The complete mitochondrial genome of the eusocial sponge-dwelling snapping shrimp *Synalpheus microneptunus*

Solomon T. C. Chak, Phillip Barden & J. Antonio Baeza

In the marine realm, eusociality is only known to have evolved within a clade of sponge-dwelling snapping shrimps in the genus *Synalpheus*. Deciphering the genomic underpinnings of eusociality in these marine shrimps has been limited by the sparse genomic resources in this genus. Here, we report, for a eusocial shrimp *Synalpheus microneptunus*, a complete mitochondrial genome (22X coverage) assembled from short Illumina 150 bp paired-end reads. The 15,603 bp long mitochondrial genome of *S. microneptunus* is AT-rich and includes 13 protein-coding genes (PCGs), 2 ribosomal RNA genes, 22 transfer RNA genes and an 834 bp intergenic region assumed to be the D-loop. The gene order is identical to that reported for most caridean shrimps and corresponds to the presumed Pancrustacean ground pattern. All PCGs showed signs of purifying selection, with $K_{\text{A}}/K_{\text{S}} < 1$ across the whole PCGs and most sliding windows within PCGs. Maximum-likelihood and Bayesian inference phylogenetic analyses of 13 PCGs and 68 terminals supports the monophyly of the Caridea and the family Alpheidae. The complete mitochondrial genome of the eusocial shrimp *Synalpheus microneptunus* will contribute to a better understanding of the selective pressures and rates of molecular evolution in marine eusocial animals.

Sponge-dwelling snapping shrimps in the genus *Synalpheus* (Decapoda: Alpheidae) are the only known clade of marine animals to have evolved eusociality, a complex social organization that is best known in terrestrial insects such as ants, bees and termites. Eusocial *Synalpheus* colonies typically consist of a single or a few queens and up to several hundred non-sterile workers of the two sexes. At least nine described species of *Synalpheus* in the West Atlantic ‘gambarelloides’ group are known to be eusocial. These species are characterized by their high reproductive skew, overlapping generations, and, in a few representatives in which behavioral observations have been made, cooperative defense of the host sponge. *Synalpheus* belonging to the ‘gambarelloides’ clade represent a relatively young lineage that radiated between ~5 and 7 Mya, yet eusociality has independently evolved at least four times in this genus. Communal living, where multiple mating pairs live in the same sponge, has also evolved multiple times in this clade from pair-forming ancestors. The social diversity, short evolutionary history, and similar ecology among *Synalpheus* shrimps make them an ideal group of marine animals to study the evolution of sociality. Moreover, despite a decade-long period of ecological dominance, eusocial *Synalpheus* shrimps have experienced recent population declines. Our knowledge of the biology of eusocial shrimps has increased substantially over the past decades. Yet, genomic resources are scarce in this group, especially when compared to that of social insects. Such a lack of genomic knowledge is limiting our understanding of behavioral innovations in sponge-dwelling snapping shrimps. Therefore, this study focuses on the development of genomic resources that are pivotal to improve our understanding of evolutionary innovations in this and other groups of crustaceans.

*Synalpheus microneptunus* is found only in reefs along the west coast of Barbados in the eastern Caribbean Sea and is the only known eusocial species in Barbados. Their colonies typically consist of <10 individuals with a single ovigerous female (i.e., the queen). They live in sponges belonging to *Neopetrosia proxima* and *N.

1Department of Biological Sciences, New Jersey Institute of Technology, Newark, NJ, 07102, USA. 2Department of Biological Sciences, 132 Long Hall, Clemson University, Clemson, SC, 29634, USA. 3Smithsonian Marine Station at Fort Pierce, 701 Seaway Drive, Fort Pierce, Florida, 34949, USA. 4Departamento de Biología Marina, Facultad de Ciencias del Mar, Universidad Católica del Norte, Larrondo 1281, Coquimbo, Chile. 5These authors contributed equally: Phillip Barden and J. Antonio Baeza. E-mail: tchak@njit.edu; jbaezam@clemson.edu
subtriangularis (previously Xestospongia) and may share the sponge host with S. belizensis, a pair-living species13. S. microneptunus comprised 25% of the total abundance of Synalpheus in Barbados, despite being one out of 14 Synalpheus species in the region. This observation mirrors the ecological dominance of eusocial species observed in Belizean coral reefs documented by a decade of field survey10. Synalpheus microneptunus is part of the S. paraneptunus species complex that shares several synapomorphic features that distinguish them from the rest of the S. gambareloides group, including sparse, unorganized setae on the non-snapping minor chela (while most other species have organized rows of setae) and the “excavated” inner surface of the fingers of the minor chela14. Phylogenetically, S. microneptunus is closely related to S. duffyi, a eusocial species that have larger colony sizes and is geographically widespread (Cuba, Florida, Jamaica, and Panama).

