Available online at www.sciencedirect.com
ScienceDirect
Gene 394 (2007) 35–44
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

The complete mitochondrial genome of the Hawaiian anchialine shrimp *Halocaridina rubra* Holthuis, 1963 (Crustacea: Decapoda: Atyidae)

Jennifer L. Ivey [a,c], Scott R. Santos [b,*]

[a] *Department of Biological Sciences, Auburn University, Auburn, AL 36849, USA*
[b] *Department of Biological Sciences and Cell and Molecular Biosciences Peak Program, Auburn University, Auburn, AL 36849, USA*
[c] *Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA*

Received 28 November 2006; received in revised form 22 December 2006; accepted 8 January 2007
Available online 26 January 2007
Received by J.G. Zhang

Abstract

Shrimp of the family Atyidae are important members of nearly all tropical (and most temperate) fresh and brackish water ecosystems in the world. To date, a complete mitochondrial genome from this important crustacean group has not been reported. Here, we present the complete mitochondrial DNA sequence of the Hawaiian atyid *Halocaridina rubra* [Holthuis, L.B., 1963. On red coloured shrimps (Decapoda, Caridea) from tropical land-locked saltwater pools. Zool. Meded.16, 261–279.] (Crustacea: Decapoda: Atyidae). The genome is a circular molecule of 16,065 bp and encodes the 37 mitochondrial genes (13 protein-coding, 22 tRNAs, and two rRNAs) typically found in the metazoa. Gene order and orientation in the *H. rubra* mitochondrial genome is syntenic with most malacostracans that have been examined to date. Of special note is the absence of the dihydrouridine (DHU) arm stem from $tRNA^{Tyr}$ and the use of CCG as an initiation codon for cytochrome oxidase subunit I (*COI*); these represent the first reported examples of such phenomena in the Malacostraca. Phylogenetic analyses utilizing complete mitochondrial sequences from other malacostracans place *H. rubra* as sister to *Macrobrachium rosenbergii*, which also belongs to the Infraorder Caridea. However, the placement of this infraorder, as well as the Infraorder Dendrobrachiata, in the phylogeny of the Decapoda varied depending on outgroup selection. Data from additional mitochondrial genomes, such as basal decapods like the Stenopodidea, should contribute to a better overall understanding of decapod phylogenetics.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Atyid; Mitochondrial DNA

1. Introduction

Mitochondria are organelles found in eukaryotic cells and are responsible for the majority of cellular ATP production (Dimijian, 2000; Hedges et al., 2001). Mitochondria contain their own genome composed of double-stranded circular molecules that typically code for 37 genes (i.e., 13 protein-coding, 22 tRNAs and two rRNAs) involved in the electron transport chain (reviewed by Ballard and Whitlock, 2004). Mitochondrial genes are commonly utilized as genetic markers in determining the evolutionary history and phylogenetic relationships of organisms. Biological characteristics such as a uniparental (usually maternal) mode of inheritance, lack of recombination, and an accelerated rate of mutation have made mitochondrial genes popular tools in molecular genetic studies (Brown et al., 1979; Birky, 1995; Ballard and Whitlock, 2004).

The Atyidae are a family of caridean shrimp characterized by chelae with hair-like setae (Bruce, 1992) and occur in nearly all tropical, and most temperate, fresh and brackish water systems of the world. Typically of small (i.e., <35 mm) size, members of this family serve key ecological roles in sediment decomposition (Pringle and Blake, 1994), processing

Abbreviations: ATP6, and *8*, ATPase subunits 6 and 8; bp, base pair (s); *COI–III*, cytochrome *c* oxidase subunits I–III; CR, control region; *CytB*, cytochrome *b*; DHU, dihydrouridine; *ND1–6*, and *4L*, NADH dehydrogenase subunits 1–6 and 4L; PCR, polymerase chain reaction; *srRNA*, and *lrRNA*, small and large subunits ribosomal RNA; *tRNA*, transfer RNA.

* Corresponding author. Department of Biological Sciences, Auburn University, 101 Life Sciences Building, Auburn University, AL 36849, USA. Tel.: +1 334 844 7410; fax: +1 334 844 1645.
E-mail address: santos@auburn.edu (S.R. Santos).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2007.01.009

of particulate organic matter (Pringle et al., 1999) and influencing the composition of benthic communities (Pringle et al., 1993; Pringle, 1996; Pringle and Hamazaki, 1998) in these habitats.

Although numerous studies have utilized mitochondrial genes to explore the phylogeography (e.g., Hurwood and Hughes, 2001; Chenoweth and Hughes, 2003; Hurwood et al., 2003; Baker et al., 2004) and evolution (e.g., Page et al., 2005; Porter et al., 2005; Page et al., in press) of the Atyidae, a complete mitochondrial genome from this crustacean group has not been reported. Here, we present the complete mitochondrial genome sequence of the endemic Hawaiian atyid *Halocaridina rubra*, Holthuis 1963 (Holthuis, 1963). These small (i.e., ~10 mm) microphagous grazers, which live up to 20 years in captivity, occupy the unique niche of anchialine habitats in the archipelago (Bailey-Brock and Brock, 1993). Anchialine habitats are classified as "bodies of haline waters, usually with restricted exposure to air, always with extensive subterranean connections to the sea, and showing noticeable marine and terrestrial influences" (Stock, 1986). The complete mitochondrial genome sequence of *H. rubra* will provide additional genetic markers for elucidating the population structure and planning conservation management strategies for the species (Santos, 2006) as well as contribute fundamental information for evolutionary studies, such as inferring the phylogenetic history of the malacostracan crustaceans.

2. Materials and methods

2.1. Sample collection and DNA extraction

Total DNA was obtained from one *H. rubra* individual collected from an anchialine habitat at Cape Hanamanioa on the island of Maui in June 2005. This individual belongs to a distinct mitochondrial lineage from those reported by Santos (2006) to occur on the Island of Hawaii (unpublished data). The specimen was collected and immediately stored in 100% acetone (Fukatsu, 1999) prior to DNA extraction as described in Santos (2006).

2.2. PCR and sequence determination

As a starting point, the mitochondrial cytochrome oxidase subunit I (*COI*) and large subunit rRNA (*lrRNA*) genes were amplified via polymerase chain reaction (PCR) using the primers LCO1490/HCO2198 (Folmer et al., 1994) and CRUST16SF/CRUST16SR, respectively (Table 1). The primers CRUST16SF and CRUST16SR were designed to target conserved regions at the 5′ and 3′ boundaries of the *lrRNA* gene. These conserved regions, separated by ~900 bp, were identified from an alignment of 30 full-length *lrRNA* genes acquired from the complete crustacean mitochondrial genomes available in GenBank as of January 2006. Approximately 10–30 ng of DNA was utilized as template and reactions were conducted in 25 μL volumes containing 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 0.001% gelatin, 2.0 mM $MgCl_2$, 200 μM dNTPs, 1 U *Taq* DNA polymerase (Eppendorf) and 0.4 μM of each primer. The following thermocycling profile was utilized for *COI*: initial denaturing step of 94 °C for 5 min, 15 cycles of 94 °C for 45 s, 40 °C for 45 s, 72 °C for 60 s; 25 cycles of 94 °C for 45 s, 55 °C for 45 s, 72 °C for 60 s and a final extension of 72 °C for 5 min. A "touchdown" profile was used for *lrRNA*: initial denaturing step of 94 °C for 4:30 min, 11 cycles of 94 °C for 45 s, 60 °C for 45 s (−1 °C per cycle), 72 °C for 60 s; 26 cycles of 94 °C for 45 s, 50 °C for 45 s, 72 °C for 60 s and a final extension of 72 °C for 3 min.

