. (2016b)*
	Palaemonidae	*Macrobrachium bullatum*	KM978918.1	**This study**
		Macrobrachium lanchesteri	NC_012217.1	L. Ngernsiri & P. Sangthong, 2016, unpublished data
		Macrobrachium nipponense	NC_015073.1	*Ma et al. (2011)*
		Macrobrachium rosenbergii	NC_006880.1	*Miller et al. (2005)*
		Palaemon carinicauda	NC_012566.1	*Shen et al. (2009)*
		Palaemon gravieri	NC_029240.1	*Kim et al. (2015b)*
		Palaemon serenus	NC_027601.1	*Gan et al. (2016b)*
Dendrobranchiata (outgroup)	Sergestidae	*Acetes chinensis*	NC_017600.1	*Kim et al. (2012)*
	Penaeidae	*Farfantepenaeus californiensis*	NC_012738.1	*Gutiérrez-Millán et al. (2002)*
		Fenneropenaeus chinensis	NC_009679.1	*Shen et al. (2007)*
		Fenneropenaeus merguiensis	NC_026884.1	*Zhang et al. (2016a)*
		Fenneropenaeus penicillatus	NC_026885.1	*Zhang et al. (2015)*
		Litopenaeus vannamei	NC_009626.1	L. Ngernsiri & P. Sangthong, 2016, unpublished data
		Marsupenaeus japonicas	NC_007010.1	*Yamauchi et al. (2004)*

(Continued)

Table 2 (continued).

Infraorder	Family	Species	Accession	Reference
		Metapenaeopsis dalei	NC_029457.1	*Kim et al. (2016a)*
		Metapenaeus ensis	NC_026834.1	*Zhang et al. (2016b)*
		Parapenaeopsis hardwickii	NC_030277.1	*Mao et al. (2016)*
		Penaeus monodon	NC_002184.1	*Wilson et al. (2000)*
	Solenoceridae	*Solenocera crassicornis*	NC_030280.1	Y. Yuan et al., 2016, unpublished data

Note:
[1] Mitogenome from taxonomically verified *T. australiensis* sample resubmitted as version two under same accession number.
[2] Neaxius acanthus from Taiwan was wrongly identified as Neaxius glyptocercus by *Lin et al. (2012)*.

Victoria, Melbourne (NMV) and the Museum and Art Gallery of the Northern Territory, Darwin (MAGNT).

Next-generation sequencing and mitogenome assembly

Purification of ethanol-preserved tissue and partial whole genome sequencing (2×75 bp for *T. australiensis* and 2×250 bp for others) was performed on the Illumina MiSeq platform as previously described (*Gan et al., 2016*), after which each mitogenome was assembled with IDBA_UD v.1.1.1 (*Peng et al., 2012*) and annotated using MITOS (*Bernt et al., 2013*). Circular mitogenome maps were drawn with BRIG v.0.9.5 (*Alikhan et al., 2011*). Summary statistics including gene boundaries and length, strand, nucleotide composition, intergenic nucleotides, and number of genes were compiled with MitoPhAST v.1.0 (*Tan et al., 2015*). Alignment of whole mitogenome sequences and calculation of pair-wise nucleotide identities were performed with SDT v.1.2 (*Muhire, Varsani & Martin, 2014*).

Gene order analysis

Along with the five mitogenomes sequenced in this study, sequences from 28 other complete mitogenomes from the three infraorders were obtained from NCBI's RefSeq database (Table 2) for comparative analyses. Arrangements of genes for each of these 33 mitogenomes were compared with all other existing decapod mitogenomes in RefSeq to identify potential novel gene orders unreported by previous studies. Mitogenomes that exhibit gene orders differing from that of the pancrustacean ground pattern (*Boore, Lavrov & Brown, 1998*) were re-annotated with MITOS (*Bernt et al., 2013*) to confirm that differences observed are not due to misannotations. Any observed misannotations (e.g., missing genes, incorrect gene boundaries) were corrected before proceeding to further comparative and phylogenetic analyses.

Phylogenetic analysis

Mitogenomes listed in Table 2 were subject to phylogenetic analysis to establish the evolutionary relationships of species within each of the infraorders to provide a framework for interpreting mitogenome gene rearrangements. MitoPhAST v.1.0 (*Tan et al., 2015*) was used to extract individual PCG amino acid sequences, and these protein sequences were then separately aligned with MAFFT v.7.222 (*Katoh & Standley, 2013*), followed by trimming with trimAl v.1.4 (*Capella-Gutiérrez, Silla-Martínez & Gabaldón,*

2009). For nucleotide level analyses, PCG nucleotide sequences were manually extracted and fed to TranslatorX v.1.1 (*Abascal, Zardoya & Telford, 2010*), which aligns nucleotide sequences guided by amino acid translations and then trimmed with Gblocks v.0.19b (*Castresana, 2000*). On the other hand, rRNA was aligned with MAFFT v.7.222 (*mafft-linsi*) (*Katoh & Standley, 2013*) and trimmed with trimAl v.1.4 (*Capella-Gutiérrez, Silla-Martínez & Gabaldón, 2009*). Finally, mitochondrial PCG and rRNA sequences were concatenated into super-alignments to make up the following datasets:

A. 13 PCG (aa) [3,591 characters]

B. 13 PCG (nt) [9,642 characters]

C. 13 PCG (aa) + 12S rRNA + 16S rRNA [5,694 characters]

D. 13 PCG (nt) + 12S rRNA + 16S rRNA [11,755 characters]

Maximum-likelihood (ML) tree inference with ultrafast bootstrap (UFBoot) branch supports (*Minh, Nguyen & von Haeseler, 2013*) was performed using IQ-TREE v.1.5.0 (*Nguyen et al., 2015*), which also implements model selection to find the best-fit partitioning scheme. Super-alignments for all datasets were partitioned based on genes. An additional analysis for Dataset B that further partitions it according to first, second and third codon positions was also performed. For Bayesian inference, the same super-alignments generated from all datasets were analysed using Exabayes v.1.4.2 (*Aberer, Kobert & Stamatakis, 2014*). For each analysis, four independent runs were carried out concurrently for five million iterations each with 25% of initial samples discarded as burn-in. Convergence of chains was checked by ensuring the average standard deviation of split frequencies (*asdsf*) is below 0.5%, considered to be good convergence according to the Exabayes user guide. Alignments, partitions and best-fit partitioning schemes for all datasets are available as Data S1.

Codon usage analysis

Codon usage (in counts) was calculated using EMBOSS v.6.5.7 (*Rice, Longden & Bleasby, 2000*) followed by minor adjustments based on the Invertebrate Mitochondrial Code (genetic code = 5). Comparisons among the three lineages of shrimps were made by applying the chi-square test to the pooled codon usage counts for species from each infraorder. Relative synonymous codon usage (RSCU) values were calculated by taking the ratio of the number of times a codon appears to the expected frequency of the codon if all synonymous codons for a same amino acid are used equally (*Sharp & Li, 1987*). Patterns of variation among individuals in RSCU values were summarized using multidimensional scaling (MDS) based on Euclidean dissimilarities implemented in XLSTAT v.2015.4.01.20978 (*Addinsoft, 2010*).