In this study, we describe the complete mitochondrial genome of the eusocial sponge-dwelling snapping shrimp S. microneptunus. Specifically, we analyze the nucleotide composition and codon usage profiles of protein coding genes (PCGs), and examine selective constraints in PCGs. We also describe the secondary structure of each identified tRNA gene, and examine the putative D-loop/control region (CR). In addition, we examine the phylogenetic position of S. microneptunus among other caridean shrimps based on mitochondrial PCGs.

**Methods**

**Field collection and sequencing.** A single specimen collected from Barbados in 200813 was used for DNA extraction and low-coverage whole genome sequencing (LC-WGS). A detailed field collection protocol has been reported previously15. We extracted genomic DNA using several walking legs from this alcohol-preserved specimen using the Qiagen DNeasy Tissue Kit (Qiagen). Extracted DNA was quantified using a Qubit 3.0 Fluorometer with the dsDNA HS assay (ThermoFisher Scientific) and visualized on 2% agarose gels. For LC-WGS, we provided 1,500 ng of genomic DNA to Novogene (Chula Vista, CA) for TruSeq PCR-free library preparation (Illumina) and 150 bp pair-end sequencing on an Illumina NovaSeq to obtain at least 1X coverage according to published genome size16. LC-WGS reads from whole-cell extraction contain a high copy number of extranuclear sequences, and it has been shown to be an efficient and economical approach to assemble complete mitochondrial genomes17.

**Mitochondrial genome assembly of Synalpheus microneptunus.** The mitochondrial genome of S. microneptunus was de novo-assembled using the NOVOPlasty pipeline v. 1.2.317. NOVOPlasty uses a
A seed-and-extend algorithm that assembles organelle genomes from WGS data, starting from a related or distant single 'seed' sequence and an optional 'bait' reference mitochondrial genome. For assembly, we used a previously published fragment of the COI gene from *S. microneptunus* (GenBank accession number KJ595111) as a seed and a kmer size of 39. We did not use a bait reference mitochondrial genome considering that there are no mitochondrial genomes from closely related (congeneric) species published and available in GenBank. Nuclear mitochondrial pseudogenes are abundant in the closely related genus *Alpheus* and may affect the assembly quality, resulting in many contigs. However, the adverse effect of mitochondrial pseudogenes is likely minimal when the contigs are being circularized.

Mitochondrial genome annotation and analysis. The newly assembled mitochondrial genome was first annotated in the MITOS web server (http://mitos.bioinf.uni-leipzig.de) using the invertebrate genetic code. Annotation curation, including start and stop codons corrections, were conducted using Expasy (https://web.expasy.org/) and MEGA X. Genome visualization was conducted with OrganellarGenomeDRAW (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html). Nucleotide composition and codon usage profiles of PCGs were analyzed. Nucleotide composition was estimated in MEGA X. Codon usage for each PCG was predicted using the invertebrate mitochondrial code in the Codon Usage web server (http://www.bioinformatics.org/sms2/codon_usage.html). tRNA genes were identified and their secondary structures were predicted in the software

| Name | Type | Start | Stop | Strand | Length (bp) | Start | Stop | Anticodon | Inter-Genic space | Overlap |
|------|------|-------|------|--------|-------------|-------|------|-----------|-----------------|---------|
| Cox1 | Coding | 1 | 1539 | + | 1539 | ACG | TAA | 5 |
| trnL2(tta) | tRNA | 1535 | 1601 | + | 67 | TAA | 18 |
| cox2 | Coding | 1620 | 2316 | + | 697 | ATA | T | 0 |
| trnK(aaG) | tRNA | 2317 | 2386 | + | 70 | TTC | 3 |
| trnD(gac) | tRNA | 2390 | 2454 | + | 65 | GTC | 9 |
| atp8 | Coding | 2464 | 2610 | + | 147 | ATA | TAG | 14 |
| atp6 | Coding | 2625 | 3278 | + | 656 | ATA | TAA | 1 |
| cox3 | Coding | 3278 | 4064 | + | 787 | ATG | T | 0 |
| trnG(gga) | tRNA | 4065 | 4131 | + | 67 | TCC | 6 |
| nad3 | Coding | 4138 | 4485 | + | 348 | ATA | TAA | 1 |
| trnA(gca) | tRNA | 4485 | 4547 | + | 63 | TGC | 5 |
| trnR(cga) | tRNA | 4553 | 4621 | + | 69 | TCG | 2 |
| trnN(aac) | tRNA | 4620 | 4684 | + | 65 | GTT | 0 |
| trnS1(aga) | tRNA | 4685 | 4742 | + | 58 | TCT | 13 |
| trnE(gaa) | tRNA | 4756 | 4823 | + | 68 | TTC | 2 |
| trnP(utc) | tRNA | 4826 | 4889 | + | 64 | GAA | 1 |
| nad5 | Coding | 4889 | 6586 | + | 1702 | ATA | TAG | 18 |
| trnH(cac) | tRNA | 6605 | 6669 | + | 65 | GTG | 1 |
| nad4 | Coding | 6669 | 8006 | + | 1337 | ATA | TAG | 4 |
| nad4l | Coding | 8003 | 8254 | + | 252 | ATG | TAA | 43 |
| trnT(aca) | tRNA | 8298 | 8358 | + | 61 | TGT | 1 |
| trnP(cca) | tRNA | 8358 | 8422 | + | 65 | TGG | 2 |
| nad6 | Coding | 8425 | 8940 | + | 516 | ATT | TAA | 1 |
| cob | Coding | 8940 | 10074 | + | 1135 | ATG | T | 0 |
| trnS2(tca) | tRNA | 10075 | 10145 | + | 71 | TGA | 2 |
| nad1 | Coding | 10144 | 11130 | + | 987 | ATA | TAG | 10 |
| trnL1(cta) | tRNA | 11121 | 11186 | + | 66 | TAG | 0 |
| rml | rRNA | 11187 | 12503 | + | 1317 | − | − | 6 |
| trnV(gta) | tRNA | 12498 | 12562 | + | 65 | TAC | 2 |
| trn5 | rRNA | 12565 | 13364 | + | 800 | − | − | 0 |
| CR | Putative | 13365 | 14198 | + | 834 | − | − | 0 |
| trnL(ata) | tRNA | 14199 | 14262 | + | 64 | GAT | 14 |
| trnQ(caa) | tRNA | 14277 | 14344 | + | 68 | TGT | 1 |
| trnM(agt) | tRNA | 14344 | 14409 | + | 66 | CAT | 15 |
| nad2 | Coding | 14395 | 15411 | + | 1017 | ATA | TAA | 2 |
| trnW(tga) | tRNA | 15410 | 15473 | + | 64 | TCA | 1 |
| trnC(tgc) | tRNA | 15473 | 15535 | − | 63 | GCA | 1 |
| trnI(tac) | tRNA | 15537 | 15603 | − | 67 | GTA | − |