Additionally, selected PCR primers from Yamauchi et al. (2004) were utilized in the amplification and sequencing of the *H. rubra* mitochondrial genome (Table 1). For example, the pairing of H13845-12S (Yamauchi et al., 2004) and CRUST16SF yielded a fragment spanning the *lrRNA* gene and the small subunit rRNA (*srRNA*) gene (Table 1). PCR conditions for this reaction were identical to the amplification of *lrRNA* (see above) with the exception of a final $MgCl_2$ concentration of 3.0 mM.

Initial sequences from *COI*, *lrRNA* and *srRNA* were utilized in the design of primers for long-PCR reactions (Table 1). Primers were selected to be 36–40 bp in length, G/C rich at the 3′-end and nested ~150–200 bp from the ends of the initial *COI*, *lrRNA* or *srRNA* gene sequences to allow

Table 1
Primer utilized in amplification and sequencing of the *Halocaridina rubra* Holthuis, 1963 (Crustacea: Decapoda: Atyidae) mitochondrial genome

Fragment	Forward primer (5′→3′)	Reverse primer (5′→3′)	Size (bp)
Cytochrome oxidase subunit I (*COI*)	LCO1490[a]: GGT CAA CAA ATC ATA AAG ATA TTG G	HCO2198[a]: TAA ACT TCA GGG TGA CCA AAA AAT CA	709
Large subunit rRNA (*lrRNA*)	CRUST16SF: TAA TTC AAC ATC GAG GTC GCA A	CRUST16SR: TTT GTA CCT TKT GTA TTA GG	969
lrRNA+small subunit rRNA (*srRNA*)	CRUST16SF: TAA TTC AAC ATC GAG GTC GCA A	H13845-12S[b]: GTG CCA GCA GCT GCG GTT A	1881
Long-PCR 1	HrCOI_12S: CGG TCA GTC AGG AGT ATG GCA ATA GCT CCT GCT AAT ACT G	Hr12S_COI: ATG AGA GCG ACG GGC GAT GTG TAC ATA TCC TAG AGC	3927
Long-PCR 2	HrCOI_16S: GAC ATG GCA TTC CCC CGA ATG AAT AAC ATA AGA TTC TGG C	H5244-CO3[b]: GTC AAT ATC AWG CDG CDG CTT CAA ATC CWA AGT GGT G	3784
Long-PCR 3	HrCOIII_16S: AAT CGC AGA CTC TGT CTA TGG GTC AAC ATT TTT CGT GGC	Hr16S_COI: CAT GTC TAT TGG GGT TTA AAT CAA TAG TCT GAC CTG C	8052

[a] Folmer et al. (1994).
[b] Yamauchi et al. (2004).

sufficient overlap for contig assembly. Amplification of large (~3.7–8 kb) fragments was done with a Triple Master Long PCR kit (Eppendorf) under the following conditions: 1× High Fidelity buffer, 0.5 mM dNTPs, 1.2 μM of each primer, 2 U enzyme mixture and 50 ng DNA template in a final volume of 50 μL. Thermal cycling profiles were as directed by the manufacturer. All PCR reactions were conducted in a PTC-100 thermocycler (MJ Research).

For amplified products of <2 kb, sequence reads were generated using the same primers as for PCR. For large fragments, products were individually sheared into ~1 kb fragments using a HydroShear (Genomic Solutions) and blunt ended with a DNATerminator repair kit (Lucigen). Sheared fragments were cloned into the pSMART vector system (Lucigen) and 12–15 clones/large fragment were sequenced using the supplied vector primers as directed by the manufacturer. Sequences were generated using Big-Dye Terminators v3.1 and read on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Ambiguities in the chromatograms were corrected by comparison to the complement DNA strand and contigs assembled with Sequencher v4.6 (Gene Codes).

2.3. Gene identification and genome analysis

Gene homologies and the boundaries of protein-coding genes were initially identified via translated query vs. translated database BLAST searches (tblastx; Altschul et al., 1990) of all open reading frames >80 bp in length determined by the program Artemis v8.0 (Rutherford et al., 2000) under the invertebrate mitochondrial genetic code and refined by alignment to homologues from the complete mitochondrial genomes of other Malacostraca (see below). Ribosomal RNAs (i.e., *lrRNA* and *srRNA*) and their boundaries were determined by a combination of nucleotide BLAST (i.e., blastn) and alignment to homologues as described above for protein-coding genes. The location and orientation of tRNAs were identified by tRNAscan-SE v1.1 (Lowe and Eddy, 1997) under default settings and source = “mito/chloroplast”. In cases where potential tRNAs were not identified by tRNAscan-SE, the mitochondrial sequence was inspected for segments forming tRNA-like secondary structures and possessing appropriate anticodons prior to being folded manually. Final sequence annotation as well as positional and A+T content analyses were conducted with the program Artemis v8.0 and the mitochondrial genome was visualized using the program CGView (Stothard and Wishart, 2005). The complete mitochondrial genome of *H. rubra* was deposited in GenBank under Accession No. DQ917432.

