MARINE BIOTECHNOLOGY

Complete Mitochondrial DNA Sequences of the Decapod Crustaceans *Pseudocarcinus gigas* (Menippidae) and *Macrobrachium rosenbergii* (Palaemonidae)

Adam D. Miller,[1] Nicholas P. Murphy,[2] Christopher P. Burridge,[1] Christopher M. Austin[1]

[1]School of Ecology and Environment, Deakin University, P.O. Box 423, Warrnambool, Victoria 3280, Australia
[2]Centre for Evolutionary Biology and Biodiversity, Department of Environmental Biology, Adelaide University, Adelaide, South Australia 5005, Australia

Received: 30 June 2004 / Accepted: 4 October 2004 / Online publication: 17 March 2005

Abstract

The complete mitochondrial DNA sequence was determined for the Australian giant crab *Pseudocarcinns gigas* (Crustacea: Decapoda: Menippidae) and the giant freshwater shrimp *Macrobrachium rosenbergii* (Crustacea: Decapoda: Palaemonidae). The *Pse. gigas* and *M. rosenbergii* mitochondrial genomes are circular molecules, 15,515 and 15,772 bp in length, respectively, and have the same gene composition as found in other metazoans. The gene arrangement of *M. rosenbergii* corresponds with that of the presumed ancestral arthropod gene order, represented by *Limulus polyphemus*, except for the position of the $tRNA^{Leu(UUR)}$ gene. The *Pse. gigas* gene arrangement corresponds exactly with that reported for another brachyuran, *Portunus trituberculatus*, and differs from the *M. rosenbergii* gene order by only the position of the $tRNA^{His}$ gene. Given the relative positions of intergenic nonoding nucleotides, the "duplication/random loss" model appears to be the most plausible mechanism for the translocation of this gene. These data represent the first caridean and only the second brachyuran complete mtDNA sequences, and a source of information that will facilitate surveys of intraspecific variation within these commercially important decapod species.

Key words: Brachyura — Caridea — mitochondrial genome — gene translocation — duplication / random loss

Correspondence to: Christopher M. Austin; E-mail: cherax@deakin.edu.au

Introduction

The mitochondrial genome, present in almost all eukaryotic cells, contains genetic information that has greatly facilitated systematic and population genetic research over the past 2 decades. Characteristics such as a relatively rapid mutation rate, maternal inheritance, and a presumed lack of intermolecular recombination have resulted in its extensive use in investigations of population structure and phylogenetic relationships at different taxonomic levels (Avise, 1994; Avise, 2000).

To date approximately 460 eukaryote complete mitochondrial DNA sequences and corresponding gene orders have been determined, with approximately 75% representing vertebrates. By comparison, crustaceans, the most morphologically diverse animal life form (Martin and Davis, 2001), are represented by only 15 complete mitochondrial sequences: 3 branchiopods (Valverde et al., 1994; Crease, 1999; Umetsu et al., 2002); one remipede (Lavrov et al., 2004), one cephalocarid (Lavrov et al., 2004), 3 maxillopods (Machida et al., 2002; Lavrov et al., 2004); one ostracod (Ogoh and Ohmiya, 2004), 5 decapod malacostracans (Hickerson and Cunningham, 2000; Wilson et al., 2000; Yamauchi et al., 2002, 2003; Miller et al., 2004); and a single member of the dubiously placed Pentastomida (Lavrov et al., 2004).

The decapods are an extremely diverse group of crustaceans with many species of commercial importance, especially the palinurid and nephropid lobsters, penaeoid shrimps, and portunid and xanthoid crabs. The Australian giant crab *Pseudocarcinus gigas* (Crustacea: Decapoda: Menippidae) is distributed throughout the southern oceanic waters of Australia and is the largest known true crab, with males reaching up to 13.6 kg and females 6 kg.

Although *Pse. gigas* has been targeted commercially in a trap fishery since 1990, biological and genetic information essential to the sustainable management of the resource, such as knowledge of population structure, is lacking. Preliminary surveys of four mitochondrial gene regions failed to detect nucleotide variation suitable for population genetic analyses (A.D. Miller and N.P. Murphy, unpublished results). Because of the current lack of available mitochondrial sequence data, these surveys were restricted to relatively conserved gene regions and thus the use of universal primers. In order to overcome these limitations, information from more rapidly evolving gene regions is required.

Availability of mtDNA information is similarly limited for another extremely important commercial decapod crustacean, the giant freshwater shrimp *Macrobrachium rosenbergii* (Crustacea: Decapoda: Palaemonidae). *Macrobrachium rosenbergii* is one of approximately 200 species that constitute the genus *Macrobrachium*. Members of this genus are highly diverse, with much uncertainty surrounding their systematic relationships (Murphy and Austin, 2002; de Bruyn et al., 2004). The geographic distribution of the genus encompasses circumtropical marine, estuarine, and fresh waters of all continents except Europe. *M. rosenbergii*, like *Pse. gigas*, is commercially important, particularly in developing regions such as the Indian subcontinent, and Southeast Asia, where it is a major species for aquaculture and indigenous fisheries. There is a clear need for population research on this species, as evidence suggests that many natural populations are being heavily depleted and the species widely translocated (New and Valenti, 2000).

In this study we report the complete nucleotide sequence and gene arrangement of the mitochondrial genome for *Pse. gigas* and *M. rosenbergii*. These data represent the first caridean and only the second brachyuran complete mtDNA sequences. The molecular descriptions of the genomes in this study do not indicate major discrepancies from those already described for other decapod crustaceans. Nevertheless, the data generated are of great importance for taxonomic and population genetic research, which provides critical information for effective management of genetic resources within the species. In a broader sense the data generated will also be of importance for genomic and phylogenetic research, given the current uncertainties associated with deep relationships within the Crustacea and Arthropoda, mitochondrial gene rearrangement mechanisms, homoplasy of gene order, and consequently the phylogenetic utility of mitochondrial gene order (Boore, 1999; Curole and Kocher, 1999).

Materials and Methods

Specimens, DNA Extraction, and Determination of Partial Sequences. *Pseudocarcinus gigas* ethanol-preserved muscle tissue came from an area off the coast of Portland in southeast Australia (38.34°S 20, 141.60°E). Whole *M. rosenbergii* specimens were obtained from a fish market in eastern Java, Indonesia. The exact coordinates of their collection site are unknown. Mitochondrion-enriched DNA extracts were obtained from muscle tissue for both species following Tamura and Aotsuka (1988).