Table 1. Arrangement and annotation of the mitochondrial genome of *Synalpheus microneptunus*. 

seed-and-extend algorithm that assembles organelle genomes from WGS data, starting from a related or distant single 'seed' sequence and an optional 'bait' reference mitochondrial genome. For assembly, we used a previously published fragment of the COI gene from *S. microneptunus* (GenBank accession number KJ595111) as a seed and a kmer size of 39. We did not use a bait reference mitochondrial genome considering that there are no mitochondrial genomes from closely related (congeneric) species published and available in GenBank. Nuclear mitochondrial pseudogenes are abundant in the closely related genus *Alpheus* and may affect the assembly quality, resulting in many contigs. However, the adverse effect of mitochondrial pseudogenes is likely minimal when the contigs are being circularized.

Mitochondrial genome annotation and analysis. The newly assembled mitochondrial genome was first annotated in the MITOS web server (http://mitos.bioinf.uni-leipzig.de) using the invertebrate genetic code. Annotation curation, including start and stop codons corrections, were conducted using Expasy (https://web.expasy.org/) and MEGA X. Genome visualization was conducted with OrganellarGenomeDRAW (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html). Nucleotide composition and codon usage profiles of PCGs were analyzed. Nucleotide composition was estimated in MEGA X. Codon usage for each PCG was predicted using the invertebrate mitochondrial code in the Codon Usage web server (http://www.bioinformatics.org/sms2/codon_usage.html). tRNA genes were identified and their secondary structures were predicted in the software
MITFI24 as implemented in the MITOS web server. tRNA secondary structure was visualized in the Forna web server (http://rna.tbi.univie.ac.at/forna)25.

We explored the selective constraints in all mitochondrial PCGs of *S. microneptunus*. Overall values of $K_A$ (the number of nonsynonymous substitutions per nonsynonymous site: $K_A = d_N = S_N/L_N$), $K_S$ (the number of synonymous substitutions per synonymous site: $K_S = d_S = S_S/L_S$), and the $K_A/K_S$ ratio (or $\omega$ or $d_N/d_S$) were estimated for each PCG in the software KaKs_calculator 2.026. $K_A$ and $K_S$ values were based on a pairwise comparison between *S. microneptunus* and *Alpheus lobidens* (GenBank accession number KP276147), a species belonging to a genus sister to *Synalpheus*27. We chose to use *A. lobidens* for the comparison because the mitochondrial genome of this species28 is best described among other *Alpheus*. Next, to identify positively selected sites along the length of each examined sequence, we also calculated the values of $K_A$, $K_S$, and $K_A/K_S$ along sliding windows of 57 bp that ‘slipped’ every 6 bp along each PCG. The $\gamma$-MYN model29 was used during calculations to account for variable mutation rates across sequence sites26. If PCGs are under no selection, positive selective constraint (purifying selection), or diversifying selection, the $K_A/K_S$ ratio is expected to be equal to 1, >1, or <1, respectively26. To confirm that the observed ratios of $K_A/K_S$ were not affected by the choice of the outgroup *Alpheus* species, we repeated the above analyses with all other species of *Alpheus* for which mitochondrial genomes were available in GenBank: *A. bellulus*, *A. distinguendus*, *A. inopinus*, and *A. randalli* (GenBank accession numbers: MH796167, NC_014883, MG551491, and MH796168, respectively).