2.4. Phylogenetic analyses

Phylogenetic analyses were performed using 18 Malacostraca complete mitochondrial sequences archived in GenBank as of Nov. 2006: *Callinectes sapidus* (AY363392), *Cherax destructor* (AY383557), *Eriocheir sinensis* (AY274302), *Geothelphusa dehaani* (AB187570), *Gonodactylus chiragra* (DQ191682), *H. rubra* (DQ917432), *Harpiosquilla harpax* (AY699271), *Ligia oceanica* (DQ442914), *Lysiosquillina maculata* (DQ191683), *Macrobrachium rosenbergii* (AY659990), *Marsupenaeus japonicus* (AP006346), *Pagurus longicarpus* (AF150756), *Panulirus japonicus* (AB071201), *Penaeus monodon* (AF217843), *Portunus trituberculatus* (AB093006), *Pseudocarcinus gigas* (AY562127), *Squilla empusa* (DQ191684), and *Squilla mantis* (AY639936). Amino acid and nucleotide sequences for all 13 protein-coding mitochondrial genes from these Malacostraca (NCBI Taxonomy ID = 6681) were acquired and parsed into individual gene datasets using the program MitoBank.pl v2.0 (Abascal et al., 2007). In addition, the complete mitochondrial genome of *Daphnia pulex* (Accession No.=AF117817; NCBI Taxonomy ID = 6669) was obtained for use as an outgroup. Amino acid sequences for each gene were individually aligned using ClustalX (Thompson et al., 1997) under default settings; these alignments were subsequently utilized as a scaffold to align the corresponding nucleotide sequences with the program tranalign from the EMBOSS v3.0 suite (Rice et al., 2000). Individual genes for both the amino acid and nucleotide datasets were then concatenated into single alignments for phylogenetic analyses. Given the potential for mutational saturation due to the level of divergence between taxa, the third position of each codon from all 13 protein-coding genes was excluded from the nucleotide dataset prior to analysis. Likewise, RNA genes (i.e., *lrRNA*, *srRNA* and tRNAs) were not utilized for phylogenetic analysis due to ambiguity in their alignments (data not shown). Alignments are available from http://www.auburn.edu/~santosr/sequencedatasets.htm or by request from the corresponding author.

Each of the datasets was analyzed using the maximum likelihood (ML) strategy implemented by the program PHYML v2.4.4 (Guindon and Gascuel, 2003). For the amino acid alignment, the mitochondrial fixed-values *R*-matrix (i.e., mtREV) model was utilized while analysis of the nucleotide dataset was conducted under the General Time Reversible model of evolution with a proportion of invariable sites and rate variation among sites (i.e., GTR+I+G) as chosen by the Akaike Information Criterion (AIC) in ModelTest v3.6 (Posada and Crandall, 1998). Support values for this model were: $-\ln L=70{,}750.41$; $K=10$; AIC=141,520.83. The proportion of invariable sites and γ shape distribution parameters of rate variation among sites for the nucleotide dataset were estimated by PHYML v2.4.4. Branch supports in the amino acid and nucleotide ML trees were estimated by bootstrap analysis of 100 replicates.

3. Results and discussion

3.1. Genome composition

The *H. rubra* mitochondrial genome is a circular molecule of 16,065 bp and encodes the 37 genes (13 protein-coding, 22 tRNAs and two rRNAs) typically found in metazoan mitochondrial genomes (Fig. 1). Overlap between genes of up to 6 bp was observed on four occasions (Table 2). Most genes (23 of 37; 62%) are encoded on the positive (+) strand while the

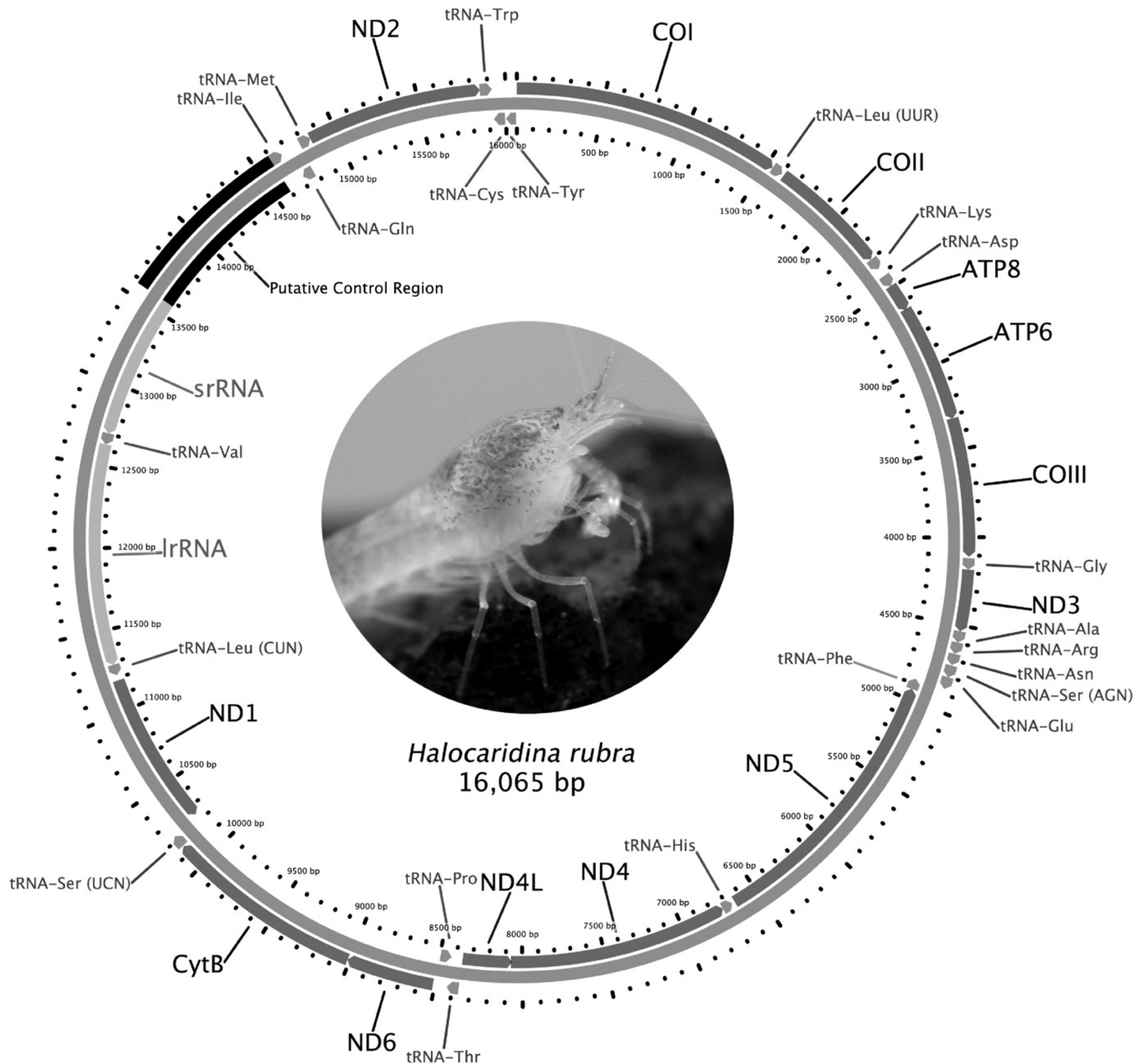

Fig. 1. The mitochondrial genome of *Halocaridina rubra* Holthuis, 1963 (Crustacea: Decapoda: Atyidae). Middle contiguous circle represents genome sequence. Genes encoded on positive (+) or negative (−) strands are depicted in outer and inner circles, respectively. Gene orientation is designated by squared end (5′) and arrow point (3′). Color codes: protein-coding gene = dark gray; tRNA = gray; ribosomal RNA = light gray. Putative control region is presented as occurring on both strands (colored in black).

remaining 14 genes are located on the negative (−) strand (Table 2). A total of 1255 bp of non-coding nucleotides occur over 19 intergenic regions (Table 2), with the largest region (1020 bp) located between the *srRNA* and *tRNA^{Ile}* genes. This region is considered to represent the putative control region of the mitochondrial genome due to its non-coding nature and positional homology among arthropods (e.g., Miller and Austin, 2006).