RESULTS

Mitogenome composition

Mitogenomes for specimen J40715 of *C. ceramica* (16,899 bp, 130× cov), *T. australiensis* (15,485 bp, 86× cov), *M. bullatum* (15,774 bp, 27× cov), *A. lobidens* (15,735 bp, 60× cov)
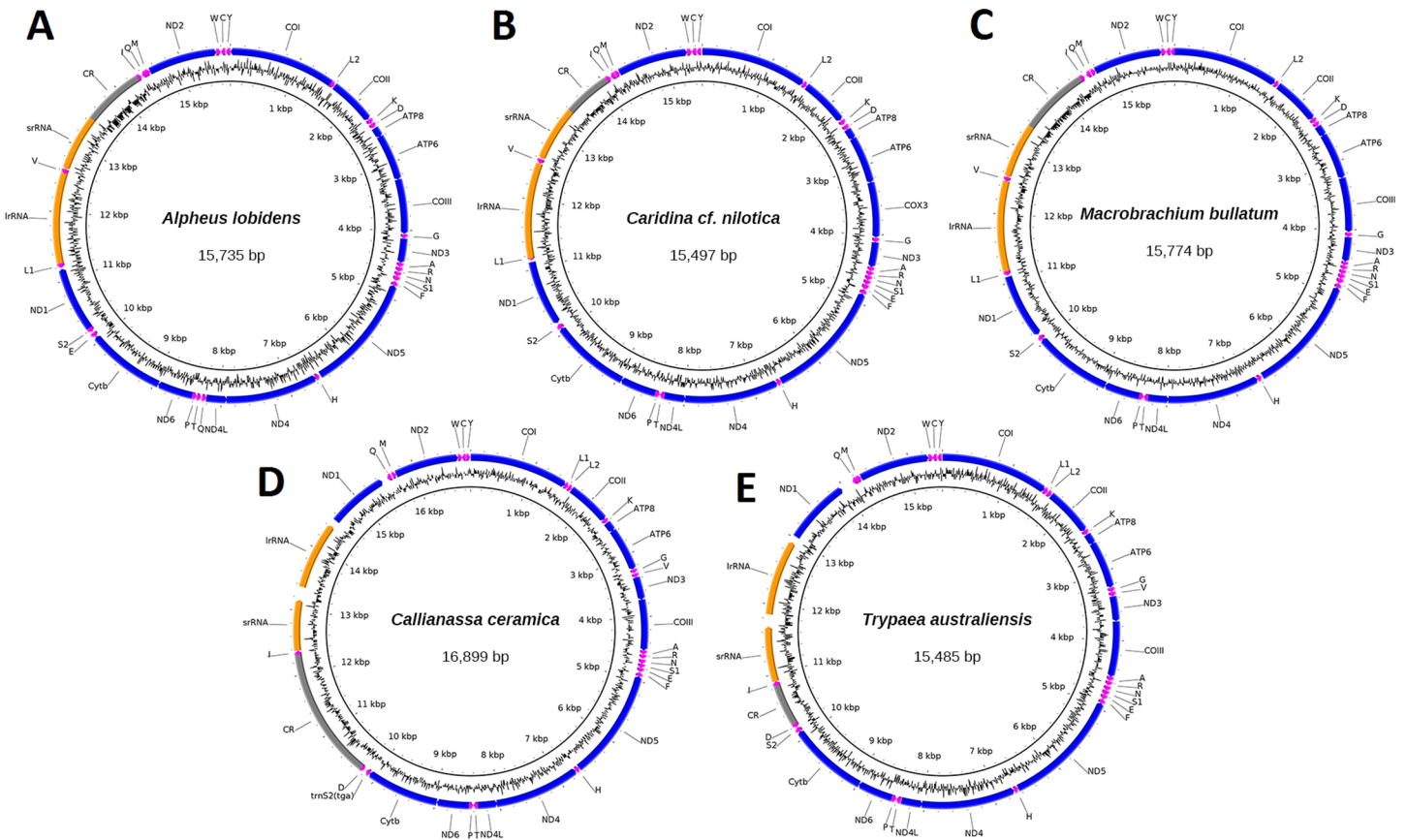
Figure 1 Circular representation of three caridean and two axiidean species. These figures show composition and order of protein-coding genes (blue), ribosomal RNAs (orange), transfer RNAs (purple) and large non-coding region (grey) for the following mitogenomes: (A). *Alpheus lobidens*, (B). *Caridina* cf. *nilotica*, (C). *Macrobrachium bullatum*, (D). *C. ceramica*, (E). *Trypaea australiensis*.

and *Caridina* cf. *nilotica* (15,497 bp, 63× cov) were assembled into complete circular sequences, annotated (Fig. 1), and deposited in GenBank with accession numbers listed in Table 1. Four of the mitogenomes contain the typical 13 PCG, two ribosomal RNA genes, 22 transfer RNA genes and one long non-coding region while *A. lobidens* has an additional *trnQ* flanked by the *ND4L* and *trnT* genes (Fig. 1). Detailed composition of each mitogenome can be found in Table S1 while general information on % AT and lengths for all mitogenomes included in this study are in Table S2. The mitogenomes are AT rich (58.9–69.7%), with *A. lobidens* having the lowest AT content, matching closely to *Alpheus distinguendus* (60.2%), the only other species of *Alpheus* having a published mitogenome sequence. Gene lengths are typical but *Callianassa* and *Trypaea* have an elevated proportion of intergenic nucleotides, with some spacers in the order of 200 bp in length. This is significantly larger than for other members of the Axiidea, but similar to spacers reported for the Gebiidea. *C. ceramica* (KU350630.1) has an unusually long control region of 2,036 bp, whereas for all the other taxa it is less than 1,000 bp, including the closely-related *T. australiensis* (587 bp). However, the control region for the conspecific *C. ceramica* (KU362925.1) is very similar (1,978 bp) and these two specimens are also very similar in terms of % AT

Table 3 List of samples for which annotations were corrected based on re-annotation with MITOS.

Accession #	Species	Genes involved[1]	Correction/edits made in this study
NC_024751.1	*Caridina gracilipes*	ND2, trnS	Added ND2, removed duplicated trnS
NC_012217.1	*Macrobrachium lanchesteri*	rrnS, rrnL	Inverted coordinates for rrnS and rrnL
NC_020025.1	*Corallianassa coutierei*	trnI, trnQ	Inverted coordinates for trnI and trnQ
NC_020351.1	*Nihonotrypaea japonica*	rrnS, rrnL	Added rrnS and rrnL coordinates

Note:
[1] Genes involved in discrepancies found between NCBI's RefSeq entry and re-annotation with MITOS.

(69.7 and 70.2%). A matrix of pair-wise identities of whole mitogenome sequence alignments can be found in Fig. S1.

Manual inspection and MITOS annotation identified multiple erroneously annotated crustacean mitogenomes in GenBank RefSeq database

The re-annotated species with gene orders divergent to that of the pancrustacean ground patterns obtained from GenBank's RefSeq database identified several that require revision. One was found to have a missing protein-coding gene and an extra *trnS*, and others inverted rRNA and tRNA coordinates and other annotation anomalies as detailed in Table 3. For our study, entries were edited based on MITOS annotations and the revised GenBank files for these entries are included as Data S2.

Whole mitogenomes are consistent with monophyly of infraorders Axiidea, Gebiidea, and Caridea

A total of 33 mitogenome sequences from the three groups of interest were utilized (Axiidea: eight, Gebiidea: five, Caridea: 20) to elucidate phylogenetic relationships, with an additional 12 Dendrobranchiata mitogenomes as an outgroup (Table 2). Trees constructed from every dataset and analysis are available in Fig. S2. All trees place *M. bullatum*, *A. lobidens*, and *Caridina* cf. *nilotica* as sister taxa to other species from their respective genera with relatively high nodal support (UFBoot ≥ 93%, PP 1.00). These trees also consistently place *C. ceramica* as sister to *T. australiensis* (UFBoot 100%, PP 1.00) within Axiidea.