Partial sequences for the *Cyt b* and *COI* mitochondrial genes of *Pse. gigas* were amplified by polymerase polymerase chain reaction (PCR) using the following primer pairs: *cyt b*.10862.F with *cyt b*.11317.R, and COIA with COIF (Palumbi and Benzie, 1991). *Cyt b* primer nomenclature was derived from corresponding localities within the *Penaeus monodon* mitochondrial DNA sequence (GenBank accession number NC_002184), and primer sequences are displayed in Table 1. For *M. rosenbergii* partial sequences were initially generated for the *IrRNA* and *COI* genes using the primers pairs 1471 and 1472 (Crandall et al., 1995) and COIA and COIF (Palumbi and Benzie, 1991), respectively.

PCR was performed using *Taq* DNA polymerase (Invitrogen) following the supplier's instructions. PCR products were purified using a QIAquick PCR purification kit (QIAGEN) and directly sequenced (see below).

Long PCR

Pseudocarcinus gigas. PCR primers designed from the sequence data obtained above were used to amplify the entire *Pse. gigas* mitochondrial genome in 2 large fragments, approximately 6.7 and 8.8 kb in size. The primer pairs were *P.gigas*.9598.F with *P.gigas*.830.R, and *P.gigas*.799.F with *P.gigas*.9233.R. PCR employed High Fidelity Platinum *Taq* DNA Polymerase (Invitrogen), following the supplier's instructions. The 8.8-kb PCR product was subjected to nested PCR using the following internal primer pairs: *ND4*.7468.F with *cyt b*.9386.R, *COIII*.3260.F with *ND4*.7697.R, and *COI*. 1038.F with *COIII*.3444.R, with PCR conditions as described above. These overlapping PCR fragments were approximately 1.8, 4.4, and 2.4 kb, respectively. Primer nomenclature given above corresponds to the relative positions in the *Pse. gigas* mitochondrial genome.

Macrobrachium rosenbergii. Using species-specific primers designed from the data obtained above, and another primer designed from a conserved

Table 1. Primers and Corresponding Sequences

Primer Name	*Primer Sequence (5′–3′)*
cyt b. 10862.F	TTA CCT TGA GGA CAA ATA TCA T
cyt b. 11317.R	CAC CTC CTA ATT TAT TAG GAA
P.gigas.9598.F	AGC CGC GGC TAG AAT AGT CC
P.gigas.830.R	GCC AAT ATA GCG TAA ATT ATA CCT AAG GTC CC
P.gigas.799.F	GGG ACC TTA GGT ATA ATT TAC GCT ATA TTG GC
P.gigas.9233.R	GAG CTA CTC TGG AGA AAG C
ND4.7468.F	ACA TGA GCT TTH GGT AAT CA
cyt b.9386.R	GAA TAT GGG CAG GGG TGA C
COIII.3260.F	CCC AAT CAC ACG GAC ATC ATC CTT AC
ND4.7697.R	CGC ATT CAG GCT GGT GTT TAG ATG TTG
COI.1038.F	CAC TGT AGG TGG ACT AAC
COIII.3444.R	AA TAT CTC GTC ATC ATT G
MR.11371.F	CAA CAT CGA GGT CGC AAA C
MR.1181.R	CAG TGA GCG ATT CCT GCG AAG ATG CCG
MR.1154.F	CGG CAT CTT CGC AGG AAT CGC TCA CTG
MR.8082.R	CTT GCT GCT TGT GAG GGG
MR.7362.F	ACA TGA GCT TTT GGT AAT CA
MR.11389.R	GTT TGC GAC CTC GAT GTT G

region of the ND4 gene shared by a variety of decapod crustaceans, the entire mitochondrial genome was amplified in three fragments, approximately 5.5, 7.0, and 4.0 kb in length. The primer pairs were MR.11371.F with MR.1181.R, MR.1154.F with MR.8082.R, and MR.7362.F with MR.11389.R. Primer nomenclature corresponds to the position within the *M. rosenbergii* mtDNA molecule.

Cloning, Sequencing, and Gene Identification. The *Pse. gigas* 6.7-kb and 4.4-kb PCR products and the three *M. rosenbergii* PCR products were gel purified and ligated into pCR® XL plasmid vector using the TOPO XL cloning kit (Invitrogen). DNA sequence data from both strands were generated from single clones using the primer walking approach (Yamauchi et al., 2003). The remaining *Pse. gigas* 1.8-kb and 2.4-kb PCR products were purified using a QIAquick PCR purification kit and sequenced directly.

Automated sequencing was performed with ABI PRISM BigDye terminator chemistry, version 3, and analysed on an ABI 3700 automated sequencer. Chromatograms were visually inspected using the computer software EditView 1.0.1 (PerkinElmer), and DNA sequences were aligned using SeqPup (Gilbert, 1997).

Protein-coding and ribosomal RNA gene sequences were initially identified using BLAST searches on GenBank, and subsequently by alignment with *Portunus trituberculatus* (GenBank accession number NC_005037) and *Pen. monodon* mtDNA sequences. Protein-coding genes were further aligned with *Por. trituberculatus* and *Pen. monodon* amino acid sequences. Amino acid sequences of *Pse. gigas* and *M. rosenbergii protein-coding* genes were inferred using the *Drosophila* translation code. The majority of the transfer RNA genes were identified using tRNAscan-SE 1.21 (Lowe and Eddy, 1997), employing the default search mode, specifying mitochondrial or chloroplast DNA as the source, and using the invertebrate mitochondrial genetic code for tRNA structure prediction. Remaining tRNA genes were identified by inspecting sequences for tRNA-like secondary structures and anticodons. The resulting *Pse. gigas* and *M. rosenbergii* sequences were deposited in GenBank under accession numbers AY562127 and AY659990, respectively.