The overall A+T content of the *H. rubra* mitochondrial genome is 63.2%. While relatively low, this value lies within the range observed for Malacostraca mitochondrial genomes (Table 3). Likewise, the A+T compositions of the 13 protein-coding, 22 tRNAs, two rRNAs, and putative control region fall within the observed range of other Malacostraca (Table 3).

3.2. Gene order

Although rearrangements of protein-coding, tRNA, and rRNA genes have been reported from the mitochondrial genomes of four malacostracan species (i.e., *P. longicarpus*, Hickerson and Cunningham, 2000; *P. trituberculatus*, Yamauchi et al., 2003; *C. destructor*, Miller et al., 2004; *P. gigas*, Miller et al., 2005), the gene organization of the *H. rubra* mitochondrial genome (Fig. 1) is consistent with the proposed

Table 2
Organization of the *Halocaridina rubra* Holthuis, 1963 (Crustacea: Decapoda: Atyidae) mitochondrial genome

Gene	Strand	Begin	End	Length	Start	Stop	Intergenic nucleotides [a]
COI	+	1	1536	1536	CCG	TAA	2
tRNA$^{Leu(UUR)}$	+	1539	1601	63			13
COII	+	1615	2299	685	ATG	T [b]	3
tRNA^{Lys}	+	2303	2370	68			50
tRNA^{Asp}	+	2421	2489	69			0
ATP8	+	2490	2648	159	ATC	TAA	−4
ATP6	+	2645	3319	675	ATA	TAA	−1
COIII	+	3319	4107	789	ATG	TAA	9
tRNA^{Gly}	+	4117	4181	65			0
ND3	+	4182	4533	352	ATT	T [b]	0
tRNA^{Ala}	+	4534	4598	65			0
tRNA^{Arg}	+	4599	4664	66			1
tRNA^{Asn}	+	4666	4730	65			3
tRNA$^{Ser(AGN)}$	+	4734	4802	69			0
tRNA^{Glu}	+	4803	4871	69			−2
tRNA^{Phe}	−	4870	4936	67			0
ND5	−	4937	6646	1710	ATA	TAA	18
tRNA^{His}	−	6665	6729	65			0
ND4	−	6730	8068	1339	ATG	T [b]	−6
ND4L	−	8062	8361	300	ATG	TAA	2
tRNA^{Thr}	+	8364	8430	67			0
tRNA^{Pro}	−	8431	8496	66			14
ND6	+	8511	9013	503	ATA	TA [b]	0
CytB	+	9014	10,150	1137	ATG	TAA	4
tRNA$^{Ser(UCN)}$	+	10,155	10,224	70			19
ND1	−	10,244	11,185	942	ATT	TAA	33
tRNA$^{Leu(CUN)}$	−	11,219	11,284	66			0
lrRNA	−	11,285	12,635	1351			0
tRNA^{Val}	−	12,636	12,704	69			0
srRNA	−	12,705	13,576	872			0
Putative CR	+/−	13,577	14,596	1020			0
tRNA^{Ile}	+	14,597	14,667	71			41
tRNA^{Gln}	−	14,709	14,777	69			8
tRNA^{Met}	+	14,786	14,852	67			0
ND2	+	14,853	15,855	1003	ATT	T [b]	0
tRNA^{Trp}	+	15,856	15,923	68			4
tRNA^{Cys}	−	15,928	15,994	67			6
tRNA^{Tyr}	−	16,001	16,060	60			5

[a] Numbers indicate the number of nucleotides to the start of the adjacent gene. Negative values correspond to overlapping nucleotides between adjacent genes.
[b] Termination codon completed via polyadenylation.

ancestral order represented by the horseshoe crab *Limulus polyphemus* (Lavrov et al., 2000; GenBank Accession No. AF216203) and is syntenic with the remaining malacostracans that have been examined. One exception in gene order between *H. rubra* and *L. polyphemus* is the position of *tRNA*$^{Leu(UUR)}$; however, this arrangement is common among arthropods in general and is hypothesized to be a pleisiomorphic character of Pancrustacea mitochondrial genomes (Boore et al., 1998).

3.3. Protein-coding genes

For the protein-coding genes, nine (i.e., *ATP6*, *ATP8*, *COI–III*, *CytB*, *ND2–3* and *ND6*) are encoded on the + strand while the remaining four (i.e., *ND1*, *ND4L* and *ND4–5*) occur on the − strand (Fig. 1; Table 2). Similar to other arthropods, a bias toward A+T in the first and third codon positions across all genes is evident in *H. rubra* (Table 4). In addition, a compositional bias toward cytosine residues in genes on the + strand and guanine residues in genes on the − strand is also apparent in the mitochondrial genome of *H. rubra* (Table 4). Again, this is a similar to what has been reported from other crustaceans (e.g., Yamauchi et al., 2003; Miller and Austin, 2006). It is hypothesized that this pattern stems from a higher incidence of spontaneous cytosine and adenine deamination on the + strand since it remains in a single-stranded state for a longer duration during DNA replication (Reyes et al., 1998).

One of four different transcription initiation codons (i.e., ATN, where N = A, C, G, or T) is utilized by 12 of the 13 protein-coding genes in the *H. rubra* mitochondrial genome (Table 2). The exception to this is *COI*; based on amino acid and nucleotide alignments of the Malacostraca, *H. rubra* employs CCG (encoding for a proline [P] residue) as the initiation codon for this particular gene. Although ACG has been reported as the *COI* transcription initiation codon from a variety of malacostracans (Hickerson and Cunningham, 2000; Wilson et al., 2000; Yamauchi et al., 2002; Miller et al., 2004; Miller et al., 2005), *H. rubra* represents, to the best of our knowledge, the first example where CCG serves as a *COI* initiation codon in this group.

A majority (8 of 13; 62%) of the protein-coding genes in the *H. rubra* mitochondrial genome possess TAA as their termination codon (Table 2). The remaining five genes (i.e., *COII*, *ND2–4*, and *ND6*) terminate with either a T or TA (Table 2). Truncated termination codons such as these have also been observed in a number of crustaceans (e.g., Yamauchi et al., 2003; Ogoh and Ohmiya, 2004; Segawa and Aotsuka, 2005; Sun et al., 2005; Miller and Austin, 2006), and it is postulated that the complete TAA termination codon is generated via post-transcriptional polyadenylation (Ojala et al., 1981).