The Bayesian-inferred phylogenetic tree in Fig. 2A, constructed based on amino acid sequences of 13 PCGs and nucleotide sequences of two rRNAs (Dataset C), shows a tree topology that is shared by most trees inferred in this study. Most nodes received maximal support from each analysis and dataset. The greatest levels of uncertainty in terms of phylogenetic placement are mostly relating to the relationships among closely related taxa, such as in the *Palaemon* (6/10 trees), *Rimicaris–Opaepele–Alvinocaris–Nautilocaris* (2/10 trees), *Upogebia–Austinogebia* (1/10 trees), *Corallianassa–Paraglypturus* (1/10 trees), and *Macrobrachium* (1/10 trees) clades. The only deeper clade with low support is the placement of Atyidae as the sister clade of Alvinocarididae within Caridea (UFBoot ≥ 90, PP 1.00). Within Axiidea, maximal support is observed for almost all nodes. The most
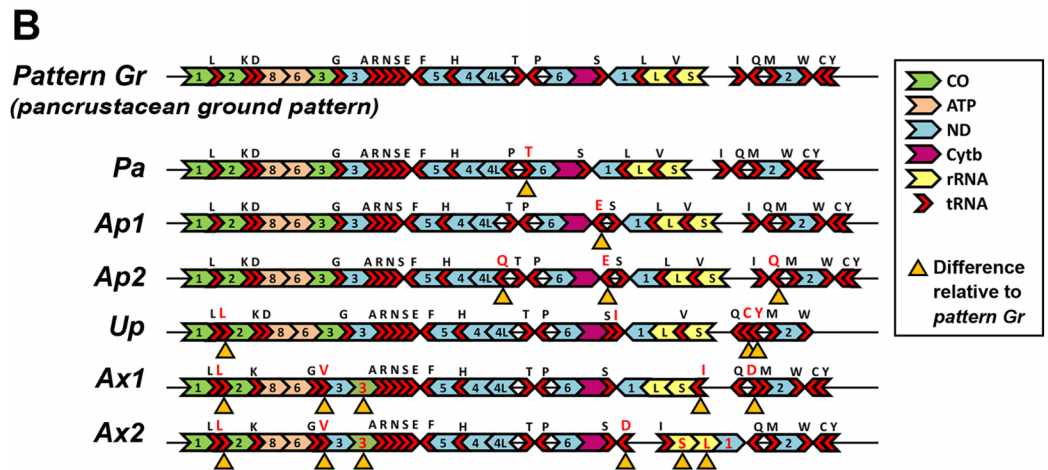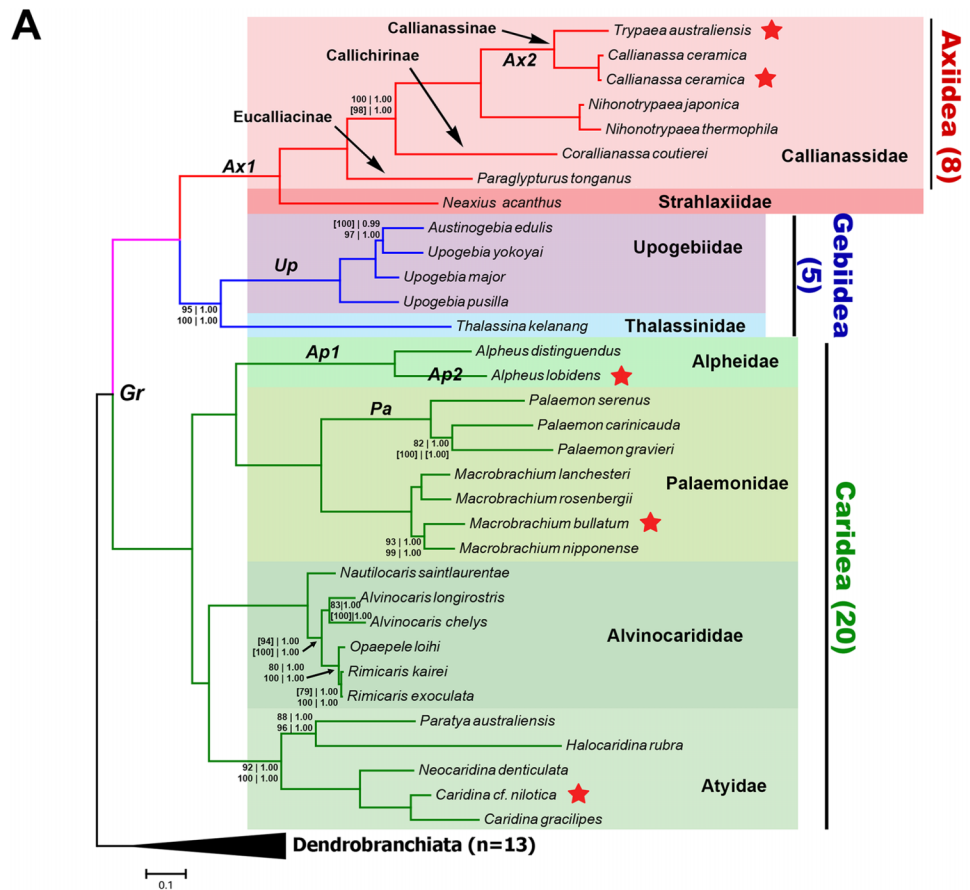
Figure 2 Phylogenetic relationships and gene order rearrangements within Axiidea, Gebiidea and Caridea. (A). Phylogenetic tree with support values indicated at each node (top, l-r: ML and BI support for PCG (aa) + 12S + 16S dataset, bottom, l-r: ML and BI support for PCG (nt) + 12S + 16S dataset). Square brackets [ ] around a value indicate that the shown topology is in conflict with that constructed by the specific dataset. If values are absent at a node, maximum support was found for that node based on all phylogenetic inference methods and datasets. Topology shown was inferred from Bayesian analysis of PCG (aa) + 12S + 16S dataset. Codes on branches (*Gr, Pa, Ap1, Ap2, Up, Ax1, Ax2*) correspond to gene order pattern listed in B while red stars indicate mitogenomes sequenced in this study. (B). Gene order of various groups. Yellow triangles under genes indicate differences in arrangement compared to the ground pancrustacean pattern.

basal split in this infraorder separates Strahlaxiidae (*Neaxius acanthus*) from Callianassidae. Within Callianassidae, the three major lineages correspond to accepted subfamilies, two represented by one species each and Callianassinae by four species. In as far as it goes, the phylogenetic placements of mitogenome sequences in Axiidea are congruent with the current classification at the family, subfamily, and generic levels (*Felder & Robles, 2009*) (Fig. 2A). These analyses of just four upogebiid species indicate that *Upogebia* may be paraphyletic with respect to *Austinogebia*. Since these two genera are nominally represented by over 120 and eight species respectively, any comment on their status is premature at this stage. The degree of divergence between the species of *Rimicaris* and *Opaepele* is small relative to the degree of divergence between congeneric species within *Macrobrachium*, *Alpheus*, *Caridina*, and *Alvinocaris*.

Deviation from the pancrustacean ground pattern is prevalent in the currently sequenced members of the Axiidea and Gebiidea

Most mitogenomes from caridean species have the pancrustacean ground pattern (pattern *Gr*, Fig. 2B). Those that differ show only minor rearrangements involving the short tRNA genes (patterns *Pa*, *Ap1*, and *Ap2*). In contrast, the mitogenomes of species within Axiidea and Gebiidea exhibit relatively substantial differences in gene order (patterns *Up*, *Ax1*, and *Ax2*) entailing rearrangements of PCG, rRNAs, and a number of tRNAs. An example is pattern *Ax2* shown in Fig. 2B, which includes the inversion of the *ND1*, *lrRNA*, *srRNA*, and *trnI* genes as a block, in addition to the inversion and translocation of *trnD* from between *trnQ* and *trnM* to a position between *trnS2* and the putative control region, as well as new placements for *trnL* and *trnV* also evident in pattern *Ax1*.

All gene order novelties relative to the pancrustacean ground pattern are consistent with the relationships depicted by the molecular phylogeny, and in several cases define taxonomic groups at different levels. In Caridea, pattern *Pa* is common to the three *Palaemon* species and pattern *Ap1* defines the two *Alpheus* species. Similarly, within Gebiidea, while *Thalassina kelanang* (family Thalassinidae) has the ground pancrustacean pattern, the other four species are all members of the family Upogebiidae and are united by novel rearrangements involving several tRNA translocations (pattern *Up*, Fig. 2B). Pattern *Ax1*, involving the rearrangements of *COIII* and several tRNAs, is shared among species of Axiidea. The elements of pattern *Ax2* that differ from pattern *Ax1* support the node joining *Callianassa* and *Trypaea*.

