Mol Biol Rep (2012) 39:9009–9021
DOI 10.1007/s11033-012-1772-z

The mitochondrial genomes of *Euphausia pacifica* and *Thysanoessa raschii* sequenced using 454 next-generation sequencing, with a phylogenetic analysis of their position in the Malacostracan family tree

Mattias L. Johansson · Angela L. Sremba · Leah R. Feinberg · Michael A. Banks · William T. Peterson

Received: 29 November 2011 / Accepted: 9 June 2012 / Published online: 26 June 2012

Abstract Euphausiid krill play a critical role in coastal and oceanic food webs, linking primary producers to upper trophic levels. In addition, some species support commercial fisheries worldwide. Despite their ecological importance, the genetics of these important species remain poorly described. To improve our understanding of the genetics of these ecological links, we sequenced the mitochondrial genomes of two species of North Pacific krill, *Euphausia pacifica* and *Thysanoessa raschii*, using long-range PCR and 454 GS Junior next-generation sequencing technology. The *E. pacifica* mitogenome (14,692 + base pairs (bp)) encodes 13 protein-coding genes (PCGs), two ribosomal RNA (rRNA) genes, and at least 22 transfer RNA (tRNA) genes. The *T. raschii* mitogenome (14,240 + bp) encodes 13 PCGs, two rRNA genes, and at least 19 tRNA genes. The gene order in both species is similar to that of *E. superba*. Comparisons between Bering Sea and Yellow Sea *E. pacifica* revealed a total of 644 variable sites. The most variable protein-coding gene were *atp8* (7.55 %, 12 of 159 sites variable), *nad4* (6.35 %, 85 variable sites) and *nad6* (6.32 %, 33 variable sites). Phylogenetic analyses to assess the phylogenetic position of the Euphausiacea, using the concatenated nucleic acid sequences of *E. pacifica* and *T. raschii* along with 46 previously published malacostracan mitogenomes, support the monophyly of the order Decapoda and indicate that the Euphausiacea share a common ancestor with the Decapoda. Future research should utilize this sequence data to explore the population genetics and molecular ecology of these species.

Keywords mtDNA · Mitochondrial genome · Phylogeny · Malacostraca · Euphausiacea

M. L. Johansson (✉) · A. L. Sremba · L. R. Feinberg · M. A. Banks
Cooperative Institute for Marine Resources Studies, Oregon State University, Newport, OR, USA
e-mail: mattias.johansson@oregonstate.edu

M. A. Banks
Coastal Oregon Marine Experiment Station, Department of Fisheries and Wildlife, Oregon State University, Corvallis, OR, USA

W. T. Peterson
Hatfield Marine Science Center, National Marine Fisheries Service, NOAA, Newport, OR, USA

Introduction

Krill are an important link in coastal and oceanic food webs, transferring energy from primary producers to higher trophic level animals, including fishes, marine mammals, and birds. Despite the central role of krill in the ecosystem, very little is known about their genetics. Two contrasting evolutionary lines are found within the North Pacific between two co-occurring and dominant genera—a species that is found across the entire North Pacific, potentially as multiple populations (*Euphausia pacifica*), and a genus that may have diverged into several species, including; *Thysanoessa spinifera*, *T. inermis*, *T. raschii* and *T. inspinata*.

Euphasia pacifica is widely distributed throughout most of the North Pacific Ocean. Across its broad range, *E. pacifica* occupies diverse habitats, including cool eutrophic upwelling regions off Baja Mexico, California, Oregon, Washington and British Columbia, as well as the downwelling environment of the Gulf of Alaska in the eastern North Pacific. In the north, it occupies shelf-break waters of the Bering Sea and oligotrophic oceanic regions across the entire North Pacific in waters north of ~40°N,

thence south through the western Pacific from Russia to China. In the western Pacific this species inhabits waters where temperatures range from sub-arctic (the Oyashio) to sub-tropical (the Kuroshio, the Japan/East Sea, and the East China and Yellow Seas) [1]. In contrast to the species *E. pacifica*, the genus *Thysanoessa* also spans much of the North Pacific, but is restricted to marginal seas and continental shelf and slope waters of the northern North Pacific and is present as several different species. Our species of interest, *T. raschii* dominates in the Bering Sea ecosystem and in shelf and slope waters off Russia [2, 3].

Most genetic research on krill to date has focused on Antarctic krill (*Euphausia superba* Dana), which, in addition to its ecological importance, is subject to a significant commercial fishery in the Southern Ocean [4]. Population genetic studies utilizing mitochondrial DNA (mtDNA) markers on *E. superba* have suggested that populations show significant genetic differentiation on a large scale [5], but that individual swarms, which are characteristic of the species, do not represent distinct genetic units [6]. The complete mitochondrial genome of *E. superba* has also been sequenced, revealing additional DNA markers for population studies, and suggesting a novel arrangement of genes in the mtDNA of the species, compared to the pancrustacean ground pattern [7, 8, 9]. The complete mitogenome of *E. superba* may have an unusually low percentage of guanine-cytosine base pairs (32 % ± 0.5 %) and therefore may be more susceptible to UV-B damage than other Antarctic organisms [10].

Next-generation sequencing technologies offer the opportunity to rapidly collect sequence data on our target species of krill. Here, we sequence and characterize the mitochondrial genomes of two species of North Pacific krill using 454 GS Junior sequencing. We report the mitochondrial genome sequence and gene organization for *E. pacifica* and *T. raschii* sampled from the Bering Sea and compare our *E. pacifica* sequence with a recently published *E. pacifica* sequence from the Yellow Sea [11]. In order to better understand the phylogenetic position of the *Euphausiacea*, phylogenetic relationships among malacostracan crustaceans were assessed using reference mitochondrial genomes from NCBI, along with the *E. pacifica* and *T. raschii* sequences from this study, and the most complete previously published *E. superba* sequence [8].

Materials and methods

Sample collection and DNA extraction

Specimens of *E. pacifica* and *T. raschii* were collected from shelf waters of the Bering Sea (57°3.012′N, 167°26.986′W) in June 2010. Specimens were preserved in 95 % ethanol. Total genomic DNA was extracted from tail muscle of six individuals of each species using a standard phenol–chloroform-isoamyl extraction protocol [12]. DNA was resuspended in TE buffer and stored at −20 °C.

Short-range PCR

In order to design species-specific long-range PCR primers for *E. pacifica* and *T. raschii*, selected regions of the mitochondrial genome (*cox1*, *cox3*, *nad5*, *srRNA*) were amplified using a combination of published primers and primers designed from *E. superba* sequences (Table 1a) [7, 8]. PCRs were conducted in a GeneAmp PCR System 9700 thermalcycler (Applied Biosystems, Life Technologies, Carlsbad, California, USA) in a total volume of 15 µl, and containing 2 µl DNA, 6.175 µl sterile nanopure water, 1.2 µl of 25 mM MgCl, 0.3 µl of 10 mM dNTPs, 0.75 µl PCR additive (2 % Tween, 5 % BSA, 10 % formamide), 0.75 µl each of forward and reverse primers, 3.0 µl of 5× GoTaq Flexi buffer (Promega, Madison, WI), and 0.075 µl of GoTaq Flexi DNA Polymerase (0.375 units, Promega). Thermal cycling profiles were as follows: initial denaturation at 94 °C for 3 min, then 25 cycles of denaturation at 94 °C for 30 s, annealing at 50–54 °C for 45 s, and extension at 72 °C for 2 min, with a final extension at 72 °C for 3 min. Prior to sequencing, PCR products were visualized on a 1 % agarose gel and cleaned using an ExoSap-IT PCR cleanup protocol (GE Healthcare, Piscataway, New Jersey). PCR products were prepared for sequencing using an ABI-Prism Big Dye terminator cycle sequencing kit v3.1 (Applied Biosystems), cleaned using a Sephadex protocol (GE Healthcare), and sequenced on an ABI 3730xl (Applied Biosystems). DNA sequences were edited and aligned using Sequencher v5.0 (Gene Codes Corporation, Ann Arbor, MI). Sequence data was used to design species-specific primers for long-range PCR in Primer3 v0.4.0 [13].

