

RESEARCH

Open Access



In silico prospecting of the mtDNA of *Macrobrachium amazonicum* from transcriptome data

Jerônimo Costa Marques-Neto¹, Gabriel Monteiro de Lima¹, Carlos Murilo Tenório Maciel^{1,2}, Bruna Ramalho Maciel², Fernando Araujo Abrunhosa³, Iracilda Sampaio⁴ and Cristiana Ramalho Maciel^{1,2*}

Abstract

Background *Macrobrachium amazonicum* is a freshwater prawn widely distributed in South America that is undergoing speciation, so the denomination “*M. amazonicum* complex” is used for it. The mitochondrial cytochrome c oxidase subunit I (COI) gene has been used to elucidate this speciation, but heteroplasmies and pseudogenes have been recorded, making separation difficult. Obtaining genes from cDNA (RNA) rather than genomic DNA is an effective tool to mitigate those two types of occurrences. The aim of this study was to assemble *in silico* the mitochondrial DNA (mtDNA) of the Amazonian coastal population of *M. amazonicum* inhabiting the state of Pará.

Results Sequences were obtained from the prawn’s transcriptome using the *de novo* approach. Six libraries of cDNA from the androgen gland, hepatopancreas, and muscle tissue were used. The mtDNA of *M. amazonicum* was 14,960 bp in length. It contained 13 protein-coding genes, 21 complete transfer RNAs, and the 12S and 16S subunits of ribosomal RNA. All regions were found on the light strand except tRNA^{Gln}, which was on the heavy strand. The control region (D-loop) was not recovered, making for a gap of 793 bp. The cladogram showed the formation of the well-defined *Macrobrachium* clade, with high support value in the established branches (91–100). The three-dimensional spatial conformation of the mtDNA-encoded proteins showed that most of them were mainly composed of major α -helices that typically shows in those proteins inserted in the membrane (mitochondrial).

Conclusions It was possible to assemble a large part of the mitochondrial genome of *M. amazonicum* *in silico* using data from other genomes deposited in GenBank and to validate it through the similarities between its COI and 16S genes and those from animals of the same region deposited in GenBank. Depositing the *M. amazonicum* mtDNA sequences in GenBank may help solve the taxonomic problems recorded for the species, in addition to providing complete sequences of candidate coding genes for use as biomarkers in ecological studies.

Keywords Freshwater prawn, Mitogenome, Coding regions, Molecular analysis

*Correspondence:

Cristiana Ramalho Maciel
macielufpa@gmail.com

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

The freshwater prawn *Macrobrachium amazonicum* (Heller, 1862) has a wide distribution across South America and adapts well to coastal and continental water bodies, including channels, lakes, rivers, and reservoirs [1–4]. Anger (2013) [5] attributes to the variations between populations a species-complex status – the *M. amazonicum* complex – which is undergoing speciation. Studies along these lines have addressed both morphological aspects and molecular biology, but the separation of this complex has not yet been fully established [6–9].

To date, only the population of the Brazilian Pantanal is considered separate due to molecular evidence, and peculiarities in its ontogenetic development and morphology [7, 9]. It is considered a new species called *M. pantanalense* [7]. Recently, it was found that specimens of *M. amazonicum* collected in the Tietê River (São Paulo state) and *M. pantanalense* collected in the Baiazinha Lagoon (Miranda, Mato Grosso do Sul state—Pantanal) could not copulate, reinforcing the speciation hypothesis for this population [10]. For the others, their speciation status is not yet known, and they continue to be grouped in the *M. amazonicum* complex. In an attempt to better describe the *M. amazonicum* complex, fragments of only two mitochondrial genes have been used: cytochrome *c* oxidase subunit I and the 16S ribosomal RNA gene [6, 9], and there is no record of a complete mitochondrial genome.

Most prawn mitochondrial genomes present in GenBank belong to other genera. Only five complete mitogenomes have been registered for *Macrobrachium*, of which three were published in the literature [11–13]. Obtaining the mitochondrial genome of the Amazon river prawn is one viable way to start answering questions about the taxonomy of the species.