Evidence of significant codon usage bias in mitochondrial genomes at the infraorder level

Figure 3 shows there is strong A+T bias in codon usage across the 33 shrimp mitogenomes. RSCU frequencies demonstrate distinct preference for codons with A or T in the third codon position compared to other synonymous codons. Counts of codons used and RSCU values can be found in Table S3. Among the 62 available codons, the four most used codons in all three infraorders are TTT (Phe), TTA (Leu), ATT (Ile), and ATA (Met), all made up of solely A and T nucleotides. Even so, this preference for A+T codons is stronger in Axiidea and Gebiidea mitogenomes and less so
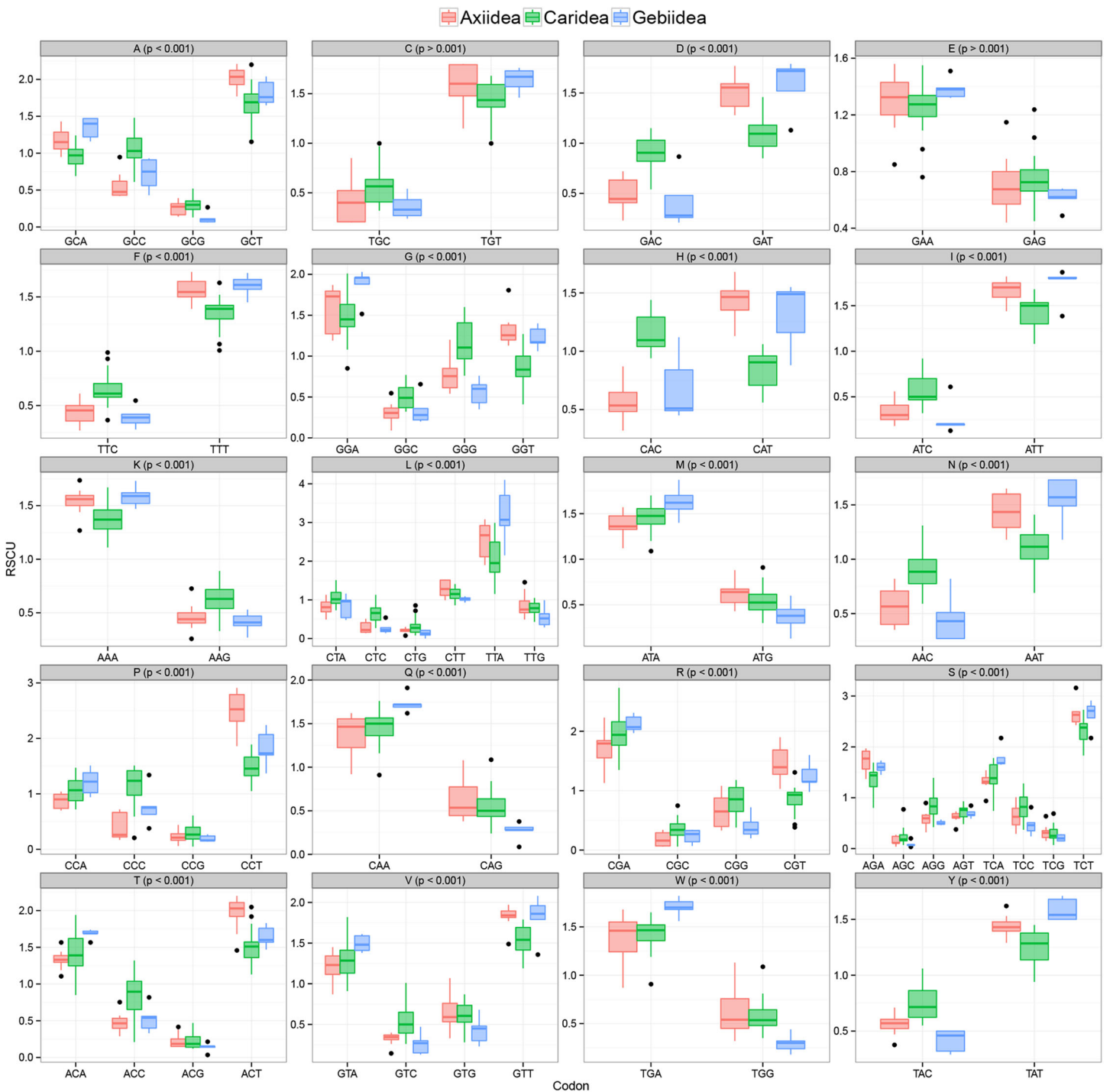
Figure 3 Relative synonymous codon usage values (RSCU) (y-axis) in protein-coding genes of mud shrimps and true shrimps. Encoded amino acid and its corresponding *p*-value (> or <0.001) is shown at the top of each box plot while synonymous codons are indicated on the x-axis.

in Caridea, most obvious for amino acids Asp (D), His (H), Asn (N), and Tyr (Y). Statistical comparisons show that, for each amino acid, there are significant differences among the three infraorders in the proportions of the codons being used (*p*-values in Fig. 3). A separate comparison for species with pattern *Ax1* and *Ax2* within the infraorder
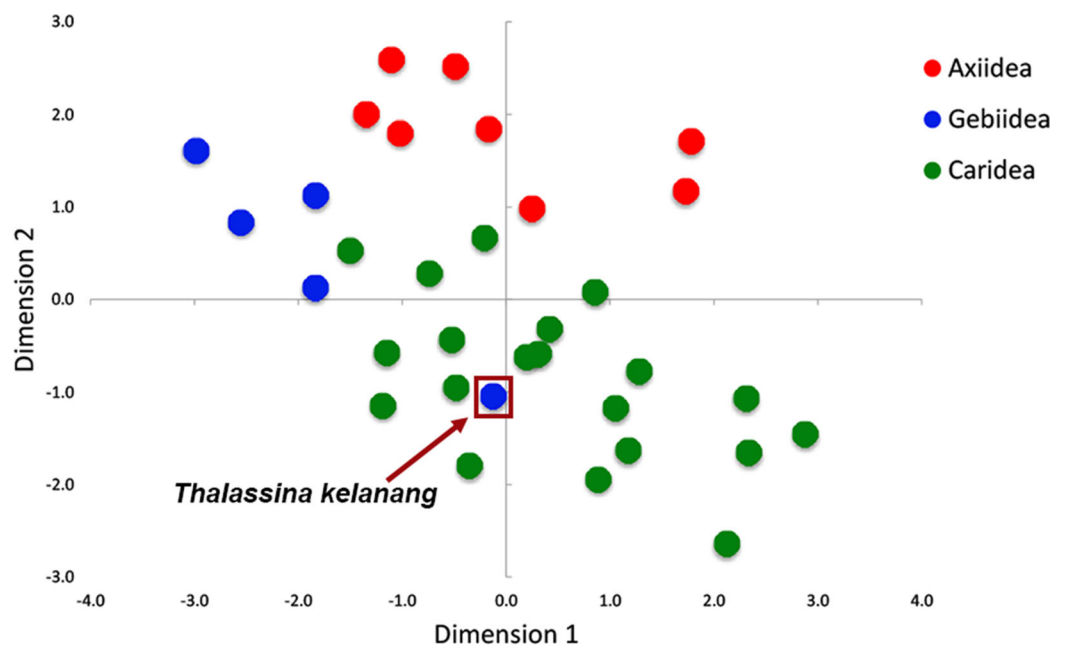
Figure 4 Patterns of variation among individuals based on RSCU values shown using multidimensional scaling (MDS) based on Euclidean dissimilarities.

Axiidea and also with those with the ground pattern (*Gr*) do not reveal any substantial difference in their codon usage bias (Fig. S3). The MDS plot shows that, for the most part, members of each infraorder cluster together and are largely distinct from the samples from the other infraorders (Fig. 4). A sample of the Gebiidea, *T. kelanang*, is a maverick, being placed well inside the Caridea cluster and remote from the other members of its infraorder. It also has a very low AT content for this group, being more similar to caridean shrimps. In this context, it is noteworthy that the species is placed as the most basal member of the Gebiidea and is also the only member of the infraorder that has the primitive pancrustacean gene order, which it shares with all other of the members of the Caridea with the exception of some species with minor derived rearrangements (Fig. 2B).

DISCUSSION

The five new decapod mitogenomes presented in this study considerably expand the number of samples of Axiidea and Caridea currently available for mitogenome-based phylogenetics and comparative mitogenomic studies. Members of the infraorder Gebiidea were also included in this analysis due to its taxonomic history of having been placed with Axiidea in the infraorder Thalassinidea, prior to obtaining recent recognition as its own separate infraorder (*Ahyong & O'Meally, 2004*; *Crandal, Harris & Fetzner, 2000*; *Robles et al., 2009*; *Timm & Bracken-Grissom, 2015*). The mitogenomic features of these taxa are generally consistent with those for the Decapoda and for the three infraorders (*Kim et al., 2013*; *Lin et al., 2012*; *Miller et al., 2005*) (Table S1). In addition, the high AT content (58.9–73.6%) observed in all of the mitogenomes utilized in this study is typical for the Crustacea and the Arthropoda (*Cameron, 2014*; *Cook, Yue & Akam, 2005*; *Lin et al., 2012*; *Shen et al., 2015*).