3.4. Ribosomal and transfer RNA genes

The mitochondrial rRNA genes of *H. rubra* are encoded on the − strand, with *lrRNA* flanked by *tRNA*$^{Leu(CUN)}$ and *tRNA*^{Val} in the 5′ and 3′ directions, respectively, and *srRNA* between *tRNA*^{Val} and the putative control region (Fig. 1). The location and orientation of these genes is relatively universal across the malacostracan mitochondrial genomes sequenced to date, with an exception being the crayfish *C. destructor* (Miller et al., 2004). As with the protein-coding genes, the lengths and A+T content of the *H. rubra* rRNA genes are consistent with other Malacostraca (Table 3).

Twenty of the 22 tRNA genes were detected by tRNAscan-SE, with cove scores of 17.5–42.3. The remaining two, *tRNA*$^{Ser(AGN)}$ and *tRNA*^{Tyr}, were identified by positional homologue to other malacostracan mitochondrial genomes. In addition, when folded manually, these candidate regions formed tRNA-like secondary structures and possessed appropriate anticodons. Overall, the anticodon usage is consistent with that reported from other Malacostraca. Thus, the mitochondrial genome of *H. rubra* harbors the 22 tRNA genes typical of metazoa mitochondria (Fig. 2).

Table 3
Genomic characteristics of 18 malacostracan mitochondrial genomes acquired from GenBank

Species	GenBank Accession	Genome characteristics		13 protein-coding genes		*lrRNA* gene		*srRNA* gene		22 tRNA genes		Putative CR	
		Length (bp)	A+T (%)	No. of amino acid	A+T (%)	Length (bp)	A+T (%)	Length (bp)	A+T (%)	Length (bp)	A+T (%)	Length (bp)	A+T (%)
Callinectes sapidus	AY363392	16,263	69.1	3708	67.1	1323	71.8	785	70.3	1466	71.6	1435	78.3
Cherax destructor	AY383557	15,895	62.4	3705	60.0	1302	67.9	917	68.3	1436	70.7	977	65.8
Eriocheir sinensis	AY274302	16,354	71.6	3718	69.0	1311	77.4	892	76.6	1477	72.5	896	83.1
Geothelphusa dehaani	AB187570	18,197	74.9	3701	71.5	1315	77.1	821	76.4	1519 [a]	75.8 [a]	514	87.2
Gonodactylus chiragra	DQ191682	16,279	67.5	3712	65.5	1365	70.5	843	68.1	1489	69.2	1368	78.4
Halocaridina rubra	DQ917432	16,065	63.2	3700	60.3	1351	68.2	872	68.8	1471	67.3	1020	78.4
Harpiosquilla harpax	AY699271	15,714	69.8	3726	68.2	1359	75.1	836	70.8	1486	69.2	801	81.1
Ligia oceanica	DQ442914	15,289	60.8	3677	60.1	1234	65.2	850	60.1	1278 [b]	65.0 [b]	737	55.8
Lysiosquillina maculata	DQ191683	16,325	63.8	3709	61.9	1347	69.7	856	67.6	1486	68.5	1319	66.9
Macrobrachium rosenbergii	AY659990	15,772	62.3	3708	60.1	1305	66.0	852	66.0	1449	64.7	931	75.7
Marsupenaeus japonicus	AP006346	15,968	66.5	3712	64.7	1367	70.5	853	67.9	1483	64.0	992	82.5
Pagurus longicarpus	AF150756	– [c]	– [c]	3698	69.6	1303	77.1	789	77.2	1458	74.1	– [c]	– [c]
Panulirus japonicus	AB071201	15,717	64.5	3715	62.6	1355	69.2	855	67.1	1484	68.9	786	70.6
Penaeus monodon	AF217843	15,984	70.6	3716	69.3	1365	74.9	852	71.6	1494	68.0	991	81.5
Portunus trituberculatus	AB093006	16,026	70.2	3715	68.8	1332	73.8	840	70.1	1468	72.0	1104	76.3
Pseudocarcinus gigas	AY562127	15,515	70.5	3784	68.9	1324	74.8	821	73.8	1460	73.2	593	80.3
Squilla empusa	DQ191684	15,828	68.4	3708	66.4	1354	74.9	834	69.9	1490	69.9	936	79.8
Squilla mantis	AY639936	15,994	70.2	3683	68.2	1383	75.9	832	70.6	1463	71.6	862	76.8

[a] 23 tRNAs.
[b] 21 tRNAs.
[c] Incomplete mitochondrial genome sequence (Hickerson and Cunningham, 2000).

Twelve of the 22 tRNA genes in the *H. rubra* mitochondrial genome are arranged in three clusters, each containing three or more tRNAs in a tandem arrangement (Fig. 1). The remaining tRNA genes are found scattered around the genome in singletons or pairs (Fig. 1). The tRNA genes range in size from 60 to 71 bp and a majority (14 of 22; 64%) are encoded on the + strand (Table 2). The secondary structures of most *H. rubra* tRNA genes are also consistent to what has been previously reported from the Malacostraca (Fig. 2). Of note includes the absence of the dihydrouridine (DHU) arm stem from $tRNA^{Ser(AGN)}$, which is characteristic of this gene across the metazoans (e.g., Yamazaki et al., 1997; Miller and Austin, 2006). Along with this, $tRNA^{Tyr}$ of the *H. rubra* mitochondrial genome also lacks a DHU arm stem (Fig. 2). To the best of our knowledge, *H. rubra* represents the first example of the absence of the DHU arm stem from $tRNA^{Tyr}$ in the Malacostraca.

3.5. Phylogenetic analyses

The concatenated alignments of amino acid and nucleotide data from the 13 protein-coding genes contained 3764 and 7530 characters (including gaps), respectively. The proportion of invariable sites and γ shape distribution parameters of rate variation among sites for the nucleotide dataset were estimated by PHYML to be 0.294 and 0.661, respectively. The ML analyses of amino acid and nucleotide data produced identical topologies, depicted in Fig. 3. Values of nodal support were typically congruent between the two trees or higher in the amino acid tree (Fig. 3).

Both the amino acid and nucleotide phylogenies indicate strong (i.e., 100%) bootstrap support for the monophyly of the Isopoda, Stomatopoda and Decapoda (Fig. 3). Within the Stomatopoda and Decapoda, most nodes had moderate to strong support. Exceptions to this include low bootstrap support in both trees for a sister-grouping of Dendrobrachiata and Palinura/Astacidea/Anomura/Brachyura as well as the placement of the Anomura basal to the Brachyura in the nucleotide tree (Fig. 3). The Infraorder Caridea, with *H. rubra* (Atyoidea) as a sister to *M. rosenbergii* (Palaemonoidea), is located at the

Table 4
Base composition (%) of the 13 protein-coding genes from the mitochondrial genome of *Halocaridina rubra* Holthuis, 1963 (Crustacea: Decapoda: Atyidae)

	A	C	G	T
All genes				
1st codon	27.6	19.3	23.6	29.6
2nd codon	17.9	22.7	15.1	44.4
3rd codon	30.2	23.5	13.3	33.0
Total	25.2	21.8	17.3	35.7
Genes encoded on + strand [a]				
1st codon	28.5	23.0	21.9	26.6
2nd codon	18.6	25.5	12.6	43.3
3rd codon	33.0	30.7	8.0	28.3
Total	26.7	26.4	14.2	32.7
Genes encoded on − strand [b]				
1st codon	25.6	10.8	27.3	36.3
2nd codon	16.2	16.4	20.5	46.8
3rd codon	24.0	7.2	25.2	43.6
Total	21.9	11.5	24.3	42.2

[a] *ATP6*, *ATP8*, *COI*, *COII*, *COIII*, *CytB*, *ND2*, *ND3*, and *ND6* genes.
[b] *ND1*, *ND4*, *ND4L*, and *ND5* genes.