Mitochondrial DNA (mtDNA) is located in the inner membrane matrix and is similar to a bacterial chromosome, being circular and highly compact [14]. It has uniparental inheritance and undergoes changes in its nucleotide composition at a relatively rapid rate [15–17]. It differs in size between eukaryotic species but usually has the same total number of regions, 37, of which 13 contain genes responsible for protein coding (cytochrome oxidase subunit I, II and III; ATP synthase subunits 6 and 8; nicotinamide adenine dinucleotide dehydrogenase (NADH) subunits 1, 2, 3, 4, 5, 6, and 4L; and cytochrome *b*), 22 encode transfer RNAs (tRNAs), and two encode ribosomal RNA subunits (one large, 16S, and one small, 12S) [18, 19]. Some of these genes are used as bioindicators of the presence of chemical compounds in the body, whereas others can be used as tools for phylogenetic analysis, arousing interest in obtaining the complete mtDNA genomes from many species [20, 21].

Molecular analyses using mitochondrial genes have helped solve some taxonomic problems, including in prawn species of the genus *Macrobrachium* Spence Bate, 1868, because these species are very diverse, have a wide distribution, and have complex evolutionary histories [5, 22]. There are records of phenotypic variations, causing problems in the separation of species within the genus [23–25].

There are several ways to obtain the complete mtDNA molecule, the most traditional of which is DNA purification, using specific primers, followed by Sanger sequencing [26]. Over the past 10 years, the number of functional genome (transcriptome) libraries of various species deposited in public databases has increased [27, 28]. Recently, the mitochondrial genome of several species was obtained through in silico prospecting in databases of mitochondrial RNA sequenced in transcriptomes, using bioinformatics tools [29, 30]. The use of transcriptome data linked to the de novo method has aroused interest as an assembly tool due to its speed and accuracy when editing mitogenomes [31]. *M. amazonicum* of the Amazonian coast has a sequenced functional genome, allowing the search for its mitochondrial genes in databases. The aim of this study is to assemble the mtDNA genome of this population using the de novo assembly approach in silico.

Methods

Sampling

The databases of six cDNA libraries were obtained from androgen gland (AG), hepatopancreas (HEP), and muscle (MUS) tissue and sequenced on the Illumina HiSeq 2500 platform (Illumina, San Diego, CA, USA). Each library was constructed from a pool of 10 animals, with two replicates each. *M. amazonicum* has evident and significant genetic divergence among populations, structuring them in three groups: I- inland waters of the Amazonian Hydrographic Region (HR); II- Paraná/Paraguay HR; and III- coastal systems of northern and northeastern Brazil [6]. The animals used were descendants of the individuals collected from the third population (01°12′37.7″S, 46°08′17.1″W).

mtDNA assembly and data analysis

The *M. amazonicum* transcriptome databases were assembled using Trinity assembler [32] without performing the filtering step that removes ribosomal genes because there are genes of this category in the mtDNA. Filtering was performed to remove low-quality sequences (<20) using FastQC [33] (which analyses the sequences from the transcriptome) and Trimmomatic 0.39 [34] (which is used for trimming bases and sequences). The assembly was performed by the de novo method