It has been suggested that evolutionary rate and the frequency of rearrangements are independent (*Gissi, Iannelli & Pesole, 2008*) and this is consistent with results depicted in Fig. 2, which indicate no obvious correlations between the size and distribution of rearrangements and branch lengths (substitution rate) within and between groups. Despite this general observation, some studies have noted an association between higher substitution rates and the occurrence of mitogenome rearrangements involving a transfer of genes between strands (e.g., *Stokkan et al., 2016*). Although it is expected that codon usage varies between major groups of organisms and among animal phyla (*Castellana, Vicario & Saccone, 2011*; *Gissi, Iannelli & Pesole, 2008*; *Prat et al., 2009*), finding substantial codon usage differences at the infraorder level is somewhat unusual. It is becoming more apparent now that there are divergent patterns in AT content among orders in insects and among the major taxonomic groups in the Malacostraca (see Table 1 in *Sun et al. (2009)* and *Shen et al. (2015)*). Studies examining patterns of codon usage in mitogenomes have failed to observe differences at finer taxonomic scales within insects and crustaceans (*Cook, Yue & Akam, 2005*; *Shen et al., 2015*; *Sun et al., 2009*).

Elevated mutational pressure is thought to be a major driver of non-random mitochondrial synonymous variation. However, selection of optimal codons for translational efficiency and genetic drift is also thought to play a role (*Castellana, Vicario & Saccone, 2011*; *Jia & Higgs, 2008*; *Prat et al., 2009*). It is tempting to speculate that the distinctive pattern of codon usage in the Axiidea, and the frequency and extent of mitogenome gene order rearrangements may be correlated with the acquisition of specialized adaptations by members of this infraorder to largely burrowing lifestyles (*Lin et al., 2012*; *Sakai, 2011*) compared to the members of the Caridea, which are mostly free living (*Bracken, De Grave & Felder, 2009*).

The phylogenetic analyses using whole mitogenomes support the monophyly of each of the three infraorders, Axiidea, Gebiidea, and Caridea, with inclusion of more comprehensive taxon sampling than previous studies (*Lin et al., 2012*; *Shen, Braband & Scholtz, 2013*; *Shen et al., 2015*; *Tan et al., 2015*). Further, the members of Axiidea have a common gene rearrangement that is a potential synapomorphy of the infraorder, and therefore also supports the monophyly of this group. A notable exception to these results is the analysis of *Shen et al. (2015)*, which supported a non-monophyletic Gebiidea, by placing *T. kelanang* inside a lineage comprising Acheleta, Polychelida, and Caridea. Other studies using a combination of mitochondrial and nuclear gene fragments (*Bracken, De Grave & Felder, 2009*; *Robles et al., 2009*) also support the monophyly of these groups, but their relationships with each other and other decapod infraorders have yet to be resolved (*Tan et al., 2015*). A caveat of our findings and those of other studies is that monophyly cannot be fully tested without comprehensive taxonomic and gene sampling and the inclusion of species from all decapod infraorders (*Timm & Bracken-Grissom, 2015*).

The contribution of mitogenomic data for more species will be particularly important for exploring and testing evolutionary relationships within Axiidea, given its lack of broad taxonomic representation on the current evolutionary tree. Two of the mitogenomes contributed in this study (*C. ceramica*, *T. australiensis*) belong to the

family Callianassidae, a diverse group of axiidean shrimps adapted to digging in soft marine sediments. Over 100 species placed in 22 genera and divided between several subfamilies are recognized from this family (*Sakai, 2011*). However, relationships among the major groups and the definition of subfamily and generic level boundaries are contentious within the Callianassidae (*Poore et al., 2014*; *Sakai, 2011*). It is noteworthy that for the Axiidea, the tree-based relationships and the mitogenome rearrangements are entirely consistent with the infraorder, family, subfamily and genus level relationships for the samples included (Fig. 2), and that the phylogenetic tree and the *Ax1* rearrangement pattern further supports the monophyly of the Axiidea.

The potential for mitochondrial rearrangement to act as "super" characters for phylogenetic estimation for the arthropod has been recognised by a number of authors (*Boore, Lavrov & Brown, 1998*; *Dowton & Austin, 1999*; *Dowton, Castro & Austin, 2002*). This study further supports this position but also notes that the distribution of rearrangements is uneven across the tree generated in this study and the larger analysis by *Tan et al. (2015)*. Thus, while novel gene order attributes act as useful phylogenetic characters for the Axiidea and Gebiidea, their usefulness appear to be limited for the Caridea, even though the age of the lineages overlap. Similarly, for other major crustaceans groups, lobsters (Astacidea and Achelata), crabs (Brachyura), and penaeid shrimps (Dendrobranchiata), rearrangements are largely absent or minor (*Shen et al., 2015*; *Tan et al., 2015*), but freshwater crayfish (Astacidea) and anomuran crabs (Anomura) are phylogenetic "hotspots" for mitogenome gene order evolution in both the number and scale of rearrangements. As pointed out by *Mao, Gibson & Dowton (2014)*, far greater sampling is required to adequately test the phylogenetic utility of observed mitogenome gene rearrangements and identification of suitable models for investigating the evolutionary and molecular drivers that shape the organization and architecture of animal mitogenomes.

In this regard, rapid and efficient methods for mitogenome sequencing using next-generation sequencing (NGS) platforms will accelerate this task (*Gan et al., 2016*; *Timmermans et al., 2016*). Unlike more classical methods utilizing long-range polymerase chain reaction (PCR) to amplify the mitogenome into a smaller number of fragments with universal primers, followed by Sanger sequencing (*Lin et al., 2012*; *Miller et al., 2004*; *Shen, Braband & Scholtz, 2013*), the primer-free and shotgun nature of NGS will likely increase the discovery of mitogenome rearrangements as it makes no assumptions about the pre-existing gene order of the species under study. In fact, the availability of a reference whole mitogenome will improve primer design and consequently the success rate of complete mitogenome recovery from members of the same genus (or family) using long-range PCR. It is envisaged with the advent of third generation sequencing technology such as PacBio (*Rhoads & Au, 2015*) and Nanopore (*Branton et al., 2008*) sequencing, problematic and repetitive regions commonly associated with the control region can be readily resolved and confirmed.

Lastly, while preparing the dataset for analysis, we identified several misannotated mitogenomes on NCBI (Table 3) although these misannotations were absent in their related publications i.e., the correct gene coordinates and orientations were

reported by the authors in their respective publications. We postulate that these discrepancies may have arisen due to errors during the submission of these mitogenomes to public databases. It is also possible, though less likely, that the mitogenomes may have been erroneously edited when they were reviewed by NCBI staff. Hence, this highlights that although there is tremendous gain from having a growing number of mitogenome submissions to public databases as molecular resources, the accuracy of annotations should not be assumed and it is prudent to re-evaluate species with any form of gene order rearrangements or related anomalies before inclusion in datasets for comparative analyses.

ACKNOWLEDGEMENTS

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Grant Disclosures

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Mun Hua Tan performed the experiments, analysed the data, wrote the paper, prepared figures and/or tables and reviewed drafts of the paper.
- Han Ming Gan performed the experiments, analysed the data, wrote the paper and reviewed drafts of the paper.
- Yin Peng Lee performed the experiments.
- Gary C. B. Poore contributed analysis tools, wrote the paper and reviewed drafts of the paper.
- Christopher M. Austin conceived and designed the experiments, analysed the data, contributed reagents/materials/analysis tools, wrote the paper and reviewed drafts of the paper.

Data Deposition

The following information was supplied regarding data availability.

The raw data is included in the manuscript as GenBank accession code. The corresponding accession code to each mitogenome reported in this study is listed in Table 1 and also in Data S3. GenBank: KU362925.1; KU350630.1; KM501040.2; KP276147.1; KU726823.1; KM978918.1.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.2982#supplemental-information.

REFERENCES

Abascal F, Zardoya R, Telford MJ. 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research* **38(suppl_2):**W7–W13 DOI 10.1093/nar/gkq291.