The presence of inverted repeats in the putative D-loop/CR of *S. microneptunus* was explored with the 'EMBOSS:einverted' web server (http://www.bioinformatics.nl/cgi-bin/emboss/einverted) using the default options30. The presence and number of microsatellites (Simple Sequence Repeats, SSRs) were investigated with the 'Microsatellite repeats finder' web server using the default options (http://insilico.ehu.es/mini_tools/microsatellites)31. The RNAstructure web server (http://rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html)32 was used to predict the lowest free energy secondary structure of the putative control region with particular attention to the presence of stem-loops.

Lastly, we examined the phylogenetic position of *S. microneptunus* among other species of caridean shrimps (Decapoda: Caridea). The newly assembled and annotated mitochondrial genome of *S. microneptunus* and those of a total of 63 other species of carideans available in the GenBank database were used for phylogenetic analyses conducted using the MitoPhAST pipeline v2.033. Phylogenetic analyses included a total of 27 different genera in the infraorder Caridea. Outgroups included one species of lobster (*Stereomastis sculpia* [Polychelida]), two...
species of stenopodid shrimps (*Stenopus hispidus* and *Spongicola levigatus* [Stenopodidea]) and two species of prawns (*Penaeus vannamei* and *P. monodon* [Penaeoidea]). MitoPhAST extracts all 13 PCG nucleotide sequences from species available in GenBank and others provided by the user (i.e., *S. microneptunus*), translates each PCG nucleotide sequence to amino acids, conducts alignments for each PCG amino acid sequence using Clustal Omega, removes poorly aligned regions with trimAl, partitions the dataset and selects best fitting models of sequence evolution for each PCG with ProtTest and uses the concatenated and partitioned PCG amino acid alignments to perform a maximum likelihood phylogenetic analysis in the software IQ-TREE. The robustness of the ML tree topology was assessed by 1,000 bootstrap reiterations of the observed data. We also optimized the resulting amino acid sequence matrix under Bayesian Inference in MrBayes v3.2.7 with the same partitioning and model scheme as the ML search in IQ-TREE. Because MrBayes does not natively support one of the best fit models of molecular evolution identified with ProtTest (MtZoa), we manually implemented the substitution rate and state frequency priors from Rota-Stabelli et al. Our search spanned 5 million generations with four chains set to default temperature. We disregarded 25% of sampled trees as burn-in and sampled every 500 cycles. We assessed effective sample size in Tracer v 1.7 and convergence of runs through the average standard deviation of split frequencies.

**Results and Discussion**

Using 51,305,421 paired-end sequences (SRA: SRX6711388), we completely assembled and circularized the mitochondrial genome of *S. microneptunus* with a coverage of 22× (GenBank accession number MN750781). The complete mitochondrial genome of *S. microneptunus* was 15,603 bp in length and comprised 13 protein-coding genes (PCGs), two ribosomal RNA genes (*rrnS* [12S ribosomal RNA] and *rrnl* [16S ribosomal RNA]), and 22 transfer RNA (tRNA) genes. The mitochondrial genome of *S. microneptunus* was compact with only a few intergenic spaces and overlaps among gene junctions (Fig. 1, Table 1). Most of the PCGs and tRNA genes were encoded on the heavy strand, while only four PCGs (in order from 5′ to 3′: *nad5*, *nad4*, *nad4L*, and *nad1*), two ribosomal RNA genes and 8 tRNA genes (*trnF, trnH, trnP, trnL1, trnV, trnQ, trnC, and trnY*) were encoded in the light strand (Fig. 1, Table 1). A single, long intergenic space of 834 bp was assumed to be the D-loop/control region (Fig. 1, Table 1). The gene order observed in *S. microneptunus* is identical to that reported for most caridean shrimps and corresponds to the presumed Pancrustacean (Hexapoda + Crustacea) ground pattern. Interestingly, the gene order observed in *S. microneptunus* is different from that reported in the closely related genus *Alpheus* and other decapod crustaceans, which remains to be addressed.
In the mitochondrial genome of *S. microneptunus*, 12 out of the 13 PCGs exhibited conventional invertebrate and Pancrustacean mitochondrial start codons (ATA, ATG, and ATC) (Table 1). Cox1 featured an alternative putative start codon (ACG) that was previously reported for other decapod crustaceans and references therein⁴⁶, but was found only in a few caridean shrimps (i.e., *Nautilocaris saintlaurentae*⁴⁷; *Macrobrachium rosenbergii*⁴⁸). Eleven PCGs ended with a complete and conventional stop codon (TAA or TAG) (Table 1). The genes *cox2*, *cox3*, and *cob* each terminated with an incomplete stop codon T. Truncated stop codons are often observed in crustacean mitochondrial genomes⁴⁳,⁴⁹ and are hypothesized to be completed via post-transcriptional polyadenylation⁵⁰.