Long-range PCR

Large fragments of the mitochondrial genome were amplified for 454 sequencing using long-range PCR [14] with primers designed from short-range sequences (Table 1b). PCRs of a total volume of 20 µl, containing 0.5–1.0 µl DNA, 11.6–12.3 µl sterile nanopure water, 0.4–0.6 µl of 10 mM dNTPs, 0.6 µl DMSO, 1 µl each of forward and reverse primers, 4 µl of 5× Phusion HF reaction buffer (Finnzymes, Thermo Scientific, Vantaa, Finland), and 0.2 µl of Phusion polymerase (0.4 units, Finnzymes), were conducted in a GeneAmp PCR System 9700 thermalcycler (Applied Biosystems). Thermal cycling profiles were as follows: initial denaturation at 98 °C for 30 s, then 30 cycles of denaturation at 98 °C for 10 s, annealing at 60–64 °C for 20 s, and extension at 72 °C for

Table 1 Primers used for PCR and sequencing

Segment	Primers	Species	Sequence	Citation
a. Short-range PCR				
cox1	Eus-*cox1*F	*E. pac/T.ras*	GGT GCA TGA GCT GGA ATA GT	[11]
	Eus-*cox1*R	*E. pac/T.ras*	TTA AGT TGT GCA CCG TGA AG	[11]
cox3	*cox3*-5F	*E. pacifica*	GAA GCT GCC GTT GCA GTA AT	
	cox3-4R	*E. pacifica*	TGA ACT GCT TGG GTG TGA TT	
nad2	Ep G1F	*E. pacifica*	TGG AGT GCC TGA TAA AAG GA	
	Ep G2R	*E. pacifica*	TGC AAA TCT GAA GAT GCA CAC	
nad5	*nad5*-2F	*E. pacifica*	GAG CTG CTA TTG CTG CAG GT	
	nad5-1R	*E. pac/T.ras*	AAA CGA CGG TCT GTG AAT CC	
	Eus-*nad5*F	*T. raschii*	TTA TGA ATT ACA GCC CCA GC	[11]
srRNA	Eus-srRNAF	*E. pac/T.ras*	TAA GAA TGA GAG CGA CGG G	[11]
	Eus-srRNAR	*E. pac/T.ras*	TTT GGC GGT GTC TTA GTC TAG	[11]
b. Long-range PCR				
cox1-cox3	Epco1-131F	*E. pacifica*	ATT AGG TGC CCC TGA TAT GG	
	Epco3-521R	*E. pacifica*	TCG CAT GAT GTG ATC AGG TT	
cox1-nad5	Trco1-46F	*T. raschii*	AGT CAC AGC TCA CGC TTT TG	
	Trnd5-67WR	*T. raschii*	GCT TTG TCT ACA TTA AGT CAA TTA GGT	
cox3-nad5	Epco3-227F	*E. pacifica*	ACC ATG ATC CAA TGA TGA CG	
	Epnd5-240R	*E. pacifica*	GGG TGA GAT GGT CTC GGT CT	
nad5-srRNA	Epnd5-88F	*E. pacifica*	ACC TCC TCG CGA TGC TAA TA	
	Epsr-330R	*E. pacifica*	TTG GCG GTG CTT AGT CTA GTT	
	Trnd5-266F	*T. raschii*	TGC TGG AAG TCA AGC AGA AA	
	Trsr-136WR	*T. raschii*	GAT CAA GGT GCA GCT AAT AAA AA	
	Trsr-182R	*T. raschii*	GGC GGT GTC TTA GTC TAG TTA GAG G	
c. Genome walking				
srRNA-cox1	Epco1-143GW	*E. pacifica*	TCA GGG GCA CCT AAT ATA AGA GGA ACA A	
	Epco1-165GW	*E. pacifica*	ATT CAT TCG AGG GAA AGC CAT ATC AGG	
	Trco1-244GW	*T. raschii*	ACC AAC ACC TCT TTC TAC AAG CCC ACT	
	AP1	*E. pac/T.ras*	GTA ATA CGA CTC ACT ATA GGG C	Clontech

1 min (3 kb fragments) or 1 min 45 s (7.5 kb fragments), followed by a final extension at 72 °C for 5 min. For each individual, PCR products for each of the fragments were visualized on 1 % agarose gels, combined in equimolar amounts, and purified using a Qiaquick PCR Cleanup Kit (Qiagen, Valencia, CA) in preparation for 454 GS Junior sequencing.

Sequencing

Mitogenome libraries were prepared from long-range PCR products for three individual samples from each species using the Rapid Library Preparation Method protocol (454 Life Sciences, Roche Applied Science, Branford, Connecticut). PCR products were treated as genomic DNA and unique RL MID adaptors were used for each individual krill sample. Libraries for each individual were combined in equimolar amounts, and subjected to emulsion PCR following the emPCR Amplification Method Manual—Lib-L (454 Life Sciences). Emulsion PCR products were then cleaned and prepared for sequencing according to the Sequencing Method Manual (454 Life Sciences) before being analyzed on a 454 GS Junior second generation sequencer (454 Life Sciences). Two runs were performed on the GS Junior sequencer. The first run consisted of PCR products from all three individuals per species of *T. raschii* and *E. pacifica* (1/12 run each), while the second run included mitochondrial PCR products from the single best-performing sample from run one for each species (1/13th run each). For each species, sequences from all three individuals from both runs were assembled using GS De Novo Assembler (454 Life Sciences) to make a composite reference sequence. All the reads from the single individual sequenced in both runs were then assembled to this