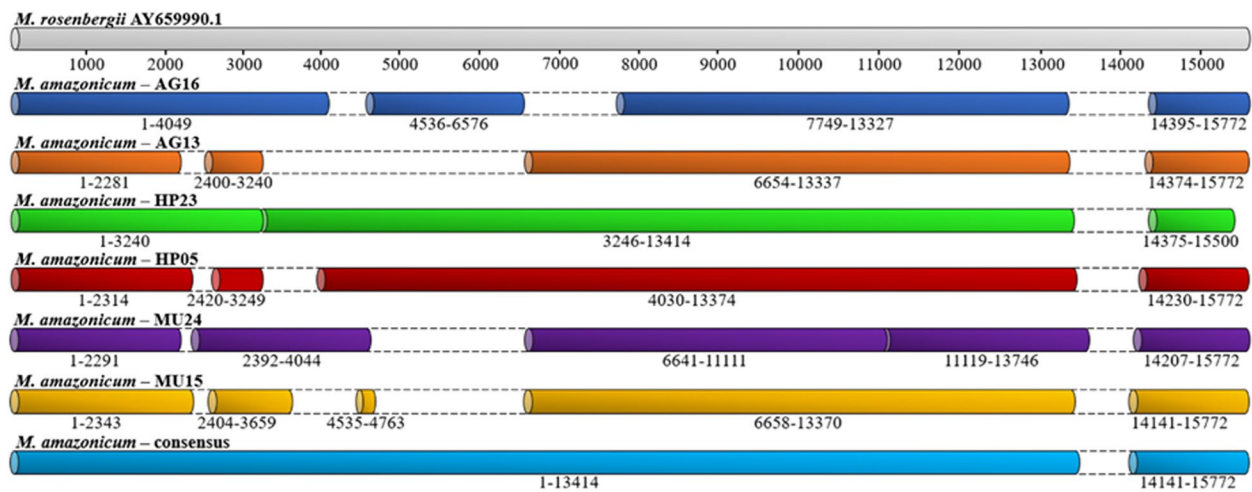


Fig. 1 Databases used in the assembly of the mitogenome of *M. amazonicum*. The genome of the reference species is shown (top, in light gray); the androgen gland (AG), hepatopancreas (HP) and muscle (MU) libraries; and the final consensus sequence (light blue) with the recovered transcripts. The bp size is shown. Unrecovered sites are represented by white space between dashed lines in the databases and in the consensus sequence

without a reference genome. Then, the mtDNA regions were searched for, using as reference the mitochondrial genome of the *M. rosenbergii* congener (NCBI AY659990.1), which is phylogenetically close to the target species.

The MEB program (Maciel, 2015 – not published) was used to search for sequences with higher similarities using the BLASTn tool, in which the six databases of the *M. amazonicum* transcriptome were compared against the 37 mtDNA sequences of *M. rosenbergii*. The results were visualized in Notepad + v. 7.8.4 [35].

The sequences identified as possible regions of the mitogenome of *M. amazonicum* were then automatically aligned in BioEdit v. 7.2.5 [36] with the aid of the ClustalW tool [37], taking the complete genome of *M. rosenbergii* as a reference. To determine the positions of the mtDNA regions identified in the Amazon river prawn, in order to construct a consensus and complete mitogenome, all sequences previously identified with the MEB program were used.

The start/stop codon and amino acid composition of the 13 protein-coding genes (PCGs) were identified by the ORFfinder tool. The content of the nitrogenous bases was automatically quantified in BioEdit v. 7.2.5. The closed circular representation of the mtDNA was created using the online tool Genome Vx (<http://wolfe.ucd.ie/GenomeVx/>) and was edited in Adobe Photoshop CC 2018 (version 19.1.6.).

Two divergence matrices were generated. In the first, ~85% of the mitogenome of seven different prawn species was used, in which the cutting position occurred at the beginning of the control region (not recovered

in the *M. amazonicum* transcriptome). In the second matrix, the sequences of the protein-coding regions were used to quantify the differences between *M. amazonicum* and five other species of the genus. This analysis was based on the model of Kimura (1980) [38] and was performed with the aid of MEGA X software [39]. The maximum likelihood tree was also generated in MEGA X using the GTR evolutionary model with 1,000 pseudoreplicates, and ~85% of the mitochondrial genome of the prawn species was also used. The three-dimensional structures of each PCG were drawn in SWISS-MODEL software [40] from the amino acid sequence using ENDscript 2.0 to identify the α -helices and β -sheets [41] and PyMOL [42] for later editing.

Results and discussion

General arrangement of the mitochondrial genome

The use of the six transcriptome databases (in silico), combined with the de novo assembly method, optimized the recovery of more than 90% of the *M. amazonicum* mitogenome (Fig. 1), yielding a 14,960-bp fragment, indicating that this is a satisfactory way to assemble mitochondrial genomes. In addition, this method offers cost savings, agility, and good coverage of regions because it can be performed by mining databases deposited in GenBank or similar databases [43–45].