Aberer AJ, Kobert K, Stamatakis A. 2014. ExaBayes: massively parallel Bayesian tree inference for the whole-genome era. *Molecular Biology and Evolution* **31(10):**2553–2556 DOI 10.1093/molbev/msu236.

Addinsoft. 2010. *XLSTAT 2015, Data Analysis and Statistics Software for Microsoft Excel.* Paris: Addinsoft.

Ahyong ST, O'Meally D. 2004. Phylogeny of the Decapoda Reptantia: resolution using three molecular loci and morphology. *The Raffles Bulletin of Zoology* **52(2):**673–693.

Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* **12(1):**1–10 DOI 10.1186/1471-2164-12-402.

Bauzà-Ribot MM, Jaume D, Juan C, Pons J. 2009. The complete mitochondrial genome of the subterranean crustacean *Metacrangonyx longipes* (Amphipoda): a unique gene order and extremely short control region. *Mitochondrial DNA* **20(4):**88–99 DOI 10.1080/19401730902964417.

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution* **69(2):**313–319 DOI 10.1016/j.ympev.2012.08.023.

Boore JL. 2006. The use of genome-level characters for phylogenetic reconstruction. *Trends in Ecology and Evolution* **21(8):**439–446 DOI 10.1016/j.tree.2006.05.009.

Boore JL, Lavrov DV, Brown WM. 1998. Gene translocation links insects and crustaceans. *Nature* **392(6677):**667–668 DOI 10.1038/33577.

Bracken HD, De Grave S, Felder DL. 2009. Phylogeny of the infraorder Caridea based on mitochondrial and nuclear genes (Crustacea: Decapoda). In: Martin JW, Crandall KA, Felder DL, eds. *Decapod Crustacean Phylogenetics.* Boca Raton: CRC Press, 281–305.

Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, Di Ventra M, Garaj S, Hibbs A, Huang X, Jovanovich SB, Krstic PS, Lindsay S, Ling XS, Mastrangelo CH, Meller A, Oliver JS, Pershin YV, Ramsey JM, Riehn R, Soni GV, Tabard-Cossa V, Wanunu M, Wiggin M, Schloss JA. 2008. The potential and challenges of nanopore sequencing. *Nature Biotechnology* **26(10):**1146–1153 DOI 10.1038/nbt.1495.

Cameron SL. 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annual Review of Entomology* **59(1):**95–117 DOI 10.1146/annurev-ento-011613-162007.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25(15):**1972–1973 DOI 10.1093/bioinformatics/btp348.

Castellana S, Vicario S, Saccone C. 2011. Evolutionary patterns of the mitochondrial genome in Metazoa: exploring the role of mutation and selection in mitochondrial protein–coding genes. *Genome Biology and Evolution* **3:**1067–1079 DOI 10.1093/gbe/evr040.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17(4):**540–552 DOI 10.1093/oxfordjournals.molbev.a026334.

Chong RA, Mueller RL. 2013. Evolution along the mutation gradient in the dynamic mitochondrial genome of salamanders. *Genome Biology and Evolution* **5(9):**1652–1660 DOI 10.1093/gbe/evt119.

Cook CE, Yue Q, Akam M. 2005. Mitochondrial genomes suggest that hexapods and crustaceans are mutually paraphyletic. *Proceedings of the Royal Society of London B: Biological Sciences* **272(1569):**1295–1304 DOI 10.1098/rspb.2004.3042.

Crandal KA, Harris DJ, Fetzner JW. 2000. The monophyletic origin of freshwater crayfish estimated from nuclear and mitochondrial DNA sequences. *Proceedings of the Royal Society of London B: Biological Sciences* **267(1453):**1679–1686 DOI 10.1098/rspb.2000.1195.

Dowton M, Austin AD. 1999. Evolutionary dynamics of a mitochondrial rearrangement "hot spot" in the Hymenoptera. *Molecular Biology and Evolution* **16(2):**298–309 DOI 10.1093/oxfordjournals.molbev.a026111.

Dowton M, Castro LR, Austin AD. 2002. Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome 'morphology'. *Invertebrate Systematics* **16(3):**345–356 DOI 10.1071/IS02003.

Dowton M, Castro RL, Campbell LS, Bargon DS, Austin DA. 2003. Frequent mitochondrial gene rearrangements at the hymenopteran nad3–nad5 junction. *Journal of Molecular Evolution* **56(5):**517–526 DOI 10.1007/s00239-002-2420-3.

Fahrein K, Talarico G, Braband A, Podsiadlowski L. 2007. The complete mitochondrial genome of *Pseudocellus pearsei* (Chelicerata: Ricinulei) and a comparison of mitochondrial gene rearrangements in Arachnida. *BMC Genomics* **8:**386 DOI 10.1186/1471-2164-8-386.

Felder DL, Robles R. 2009. Molecular phylogeny of the family Callianassidae based on preliminary analyses of two mitochondrial genes. In: Martin JW, Crandall KA, Felder DL, eds. *Decapod Crustacean Phylogenetics Crustacean Issues.* Vol. 18. Boca Raton: CRC Press, 327–342.

Gan HY, Gan HM, Lee YP, Austin CM. 2016. The complete mitochondrial genome of the bass yabby *Trypaea australiensis* Dana 1852, (Crustacea; Decapoda; Callianassidae) – a new gene order for the Decapoda. *Mitochondrial DNA Part A* **27:**3985–3986 DOI 10.3109/19401736.2014.989516.

Gan HY, Gan HM, Tan MH, Lee YP, Austin CM. 2016b. 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* **27:**4099–4100 DOI 10.3109/19401736.2014.1003862.

García-Machado E, Pempera M, Dennebouy N, Oliva-Suarez M, Mounolou J-C, Monnerot M. 1999. Mitochondrial genes collectively suggest the paraphyly of Crustacea with respect to Insecta. *Journal of Molecular Evolution* **49(1):**142–149 DOI 10.1007/pl00006527.

Gissi C, Iannelli F, Pesole G. 2004. Complete mtDNA of *Ciona intestinalis* reveals extensive gene rearrangement and the presence of an atp8 and an extra trnM gene in ascidians. *Journal of Molecular Evolution* **58(4):**376–389 DOI 10.1007/s00239-003-2559-6.

Gissi C, Iannelli F, Pesole G. 2008. Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity* **101(4):**301–320 DOI 10.1038/hdy.2008.62.

Gutiérrez-Millán EL, Peregrino-Uriarte BA, Sotelo-Mundo R, Vargas-Albores F, Yepiz-Plascencia G. 2002. Sequence and Conservation of a rRNA and tRNAVal Mitochondrial Gene Fragment from *Penaeus californiensis* and Comparison with *Penaeus vannamei* and *Penaeus stylirostris*. *Marine Biotechnology* **4:**392–398 DOI 10.1007/s10126-002-0008-y.

Ivey JL, Santos SR. 2007. The complete mitochondrial genome of the Hawaiian anchialine shrimp *Halocaridina rubra* Holthuis, 1963 (Crustacea: Decapoda: Atyidae). *Gene* **394:**35–44 DOI 10.1016/j.gene.2007.01.009.

Jia W, Higgs PG. 2008. Codon usage in mitochondrial genomes: distinguishing context-dependent mutation from translational selection. *Molecular Biology and Evolution* **25(2):**339–351 DOI 10.1093/molbev/msm259.

Jiang ZJ, Castoe TA, Austin CC, Burbrink FT, Herron MD, McGuire JA, Parkinson CL, Pollock DD. 2007. Comparative mitochondrial genomics of snakes: extraordinary substitution rate dynamics and functionality of the duplicate control region. *BMC Evolutionary Biology* **7:**123 DOI 10.1186/1471-2148-7-123.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30(4):**772–780 DOI 10.1093/molbev/mst010.

Kilpert F, Podsiadlowski L. 2006. The complete mitochondrial genome of the common sea slater, *Ligia oceanica* (Crustacea, Isopoda) bears a novel gene order and unusual control region features. *BMC Genomics* **7:**241 DOI 10.1186/1471-2164-7-241.

Kim S, Ahn D-H, Park J-K, Kim S-J, Choi H-G, Min G-S. 2013. The complete mitochondrial genome of the Japanese ghost shrimp *Nihonotrypaea japonica* (Crustacea, Decapoda, Axiidea). *Mitochondrial DNA* **24(3):**243–245 DOI 10.3109/19401736.2012.752480.