Table 2 NCBI Genbank sequence used for phylogenetic analyses

Accession number	Species	Citation
NC_014883	*Alpheus distinguendus*	[31]
NC_006281	*Callinectes sapidus*	[32]
NC_014492	*Caprella mutica*	[33]
NC_014687	*Caprella scaura*	[34]
NC_013246	*Charybdis japonica*	[35]
NC_011243	*Cherax destructor*	[36]
NC_013976	*Eophreatoicus* sp. 14 FK-2009	[37]
NC_011598	*Eriocheir hepuensis*	Wang et al. unpublished
NC_011597	*Eriocheir japonica*	Wang et al. unpublished
NC_006992	*Eriocheir sinensis*	[38]
EU583500	*Euphausia superba*	[8]
NC_012566	*Exopalaemon carinicauda*	[39]
NC_012738	*Farfantepenaeus californiensis*	[40]
NC_009697	*Fenneropenaeus chinensis*	[32]
NC_013713	*Gandalfus yunohana*	[41]
NC_007379	*Geothelphusa dehaani*	[42]
NC_007442	*Gonodactylus chiragra*	Swinstrom et al. unpublished
NC_008413	*Halocaridina rubra*	[43]
NC_006916	*Harpiosquilla harpax*	[44]
NC_015607	*Homarus americanus*	Kim et al. unpublished
NC_008412	*Ligia oceanica*	[45]
NC_012060	*Litopenaeus stylirostris*	[40]
NC_009626	*Litopenaeus vannamei*	Muhlia-Almazan et al. unpublished
NC_007443	*Lysiosquillina maculata*	Swinstrom et al. unpublished
NC_012217	*Macrobrachium lanchesteri*	Ngernsiri and Sangthong unpublished
NC_015073	*Macrobrachium nipponense*	Ma et al. unpublished
NC_006880	*Macrobrachium rosenbergii*	[46]
NC_007010	*Marsupenaeus japonicus*	[47]
NC_013032	*Metacrangonyx longipes*	[48]
NC_013819	*Onisimus nanseni*	Lee unpublished
NC_014342	*Oratosquilla oratoria*	[49]
NC_003058	*Pagurus longicarpus*	[50]
NC_004251	*Panulirus japonicus*	[51]
NC_014854	*Panulirus ornatus*	[31]
NC_014339	*Panulirus stimpsoni*	[29]
NC_002184	*Penaeus monodon*	[52]
NC_005037	*Portunus trituberculatus*	[53]
NC_006891	*Pseudocarcinus gigas*	[46]
NC_012569	*Scylla olivacea*	Sangthong unpublished
NC_012572	*Scylla paramamosain*	Sangthong unpublished
NC_012565	*Scylla serrata*	Sangthong unpublished
NC_012567	*Scylla tranquebarica*	Sangthong unpublished
NC_011013	*Shinkaia crosnieri*	[54]
NC_007444	*Squilla empusa*	Swinstrom et al. unpublished
NC_006081	*Squilla mantis*	[55]
NC_013480	*Xenograpsus testudinatus*	[56]

reference using GS Reference Mapper (454 Life Sciences) to give a near-complete single-individual sequence.

Genome walking

A portion of the *srRNA-cox1* fragment for each species (and the same individual as sequenced above) was sequenced by genome walking using a Genomewalker Universal Kit (Clontech Laboratories, Inc., Mountain View, CA) following standard protocols and using primers designed from *E. pacifica* and *T. raschii* long PCR sequences (Table 1c). Sequencing followed the same protocol as detailed above for short-range PCR.

Mitogenome sequence assembly and annotation

Final sequences were aligned, assembled, and edited using Sequencher v5.0 (Gene Codes). The mitochondrial genomes for each species were annotated using DOGMA [15] and through comparisons with published *E. superba* annotations (accession #AB084378 [7]; accession #EU583500 [8]). Gene maps for both species were drawn using OGDRAW [16]. Nearly complete mitochondrial genome sequences were submitted to GenBank for *E. pacifica* (accession #JN713149) and *T. raschii* (accession #JN713150).

Sequence comparison with Yellow Sea *E. pacifica*

To characterize sequence variation in *E. pacifica*, we compared our sequence with another recently published for a sample from the Yellow Sea [11]. Sequences were aligned using Sequencher v5.0 (Gene Codes), and number of varying sites enumerated for each protein coding, tRNA, and rRNA gene.

Phylogenetic reconstruction

The nucleic acid sequences from all 13 protein-coding genes (PCGs) of the *E. pacifica* and *T. raschii* mitochondrial genomes were compared to 45 malacostracan NCBI reference mitochondrial genome sequences and the near-complete *E. superba* mitochondrial genome sequence (Table 2) in a phylogenetic reconstruction. Sequences were aligned using Clustal W [17] as implemented in the BioEdit sequence alignment editor (Ibis Biosciences, Carlsbad, California, USA). The best fitting model of sequence evolution was determined to be the GTR + I + G model using a nested likelihood ratio test in jModeltest v0.1.1 [18, 19]. Phylogenetic relationships were estimated using a maximum-likelihood approach implemented in RAxML BlackBox [20] and Bayesian inference analyses implemented in MrBayes v.3.1 [21]. Statistical support for the maximum likelihood analysis was estimated using 1,000 bootstrap replicates. For the Bayesian analysis, Markov Chain Monte Carlo (MCMC) analyses were run for 2,000,000 generations with trees sampled every 1,000 generations. After omitting the first 25 % of saved trees (burn-in), the remaining 1,500 sampled trees were used to estimate a 50 % majority rule consensus tree. Consensus trees were visualized using Archaeopteryx v.0.957 beta [22]. Trees were outgroup rooted with six species from the family Stomatopoda.

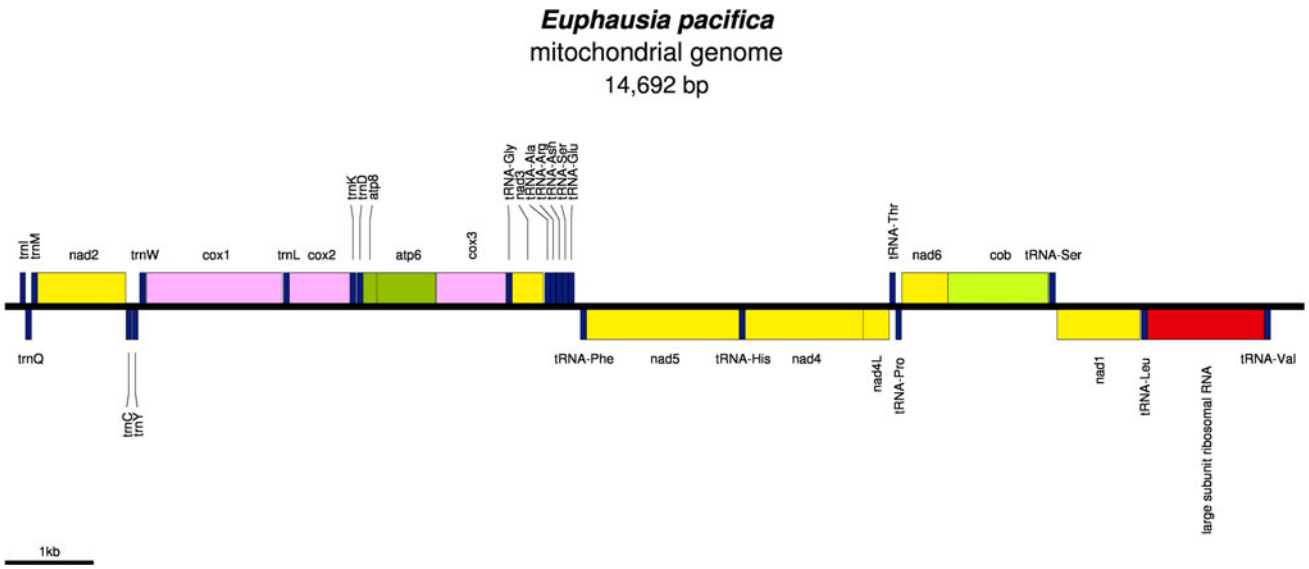

Fig. 1 Gene map of the mitochondrial genome of *Euphausia pacifica*. PCGs above the *midline* are transcribed from *left* to *right* (*heavy strand*), while genes below the *midline* are transcribed from *right* to *left* (*light strand*)

Results and discussion

Sequence analysis

We amplified the mitogenomes in four (*E. pacifica*) and three (*T. raschii*) long-range PCR fragments. Short PCRs for the *cox3* region were unsuccessful in *T. raschii*, and attempts to amplify a portion of the *cob* gene failed in both target species. The long range PCR products for the *srRNA-cox1* fragment failed to amplify with any primers tested. We attempted to sequence complete mitogenomes for three *E. pacifica* and three *T. raschii* individuals. Due to low sequence coverage from our initial run, we characterized the mitogenomes of each species by combining all the sequence data from both sequencer runs for each species in a consensus sequence to form a scaffold, and then assembled the sequence data for a single individual from both sequencer runs, along with genome-walking sequences, to