The total adenine and thymine (A+T) content was higher (61.33%) than that of guanine and cytosine G+C (38.67%) in the mtDNA. The percentages of the four bases in the whole mitogenome and its parts are listed in Table 1. The A+T content was similar to that of the mitogenome of the congener species used for

Table 1 Percentage of nitrogenous bases by regions in the mitochondrial genome of *M. amazonicum*

<i>M. amazonicum</i>	Size (bp)	A%	C%	G%	T%	A + T%	G + C%
Mitogenome	14,960	36.21	25.49	13.18	25.12	61.33	38.67
PCGs	11,196	36.02	26.97	12.80	24.21	60.23	39.77
tRNAs	1418	34.56	20.58	14.63	30.24	64.80	35.20
rRNA	2143	39.41	21.39	10.18	29.03	68.44	31.57

A, C, G, and T contents in the whole mitochondrial genome of *M. amazonicum* and in its protein-coding regions (PCGs), transfer RNAs (tRNAs), and ribosomal RNAs (rRNA). The noncoding spaces are accounted for only in the value of the complete genome, and gene overlaps were removed from the final values

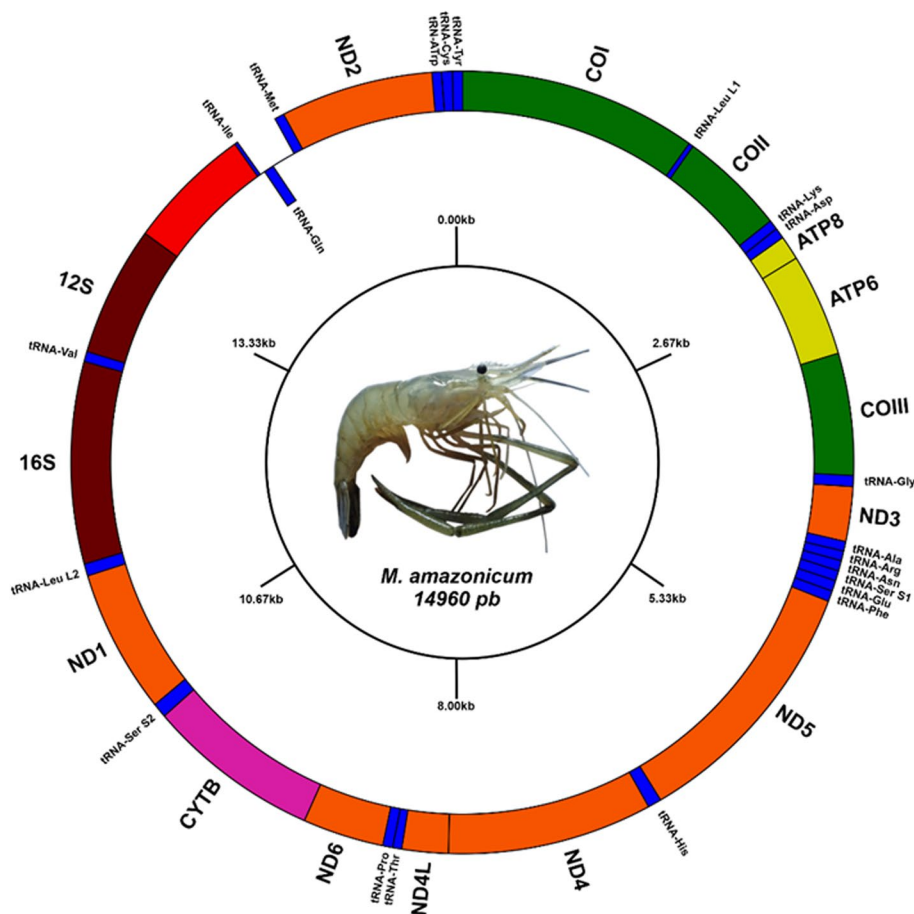


Fig. 2 Gene map showing the organization of the mitogenome regions of *M. amazonicum*. The three cytochrome c oxidase (COI/II/III) genes are in green, ATP6 and ATP8 in yellow, the six NADH subunits (1, 2, 3, 4, 4L, 5, 6) in orange, cytochrome *b* in pink, the two ribosomal RNA subunits in brown (16S and 12S), and the transfer RNAs in blue. The latter are differentiated by the three-letter codes corresponding to the amino acid name. The heavy strand is denoted by the position of tRNA^{Gln}. The light strand has most of the genes. The noncoding sites are spaced and not filled. The middle circle indicates the kilobase (kb) values, and the inner circle indicates the total number of bases recovered from the mtDNA of *M. amazonicum*

comparison, suggesting it may be a characteristic of this genus [11–13].