The mitochondrial genome of *S. microneptunus* contained an A + T bias with an overall base composition of A = 36.6%, T = 28.0%, C = 23.9%, and G = 11.4% at the heavy strand. This A + T bias is within the known range reported for mitochondrial genomes in caridean shrimps⁵¹. The most frequently used codons found in the PCGs

---

**Figure 4.** Maximum-likelihood phylogenetic tree based on the amino acid alignments of 13 protein coding genes in the mitochondrial genome of *Synalpheus microneptunus* and 43 caridean shrimps. Outgroups included one species of lobster (*Stereomastis sculpta*), two species of stenopodid shrimps (*Stenopus hispidus* and *Spongicola levigatus*) and two species of prawns (*Peneaus vannamei* and *P. monodon*). Number at each node represents bootstrap values. The analysis included a total of 68 terminals, 3636 amino acid characters, and 1864 parsimony informative sites. The tree was drawn using Mesquite v3.6⁶³.
of S. microneuptus were UUA (Leu, \( N = 223 \) times used), UUU (Phe, \( N = 186 \)), and AUU (Ile, \( N = 182 \)). Least frequently used codons (excluding termination codons) included CCG (Pro, \( N = 9 \)), CGG (Arg, \( N = 8 \)), and AGC (Ser, \( N = 8 \)) (Suppl. Mat. Table S1).

All PCGs in the mitochondrial genome of S. microneuptus exhibited overall KA/KS ratios << 1 (Fig. 2). This indicates that these PCGs are generally evolving under purifying selection. Examination of KA/KS ratios in sliding windows across the length of each PCG further indicated that purifying selection is acting along most of the length of each PCG (Suppl. Mat. Fig. S1). The results were very similar when the above analyses were performed using other species of Alpheus as outgroups (Suppl. Mat. Figs. S2–S6), confirming the general pattern of purifying selection in mitochondrial PCGs. Selective pressure in mitochondrial PCG has been poorly studied in crustaceans but a similar pattern of widespread purifying selection in mitochondrial PCGs has been observed in other arthropods, including decapod crustaceans, and references therein46,51. Interestingly, regardless of Alpheus outgroup, the genes atp8 and nad6 exhibited higher \( K_s/K_a \) ratios than other genes, but with values lower than 1. These two genes were also found to have higher \( K_s/K_a \) ratios than other mitochondrial genes between two Alpheus species52. This suggests that selective pressures may tend to be relaxed in these genes across Alpheidae. However, broad investigations of

![Figure 5](https://www.nature.com/scientificreports/)
the selective pressures in mitochondrial genes across Caridean species, or crustaceans in general, are sparse. It is possible that eusociality may drive changes in the rate of evolution in mitochondrial genes due to prolonged longevity in the queens, longer generation time, and reduced effective population size. Whether the higher Ks/Ka ratios observed in a few PCGs in *S. microneptunus* are driven by eusociality or result from other unknown (e.g., ecological) differences between *Synalpheus* and *Alpheus* remains to be investigated in further comparative analyses.

The mitochondrial genome of *S. microneptunus* encoded tRNA genes that ranged in length from 58 (tRNA-Ser) to 71 (tRNA-Ser) bp. All tRNA genes, except tRNA-Ser, exhibited a standard ‘cloverleaf’ secondary structure as predicted by MIFTI (Fig. 3). In the tRNA-Ser gene, the stem and loop of the pseudouridine arm (T-arm) was missing. Complete (stem and loop) or partial (loop only) tRNA arm deletions are known to occur in other decapod crustaceans, and references therein and the function of these tRNA may be complemented by elongation factors.

The *rrnS* and *rrnL* genes identified in the mitochondrial genome of *S. microneptunus* were 800 and 1371 nucleotides long, respectively. These genes were located close to each other between tRNA-L1 and the putative D-loop/CR, but separated by tRNA-V (Fig. 1, Table 1). As shown to occur in other crustaceans, including caridean shrimps, the two genes were highly A + T biased. The overall base composition of the *rrnL* gene was A = 29.4%, T = 39.9%, C = 8.2%, and G = 22.5%. In turn, that of the *rrnS* gene was A = 29.2%, T = 38.8%, C = 8.8%, and G = 23.1%.