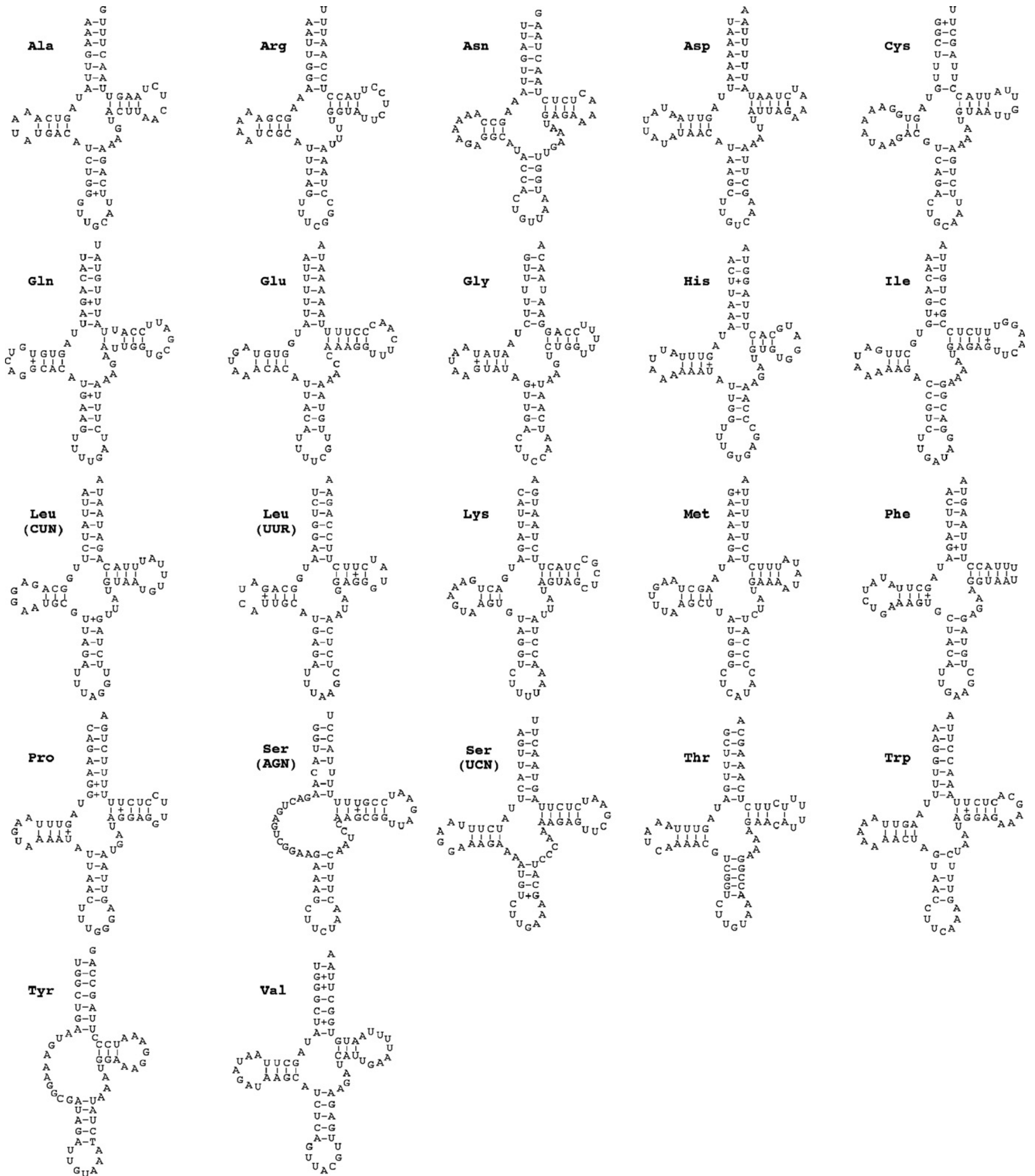

Fig. 2. Proposed secondary structures for the 22 mitochondrial tRNAs of *Halocaridina rubra* Holthuis, 1963 (Crustacea: Decapoda: Atyidae). Watson–Crick and GT bonds are denoted by "−" and "+", respectively.

base of the decapods with moderate (83–85%) bootstrap support.

The placement of the Caridea at the base of the decapods in our phylogeny contrasts with previous reports based on morphological (reviewed by Schram, 2001) and molecular (Porter et al., 2005; Miller and Austin, 2006) data proposing that the Dendrobrachiata occupy this basal position. In their analysis of whole mitochondrial genome data, Miller and Austin (2006) found that the selection of particular taxa for use as an outgroup lead to changes in topology and nodal support