All 13 regions encoding common proteins in eukaryotes were found, as were 21 complete tRNAs and two rRNA subunits (12S and 16S) (Fig. 2). The noncoding control region (D-loop) was not recorded in any of the

six transcriptomes analysed, and nor were the 41 bases at the beginning of the tRNA^{Ile} that followed the site of the D-loop, together making up an unrecovered 793-bp portion (Fig. 1). Moreira et al. (2015) [46] previously reported the absence of the D-loop in an assembly made from transcriptomic data in a fish species of the family

Loricariidae. According to those authors, this was due to the low number of supporting reads of the region in the transcriptomic data, which was directly related to the role of the D-loop in replication/transcription, culminating in its absence from mature transcripts. Therefore, our results do not deviate from the pattern identified in the literature; however, this region can be recovered using specific primers in short-range PCR and subsequently sequenced by the Sanger method [47].

Region locations and characteristics

The light strand encoded all regions except tRNA^{Gly}, which was located in the heavy strand (H) (Fig. 1 and Table 2). As described by Anderson et al. (1981) [48], this condition is related to the amounts of G and T in each of the strands; the more of these two bases there are, the heavier it will be. This difference in the percentage of bases between the strands is attributed to the replication process because this step occurs asynchronously, which facilitates the emergence of mutations when nucleotides are added or removed at the end of the process [49, 50]. In the present study, the percentage of bases revealed a prevalence of A and C in the mtDNA of *M. amazonicum*, explaining the presence of most regions in the light strand.

Regarding the intergenic regions, seven gene overlaps were found between regions, the largest 29 nt (tRNA^{Ile} – tRNA^{Gln}) and the smallest 1 nt (Table 2). Small non-coding spaces were observed throughout the mitogenome, the most extensive being between the tRNA^{Gln} and tRNA^{Met} regions (181 nt) and the smallest located between tRNA^{Ile} and tRNA^{Gln} (28 nt) (Table 2). Both overlaps and noncoding spaces are characteristics observed in the mtDNA of many crustaceans [51–53] but are not exclusive to the phylum. The lack of synchrony of the strands during the replication process, generating nucleotide sequences that do not encode any transcript, is well documented in other groups [54–56]. In addition, the small size of these fragments indicates that they do not play an important role in the mtDNA, unlike the control region, which has the key role of holding the necessary promoters for initiating transcription/replication [57].

The total length of all PCGs was 11,196 bp, and the A + T content of this region was 60% (Table 1). The longest gene was ND5: 1707 bp, located between tRNA^{Phe} and tRNA^{His}, while the shortest was ATP8: 168 bp, located between tRNA^{Asp} and ATP6. The other 11 genes were ~300 to 1560 bp (Table 2). This pattern has also been observed in the mitogenomes of the five species of the genus *Macrobrachium*, in which the largest and smallest regions were also ND5 and ATP8, respectively. The lengths of these two genes were also similar between

these species, as were the sites that flanked these genes; this was not observed in other shrimp genera, corroborating the existence of a pattern particular to the genus *Macrobrachium* [11, 12, 58].

The space occupied by the tRNA genes was 1418 bp. The content of the A + T bases was higher (65%) than that of G + C (35%) in these regions (Table 1). The size of most tRNA genes ranged from ~60 to 78 bp, and the smallest was 30 bp, that of tRNA^{Leu (L1)} (Tables 1 and 2). When comparing this tRNA with that of the reference species, we observed that the sizes were similar; however, when searching for the amino acid composition of the region preceding this site (COI), we found that part of what was considered Leu L1 included the amino acid triplet equivalent to the stop codon of COI, so its total number of nucleotides was lower than that of *M. rosenbergii*.