Results and Discussion

Genome Composition. The mitochondrial genomes of *Pse. gigas* and *M. rosenbergii* are circular molecules 15,515 bp and 15,772 bp in length, respectively, and both contain the typical gene content found in other metazoans: 13 protein-coding, 2 rRNA, and 22 tRNA genes (Figure 1; Table 2). Two and five gene pairs were found overlapping by up to 7 bp in the *Pse. gigas* and *M. rosenbergii* mitochondrial genomes, respectively (Table 2), a characteristic that has been reported for other animal mtDNAs (Wolstenholme, 1992). The majority-strand (α) of both *Pse. gigas* and *M. rosenbergii* encodes 23 genes, while the minority-strand (β) encodes 14 genes (Table 2). The nucleotide composition of the *Pse. gigas* β-strand is 5,501 A (35.5%), 5,433 T (35.0%), 1,676 C (10.8%), and 2,905 G (18.7%). The nucleotide composition of the *M. rosenbergii* β-strand is 4,480 A (26.5%), 5,646 T (35.8), 2,113 C (13.4%), and 3,833 G (24.3%). Gene lengths md A+T base compositions of the *Pse. gigas* and *M. rosenbergii* β-strands, protein-coding, rRNA, and tRNA genes, as well as the

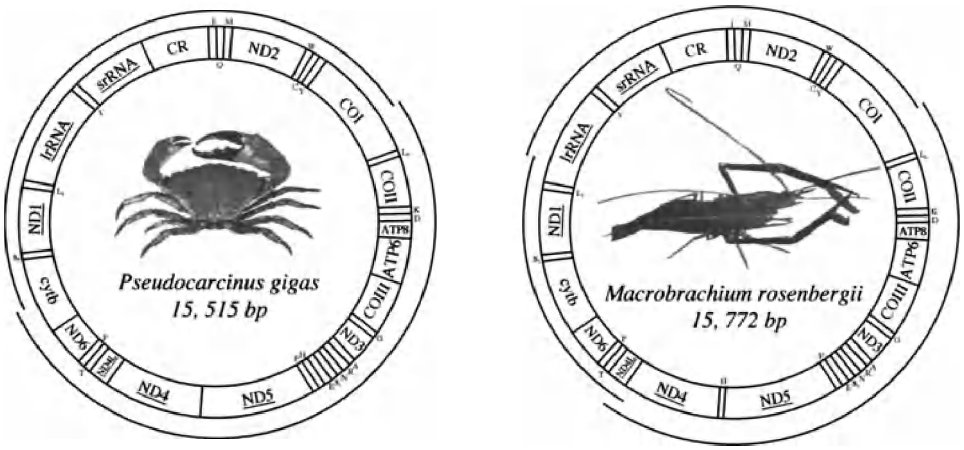

Fig. 1. Gene map of the *Pseudocarcinus gigas* and *Macrobrachium rosenbergii* mitochondrial genomes. COI-III indicates cytochrome *c* oxidase subunits 1–3; cyt b, cytochrome *b*; ATP6–8, ATPase subunits 6 and 8; ND1–6/4L, NADH dehydrogenise subunits 1–6/4L. Transfer RNA genes are designated by single-letter amino acid codes except those encoding leucine and serine, which are labeled L_1 ($tRNA^{Leu(CUN)}$), L_2 ($tRNA^{Leu(UUR)}$), S_1 ($tRNA^{Ser(AGN)}$) and S_2 ($tRNA^{Ser(UGN)}$). Protein-coding and RNA genes are transcribed in a clockwise direction except those indicated by underlining, and tRNAs depicted on the inside of the molecule. Arcs on the outside of the gene maps denote the regions amplified by long PCR.

putative control regions, are displayed in Tables 2 and 3. Chi-square tests indicated that A+T compositions of *Pse.gigas* and *M. rosenbergii* differed significantly ($P < 0.05$) from various decapod (Table 3).

A total of 745 noncoding nucleotides are evident in the *Pse. gigas* mitochondrial genome, with 152 bp at 17 intergenic regions and a 593-bp noncoding region (Table 2). For *M. rosenbergii* we found 1,057 bp of noncoding nucleotides spread over 13 intergenic regions including one large noncoding region 931 bp in length. In both cases we propose that the large noncoding region found represents the putative control region on the basic of its relative position between the *srRNA* and $tRNA^{Ile}$, typical of arthropods, and sequence characteristics (A+T-rich, noncoding). The *Pse. gigas* putative control region is notably shorter than that reported for other decapod crustaceans (Table 3); however, control region length variations are evident among crustaceans (*Triops cancriformis* = 467 bp; *Artemia franciscana* = 1,822 bp) (Valverde et al., 1994; Umetsu et al., 2002). In contrast, no substantial length differences of *Pse. gigas* or *M. rosenbergii* mitochondrial genes were observed when compared with those reported for other decapod crustaceans (Table 3).

Gene Order. The order and transcriptional orientation of *M. rosenbergii* mitochondrial genes is the same as that displayed by the putative ancestral arthropod gene order depicted by *Limulus polyphemus* (GenBank accession number NC_003057) with the exception of the $tRNA^{Leu(UUR)}$ gene. The position of the *M. rosenbergii* $tRNA^{Leu(UUR)}$ is consistent with that reported for other crustacean species (Figure 1). The *Pse. gigas* mitochondrial gene order is identical to that of *M. rosenbergii* with the exception of the relative position of $tRNA^{His}$. Typically this gene lies between *ND4* and *ND5*; however, as also displayed by another brachyuran, *Por. trituberculatus* (Yamauchi et al, 2003), it is located between the $tRNA^{Glu}$ and the $tRNA^{Phe}$ genes. We speculate that the relative position of $tRNA^{His}$ is a potential genomic synapomorphy for the Brachyura.

Although the mechanisms responsible for mtDNA gene rearrangements are still uncertain, one of the most widely documented and accepted mechanisms is the duplication / random loss model (Levinson and Gutman, 1987; Moritz and Brown, 1987; Macey et al., 1997, Boore, 2000). This involves the tandem duplication of gene regions, most widely considered a result of slipped-strand mispairing during replication, followed by the deletion of one of the duplicated gene regions. This is a plausible mechanism for the translocation of the *Pse. gigas* and *Por. trituberculatus* $tRNA^{His}$ gene, most likely representing only a single-step rearrangement process. In further support of the duplication / random loss mechanism, incomplete gene deletions are evident in the *Pse. gigas* mitochondrial genome, with 21 and 46 unassignable nucleotides displayed at the $tRNA^{Glu}$–$tRNA^{His}$ and *ND4–ND5* gene boundaries, respectively (Table 2). Although these intergenic sequences bear no homology to candidate ancestral genes, the homology may have been lost due to mutation events as a consequence of freedom from selective constraints. Yamauchi et al. (2003) give a detailed representation of the possible duplication / random loss model and translocation of the $tRNA^{His}$ gene.