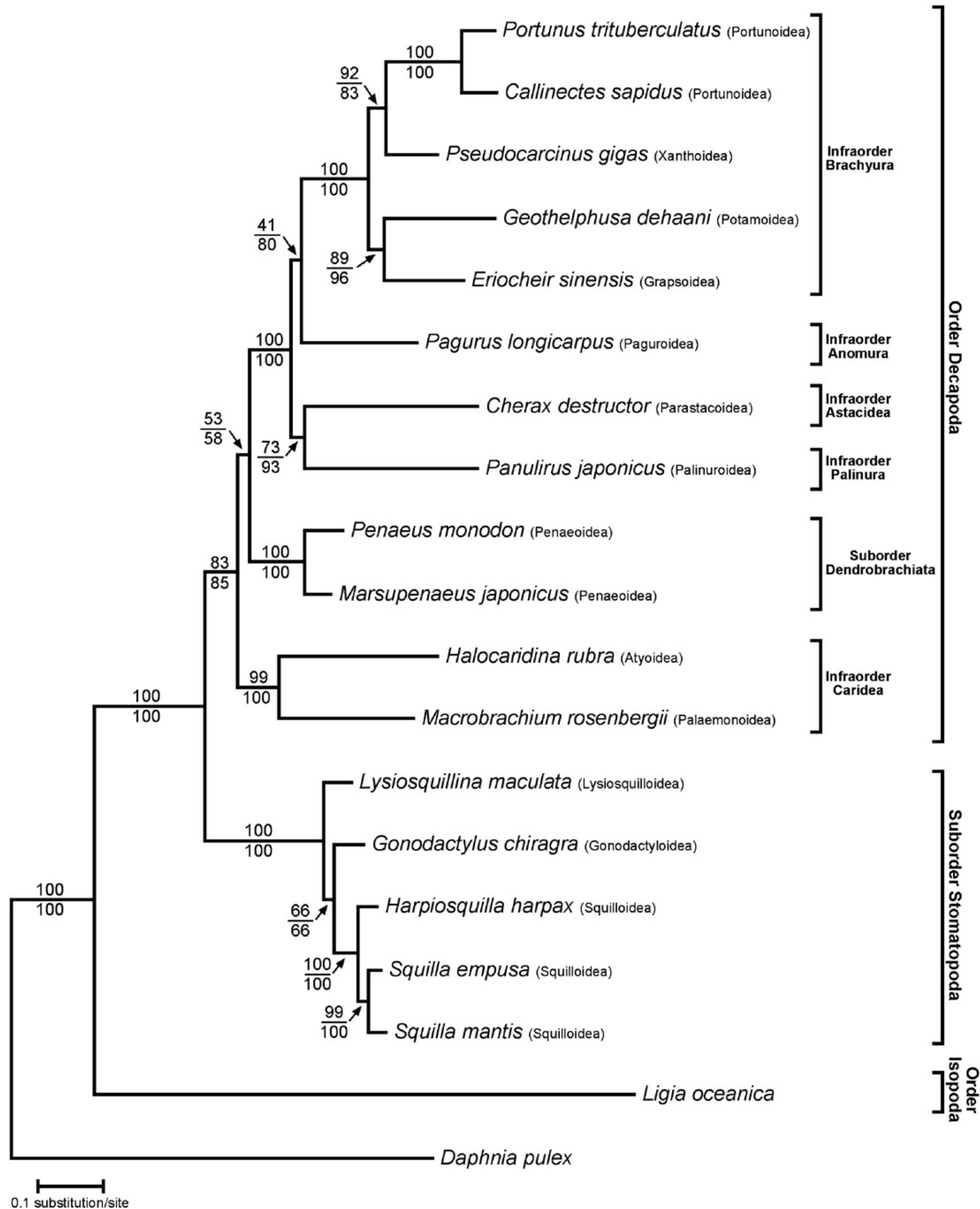

Fig. 3. Proposed phylogeny of the Malacostraca derived from maximum likelihood analyses of nucleotide and amino acid data taken from 19 complete mitochondrial genomes. Presented tree is based on nucleotide sequence data (ln likelihood = −70,784.03) from the 13 protein-coding genes that were utilized (see Section 2.4 for details). The ln likelihood of the amino acid tree = −68,894.25. Values above and below a line represent nodal support from 100 bootstrap replicates of the nucleotide and amino acid data sets, respectively. Superfamily designations are presented in parentheses following species names. The cladoceran *Daphnia pulex* is included as the outgroup.

in parts of their proposed phylogeny for the Decapoda. The relationships that became unstable occurred in the most basal decapod lineages, and in one case, the simultaneous use of maxillopod and non-decapod malacostracans as outgroups leads to a topology suggesting the carideans are at the base of the Decapoda (Miller and Austin, 2006), which is similar to our phylogeny. It has been proposed that non-decapod malacostracans, rather than more evolutionary distant relatives such as the maxillopods and branchiopods, are more appropriate outgroups when exploring the interrelationships of the Decapoda using whole mitochondrial genome data (Miller and Austin, 2006).

To test this hypothesis, additional phylogenies based on the above nucleotide and amino acid datasets were constructed, but utilizing the stomatopod *L. maculata* as an outgroup. In this case, the relationships between the infraorders Palinura/Astacidea/Anomura/Brachyura were identical to those depicted in Fig. 3 and to what has been previously reported

(amino acid tree [Fig. 4] of Miller and Austin, 2006), suggesting stability in this portion of the phylogeny. However, both phylogenies also suggested 1) the Caridea and Dendrobrachiata to be sister-groups, albeitly with low bootstrap support from both nucleotide and amino acid data (37% and 69%, respectively), and; 2) the Caridea/Dendrobrachiata to be sister to the Palinura/Astacidea/Anomura/Brachyura (data not shown). Thus, our analyses support the conclusion of Miller and Austin (2006), namely that outgroup selection influences phylogenetic reconstruction and placement of the basal decapod lineages (i.e., Caridea and Dendrobrachiata) when whole mitochondrial genome data are utilized. Additional mitochondrial genomes from basal decapods such as the Stenopodidea (Porter et al., 2005) should contribute to resolving this situation and lead to a more complete consensus on the phylogeny of the Decapoda.

Acknowledgements

We would like to thank Mike N. Yamamoto and Tom Iwai (State of Hawai'i Division of Aquatic Resources — Oahu) for assistance in acquiring samples and the photograph of *Halocaridina rubra* used in Fig. 1 as well as two anonymous reviewers for valuable suggestions that improved this work. This is contribution #19 to the Auburn University (AU) Marine Biology Program.

References

Abascal, F., Posada, D., Zardoya, R., 2007. MtArt: a new model of amino acid replacement for Arthropoda. Mol. Biol. Evol. 24, 1–5.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Bailey-Brock, J.H., Brock, R.E., 1993. Feeding, reproduction, and sense organs of the Hawaiian anchialine shrimp *Halocaridina rubra*. Pac. Sci. 47, 338–355.

Baker, A.M., Hurwood, D.A., Krogh, M., Hughes, J.M., 2004. Mitochondrial DNA signatures of restricted gene flow within divergent lineages of an atyid shrimp (*Paratya australiensis*). Heredity 93, 196–207.

Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13, 729–744.

Birky, C.W., 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proc. Natl. Acad. Sci. U. S. A. 92, 11331–11338.

Boore, J.L., Lavrov, D.V., Brown, W.M., 1998. Gene translocation links insects and crustaceans. Nature 392, 667–668.

Brown, W.M., George, M., Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. Genetics 76, 1967–1971.

Bruce, A.J., 1992. *Pycnisia raptor*, a new genus and species of predatory troglobic shrimp (Crustacea: Decapoda: Atyidae) from Northern Australia. Invertebr. Taxon 6, 553–566.

Chenoweth, S.F., Hughes, J.M., 2003. Speciation and phylogeography in *Caridina indistincta*, a complex of freshwater shrimps from Australian heathland streams. Mar. Freshw. Res. 54, 807–812.

Dimijian, G.G., 2000. Evolving together: the biology of symbiosis, Part 2. BUMC Proc. 13, 381–390.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294–299.