The 834 bp long intergenic region assumed to be the D-loop/CR was located between the 12 S ribosomal RNA and tRNA-I (Fig. 1) in *S. microneptunus*. The region was heavily A + T rich with an overall base composition: A = 42.0%, T = 37.5%, C = 15.2%, and G = 5.3%. Visual examination of this non-coding region revealed multiple monocistronic adenine and thymine repeats along the entire stretch of this intergenic region. The region has an imperfect inverted repeat located in positions 176–221 and 227–275 (detected by EMBOSS) and multiple AT-rich dinucleotide and trinucleotide microsatellites along the entire stretch of the CR (detected by microsatellite Repeat Finder) (Suppl. Mat. Fig. S3). The secondary structure prediction analysis in RNASTRUCTURE (assuming 27 °C and other default options) resulted in a lowest free energy configuration (change in Gibbs free energy ΔG = −104 kcal/mol) that featured various stem-loop structures interspersed along the length of the region (Suppl. Mat. Fig. S7). Only a few studies have characterized the putative D-Loop/CR in crustaceans and references therein. In some species this long non-coding region appears to be relatively well organized (i.e., in the non-decapod branchiopod genus *Daphnia* and in the decapod Chinese spiny lobster *Panulirus argus*) and here in *S. microneptunus*, there is no obvious organization in the D-Loop/CR.

The ML and BI phyllogenetic trees (68 terminals, 3636 amino acid characters, and 1864 parsimony informative sites) support the monophyly of the Caridea and placed *S. microneptunus* in a monophyletic clade (family Alpheidae) sister to representatives from the genus *Alpheus*. The above relationship supports the monophyly of the family Alpheidae in agreement with results from previous phylogenetic studies using a combination of partial mitochondrial and nuclear genes (Figs. 4–5) or using mitochondrial PCGs but with a more limited sample of caridean shrimps. Additional well supported clades within the Caridea included the families Alvinocaridae, Atyidae, Palaemonidae, and Pandalidae. While the monophyly of these caridean families was supported in both ML and BI analyses, the relationships among families was found to be sensitive to optimality criteria. These differences are reflected in the low support we recovered for the node leading to Atyidae (Palaemonidae + Alpheidae) + Pandalidae in the ML tree. Our BI analysis recovered a monophyletic (Palaemonidae + Alpheidae) clade; however all other inter-familial relationships are recovered as distinct from the ML topology, and with high support. The sister relationship herein observed between the families Palaemonidae and Alpheidae was also supported by a recent phylogenomic study. The same phylogenomic analysis resolved Atyidae as the sister group to all other caridean taxa, however, none of our analyses recover this relationship. Our results suggest that mitochondrial genomes contain enough phylogenetic information to delineate monophyly of higher taxa within the Caridea (at superfamily and family levels), but the relationships among monophyletic clades may be sensitive to marker choice and reconstruction methodology.

**Conclusions**

This study assembled and analyzed the first mitochondrial genome of a eusocial marine invertebrate, the caridean shrimp *S. microneptunus*. The complete mitochondrial genome of *S. microneptunus* will enhance the genomic resources in the only known group of eusocial animals in the sea and allow further investigations of the relationship between complex social behaviors (i.e., eusociality and communal living) and selective pressures and rates of molecular evolution in mitochondrial genomes.

**Data availability**

Data is available at GenBank (MN750781 and SRX6711388).

Received: 18 December 2019; Accepted: 13 April 2020;
Published online: 08 May 2020

**References**

1. Hultgren, K. M., Duffy, J. E. & Rubenstein, D. R. In Comparative Social Evolution (eds. Rubenstein, D.R. & Abbot, P.) 224–249 (Cambridge University Press, 2017).
2. Duffy, J. E. Eusociality in a coral-reef shrimp. Nature 381, 512–514, https://doi.org/10.1038/381512a0 (1996).
3. Sherman, P. W., Lacey, E. A., Reeve, H. K. & Keller, L. The eusociality continuum. Behav. Ecol. 6, 102–108 (1995).
4. Chak, S. T. C., Duffy, J. E. & Rubenstein, D. R. Reproductive skew drives patterns of sexual dimorphism in sponge-dwelling snapping shrimps. Proc. R. Soc. B 282, 20150342, https://doi.org/10.1098/rspb.2015.0342 (2015).
41. Qian, G.
42. Shen, X., Li, X., Sha, Z., Yan, B. & Xu, Q. Complete mitochondrial genome of the Japanese snapping shrimp
39. Rota-Stabelli, O., Yang, Z. & Telford, M. J. MtZoa: A general mitochondrial amino acid substitutions model for animal evolutionary
38. Ronquist, F.
34. Sievers, F.
32. Reuter, J. S. & Mathews, D. H. RNAstructure: software for RNA secondary structure prediction and analysis.
22. Lohse, M., Drechsel, O., Kahlau, S. & Bock, R. OrganellarGenomeDRAW–a suite of tools for generating physical maps of plastid and
25. Kerpedjiev, P., Hammer, S. & Hofacker, I. L. Forna (force-directed. RNA): Simple and effective online RNA secondary structure
31. Bikandi, J., Millan, R. S., Rementeria, A. & Garaizar, J. In silico analysis of complete bacterial genomes: PCR, AFLP-PCR and
15. Macdonald, K. S., Ríos, R. & Duffy, J. E. Biodiversity, host specificity, and dominance by eusocial species among sponge-dwelling
13. Hultgren, K. M., MacDonald, K. S. & Duffy, E. Sponge-dwelling snapping shrimps (Alpheidae: Synalpheus) and Barbados, West
12. Gaynor, K. M.
10. Duffy, J. E. & Macdonald, K. S. Kin structure, ecology and the evolution of social organization in shrimp: a comparative analysis.
Plos ONE 8, e54637, https://doi.org/10.1371/journal.pone.0054637 (2013).
Gnathophausia gigas Briggs, 1869
Herring: Clupeidae, Idreidae