The two ribosomal RNA subunit genes had a size of 2143 bp, the 16S gene being longer than the 12S (1297 and 846 bp, respectively). These were separated by tRNA^{Val}. Both were located on the light strand and had 68% A + T content (Tables 1 and 2). These findings were similar to those in other *Macrobrachium* mitogenomes, reinforcing the idea of a pattern in the distribution of the genes thus far observed for the genus. These similarities suggest the absence of the gene rearrangement phenomenon for *Macrobrachium*, so the regions are aligned at the same position between species, suggesting that there are no cases of mispairing between the strands [59]. To date, there are only five congeneric species with fully sequenced mtDNA genomes, including the one here. When these data are expanded, we can determine whether these traits are actually conserved throughout the genus.

Start and stop codons

The start codon ATG, encoding the amino acid methionine, was present at the beginning of eight genes, and the start of transcription is attributed to this triplet, as it is usually considered the start codon of the transcription process, both in vertebrates and in invertebrates [55, 60–62]. Alternative start codons (threonine, leucine, tyrosine, and isoleucine) were recorded in the other genes (Table 2), threonine starting two of the genes: COI and ND6 (ACx codons). Records of these alternative start codons are frequent in the mitochondrial genomes of malacostracans, including the *Macrobrachium* sequences used for comparison in the present study [11, 12, 58, 63–68]. In addition, the consistency of these data between *M. amazonicum* and the six replicate libraries used in the assembly of the mtDNA was verified, showing that the start codons were not random phenomena due to possible inaccuracies in the in silico

Table 2 Gene components of mitochondrial DNA of *M. amazonicum*

Gene	Position	Size (bp)	Codon (aa)		Strand (L=light, H=heavy)	aa ¹	In ²
			Start (aa)	Stop			
COI	1–1560	1560	ACG (T)	TAG	L	519	0
tRNA ^{Leu (L1)}	1561–1590	30			L		0
COI	1591–2286	696	CTT (L)	TAA	L	231	0
tRNA ^{Lys}	2285–2354	68			L		0
tRNA ^{Asp}	2355–2415	61			L		0
ATP8	2416–2583	168	TAT (Y)	TAA	L	55	-7
ATP6	2577–3251	675	ATG (M)	TAA	L	224	4
COIII	3256–4044	789	ATG (M)	TAA	L	262	0
tRNA ^{Gly}	4045–4115	71			L		0
ND3	4116–4481	366	ATG (M)	TAG	L	121	-10
tRNA ^{Ala}	4472–4534	63			L		-1
tRNA ^{Arg}	4534–4599	66			L		1
tRNA ^{Asn}	4601–4665	65			L		0
tRNA ^{Ser (S1)}	4666–4732	67			L		2
tRNA ^{Glu}	4733–4801	69			L		-1
tRNA ^{Phe}	4801–4865	65			L		0
ND5	4866–6572	1707	ATG (M)	TAA	L	568	0
tRNA ^{His}	6573–6655	83			L		0
ND4	6656–7983	1335	ATG (M)	TAA	L	452	-7
ND4L	7984–8283	300	ATG (M)	TAA	L	99	0
tRNA ^{Thr}	8284–8338	55			L		-1
tRNA ^{Pro}	8338–8403	66			L		0
ND6	8404–8931	528	ACT (T)	TAA	L	175	-1
Cytb	8931–10067	1137	ATG (M)	TGA	L	378	0
tRNA ^{Ser (S2)}	10068–10158	85			L		2
ND1	10153–11091	939	ATG (M)	TAG	L	312	0
tRNA ^{Leu (L2)}	11092–11169	78			L		0
16S	11170–12466	1297			L		0
tRNA ^{Val}	12467–12,532	66			L		0
12S	12533–13378	846			L		--
D-loop	--	--			L		--
tRNA ^{Ile*}	13403–13424	22			L		-29
tRNA ^{Gln}	13452–13518	67			H		181
tRNA ^{Met}	13698–13767	70			L		0
ND2	13768–14763	996	ATT(I)	TAA	L	331	-2
tRNA ^{Trp}	14762–14830	69			L		1
tRNA ^{Cys}	14832–14894	63			L		0
tRNA ^{Tyr}	14895–14960	66			L		0