Fukatsu, T., 1999. Acetone preservation: a practical technique for molecular analysis. Mol. Ecol. 8, 1935–1945.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Hedges, S.B., Chen, H., Kumar, S., Wang, D. Y.-C., Thompson, A.S., Watanabe, H., 2001. A genomic timescale for the origin of eukaryotes. BMC Evol. Biol. 1, 4.

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

Holthuis, L.B., 1963. On red coloured shrimps (Decapoda, Caridea) from tropical land-locked saltwater pools. Zool. Meded. 16, 261–279.

Hurwood, D.A., Hughes, J.M., 2001. Nested clade analysis of the freshwater shrimp, *Caridina zebra* (Decapod, Atyidae), from north-eastern Australia. Mol. Ecol. 10, 113–125.

Hurwood, D.A., Hughes, J.M., Bunn, S.E., Cleary, C., 2003. Population structure in the freshwater shrimp (*Paratya australiensis*) inferred from allozymes and mitochondrial DNA. Heredity 90, 64–70.

Lavrov, D.V., Boore, J.L., Brown, W.M., 2000. The complete mitochondrial DNA sequence of the horseshoe crab *Limulus polyphemus*. Mol. Biol. Evol. 17, 813–824.

Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.

Miller, A.D., Austin, C.M., 2006. The complete mitochondrial genome of the mantid shrimp *Harpiosquilla harpax*, and a phylogenetic investigation of the Decapoda using mitochondrial sequences. Mol. Phylogenet. Evol. 38, 565–574.

Miller, A.D., Nguyen, T.T., Burridge, C.P., Austin, C.M., 2004. Complete mitochondrial DNA sequence of the Australian freshwater crayfish, *Cherax destructor* (Crustacea: Decapoda: Parastacidae): a novel gene order revealed. Gene 331, 65–72.

Miller, A.D., Murphy, N.P., Burridge, C.P., Austin, C.M., 2005. Complete mitochondrial DNA sequences of the decapod crustaceans *Pseudocarcinus gigas* (Menippidae) and *Macrobrachium rosenbergii* (Palaemonidae). Mar. Biotechnol. 7, 339–349.

Ogoh, K., Ohmiya, Y., 2004. Complete mitochondrial DNA sequence of the sea-firefly, *Vargula hilgendorfii* (Crustacea, Ostracada) with duplicate control regions. Gene 327, 131–139.

Ojala, D., Montoya, J., Attardi, G., 1981. tRNA punctuation model of RNA processing in human mitochondria. Nature 290, 470–474.

Page, T.J., Baker, A.M., Cook, B.D., Hughes, J.M., 2005. Historical transoceanic dispersal of a freshwater shrimp: the colonization of the South Pacific by the genus *Paratya* (Atyidae). J. Biogeogr. 32, 581–593.

Page, T.J., von Rintelen, K., Hughes, J.M., in press. An island in the stream: Australia's place in the cosmopolitan world of Indo-West Pacific freshwater shrimp (Decapoda: Atyidae: *Caridina*). Mol. Phylogenet. Evol. doi:10.1016/j.ympev.2006.08.007.

Porter, M.L., Pérez-Losada, M., Crandall, K.A., 2005. Model-based multi-locus estimation of decapod phylogeny and divergence times. Mol. Phylogenet. Evol. 37, 355–369.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Pringle, C.M., 1996. Atyid shrimps (Decapoda: Atyidae) influence the spatial heterogeneity of algal communities over different scales in tropical montane streams, Puerto Rico. Freshw. Biol. 35, 125–140.

Pringle, C.M., Blake, G.A., 1994. Quantitative effects of atyid shrimp (Decapoda, Atyidae) on the depositional environment in a tropical stream — use of electricity for experimental exclusion. Can. J. Fish. Aquat. Sci. 51, 1443–1450.

Pringle, C.M., Hamazaki, T., 1998. The role of omnivory in a neotropical stream: separating diurnal and nocturnal effects. Ecology 79, 269–280.

Pringle, C.M., Blake, G.A., Covich, A.P., Buzby, K.M., Finley, A., 1993. Effects of omnivorous shrimp in a montane tropical stream — sediment removal, disturbance of sessile invertebrates and enhancement of understory algal biomass. Oecologia 93, 1–11.

Pringle, C.M., Hemphill, N., McDowell, W.H., Bednarek, A., March, J.G., 1999. Linking species and ecosystems: different biotic assemblages cause interstream differences in organic matter. Ecology 80, 1860–1872.

Reyes, A., Gissi, C., Pesole, G., Saccone, C., 1998. Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. Mol. Biol. Evol. 15, 957–966.

Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 16, 276–277.

Rutherford, K., et al., 2000. Artemis: sequence visualisation and annotation. Bioinformatics 16, 944–945.

Santos, S.R., 2006. Patterns of genetic connectivity among anchialine habitats: a case study of the endemic Hawaiian shrimp *Halocaridina rubra* on the Island of Hawaii. Mol. Ecol. 15, 2699–2718.

Schram, F.R., 2001. Phylogeny of decapods: moving towards a consensus. Hydrobiologia 449, 1–20.

Segawa, R.D., Aotsuka, T., 2005. The mitochondrial genome of the Japanese freshwater crab, *Geothelphusa dehaani* (Crustacea: Brachyura): evidence for its evolution via gene duplication. Gene 355, 28–39.

Stock, J.H., 1986. The concept “anchialine” reconsidered. Stygologia 2, 1–2.

Stothard, P., Wishart, D.S., 2005. Circular genome visualization and exploration using CGView. Bioinformatics 21, 537–539.

Sun, H., Zhou, K., Song, D., 2005. Mitochondrial genome of the Chinese mitten crab *Eriocheir japonica sinenesis* (Brachyura: Thoracotremata: Grapsoidea) reveals a novel gene order and two target regions of gene rearrangements. Gene 349, 207–217.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignments aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.

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? Mol. Biol. Evol. 17, 863–874.

Yamauchi, M.M., Miya, M.U., Nishida, M., 2002. Complete mitochondrial DNA sequence of the Japanese spiny lobster, *Panulirus japonicus* (Crustacea: Decapoda). Gene 295, 89–96.

Yamauchi, M.M., Miya, M.U., Nishida, M., 2003. Complete mitochondrial DNA sequence of the swimming crab, *Portunus trituberculatus* (Crustacea: Decapoda: Brachyura). Gene 311, 129–135.

Yamauchi, M.M., Miya, M.U., Machida, R.J., Nishida, M., 2004. PCR-based approach for sequencing mitochondrial genomes of decapod crustaceans, with a practical example from the kuruma prawn (*Marsupenaes japonicus*). Mar. Biotechnol. 6, 419–429.

Yamazaki, N., et al., 1997. Evolution of pulmonate gastropod mitochondrial genomes: comparisons of gene organizations of *Euhadra*, *Cepaea*, and *Albinaria* and implications of unusual tRNA secondary structures. Genetics 145, 749–758.