Table 3 Mitochondrial gene profile of *E. pacifica* (accession no. JN713149)

Feature	Position from–to	Strand	Size (bp)	Codon		Intergenic nucleotides (bp)
				Start	Stop	
nCR[a]	1–166	H	166			
trnI	167–231	H	65			0
trnQ	229–297	L	69			−3
trnM	297–364	H	68			−1
nad2	365–1366	H	1002	ATT	TAA	–1
trnC	1366–1431	L	66			1
trnY	1439–1504	L	66			7
trnW	1521–1590	H	70			16
cox1	1597–3135	H	1539	ACG	TAA	6
trnL1(CUN)	3147–3212	H	66			11
cox2	3213–3900	H	688	ATA	T-	0
trnK	3901–3969	H	69			0
trnD	3979–4045	H	67			9
atp8	4046–4204	H	159	ATC	TAA	0
atp6	4198–4872	H	675	ATG	TAA	−7
cox3	4872–5664	H	793	ATG	T-	−1
trnG	5665–5731	H	67			0
nad3	5732–6085	H	354	ATT	TAA	0
trnA	6100–6164	H	65			15
trnR	6165–6230	H	66			0
trnN	6233–6299	H	67			2
trnS1(AGN)	6300–6366	H	67			0
trnE	6368–6436	H	69			1
trnF	6503–6572	L	70			66
nad5	6572–8302	L	1731	GTG	TAA	−1
trnH	8303–8367	L	65			0
nad4	8368–9705	L	1338	ATG	TAA	0
nad4L	9699–9998	L	300	ATG	TAA	−7
trnT	10001–10067	H	67			2
trnP	10069–10135	L	67			1
nad6	10139–10660	H	522	ATC	TAA	3
cob	10660–11796	H	1137	ATG	TAA	−1
trnS2(UCN)	11807–11877	H	71			10
nad1	11895–12833	L	939	ATA	TAA	17
trnL2(UUR)	12849–12914	L	66			15
lrRNA	12916–14239	L	1324			1
trnV	14238–14309	L	72			−2
srRNA[a]	14310–14692	L	383			0

"−" Termination codon is completed via polyadenylation

[a] Partial sequence

arrive at our final sequences. Thus, our final sequences consist solely of data from a single individual for each species.

The characterized mitochondrial genomes of *E. pacifica* and *T. raschii* are greater than 14,692 and 14,240 bp in length, respectively. Our initial sequencer run returned 13,986 reads and 1,832,909 bp of *E. pacifica* data (131 bp average read length) and 4,746 reads and 614,405 bp of *T. raschii* data (129 bp average read length) divided among three samples per species (1/12th plate per sample). Our second sequencer run returned 6,445 reads and 2,645,853 bp of *E. pacifica* data (410 bp average read length) and 2,752 reads and 1,083,297 bp of *T. raschii* data (393 bp average read length) for a single individual per species (1/13th plate per sample). Including data from both runs for the repeated samples, the *E. pacifica* mitogenome was sequenced to an average depth of 170.9 and the T. raschii mitogenome was sequenced to an average depth of 68.5. The mitochondrial genomes for *E. pacifica* and *T. raschii* encode at least 37 and 34 genes, respectively. This includes 13 protein-coding genes (PCG) in each species, two rRNA genes, 22 tRNAs in *E. pacifica*, and 19 tRNAs in *T. raschii*. Based on published sequences from *E. superba* and other malacostracan crustaceans, we predict that one transfer RNA and the noncoding region remain to be sequenced in *E. pacifica* and that four tRNAs and the noncoding region remain in *T. raschii*. The overall A + T content was 70.1 % for *E. pacifica*, and 71.7 % in *T. raschii*. These values are comparable to the range of 60.0–71.5 % described in other malacostracan crustaceans [23]. Both *E. pacifica* and *T. raschii* show a very similar gene arrangement to previously published *E. superba* sequences [7, 8, 9]. In contrast to the *E. superba* mitogenome, a change at the anticodon means that the first tRNA 3′ of the non-coding region in *E. pacifica* is a tRNA-Ile, rather than a tRNA-Asn. Otherwise, the gene orders are identical between the three species.

In *E. pacifica*, nine PCGs are encoded on the heavy strand (*nad2*, *cox1*, *cox2*, *atp8*, *atp6*, *cox3*, *nad3*, *nad6*, and *cob*), along with 14 tRNAs (Fig. 1, Table 3). The remaining four PCGs (*nad5*, *nad4*, *nad4L*, *nad1*), seven tRNAs, and both rRNAs, are encoded on the light strand. In *T. raschii*, the same pattern is evident (Fig. 2, Table 4).

The *E. pacifica* (Table 3) and *T. raschii* (Table 4) mitogenomes use several alternative start codons in addition to ATG. In *E. pacifica*, five of the thirteen PCGs start with ATG (*atp6*, *cox3*, *nad4*, *nad4L*, *cob*). Two genes utilize ATT as their start codon (*nad2*, *nad3*), two utilize ATA (*cox2*, *nad1*), and two utilize ATC (*atp8*, *nad6*). *Cox1* begins with ACG, and *nad5* with GTG. Six of the 12 completely sequenced PCGs in *T. raschii* begin with ATG (*cox1*, *atp6*, *cox3*, *nad4*, *nad4L*, *cob*). In the case of *cox1*, this ATG is out of frame with the rest of the gene. This frame shift may have resulted from PCR or sequencing error. Alternatively, initiation of translation may occur through the reading of an ATGT quadruplet, as has been previously suggested in the genus *Drosophila* [24, 25]. Or, initiation may occur through editing of the mRNA transcript to align the ATG with the rest of the sequence [26], or through an alternative (TGT) start codon [27]. Two genes use ATA start codons (*cox2*, *nad6*). The remaining four genes use ATT (*atp8*), ATC (*nad3*), GTG (*nad5*), and

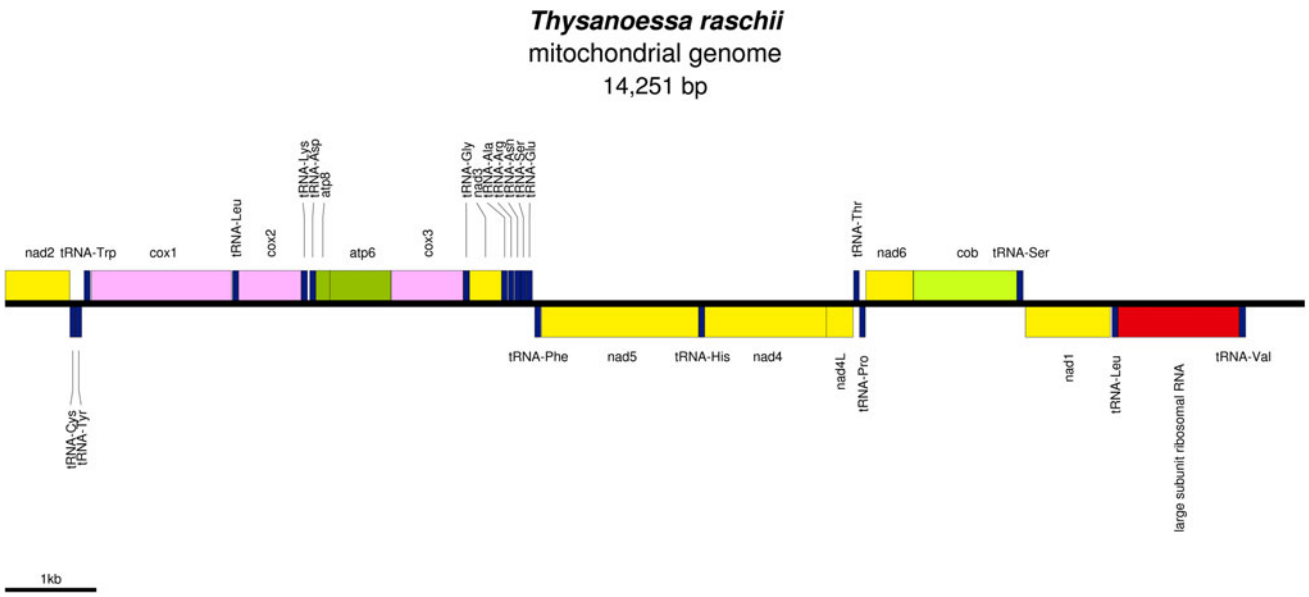

Fig. 2 Gene map of the mitochondrial genome of *Thysanoessa raschii*. PCGs above the *midline* are transcribed from *left* to *right* (*heavy strand*), while genes below the *midline* are transcribed from *right* to *left* (*light strand*)