Listed are the position; size (bp); codons, in which the amino acid triplet is followed by a letter in parentheses indicating the first amino acid in each sequence: M (methionine), T (threonine), L (leucine), D (aspartic acid), I (isoleucine), Y (tyrosine); strand (light or heavy); number of amino acids (aa¹) and intergenic nucleotides (In²), where positive values indicate noncoding spaces in the mitogenome and negative values indicate the overlap of two regions. The asterisk (*) indicates the incomplete tRNA

assembly. Kearse and Wilusz (2017) [69] noted the relevance of this type of event, since translation can be initiated by amino acids other than methionine, which is essential for the regulation of essential processes of protein synthesis.

All 13 PCGs showed functional stop codons (TAG and TAA) (Table 2). This indicates that the release factor (eRF1), which triggers a series of processes that punctuate the end of translation [70], acts on all genes of the *M. amazonicum* mitogenome.

Table 3 Divergence matrix between species of *Macrobrachium* and *Caridian* prawn

	1	2	3	4	5	6	7	8
<i>M. amazonicum</i>								
<i>M. rosenbergii</i> AY659990.1	0.21							
<i>M. rosenbergii</i> KY865098.1	0.21	0.08						
<i>M. nipponense</i> HQ830201.1	0.21	0.21	0.21					
<i>M. lanchesteri</i> FJ797435.1	0.22	0.17	0.17	0.20				
<i>M. bullatum</i> KM978918.1	0.21	0.21	0.21	0.18	0.21			
<i>E. carinicauda</i> EF560650.1	0.35	0.36	0.36	0.35	0.35	0.35		
<i>R. variabilis</i> MH714460.1	0.36	0.37	0.37	0.35	0.36	0.36	0.39	

Mitogenome of *M. amazonicum* with five other congeners + two species of Caridean prawn, using ~85% of the mitochondrial genome of each species (cut after the control region, which was not recovered for *M. amazonicum*)

Divergence matrices and phylogenetic analysis

In the divergence matrix constructed with ~85% of the total mitochondrial genome, a divergence pattern within the genus *Macrobrachium* was observed, defined as 17–22%, with the lowest divergence values between species occurring in *M. rosenbergii* vs. *M. lanchesteri* (17%) and *M. nipponense* vs. *M. bullatum* (18%). *M. amazonicum* showed the highest divergence levels of all the species of the group, with 21% divergence from *M. rosenbergii*, *M. nipponense*, and *M. bullatum* and 22% divergence from *M. lanchesteri* (Table 3). These values were close to those found in the divergence matrix constructed with the coding genes alone, showing that the genetic distance between *M. amazonicum* and the other five congeners fits within the parameters considered to distinguish species within this genus [23, 24, 71]. Of the coding genes, ND6 (30%) showed the greatest divergence, and ND4L and COI the lowest (18% to 19%) (Table 4).

Population studies carried out among the three populations/clades identified for freshwater prawn, using partial regions of the mitochondrial COI gene, showed nucleotide divergences of around 0–3.3% and 0–0.1% for the COI and 16S rRNA genes, while divergences recorded at the interspecific level were 4.8–14.7% and 13.3–19.9% [6], respectively. Previous studies carried out in the *Macrobrachium* complex suggest a minimum divergence above

10% of the COI gene to distinguish species at the molecular level, identifying variations of 0–3.2% and 0–12.6% at the population level, for 16S rRNA and COI [23, 24], corroborating the findings for the three groupings of *M. amazonicum* in the literature [6].

Work carried out between different populations of the genus using partial regions of the mitochondrial COI gene show divergence percentages of 0 to ~10% [23, 24, 71], like those recorded for *M. rosenbergii* [72–75], populations of *M. nipponense* [76], *M. australiense* [77] and *M. amazonicum* [6, 78]. However, there is a complex of cryptic species recorded for the genus, necessitating caution when drawing conclusions [24, 79–82].