References

18. Williams, S. T. & Knowlton, N. Mitochondrial pseudogenes are pervasive and often insidious in the snapping shrimp genus Alpheus.
Mol. Biol. Evol. 18, 1484–1493 (2001).
19. Stothard, P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences.
BioTechniques 59, 393–402, https://doi.org/10.1139/gen-2015-0206 (2016).
20. Williams, S. T. & Knowlton, N. Mitochondrial pseudogenes are pervasive and often insidious in the snapping shrimp genus Alpheus.
Mol. Biol. Evol. 18, 1484–1493 (2001).
17. Dierckxsens, N., Mardulyn, P. & Smits, G. NOVOPlasty: de novo assembly of organelle genomes from whole genome data.
16. Jeffery, N. W., Hultgren, K., Chak, S. T. C., Gregory, T. R. & Rubenstein, D. R. Patterns of genome size variation in snapping shrimp.
15. Macdonald, K. S., Ríos, R. & Duffy, J. E. Biodiversity, host specificity, and dominance by eusocial species among sponge-dwelling
14. Hultgren, K. M. & Brandt, A. Taxonomy and phylogenetics of the species-complex (Decapoda: Alpheidae), with a description of two new
13. Hultgren, K. M., MacDonald, K. S. & Duffy, E. Sponge-dwelling snapping shrimps (Alpheidae: Synalpheus) of Barbados, West
Indies, with a description of a new eusocial species. Zootaxa 2834, 1–16 (2011).
12. Gaynor, K. M.
11. Hultgren, K. M. & Rubenstein, D. R. Decline and local extinction of Caribbean eusocial shrimp. PLoS ONE 8, e54637, https://doi.org/10.1371/journal.pone.0054637 (2013).
10. Duffy, J. E. & Macdonald, K. S. Kin structure, ecology and the evolution of social organization in shrimp: a comparative analysis.
Plos ONE 8, 757–764, https://doi.org/10.1371/journal.pone.00252-5 (2002).
9. Chak, S. T. C., Rubenstein, D. R. & Duffy, J. E. Social control of reproduction and breeding monopolization in the eusocial snapping
shrimp Synalpheus elisabethae. The American Naturalist 186, 660–668, https://doi.org/10.1086/683132 (2015).
8. Morrison, C. L., Rios, R. & Duffy, J. E. Phylogenetic evidence for an ancient rapid radiation of Caribbean sponge-dwelling snapping
shrimps (Synalpheus). Mol Phylogenet Evol 30, 563–581, https://doi.org/10.1016/S1055-7903(03)00225-5 (2004).
7. Duffy, J. E., Morrison, C. L. & Macdonald, K. S. Colony defense and behavioral differentiation in the eusocial shrimp Synalpheus
regalis. Behav. Ecol. Sociobiol. 51, 488–495, https://doi.org/10.1007/s00265-002-0456-5 (2002).
6. Tóth, E. & Duffy, J. E. Coordinated group response to nest intruders in social shrimp. Biol. Lett. 1, 49–52, https://doi.org/10.1098/
rbl.2004.0237 (2005).
5. Chak, S. T. C., Rubenstein, D. R. & Duffy, J. E. Social control of reproduction and breeding monopolization in the eusocial snapping
shrimp Synalpheus elisabethae. The American Naturalist 186, 660–668, https://doi.org/10.1086/683132 (2015).
63. Mesquite: a modular system for evolutionary analysis v. 3.61 (2019).
60. Liu, Y. & Cui, Z. Complete mitochondrial genome of the Chinese spiny lobster Panulirus stimpsoni (Crustacea: Decapoda): genome characterization and phylogenetic considerations. Mol. Biol. Evol. 107–112, https://doi.org/10.1093/molbev/msw097 (2018).
62. Wolfe, J. M. & Cui, Z. Complete mitochondrial genome of the Chinese spiny lobster Panulirus stimpsoni (Crustacea: Decapoda): genome characterization and phylogenetic considerations. Mol. Biol. Evol. 107–112, https://doi.org/10.1093/molbev/msw097 (2018).
64. Baer, J. A. The complete mitochondrial genome of the Caribbean spiny lobster Panulirus argus. Scientific Reports 8, 17690, https://doi.org/10.1038/s41598-018-36132-6 (2018).
67. Kim, S.-J., Pak, S. I. & Ju, S.-I. Mitochondrial genome of the hydrothermal vent shrimp Nautilocaris saintlaurentae (Crustacea: Caridea: Alvinocaridae). Mitochondrial DNA 26, 127–128, https://doi.org/10.1080/19401736.2013.815169 (2015).