Table 2. Mitochondrial Genome Profiles of *Pseudocarcinus gigas* and *Macrobrachium rosenbergii*

Pseudocarcinus gigas						*Macrobrachium rosenbergii*					
			Codon						Codon		
Feature	*Position*[a] *numbers*	*Size (bp)*	*Start*	*Stop*	*Intergenic*[b] *nucleotides*	*Feature*	*Position numbers*[a]	*Size (bp)*	*Start*	*Stop*	*Intergenic nucleotides*[b]
COI	1–1530	1530	ATG	T[c]		*COI*	1–1535	1535	ACG	TA[c]	
tRNA$^{Leu(UUR)}$	1535–1598	64			4	*tRNA*$^{Leu(UUR)}$	1536–1599	64			0
COII	1605–2291	687	ATG	TAA	6	*COII*	1602–2289	688	AAT	T[c]	2
tRNA^{Lys}	2294–2359	66			2	*tRNA*^{Lys}	2290–2357	68			0
tRNA^{Asp}	2361–2426	66			1	*tRNA*^{Asp}	2358–2423	66			0
ATP8	2427–2585	159	ATG	TAG	0	*ATP8*	2424–2582	159	ATC	TAA	0
ATP6	2579–3252	674	ATT	TA[c]	−7	*ATP6*	2576–3248	673	ATG	T[c]	−7
COIII	3253–4046	794	ATG	T[c]	0	*COIII*	3249–4043	795	ACT	TAA	0
tRNA^{Gly}	4047–4111	65			0	*tRNA*^{Gly}	4050–4114	65			6
ND3	4112–4465	354	ATT	TAA	0	*ND3*	4115–4468	354	ATG	TAA	0
tRNA^{Ala}	4468–4533	66			2	*tRNA*^{Ala}	4469–4531	63			0
tRNA^{Arg}	4540–4604	65			6	*tRNA*^{Arg}	4531–4592	62			−1
tRNA^{Asn}	4605–4672	68			0	*tRNA*^{Asn}	4594–4658	65			1
tRNA$^{Ser(AGN)}$	4697–4758	62			24	*tRNA*$^{Ser(AGN)}$	4659–4725	67			0
tRNA^{Glu}	4759–4829	71			0	*tRNA*^{Glu}	4728–4796	69			2
tRNA^{His}	(4851–4914)	64			21	*tRNA*^{Phe}	(4796–4861)	66			−1
tRNA^{Phe}	(4915–4979)	65			0	*ND5*	(4862–6568)	1707	ATG	TAA	0
ND5	(4987–6711)	1725	ATG	TAG	7	*tRNA*^{His}	(6587–6650)	64			18
ND4	(6759–8093)	1335	ATG	TAG	46	*ND4*	(6653–7987)	1335	ATG	TAA	2
ND4L	(8087–8389)	303	ATG	TAA	−7	*ND4L*	(7981–8280)	300	ATG	TAA	−7
tRNA^{Thr}	8392–8455	64			2	*tRNA*^{Thr}	8283–8347	65			2
tRNA^{Pro}	(8456–8520)	65			0	*tRNA*^{Pro}	(8347–8412)	66			−1
ND6	8523–9028	506	ATG	TA[c]	2	*ND6*	8414–8928	515	ATC	TA[c]	1
Cyt b	9029–10164	1136	ATG	TA[c]	0	*Cyt b*	8929–10060	1132	ATG	T[c]	0
tRNA$^{Ser(UCN)}$	10165–10232	68			0	*tRNA*$^{Ser(UCN)}$	10061–10129	69			0
ND1	(10245–11210)	966	AGA	TAA	12	*ND1*	(10150–11088)	939	ATA	TAG	20
tRNA$^{Leu(CUN)}$	(11219–11286)	68			8	*tRNA*$^{Leu(CUN)}$	(11121–11184)	64			32
lrRNA	(11287–12610)	1324			0	*lrRNA*	(11185–12489)	1305			0
tRNA^{val}	(12611–12683)	73			0	*tRNA*^{val}	(12490–12556)	67			0
srRNA	(12684–13504)	821			0	*srRNA*	(12557–13408)	852			0
CR	13505–14097	593			0	CR	13409–14339	931			0
tRNA^{Ile}	14098–14163	66			0	*tRNA*^{Ile}	14340–14406	67			0
tRNA^{Gln}	(14164–14232)	69			0	*tRNA*^{Gln}	(14435–14502)	68			28
tRNA^{Met}	14237–14302	66			4	*tRNA*^{Met}	14514–14581	68			11
ND2	14303–15311	1009	ATG	T[c]	0	*ND2*	14582–15575	994	ATT	T[c]	0
tRNA^{Trp}	15312–15379	68			0	*tRNA*^{Trp}	15576–15644	69			0
tRNA^{Cys}	(15384–15448)	65			4	*tRNA*^{Cys}	(15646–15709)	64			1
tRNA^{Tyr}	(15450–15515)	66			1	*tRNA*^{Tyr}	(15710–15772)	63			0

[a]Parentheses denote that the gene is encoded on the β-strand.
[b]Numbers correspond to the nucleotides separating genes. Negative numbers indicate overlapping nucleotides between adjacent genes.
[c]Truncated termination codon.

Table 3. Genomic Characteristics of Decapod Crustacean mtDNAs[a]

Species	GenBank accession number	β-strand		13 Protein-coding		lrRNA gene		srRNA gene		22 tRNA genes		Putative control region	
		Length (bp)	A + T (%)	No. of amino acids	A + T (%)	Length (bp)	A + T (%)	Length (bp)	A + T (%)	Length (bp)	A + T (%)	Length (bp)	A + T (%)
Pseudocarcinus gigas	AY562127	15,515	70.5†	3734	68.9†	1324	74.8†	821	73.8†	1460	73.2†	593	80.3†
Macrobrachium rosenbergii	AY659990	15,772	62.3*	3708	60.1*	1305	66.0*	852	66.0*	1449	64.7*	931	75.7*
Portunus trituberculatus	NC_005037	16,026	70.2†	3715	68.8†	1332	73.8†	840	70.1	1468	72.0†	1104	76.3
Cherax destructor	AY383557	15,895	62.4*	3705	60.0*	1302	67.9*	917	68.3*	1436	70.7†	977	65.8*†
Penaeus monodon	NC_002184	15,984	70.6†	3716	69.3†	1365	74.9†	852	71.6†	1494	68.0*	991	81.5†
Panulirus japonicus	NC_004251	15,717	64.5*†	3715	62.6*†	1355	69.2*	855	67.1*	1484	68.9*†	786	70.6*†
Pagurus longicarpus[b]	NC_003058	—	—	3698	69.6†	1303	77.1†	789	77.2†	1458	74.1†	—	—

[a]Comparisons of A + T compositions among taxa for each of the mitochondrial regions listed have were performed using χ^2 tests. A+T compositions of taxa differing significantly ($P < 0.05$) from *P. gigas* and *M. rosenbergii* are denoted by * and †, respectively

[b]Incomplete mtDNA sequence (Hickerson and Cunningham, 2000).