Kim G, Yoon T-H, Kang H-E, Park WG, Kim HJ, Choi J-H, Kim H-W. 2016a. Complete mitochondrial genome of Kishi velvet shrimp, *Metapenaeopsis dalei* (Rathbun, 1902) (Crustacea: Decapoda: Penaeidae) in the East China Sea. *Mitochondrial DNA Part B* **1:**120–121 DOI 10.1080/23802359.2016.1143341.

Kim S-J, Kim J, Ahn D-H, Ju S-J, Min G-S, Kim S. 2016b. Complete mitochondrial genome of the hydrothermal vent ghost shrimp *Paraglypturus tonganus* (Crustacea, Axiidea, Callianassidae). *Mitochondrial DNA Part A* **27:**1048–1049 DOI 10.3109/19401736.2014.928860.

Kim S-J, Pak SJ, Ju S-J. 2015a. Mitochondrial genome of the hydrothermal vent shrimp *Nautilocaris saintlaurentae* (Crustacea: Caridea: Alvinocarididae). *Mitochondrial DNA* **26:**127–128 DOI 10.3109/19401736.2013.815169.

Kim S-T, Kim G, Yoon T-H, Kim A-R, Kim HJ, Kim H-W. 2015b. Complete mitochondrial genome of *Palaemon gravieri* (Yu, 1930) (Crustacea: Decapoda: Palaemonidae). *Mitochondrial DNA* 1–2 DOI 10.3109/19401736.2015.1118078.

Kim S, Park M-H, Jung J-H, Ahn D-H, Sultana T, Kim S, Park J-K, Choi H-G, Min G-S. 2012. The mitochondrial genomes of *Cambaroides similis* and *Procambarus clarkii* (Decapoda: Astacidea: Cambaridae): the phylogenetic implications for Reptantia. *Zoologica Scripta* **41(3):**281–292 DOI 10.1111/j.1463-6409.2012.00534.x.

Li X, Huang Y, Lei F. 2015. Comparative mitochondrial genomics and phylogenetic relationships of the *Crossoptilon* species (Phasianidae, Galliformes). *BMC Genomics* **16**:42 DOI 10.1186/s12864-015-1234-9.

Lin F-J, Liu Y, Sha Z, Tsang L, Chu K, Chan T-Y, Liu R, Cui Z. 2012. Evolution and phylogeny of the mud shrimps (Crustacea: Decapoda) revealed from complete mitochondrial genomes. *BMC Genomics* **13(1):**631 DOI 10.1186/1471-2164-13-631.

Ma K, Feng J, Lin J, Li J. 2011. The complete mitochondrial genome of *Macrobrachium nipponense*. *Gene* **487**:160–165 DOI 10.1016/j.gene.2011.07.017.

Mao M, Gibson T, Dowton M. 2014. Evolutionary dynamics of the mitochondrial genome in the Evaniomorpha (Hymenoptera)—a group with an intermediate rate of gene rearrangement. *Genome Biology and Evolution* **6(7):**1862–1874 DOI 10.1093/gbe/evu145.

Mao Z-C, Liu P, Duan Y-F, Li J, Chen P, Meng X-L. 2016. Sequencing of complete mitochondrial genome of sword prawn *Parapenaeopsis hardwickii* (Miers) (Crustacea, Decapoda, Penaeidae). *Mitochondrial DNA Part B* **1**:259–260 DOI 10.1080/23802359.2016.1159927.

Miller AD, Murphy NP, Burridge CP, Austin CM. 2005. Complete mitochondrial DNA sequences of the decapod crustaceans *Pseudocarcinus gigas* (Menippidae) and *Macrobrachium rosenbergii* (Palaemonidae). *Marine Biotechnology* **7(4):**339–349 DOI 10.1007/s10126-004-4077-8.

Miller AD, Nguyen TTT, Burridge CP, Austin CM. 2004. Complete mitochondrial DNA sequence of the Australian freshwater crayfish, *Cherax destructor* (Crustacea: Decapoda: Parastacidae): a novel gene order revealed. *Gene* **331**:65–72 DOI 10.1016/j.gene.2004.01.022.

Minh BQ, Nguyen MAT, von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution* **30(5):**1188–1195 DOI 10.1093/molbev/mst024.

Muhire BM, Varsani A, Martin DP. 2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS ONE* **9(9):**e108277 DOI 10.1371/journal.pone.0108277.

Murphy NP, Austin CM. 2004. Multiple origins of the endemic Australian *Macrobrachium* (Decapoda : Palaemonidae) based on 16S rRNA mitochondrial sequences. *Australian Journal of Zoology* **52(5):**549–559 DOI 10.1071/ZO03062.

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **32(1):**268–274 DOI 10.1093/molbev/msu300.

Oliveira DCSG, Raychoudhury R, Lavrov DV, Werren JH. 2008. Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). *Molecular Biology and Evolution* **25(10):**2167–2180 DOI 10.1093/molbev/msn159.

Page TJ, von Rintelen K, Hughes JM. 2007. An island in the stream: Australia's place in the cosmopolitan world of Indo-West Pacific freshwater shrimp (Decapoda: Atyidae: Caridina). *Molecular Phylogenetics and Evolution* **43(2):**645–659 DOI 10.1016/j.ympev.2006.08.007.

Peng Y, Leung HCM, Yiu SM, Chin FYL. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **28(11):**1420–1428 DOI 10.1093/bioinformatics/bts174.

Poore GCB, Ahyong ST, Bracken-Grissom HD, Chan T-Y, Chu KH, Crandall KA, Dworschak PC, Felder DL, Feldmann RM, Hyžný M, Karasawa H, Lemaitre R, Komai T, Li X, Mantelatto FL, Martin JW, Ngoc-Ho N, Robles R, Schweitzer CE, Tamaki A, Tsang LM, Tudge CC. 2014. On stabilising the names of the infraorders of thalassinidean shrimps, Axiidea

de Saint Laurent, 1979 and Gebiidea de Saint Laurent, 1979 (Decapoda). *Crustaceana* **87(10):**1258–1272 DOI 10.1163/15685403-00003354.

Poulsen JY, Byrkjedal I, Willassen E, Rees D, Takeshima H, Satoh TP, Shinohara G, Nishida M, Miya M. 2013. Mitogenomic sequences and evidence from unique gene rearrangements corroborate evolutionary relationships of myctophiformes (Neoteleostei). *BMC Evolutionary Biology* **13**:111 DOI 10.1186/1471-2148-13-111.

Prat Y, Fromer M, Linial N, Linial M. 2009. Codon usage is associated with the evolutionary age of genes in metazoan genomes. *BMC Evolutionary Biology* **9**:285 DOI 10.1186/1471-2148-9-285.

Qian G, Zhao Q, Wang AN, Zhu LIN, Zhou K, Sun H. 2011. Two new decapod (Crustacea, Malacostraca) complete mitochondrial genomes: bearings on the phylogenetic relationships within the Decapoda. *Zoological Journal of the Linnean Society* **162(3):**471–481 DOI 10.1111/j.1096-3642.2010.00686.x.

Rhoads A, Au KF. 2015. PacBio sequencing and its applications. *Genomics, Proteomics and Bioinformatics* **13(5):**278–289 DOI 10.1016/j.gpb.2015.08.002.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European molecular biology open software suite. *Trends in Genetics* **16(6):**276–277 DOI 10.1016/s0168-9525(00)02024-2.

Robles R, Tudge CC, Dworschak PC, Poore GCB, Felder D. 2009. Molecular phylogeny of the Thalassinidea based on nuclear and mitochondrial genes. In: Martin JW, Crandall KA, Felder DL, eds. *Crustacean Issues Vol. 18: Decapod Crustacean Phylogenetics.* Bocan Raton: CRC Press, 309–326.

Rota-Stabelli O, Lartillot N, Philippe H, Pisani D. 2013. Serine codon-usage bias in deep phylogenomics: pancrustacean relationships as a case study. *Systematic Biology* **62(1):**121–133 DOI 10.1093/sysbio/sys077.

Sakai K. 2011. Axioidea of the world and a reconsideration of the Callianassoidea (Decapoda, Thalassinidea, Callianassida). *Crustaceana Monographs* **13**:1–616.