68. Miller, A. D., Murphy, N. P., Burridge, C. P. & Austin, C. M. Complete Mitochondrial DNA Sequences of the Decapod Crustaceans Pseudocarcinus gigas (Menippeidae) and Macrobrachium rosenbergii (Palaemonidae). Mar. Biotechnol. 7, 339–349, https://doi.org/10.1007/s10126-004-0077-8 (2005).
65. Ivey, J. L. & Santos, S. R. The complete mitochondrial genome of the Hawaiian anchialine shrimp Halocaridina rubra Holthuis, 1963 (Crustacea: Decapoda: Atyidae). Gene 394, 35–44, https://doi.org/10.1016/j.gene.2007.01.009 (2007).
66. Beckenbach, A. Nums and mitochondrial pseudogenes. Myrmecological News 12, 217–218 (2009).
69. Li, T. et al. A mitochondrial genome of Rhyparochromidae (Hemiptera: Heteroptera) and a comparative analysis of related mitochondrial genomes. Scientific Reports 6, 35175, https://doi.org/10.1038/srep35175 (2016).
70. Keller, L. & Genoud, M. Extraordinary lifespans in ants: a test of evolutionary theories of ageing. Nature 389, 958–960, https://doi.org/10.1038/40130 (1997).
71. Schmidt, C. M., Jarvis, J. U. M. & Bennett, N. C. The long-lived queen: reproduction and longevity in female eusocial Damaraland mole-rats (Fukomys damarensis). Afr. Zool. 48, 193–196, https://doi.org/10.1080/15627020.2013.11407583 (2013).
72. Thorne, B. L., Breisch, N. L. & Haverty, M. I. Longevity of kings and queens and first time of production of fertile progeny in Zootermopsis angusticollis colonies with different reproductive structures. Journal of Animal Ecology 71, 1030–1041, https://doi.org/10.1111/j.1365-2656.2002.00666.x (2002).
73. Ingram, C. M., Troendle, N. J., Gill, C. A., Braule, S. & Honeycutt, R. L. Challenging the inbreeding hypothesis in a eusocial mole-rat. Mol. Ecol. 24, 4848–4865, https://doi.org/10.1111/mec.13358 (2015).
74. Romiguier, J. et al. Population genomics of eusocial insects: the costs of a vertebrate-like effective population size. J. Evol. Biol. 27, 593–603, https://doi.org/10.1111/jeb.12331 (2014).
75. Bromham, L. & Leyss, R. Sociality and the rate of molecular evolution. Mol. Biol. Evol. 22, 1393–1402, https://doi.org/10.1093/molbev/msl133 (2005).
76. Watanabe, Y.-I., Suematsu, T. & Ohtsuki, T. Losing the stem-loop structure from metazoan mitochondrial tRNAs and co-evolution of interacting factors. Frontiers in Genetics 5, 109, https://doi.org/10.3389/fgen.2014.00109 (2014).
77. Kuhn, K., Streit, B. & Schwenk, K. Conservation of structural elements in the mitochondrial control region of Daphnia. Gene 420, 107–112, https://doi.org/10.1016/j.gene.2008.05.029 (2008).
78. Liu, Y. & Cui, Z. Complete mitochondrial genome of the Chinese spiny lobster Panulirus stimpsoni (Crustacea: Decapoda): genome characterization and phylogenetic considerations. Molecular Biology Reports 38, 403–410, https://doi.org/10.1007/s11033-010-0122-2 (2011).
79. Palero, F., Crandall, K. A., Abelló, P., Macpherson, E. & Pascual, M. Phylogenetic relationships between spiny, slipper and coral lobsters (Crustacea, Decapoda, Achelata). Mol. Phylog. Evol. 50, 152–162, https://doi.org/10.1016/j.ympev.2008.10.003 (2009).
80. Wolfe, J. M. et al. A phylogenomic framework, evolutionary timeline and genomic resources for comparative studies of decapod crustaceans. Proc. R. Soc. B 286, 20190079, https://doi.org/10.1098/rspb.2019.0079 (2019).

Acknowledgements
J.A.B. thanks Dr. Vincent P. Richards for bioinformatics support during the development of this project. S.T.T.C. thanks Gedaliah Wolosh for bioinformatics support and staff of IST Academic and Research Computing Systems at NJIT for support in the HPC system.

Author contributions
S.T.T.C., J.A.B. and P.B. conceived, designed and supervised the study, analyzed data, and wrote the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-64269-w.

Correspondence and requests for materials should be addressed to S.T.T.C. or J.A.B.

Reprints and permissions information is available at www.nature.com/reprints.

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

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

© The Author(s) 2020