Protein-Coding Genes. Consistent for both the *Pse. gigas* and *M. rosenbergii* mitochondrial genomes, the *ATP6* and *ATP8*, *COI-III*, *Cyt b*, *ND2*, *ND3*, and *ND6* genes are encoded by the α-strand, while *ND1*, *ND4*, *ND4L*, and *ND5* are encoded by the β-strand (Table 2). A / T base compositional bias is evident at the first and third codon positions (Table 4). This bias is comparable to that reported for other crustaceans, although the third codon bias for other arthropods has been reported to be much greater (Crease, 1999). Bias to cytosine was found to be greater on the α-strand than on the β-strand, and concomitantly the guanine composition was greater on the α-strand in comparison with the β-strand (Table 4). In mammals this asymmetry is significantly correlated with the duration of single-stranded state of the "heavy-stranded" genes during mtDNA replication. During this time the spontaneous deamination of cytosine and adenine in the heavy-strand occurs owing to preferential exposure to hydrolytic and oxidative damage, and consequential susceptibility to mutation events (Reyes et al., 1998).

Translation initiation and termination codons of the 13 protein-coding genes in *Pse. gigas* and *M. rosenbergii* are summarized in Table 2. The *Pse. gigas* initiation codons inferred for 12 of the 13 genes are ATN, which is typical for metazoan mitochondria (Wolstenholme, 1992). We suggest the putative initiation codon AGA is used for the *ND1* gene on the basis of decapod sequence alignments. Open reading frames of the *Pse. gigas* protein-coding genes were terminated with the typical TAA or TAG codons for all genes except for *COI-III*, *ATP6*, *ND2*, *6*, and *Cyt b*. We suggest that these genes are characterized by truncated termination codons, either TA or T, with the production of the TAA termini being created by posttranscriptional polyadenylation (Ojala et al., 1981).

The *M. rosenbergii* initiation codons inferred for 10 of the 13 genes are ATN. Exceptions include the *COI-III* genes, which employ ACG, AAT, and ACT initiation codons, respectively. Again, initiation codons were estimated via sequence alignments when the typical initiation codons were absent. Open reading frames of the *M. rosenbergii* protein-coding genes were terminated with TAA or TAG for seven of the 13 genes. Again we suggest the remaining genes are characterized by truncated termination codons, with the production of the TAA termini being created by posttranscriptional polyadenylation (Ojala et al., 1981).

The data suggest that for both *Pse. gigas* and *M. rosenbergii* protein-coding genes, a single overlapping reading frame is evident on the α-strand

Table 4. Base Composition (%) of the 13 Protein-Coding Genes for Mitochondral Genomes of *Pseudocarcinus gigas* and *Macrobrachium rosenbergii*[a]

	A	*C*	*G*	*T*
Pseudocarcinus gigas				
All genes				
1st	29.6	14.6	22.2	33.6
2nd	18.5	21.2	15.1	45.2
3rd	37.2	11.8	8.7	42.3
Total	28.4	15.9	15.3	40.4
Genes encoded on α -strand[b]				
1st	29.6	18.5	20.3	31.6
2nd	18.1	24.8	12.5	44.6
3rd	37.7	16.7	5.2	40.4
Total	28.5	20.0	12.6	38.9
Genes encoded on β -strand[c]				
1st	29.7	8.5	25.1	36.7
2nd	19.1	15.6	19.1	46.2
3rd	36.4	4.0	14.1	45.5
Total	28.4	9.4	19.4	42.8
Macrobrachium rosenbergii				
All genes				
1st	27.6	18.7	25.3	28.4
2nd	17.6	23.2	16.3	42.9
3rd	32.7	22.7	13.4	31.2
Total	26.0	21.5	18.3	34.2
Genes encoded on α -strand[b]				
1st	30.0	23.0	23.3	23.7
2nd	18.3	26.7	13.8	41.2
3rd	41.6	31.4	7.0	20.0
Total	30.0	27.0	14.7	28.3
Genes encoded on β -strand[c]				
1st	23.6	11.8	28.5	36.1
2nd	16.5	17.7	20.2	45.6
3rd	18.6	8.7	23.6	49.1
Total	19.6	12.7	24.1	43.6

[a]Chi-square tests indicated that base composition at each codon and across strands were heterogeneous ($P < 0.001$).
[b]*COI, COII, COIII, ATP6, ATP8, Cyt b, ND2, ND3,* and *ND6* genes.
[c]*ND1, ND4, ND4L,* and *ND5* genes.

and the β-strand (*ATP8* and *ATP6* share seven nucleotides; *ND4* and *ND4L* share 7 nucleotides). Overlapping nucleotides at these gene boundaries are common amongt metazoans (Wolstenholme, 1992), and have gained support from surveys of bicistronic transcripts and corresponding protein characteristics (Ojala et al., 1981; Fearnley and Walker, 1986).

Transfer RNA Genes. For the *Pse. gigas* and *M. rosenbergii* mitchondrial genomes 22 tRNA genes were identified on the basis of their respective anticodons and secondary structures (Figures 2 and 3). These tRNAs correspond to the standard set found in other metazoan mtDNAs. Gene lengths and anticodon sequences were congruent with those described for other crustaceans. The anticodon sequences are identical to that reported for *L. polyphemus* (Lavrov et al., 2000), with the exception of $tRNA^{Lys}$ and $tRNA^{Ser(AGN)}$. The *Pse. gigas* and *M. rosenbergii* $tRNA^{Lys}$ *and* $tRNA^{Ser(AGN)}$ genes, like those of other crustacean species, possess UUU and UCU anticodons, respectively, whereas *L. polyphemus* utilizes CUU and GCU.

Invertebrate tRNA anticodon sequences are generally conserved, although variation at the third wobble position of the $tRNA^{Lys}$ and $tRNA^{Ser(AGN)}$ genes is not uncommon (Beard et al., 1993; Boore and Brown, 1994; Yamauchi et al., 2002). We also found the DHU arm stem of the $tRNA^{Ser(AGN)}$ gene in both genomes to be absent; however, this characteristic is typical of metazoan mtDNAs (Jacobs et al., 1988; Hatzoglou et al., 1995; Helfenbein et al., 2001).