Shao R, Barker SC, Mitani H, Aoki Y, Fukunaga M. 2005. Evolution of duplicate control regions in the mitochondrial genomes of metazoa: a case study with Australasian Ixodes ticks. *Molecular Biology and Evolution* **22(3):**620–629 DOI 10.1093/molbev/msi047.

Sharp PM, Li W-H. 1987. The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Research* **15(3):**1281–1295 DOI 10.1093/nar/15.3.1281.

Shen H, Braband A, Scholtz G. 2013. Mitogenomic analysis of decapod crustacean phylogeny corroborates traditional views on their relationships. *Molecular Phylogenetics and Evolution* **66(3):**776–789 DOI 10.1016/j.ympev.2012.11.002.

Shen X, Ren J, Cui Z, Sha Z, Wang B, Xiang J, Liu B. 2007. The complete mitochondrial genomes of two common shrimps (*Litopenaeus vannamei* and *Fenneropenaeus chinensis*) and their phylogenomic considerations. *Gene* **403**:98–109 DOI 10.1016/j.gene.2007.06.021.

Shen X, Sun MA, Wu Z, Tian M, Cheng H, Zhao F, Meng X. 2009. The complete mitochondrial genome of the ridgetail white prawn *Exopalaemon carinicauda* Holthuis, 1950 (Crustacean: Decapoda: Palaemonidae) revealed a novel rearrangement of tRNA genes. *Gene* **437**:1–8 DOI 10.1016/j.gene.2009.02.014.

Shen X, Tian M, Yan B, Chu K. 2015. Phylomitogenomics of Malacostraca (Arthropoda: Crustacea). *Acta Oceanologica Sinica* **34(2):**84–92 DOI 10.1007/s13131-015-0583-1.

Shoemaker DD, Dyer KA, Ahrens M, McAbee K, Jaenike J. 2004. Decreased diversity but increased substitution rate in host mtDNA as a consequence of Wolbachia endosymbiont infection. *Genetics* **168(4):**2049–2058 DOI 10.1534/genetics.104.030890.

Stokkan M, Jurado-Rivera JA, Juan C, Jaume D, Pons J. 2016. Mitochondrial genome rearrangements at low taxonomic levels: three distinct mitogenome gene orders in the genus *Pseudoniphargus* (Crustacea: Amphipoda). *Mitochondrial DNA Part A* **27(5):**3579–3589 DOI 10.3109/19401736.2015.1079821.

Sun Z, Wan D-G, Murphy RW, Ma L, Zhang X-S, Huang D-W. 2009. Comparison of base composition and codon usage in insect mitochondrial genomes. *Genes and Genomics* **31(1):**65–71 DOI 10.1007/bf03191139.

Tan MH, Gan HM, Schultz MB, Austin CM. 2015. MitoPhAST, a new automated mitogenomic phylogeny tool in the post-genomic era with a case study of 89 decapod mitogenomes including eight new freshwater crayfish mitogenomes. *Molecular Phylogenetics and Evolution* **85:**180–188 DOI 10.1016/j.ympev.2015.02.009.

Timm L, Bracken-Grissom HD. 2015. The forest for the trees: evaluating molecular phylogenies with an emphasis on higher-level Decapoda. *Journal of Crustacean Biology* **35(5):**577–592 DOI 10.1163/1937240X-00002371.

Timmermans MJTN, Viberg C, Martin G, Hopkins K, Vogler AP. 2016. Rapid assembly of taxonomically validated mitochondrial genomes from historical insect collections. *Biological Journal of the Linnean Society* **117(1):**83–95 DOI 10.1111/bij.12552.

Yamauchi MM, Miya MU, Machida RJ, Nishida M. 2004. PCR-based approach for sequencing mitochondrial genomes of decapod crustaceans, with a practical example from kuruma prawn (*Marsupenaeus japonicus*). *Marine Biotechnology* **6:**419–429 DOI 10.1007/s10126-003-0036-2.

Yang C-H, Tsang LM, Chu KH, Chan T-Y. 2012. Complete mitogenome of the deep-sea hydrothermal vent shrimp *Alvinocaris chelys* Komai and Chan, 2010 (Decapoda: Caridea: Alvinocarididae). *Mitochondrial DNA* **23:**417–419 DOI 10.3109/19401736.2012.710212.

Yang EC, Lee J, An SM, Choi DH, Noh JH. 2016. Complete mitochondrial genome of *Upogebia yokoyai* (Decapoda, Crustacea) from Jejudo, Korea. *Mitochondrial DNA Part A* **27:**2139–2140 DOI 10.3109/19401736.2014.982595.

Yang J-S, Lu B, Chen D-F, Yu Y-Q, Yang F, Nagasawa H, Tsuchida S, Fujiwara Y, Yang W-J. 2013. When did decapods invade hydrothermal vents? Clues from the Western Pacific and Indian Oceans. *Molecular biology and evolution* **30:**305–309 DOI 10.1093/molbev/mss224.

Yu Y-Q, Liu X-L, Li H-W, Lu B, Fan Y-P, Yang J-S. 2015. The complete mitogenome of the Atlantic hydrothermal vent shrimp *Rimicaris exoculata* Williams & Rona 1986 (Crustacea: Decapoda: Alvinocarididae). *Mitochondrial DNA* **1–3** DOI 10.3109/19401736.2015.1007291.

Yu Y-Q, Yang W-J, Yang J-S. 2014. The complete mitogenome of the Chinese swamp shrimp *Neocaridina denticulata sinensis* Kemp 1918 (Crustacea: Decapoda: Atyidae). *Mitochondrial DNA* **25:**204–205 DOI 10.3109/19401736.2013.796465.

Xu G, Du F, Nie Z, Xu P, Gu R. 2016. Complete mitochondrial genome of *Caridina nilotica gracilipes*. *Mitochondrial DNA Part A* **27:**1249–1250 DOI 10.3109/19401736.2014.945541.

Whittle CA, Extavour CG. 2015. Codon and amino acid usage are shaped by selection across divergent model organisms of the Pancrustacea. *G3: Genes Genomes Genetics* **5(11):**2307–2321 DOI 10.1534/g3.115.021402.

Wilson K, Cahill V, Ballment E, Benzie J. 2000. The complete sequence of the mitochondrial genome of the crustacean *Penaeus monodon*: are malacostracan crustaceans more closely related to insects than to branchiopods? *Molecular Biology and Evolution* **17:**863–874.

Zhang D, Gong F, Liu T, Guo H, Zhang N, Zhu K, Jiang S. 2015. Shotgun assembly of the mitochondrial genome from *Fenneropenaeus penicillatus* with phylogenetic consideration. *Marine Genomics* **24(Part 3):**379–386 DOI 10.1016/j.margen.2015.09.005.

Zhang D, Huang J, Zhou F, Gong F, Jiang S. 2016a. The complete mitochondrial genome of banana shrimp *Fenneropenaeus merguiensis* with phylogenetic consideration. *Mitochondrial DNA Part A* **27(4):**2606–2607 DOI 10.3109/19401736.2015.1041116.

Zhang D, Liu T, Gong F, Jiang S. 2016b. Shotgun assembly of the first mitochondrial genome of Metapenaeus (*Metapenaeus ensis*) with phylogenetic consideration. *Mitochondrial DNA Part A* **27(5):**3689–3690 DOI 10.3109/19401736.2015.1079857.

Tan et al. (2017), *PeerJ*, DOI 10.7717/peerj.2982

23/23

Author/s:
Tan, MH; Gan, HM; Lee, YP; Poore, GCB; Austin, CM

Title:
Digging deeper: new gene order rearrangements and distinct patterns of codons usage in mitochondrial genomes among shrimps from the Axiidea, Gebiidea and Caridea (Crustacea: Decapoda)

Date:
2017-03-01

Citation:
Tan, M. H., Gan, H. M., Lee, Y. P., Poore, G. C. B. & Austin, C. M. (2017). Digging deeper: new gene order rearrangements and distinct patterns of codons usage in mitochondrial genomes among shrimps from the Axiidea, Gebiidea and Caridea (Crustacea: Decapoda). PEERJ, 5 (3), https://doi.org/10.7717/peerj.2982.

Persistent Link:
http://hdl.handle.net/11343/270700

File Description:
Published version
License:
CC BY