GTA (*nad1*). This diversity of start codons is similar to that found in *E. superba* [8], with only minor changes in the two study species.

Eleven of the PCGs in the *E. pacifica* mitogenome end with a TAA stop codon (Table 3). The remaining two (*cox2*, *cox3*) have incomplete (T-) stop codons that are completed via posttranslational polyadenylation [28]. Ten of the open reading frames in the *T. raschii* mitogenome end with the TAA stop codon (Table 4). One protein ends with a TAG stop codon (*nad4*), and the two others end with incomplete (T-) stop codons (*cox2*, *cox3*). The pattern of stop codons in *E. pacifica* and *T. raschii* is similar to that reported in *E. superba*, where three PCGs (*nad2*, *cox2*, *cox3*) end in T- stop codons, and one (*nad1*) ends in a TAG [8].

We sequenced 19 and 22 of a predicted 23 tRNA genes in *E. pacifica* and *T. raschii*, respectively. These tRNAs

Table 4 Mitochondrial gene profile of *T. raschii* (accession no. JN713150)

Feature	Position from–to	Strand	Size (bp)	Codon		Intergenic nucleotides (bp)
				Start	Stop	
nCR[a]						
trnN[a]						
trnQ[a]						
trnM[a]						
nad2[b]	1–709	H	709	[a]	TAA	[a]
trnC	709–773	L	65			−1
trnY	775–840	L	66			1
trnW	865–933	H	69			24
cox1	944–2483	H	1540	ATG[c]	TAA	10
trnL1(UAG)	2493–2558	H	66			9
cox2	2559–3246	H	688	ATA	T-	0
trnK	3247–3315	H	69			0
trnD	3339–3404	H	66			23
atp8	3405–3563	H	159	ATT	TAA	0
atp6	3557–4231	H	675	ATG	TAA	−7
cox3	4231–5023	H	793	ATG	T-	−1
trnG	5024–5090	H	67			0
nad3	5091–5444	H	354	ATC	TAA	0
trnA	5445–5509	H	65			1
trnR	5516–5580	H	65			6
trnN	5586–5650	H	65			5
trnS1(AGN)	5650–5716	H	67			−1
trnE	5718–5782	H	65			1
trnF	5806–5874	L	69			23
nad5	5874–7604	L	1731	GTG	TAA	0
trnH	7605–7669	L	65			0
nad4	7670–9007	L	1338	ATG	TAG	0
nad4L	9001–9300	L	300	ATG	TAA	−7
trnT	9303–9366	H	64			2
trnP	9368–9434	L	67			1
nad6	9438–9959	H	522	ATA	TAA	3
cob	9959–11095	H	1137	ATG	TAA	−1
trnS2(UCN)	11095–11164	H	70			−1
nad1	11182–12120	L	939	GTA	TAA	17
trnL2(UUR)	12137–12202	L	66			16
lrRNA	12203–13532	L	1330			0
trnV	13533–13604	L	72			0
srRNA[b]	13605–14251	L	659			0
trnI[a]						

"−" Indicates that termination codon is completed via polyadenylation

[a] Not sequenced

[b] Partial sequence

[c] Out of frame with the rest of the gene

Table 5 Variation within protein-coding sequences

Gene	Length (bp)	Variable sites	% Sites variable in sequence	Variable sites affecting AA sequence	% Sites causing AA sequence change
atp6	675	23	3.41	1	0.15
atp8	159	12	7.55	5	3.14
cob	1137	58	5.1	7	0.62
cox1	1539	67	4.35	8	0.52
cox2	688	26	3.78	5	0.73
cox3	793	29	3.66	6	0.76
nad1	939	49	5.22	4	0.43
nad2	1002	54	5.39	5	0.50
nad3	354	9	2.54	1	0.28
nad4	1338	85	6.35	14	1.05
nad4L	300	8	2.67	1	0.33
nad5	1731	93	5.37	18	1.04
nad6	522	33	6.32	5	0.96

ranged in size between 64 and 72 base pairs, the same as previously reported for *E. superba* [7, 8, 9]. Sequences and anticodon nucleotides were similar to those seen in *E. superba*, with the notable exception that the anticodon in the first tRNA 3′ of the non-coding region changes from *trnN* to *trnI* in *E. pacifica*. Both the large and small subunit tRNAs are encoded on the light strand in both species, and separated by a *trnV*. This is identical to the pattern identified in *E. superba* [7].

Comparison with Yellow Sea E. pacifica

A total of 644 variable sites were found in 14,692 bp of overlapping sequence between Bering Sea and Yellow Sea *E. pacifica*. Within protein-coding sequences (Table 5), the percentage of variable sites per gene varied from 2.54 % in *nad3* to 7.55 % in *atp8*. High variability in *atp8* was previously identified in *E. superba* [7]. However, the small size of *atp8* means that the high percentage of variable sites represents relatively few sequence changes (12 changes in 159 bp). The two genes with the next highest percentage of variability are *nad4*, with 85 variable sites in 1,338 bp (6.35 %), and *nad6*, with 33 variable sites in 522 bp (6.32 %), while the gene with the highest number of variable sites was *nad5*, with 93 variable sites in 1,731 bp (5.37 %). Within tRNA sequences (Table 6), the percentage of variable sites per coding sequence ranged from 0 to 4.55 %, or up to three variable sites in 66 bp of sequence. We identified 29 variable sites in the 1,324 bp (2.19 %) of the large subunit ribosomal RNA, and four variable sites in 383 bp (1.04 %) of the small subunit rRNA (Table 6).