Ribosomal RNA Genes. The *Pse. gigas* and *M. rosenbergii lrRNA* gene separates the $tRNA^{Leu\ (CUN)}$ and $tRNA^{Val}$ genes, while the *srRNA* gene separates $tRNA^{Val}$ and the putative control region, with both rRNA genes encoded by the β-strand. The arrangement of the rRNA genes in *Pse. gigas* and *M. rosenbergii* is typical of arthropods sequenced to date, with only a few exceptions including another decapod

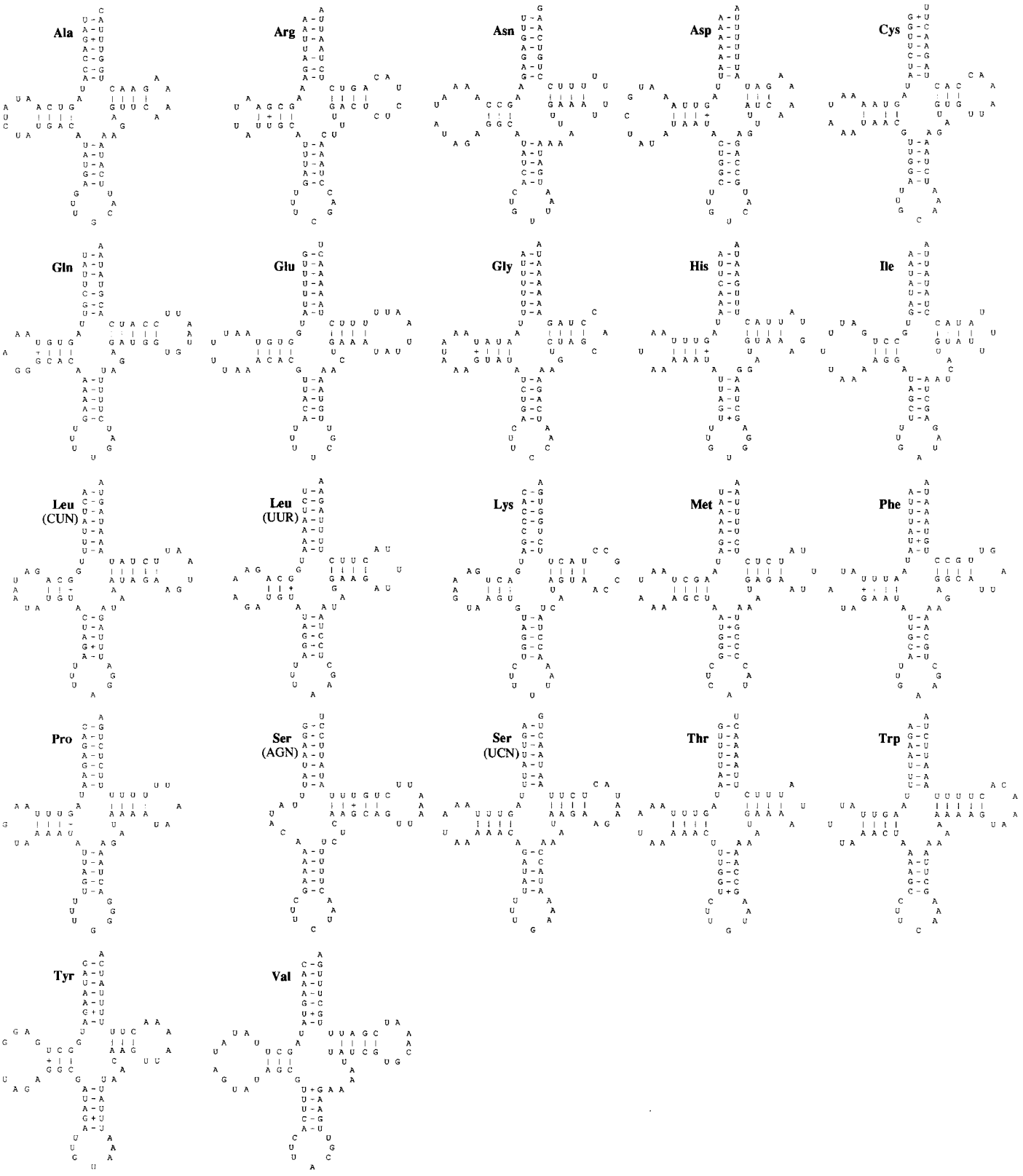

Fig. 2. Putative secondary structures for the 22 tRNA genes of the *Pse. gigas* mitochondrial genome. Watson-Crick and GT bonds are denoted by a dash and plus symbol, respectively.

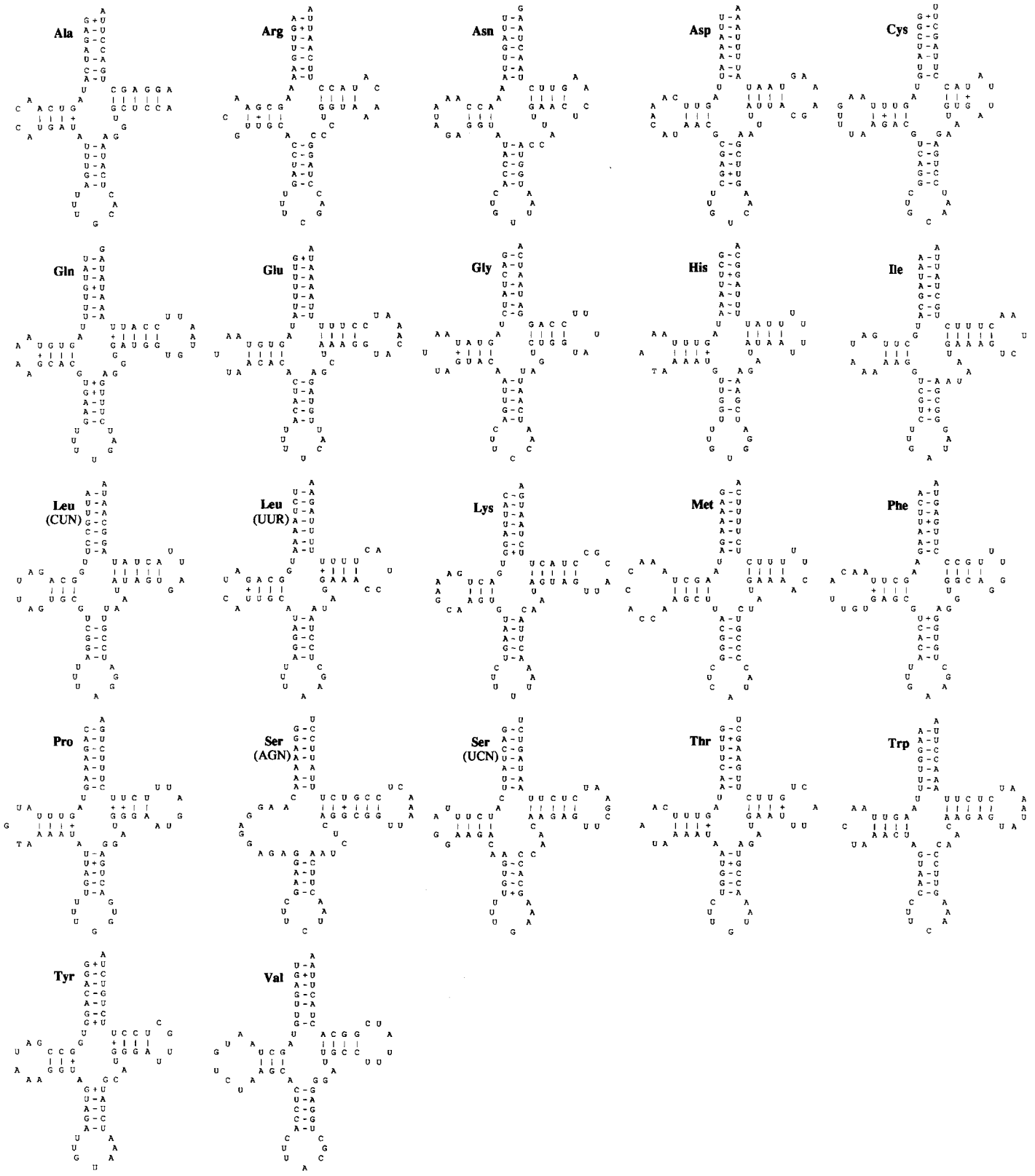

Fig. 3. Putative secondary structures for the 22 tRNA genes of the *M. rosenbergii* mitochondrial genome. Watson-Crick and GT bonds are denoted by a dash and plus symbol, respectively.

crustacean (Evans and Lopez, 2002; Shao and Barker, 2003; Miller et al., 2004). The rRNA gene boundaries were estimated via nucleotide sequence alignments with *Por. trituberculatus* and *Pen. monodon*.

Potential Sequence Utility. The *Pse. gigas* and *M. rosenbergii* mtDNA sequences bring the total number of complete crustacean mitochondrial genome sequences to 17. Both the nucleotide and amino acid sequence data from the species in this study will prove valuable for phylogenetic studies of deep crustacean and arthropod relationships. Recent studies of phylogenetic relationships among major arthropod lineages using complete mitochondrial genome sequences suggest that insufficient taxon sampling is hindering the reconstruction of reliable phytogenies. Specifically, taxonomic sampling is often limited to a single member of a major evolutionary lineage, and as a consequence constructed phylogenies are vulnerable to insufficient phylogenetic signal and long branch attraction (Delsuc et al., 2003; Nardi et al., 2003). Therefore we can expect that as the number of complete arthropod mtDNA sequences grows, so will our confidence in mitochondrial phylogenies for this group and, more broadly, our understanding of mitochondrial genome evolution (i.e.; gene rearrangements).

DNA sequences from different regions within the metazoan mitochondrial genome have proven to be powerful genetic markers for resolving population structure (Hillis et al., 1996). Given the commercial importance of *Pse. gigas* and *M. rosenbergii*, and the potential for overexploitation within the wild fisheries, information on population structure is vital for the implementation of sustainable management strategies. DNA sequence data from the *ATP8*, *ND2*, and *ND6* protein-coding genes have potential to provide valuable information for the elucidation of stock structure within these species, given that these genes demonstrate high nucleotide substitution rates in crustacean mitochondrial genomes (Machida et al., 2004). DNA sequence data from the control region would be desirable as it is also characterizsed by high levels of variability and susceptibility to genetic drift (Avise, 2000). The sequence data generated in this study will specifically facilitate population-level research through the development of PCR primers for the survey of mtDNA sequence variation.

Acknowledgments

The authors thank Andrew Levings for providing *Pse. gigas* tissue samples and Arthur Mangos, from the Institute of Medical and Veterinary Sciences, for sequencing assistance. We also acknowledge Renfu Shao, Jeffrey Boore, and Jody Martin for their valuable comments on this research. Finally, we thank our colleagues from the Molecular Ecology and Biodiversity Laboratory, Deakin University Warrnambool, for their constant support and advice throughout this project. Adam Miller was supported by a Deakin University Postgraduate Scholarship, and funding for this research was provided by Deakin University's Central Research Grant Scheme and the School of Ecology and Environment. We appreciate the helpful comments of two anonymous referees, which consequently improved our manuscript.

References

1. Avise JC (1994) *Molecular Markers, Natural History and Evolution.* (New York, NY: Chapman and Hall)
2. Avise JC (2000) *Phylogeography: The History and Formation of Species.* (Cambridge, Mass: Harvard University Press)
3. Beard CB, Hamm DM, Collins FH (1993) The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. Insect Mol Biol 2, 103–124
4. Boore JL (1999) Animal mitochondrial genomes. Nucleic Acids Res 27, 1767–1780
5. Boore JL (2000) The duplication/random loss model for gene rearrangement exemplified by mitochondrial genomes of deuterostome animals. In: *Comparitive Genomics*, Sankoff D, Nadeau JH, eds. (The Netherlands: Kluwer Academic Publishers) pp 133–147
6. Boore JL, Brown WM (1994) Complete DNA sequence of the mitochondrial genome of the black chiton, *Katharina tunicata*. Genetics 138, 423–443
7. Crandall KA, Lawler SH, Austin CM (1995) A preliminary examination of the molecular phylogenetic relationships of some crayfish genera from Australia (Decapoda:Parastacidae). Freshwater Crayfish 10, 18–30
8. Crease TJ (1999) The complete sequence of the mitochondrial genome of *Daphnia pulex* (Cladocera: Crustacea). Gene 233, 89–99
9. Curole JP, Kocher TD (1999) Mitogenomics: digging deeper with complete mitochondrial genomes. Trends Ecol Evol 14, 394–398
10. de Bruyn M, Wilson JA, Mather PB (2004) Huxley's line demarcates extensive genetic divergence between eastern ana western forms of the giant freshwater prawn, *Macrobrachium rosenbergii*. Mol Phylogenet Evol 30, 251–257
11. Delsuc F, Phillips MJ, Penny D (2003) Comment on "hexapod origins: monophyletic or paraphyletic?" Science 301, 1482d
12. Evans JD, Lopez DL (2002) Complete mitochondrial DNA sequence of the important honey bee pest, *Varroa destructor* (Acari : Varroidae). Exp Appl Acarol 27, 69–78