Phylogenetic Analysis

The ML (Fig. 3) and Bayesian inference (Fig. 4) trees returned identical phylogenetic reconstructions consistent with previous research [8, 29]. Within the subclass Eumalacostraca, the orders Isopoda and Amphipoda are identified as the nearest neighbor to the Euphausiacea plus Decapoda, and the order Euphausiacea share a common ancestor with

Table 6 Variation within tRNA and rRNA sequences

RNA	Length (bp)	Variable sites	% Sites variable in sequence
lrRNA	1324	29	2.19
srRNA[a]	383	4	1.04
trnA	65	2	3.08
trnC	66	3	4.55
trnD	67	0	0.00
trnE	69	2	2.90
trnF	70	0	0.00
trnG	67	0	0.00
trnH	65	1	1.54
trnI	65	0	0.00
trnK	69	1	1.45
trnL1(CUN)	66	0	0.00
trnL2(UUR)	66	0	0.00
trnM	68	0	0.00
trnN	67	2	2.99
trnP	67	2	2.99
trnQ	69	0	0.00
trnR	66	3	4.55
trnS1(AGN)	67	1	1.49
trnS2(UCN)	71	3	4.23
trnT	67	2	2.99
trnV	72	0	0.00
trnW	70	1	1.43
trnY	66	1	1.52

[a] Partial sequence

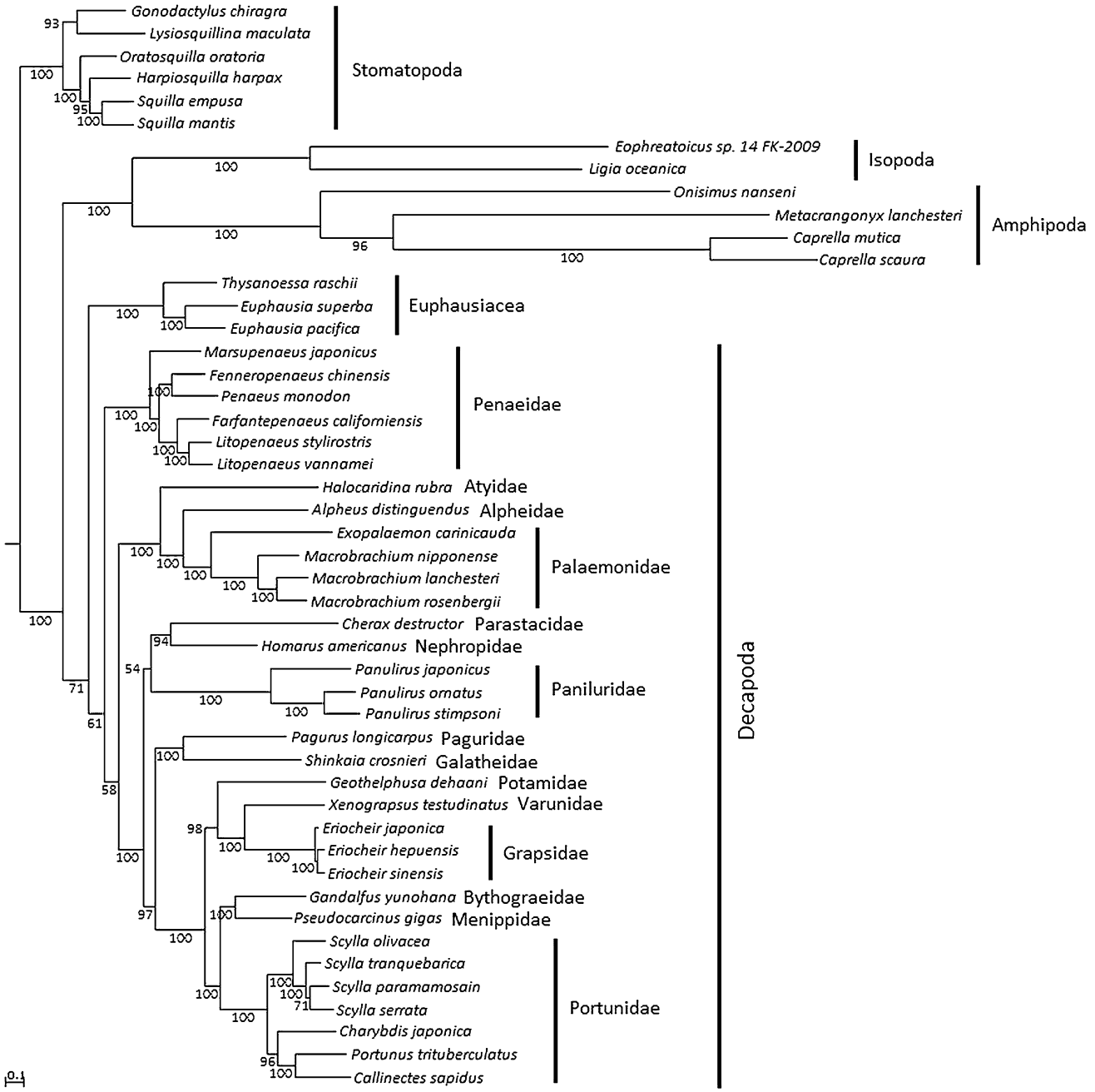

Fig. 3 Maximum likelihood phylogenetic tree of concatenated mitochondrial PCGs from 48 malacostraca crustacean species. Statistical support indicated by bootstrap values below the branch (%). Order to which species belong is indicated to the right of the species name. For species within the Decapoda, family is also indicated

the monophyletic Decapoda. Within the Decapoda, the family Penaeidae is basal to the remaining decapods.

Below the suborder level, the phylogenetic reconstructions illustrate inconsistencies with the current classification of crustaceans. In both ML and Bayesian reconstructions, the infraorder Caridea is shown as the nearest neighbor to a large, well-supported clade Reptantia, containing the infraorders Astacidea, Achelata, Anomura, and Brachyura. This suggests that infraorders within the suborder Pleocyemata do not exist at an equivalent phylogenetic level. Within the Reptantia, however, the previously mentioned infraorders fall into a recognizable order. The Achelata, represented by the family Paniluridae, is sister to the Astacidea, which contains the Parastacidae and Nephropidae. The Achelata/Astacidea clade is sister to a "crab clade", containing the Anomura and Brachyura.