13. Fearnley IM, Walker JE (1986) Two overlapping genes in bovine mitochondrial DNA encode membrane components of ATP synthase. EMBO J 5, 2003–2008
14. Gilbert, DG (1997) SeqPup software, Indiana University
15. Hatzoglou E, Rodakis GC, Lecanidou R (1995) Complete sequence and gene organization of the mitochondrial genome of the land snail. *Albinaria coerulea*. Genetics 140, 1353–1366
16. Helfenbein KG, Brown WM, Boore JL (2001) The complete mitochondrial genome of the articulate brachiopod *Terebratalia transversa*. Mol Biol Evol 18, 1734–1744
17. Hickerson MJ, Cunningham CW (2000) Dramatic mitochondrial gene rearrangements in the hermit crab *Pagurus longicarpus* (Crustacea, Anomura). Mol Biol Evol 17, 639–644
18. Hillis DM, Moritz C, Mable BK (1996) Molecular Systematics 2nd ed. (Sunderland, Mass: Sinauer Associates)
19. Jacobs HT, Elliott DJ, Veerabhadracharya BM, Farquharson A (1988) Nucleotide sequence and gene organization of sea urchin mitochondrial DNA. J Mol Biol 202, 185–217
20. Lavrov DV, Boore JL, Brown WM (2000) The complete mitochondrial DNA sequence of the horseshoe crab *Limulus polyphemus*. Mol Biol Evol 17, 813–824
21. Lavrov DV, Brown WM, Boore JL (2004) Phylogenetic position of the Pentastomida and (pan)crustacean relationships. Proc R Soc Lond B 271, 537–544
22. Levinson G, Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4, 203–221
23. Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25, 955–964
24. Macey JR, Larson A, Ananjeva NB, Fang Z, Papenfuss TJ (1997) Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. Mol Biol Evol 14, 91–104
25. Macey JR, Schulte JA, Larson A, Papenfuss TJ (1998) Tandem duplication via light-strand synthesis may provide a precursor for mitochondrial genomic rearrangement. Mol Biol Evol 15, 71–75
26. Machida RJ, Miya MU, Nishida M, Nishida S (2002) Complete mitochondrial DNA sequence of *Tigriopus japonicus* (Crustacea: Copepoda). Mar Biotechnol 4, 406–417
27. Machida RJ, Miya MU, Yamaguchi MM, Nishida S (2004) Organization of the mitochondrial genome of Antarctic krill *Euphausia superba* (Crustacea: Malacostraca). Mar Biotechnol 6, 238–250
28. Martin JW, Davis GE (2001) *An Updated Classification of the Recent Crustacea*. (Los Angeles, Calif.: Natural History Museum of Los Angeles County)
29. 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
30. Moritz C, Brown WM (1987) Tandem duplications in animal mitochondrial DNAs: variation in incidence and gene content among lizards. Proc Natl Acad Sci U S A 84, 7183–7187
31. Murphy NP, Austin CM (2002) A preliminary study of the 16S rRNA sequence variation in Australian *Macrobrachium* shrimps (Palaemonidae: Decapoda) reveals inconsistencies in their current classification. Invert Syst 16, 697–701
32. Nardi F, Spinsanti G, Boore JL, Carapelli A, Dallai R, Frati F (2003) Hexapod origins: monophyletic or paraphyletic. Science 299, 1887–1889
33. New M B, Valenti W C (2000) *Freshwater Prawn Culture*. (London, UK: Blackwell Science Ltd)
34. Ogoh K, Ohmiya Y (2004) Complete mitochondrial DNA sequence of the sea-firefly, *Vargula hilgendorfii* (Crustacea, Ostracoda) with duplicate control regions. Gene 327, 131–139
35. Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing hi human mitchondria. Nature 290, 470–474
36. Palumbi SR, Benzie J (1991) Large mitochondrial DNA differences between morphologically similar Penaeid shrimp. Mol Mar Biol Biotechnol 1, 27–34
37. Reyes A, Gissi C, Pesole G, Saccone C (1998) Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. Mol Biol Evol 18, 957–966
38. Shao R, Barker SC (2003) The highly rearranged mitochondrial genome of the plague thrips, *Thrips imaginis* (Insecta: Thysanoptera): convergence of two novel gene boundaries and an extraordinary arrangement of rRNA genes. Mol Biol Evol 20, 362–370
39. Tamura K, Aotsuka T (1988) Rapid isolation method of animal mitochondrial DNA by the alkaline lysis procedure. Biochem Genet 26, 815–819
40. Umetsu K, Iwabuchi N, Yuasa I, Saitou N, Clark PP, Boxshall G, Osawa M, Igarashi K (2002) Complete mitochondrial DNA sequence of a tadpole shrimp (*Triops cancriformis*) and analysis of museum samples. Electrophoresis 23, 4080–4084
41. Valverde JR, Batuecas B, Moratilla C, Marco R, Garesse R (1994) The complete mitochondrial DNA sequence of the crustacean *Artemia franciscana*. J Mol Evol 39, 400–408
42. Wilson K, Cahill V, Bailment 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
43. Wolstenhome DR (1992) Animal mitochondrial DNA: structure and evolution. Int Rev Cyt 141, 173–216
44. Yamauchi MM, Miya M, Nishida M (2002) Complete mitochondrial DNA sequence of the Japanese spiny lobster, *Panulirus japonicus* (Crustacea: Decapoda). Gene 295, 89–96
45. Yamauchi MM, Miya MU, Nishida M (2003) Complete mitochondrial DNA sequence of the swimming crab, *Portunus trituberculatus* (Crustacea: Decapoda: Brachyura). Gene 311, 129–135