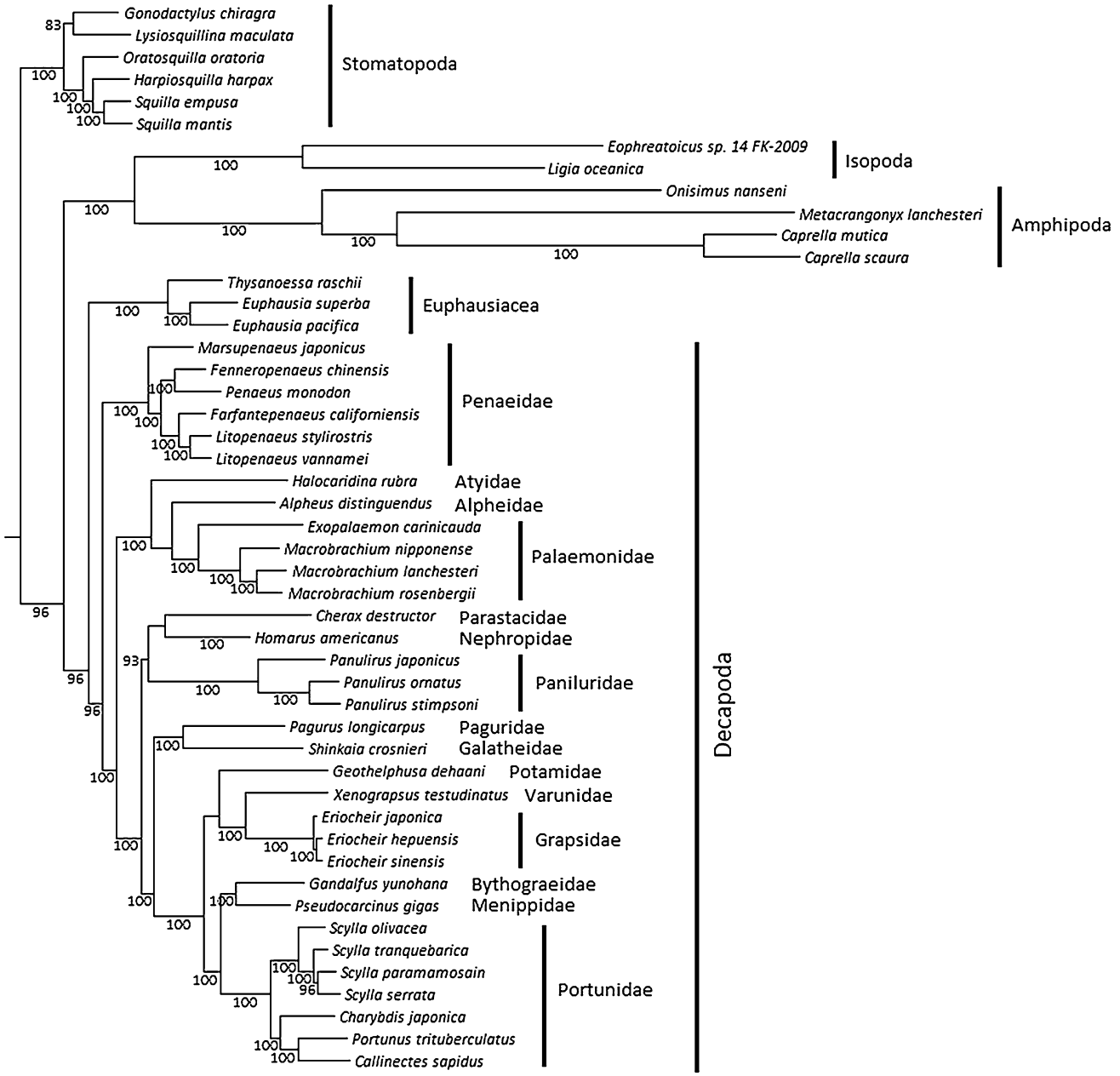

Fig. 4 Bayesian inference phylogenetic tree of concatenated mitochondrial PCGs from 48 malacostraca crustacean species. Statistical support indicated by bootstrap values below the branch (%).Order to which species belong is indicated to the right of the species name. For species within the Decapoda, family is also indicated

These relationships agree with previously published results based on nuclear [30] and mitochondrial [29] genes. Below the infraorder level, some families appear to be paraphyletic in this analysis. The families Atyidae, Alpheidae, Potamidae, and Varunidae all fail to form monophyletic clades. Whether this is an artifact of their being represented by a single species each, or provides evidence of inconsistencies within the current classification, remains to be resolved.

Conclusion

In this study we used next-generation sequencing techniques to characterize the mitochondrial genomes of two ecologically-important species of North Pacific krill. We sequenced near-complete mitochondrial genomes for *E. pacifica* and *T. raschii*, containing approximately 91 % of the sequence for both species, and including all of the PCGs (*nad2* is partial in *T. raschii*), most tRNAs, and

sequences from both rRNAs. Gene order and start and stop codon usage was similar to that previously described in *E. pacifica* [11] and *E. superba* [7, 8, 9]. Overall, variability between Bering Sea and Yellow Sea *E. pacifica* was 4.38 % (644 variable sites in 14,692 bp of sequence), and was highest in *atp8* (7.55 %) and lowest in *nad3* (2.54 %). Our phylogenetic reconstructions supported the previously published phylogenies for malacostracan crustaceans, and suggest that the Euphausiacea share a common ancestor with the Decapoda. This work adds two new euphausiid mitochondrial genome sequences to the single *E. pacifica* and two *E. superba* sequences previously published. An important next step will be to utilize the data generated by this study to develop mitochondrial molecular markers to explore population structure within this ecologically important group of species.

Acknowledgments This work was supported by a NOAA award to CIMRS (Award No. NA06NMF4550286). The authors also gratefully acknowledge O. Drechsel for assistance with constructing genome maps with OGDRAW, B. Slikas, D. Steel, A. Alexander, and D. Jacobson for assistance with 454 sequencing, and T. Shaw, J. Peterson, J. Menkel, and J. Fisher for krill sample collection. The manuscript was greatly improved by the advice of two anonymous reviewers.

References

1. Brinton E (1962) The distribution of Pacific euphausiids. Bull Scripps Inst Oceanogr Univ Calif 8:51–269
2. Mauchline J, Fisher LR (1969) The biology of euphausiids. Adv Mar Biol 7:1–454
3. Timofeev SF (1993) Distribution and age composition of euphausiids in waters around the Spitsbergen Archipelago. Oceanology 33:105–109
4. De Pitta C, Bertolucci C, Mazzotta GM, Bernante F, Rizzo G, De Nardi B, Pallavicini A, Lanfranchi G, Costa R (2008) Systematic sequencing of mRNA from the Antarctic krill (*Euphausia superba*) and first tissue specific transcriptional signature. BMC Genomics 9:45–58
5. Zane L, Ostellari L, Maccatrozzo L, Battaglia B, Patarnello T (1998) Molecular evidence for genetic subdivision of Antarctic krill (*Euphausia superba* Dana) populations. Proc Royal Soc B 265:2387–2391
6. Goodall-Copestake WP, Pérez-Espona S, Clark MS, Murphy EJ, Seear PJ, Tarling GA (2010) Swarms of diversity at the gene cox1 in Antarctic krill. Heredity 104:513–518
7. Machida RJ, Masaki UM, Mitsugu MY, Mutsumi N, Nishida S (2004) Organization of the mitochondrial genome of Antarctic Krill *Euphausia superba* (Crustacea: Malacostraca). Mar Biotechnol 6:238–250. doi:10.1007/s10126-003-0016-6
8. Shen X, Wang H, Ren J, Tian M, Wang M (2010) The mitochondrial genome of *Euphausia superba* (Prydz Bay) (Crustacea: Malacostraca: Euphausiacea) reveals a novel gene arrangement and potential molecular markers. Mol Biol Rep 37(2):771–784. doi:10.1007/s11033-009-9602-7
9. Johansson M, Duda E, Sremba A, Banks M, Peterson W (2012) Assessing population-level variation in the mitochondrial genome of *Euphausia superba* using 454 next-generation sequencing. Mol Biol Rep 39(5):5755–5760. doi: 10.1007/s11033-011-1385-y
10. Jarman S, Elliott N, Nicol S, McMinn A, Newman S (1999) The base composition of the krill genome and its potential susceptibility to damage by UV-B. Antarct Sci 11:23–26
11. Shen X, Wang H, Wang M, Liu B (2011) The complete mitochondrial genome sequence of *Euphausia pacifica* (Malacostraca: Euphausiacea) reveals a novel gene order and unusual tandem repeats. Genome 54:911–922
12. Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
13. Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, pp 365–386
14. Cheng S, Chang SY, Gravitt P (1994) Long PCR. Nature 369:684–685. doi:10.1038/369684a0
15. Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20(17): 3252–3255
16. Lohse M, Drechsel O, Bock R (2007) OrganellarGenomeDraw (OGDRAW)—a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Curr Genet 52(5–6):267–274
17. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680
18. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704
19. Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25:1253–1256
20. Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web-servers. Syst Biol 75(5):758–771
21. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: bayesian phylogenetic inference under mixed models. Bioinformatics 19(12):1572–1574
22. Han MV, Zmasek CM (2009) phyloXML: XML for evolutionary biology and comparative genomics. BMC Bioinf 10:356
23. Shen X, Ren J, Cui Z, Sha Z, Wang B, Xiang J, Liu B (2007) The complete mitochondrial genomes of two common shrimps (*Litopenaeus vannamei* and *Fenneropenaeus chinensis*) and their phylogenomic considerations. Gene 403(1–2):98–109
24. De Bruijn MH (1983) *Drosophila melanogaster* mitochondrial DNA, a novel organization and genetic code. Nature 304(5923): 234–241
25. Clary DO, Wolstenholme DR (1983) Genes for cytochrome c oxidase subunit 1, URF2, and three tRNAs in *Drosophila* mitochondrial DNA. Nucleic Acids Res 11(19):6859–6872
26. Fox TD (1987) Natural variation in the genetic code. Ann Rev Genet 21:67–91
27. Park DS, Suh SJ, Oh HW, Hebert PD (2010) Recovery of the mitochondrial COI barcode region in diverse Hexapoda through tRNA-based primers. BMC Genomics 11:423
28. Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. Nature 590(5806): 470–474
29. Liu Y, Cui Z (2011) Complete mitochondrial genome of the Chinese spiny lobster *Panulirus stimpsoni* (Crustacea: Decapoda): genome characterization and phylogenetic considerations. Mol Biol Rep 38(1):403–410
30. Tsang LM, Ma KY, Ahyong ST, Chan T-Y, Chu KH (2008) Phylogeny of Decapoda using two nuclear protein-coding genes: origin and evolution of the Reptantia. Mol Phylogenet Evol 48: 359–368
31. Qian GH, Zhao Q, Wang A, Zhu L, Zhou K, Sun H (2011) Two new decapod (Crustacea, Malacostraca) complete mitochondrial

genomes: bearings on the phylogenetic relationships within the Decapoda. Zool J Linn Soc 162(3):471–481
32. Place AR, Feng X, Steven CR, Fourcade HM, Boore JL (2005) Genetic markers in blue crabs (*Callinectes sapidus*) II: complete mitochondrial genome sequence and characterization of variation. J Exp Mar Biol Ecol 319(1–2):15–27
33. Kilpert F, Podsiadlowski L (2010) The mitochondrial genome of the Japanese skeleton shrimp *Caprella mutica* (Amphipoda: Caprellidea) reveals a unique gene order and shared apomorphic translocations with Gammaridea. Mitochondrial DNA 21(3–4): 77–86
34. Ito A, Aoki MN, Yokobori S, Wada H (2010) The complete mitochondrial genome of *Caprella scaura* (Crustacea, Amphipoda, Caprellidea), with emphasis on the unique gene order pattern and duplicated control region. Mitochondrial DNA 21(5): 183–190
35. Liu Y, Cui Z (2010) Complete mitochondrial genome of the Asian paddle crab *Charybdis japonica* (Crustacea: Decapoda: Portunidae): gene rearrangement of the marine brachyurans and phylogenetic considerations of the decapods. Mol Biol Rep 37(5):2559–2569
36. Miller AD, Nguyen TT, 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
37. Kilpert F, Podsiadlowski L (2010) The Australian fresh water isopod (Phreatoicidea: Isopoda) allows insights into the early mitogenomic evolution of isopods. Comp Biochem Physiol Part D Genomics Proteomics 5(1):36–44
38. Sun H, Zhou K, Song D (2005) Mitochondrial genome of the Chinese mitten crab *Eriocheir japonica sinensis* (Brachyura: Thoracotremata: Grapsoidea) reveals a novel gene order and two target regions of gene rearrangements. Gene 349:207–217
39. Shen X, Sun M, Wu Z, Tian M, Cheng H, Zhao F, Meng X (2009) The complete mitochondrial genome of the ridgetail white prawn *Exopalaemon carinicauda* Holthuis, 1950 (Crustacean: Decapoda: Palaemonidae) revealed a novel rearrangement of tRNA genes. Gene 437(1–2):1–8
40. Peregrino-Uriarte AB, Varela-Romero A, Muhlia-Almazan A, Anduro-Corona I, Vega-Heredia S, Gutierrez-Millan LE, De la Rosa-Velez J, Yepiz-Plascencia G (2009) The complete mitochondrial genomes of the yellowleg shrimp *Farfantepenaeus californiensis* and the blue shrimp *Litopenaeus stylirostris* (Crustacea: Decapoda. Comp Biochem Physiol Part D Genomics Proteomics 4(1):45–53
41. Yang J-S, Nagasawa H, Fujiwara Y, Tsuchida S, Yang W-J (2010) The complete mitogenome of the hydrothermal vent crab *Gandalfus yunohana* (Crustacea: Decapoda: Brachyura): a link between the Bythograeoidea and Xanthoidea. Zool Scr 39(6): 621–630
42. Segawa RD, 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
43. Ivey JL, Santos SR (2007) The complete mitochondrial genome of the Hawaiian anchialine shrimp *Halocaridina rubra* Holthuis, 1963 (Crustacea: Decapoda: Atyidae). Gene 394(1–2):35–44
44. Miller AD, Austin CM (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(3):565–574
45. Kilpert F, Podsiadlowski L (2006) The complete mitochondrial genome of the common sea slater, *Ligia oceanica* (Crustacea, Isopoda) bears a novel gene order and unusual control region features. BMC Genomics 7:241
46. Miller AD, Murphy NP, Burridge CP, Austin CM (2005) Complete mitochondrial DNA sequences of the decapod crustaceans *Pseudocarcinus gigas* (Menippidae) and *Macrobrachium rosenbergii* (Palaemonidae). Mar Biotechnol 7(4):339–349
47. Yamauchi MM, Miya MU, Machida RJ, Nishida M (2004) PCR-based approach for sequencing mitochondrial genomes of decapod crustaceans, with a practical example from kuruma prawn (*Marsupenaeus japonicus*). Mar Biotechnol 6(5):419–429
48. Bauza-Ribot MM, Jaume D, Juan C, Pons J (2009) The complete mitochondrial genome of the subterranean crustacean *Metacrangonyx longipes* (Amphipoda): a unique gene order and extremely short control region. Mitochondrial DNA 20(4):88–99
49. Liu Y, Cui Z (2010) The complete mitochondrial genome of the mantid shrimp *Oratosquilla oratoria* (Crustacea: Malacostraca: Stomatopoda): novel non-coding regions features and phylogenetic implications of the Stomatopoda. Comp Biochem Physiol Part D Genomics Proteomics 5(3):190–198
50. Hickerson MJ, Cunningham CW (2000) Dramatic mitochondrial gene rearrangements in the hermit crab *Pagurus longicarpus* (Crustacea, anomura). Mol Biol Evol 17(4):639–644
51. Yamauchi M, Miya M, Nishida M (2002) Complete mitochondrial DNA sequence of the Japanese spiny lobster, *Panulirus japonicus* (Crustacea: Decapoda). Gene 295(1):89–96
52. 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(6):863–874
53. Yamauchi MM, Miya MU, Nishida M (2003) Complete mitochondrial DNA sequence of the swimming crab, *Portunus trituberculatus* (Crustacea: Decapoda: Brachyura). Gene 311:129–135
54. Yang JS, Nagasawa H, Fujiwara Y, Tsuchida S, Yang WJ (2008) The complete mitochondrial genome sequence of the hydrothermal vent galatheid crab *Shinkaia crosnieri* (Crustacea: Decapoda: Anomura): a novel arrangement and incomplete tRNA suite. BMC Genomics 9:257
55. Cook CE, Yue Q, Akam M (2005) Mitochondrial genomes suggest that hexapods and crustaceans are mutually paraphyletic. Proc Biol Sci 272(1569):1295–1304
56. Ki JS, Dahms HU, Hwang JS, Lee JS (2009) The complete mitogenome of the hydrothermal vent crab *Xenograpsus testudinatus* (Decapoda, Brachyura) and comparison with brachyuran crabs. Comp Biochem Physiol Part D Genomics Proteomics 4(4): 290–299