Another difficulty for phylogenetic studies and species separation in *Macrobrachium* is the existence of synonymies, heteroplasmies, and pseudogenes, which are relatively frequent [24, 83–85]. Numts are copies of the mtDNA in the nuclear DNA [86], and mitochondrial heteroplasmy are more than one mitochondrial genome in the same organism [87]. Numts not only make sequence analysis difficult, but sometimes the sequences may be erroneously adopted as genuine mtDNA sequences. Iketani et al. (2021) [85] frequently detected these events in individuals of the *M. amazonicum* complex, demonstrating the need to use more than one gene for analyses between different populations and the need for caution

Table 4 Divergence matrix of the 13 mitochondrial protein-coding genes identified in *M. amazonicum*

	<i>M. amazonicum</i>												
	COI	COII	ATP8	ATP6	COIII	ND3	ND5	ND4	ND4L	ND6	Cytb	ND1	ND2
<i>M. rosenbergii</i> AY659990.1	0.19	0.18	0.26	0.22	0.22	0.33	0.26	0.26	0.19	0.33	0.21	0.19	0.29
<i>M. rosenbergii</i> KY865098.1	0.19	0.20	0.28	0.20	0.21	0.26	0.26	0.25	0.18	0.31	0.21	0.20	0.29
<i>M. nipponense</i> HQ830201.1	0.18	0.20	0.26	0.26	0.19	0.27	0.27	0.24	0.19	0.32	0.17	0.21	0.32
<i>M. lanchesteri</i> KM978918.1	0.21	0.21	0.37	0.27	0.22	0.30	0.26	0.23	0.19	0.34	0.22	0.22	0.32
<i>M. bullatum</i> FJ797435.1	0.17	0.20	0.37	0.22	0.19	0.27	0.27	0.25	0.20	0.37	0.21	0.19	0.31

The genes of *M. amazonicum* were compared with the same genes of the five species of the genus that have a complete mitochondrial genome deposited in GenBank. The regions are in the published mitogenomes of the listed species. The adopted values range from 0 to 1: the closer to 1, the greater the divergence

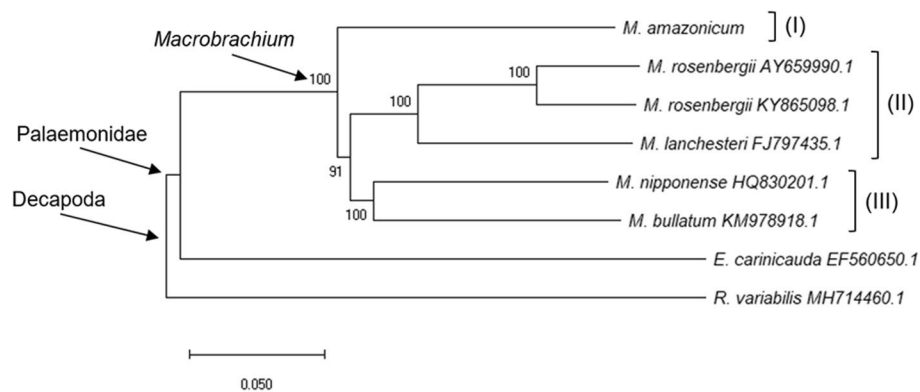


Fig. 3 Cladogram showing the relationships of the group Decapoda based on mitogenomic data with 1,000 pseudo replicates. The numbers in the branches indicate the bootstrap values. The GenBank accession code is next to each species

in drawing conclusions based exclusively on the divergences recorded for the COI gene. To circumvent the effect of Numts, the use of mtDNA-rich tissue, mtDNA enrichment, dilution of template DNA, protein-coding regions, long PCR, or cDNA analysis has been recommended [85, 88]. Kang et al. (2016) [89], when using cDNA (RNA) sequences from the orthopteran species *Anapodisma miramae*, they found that heteroplasmies and pseudogenes, common in the species, had practically disappeared. In the present study, all mtDNA genes found originated from RNA purifications, increasing the reliability of the results.

When the regions of the COI and 16S rRNA genes recorded in the mitogenome assembled in the present study were compared with the sequences generated by Vergamini et al. (2011) [6], we found ~99 to 100% similarity with the haplotypes corresponding to the population of Abaetetuba, Pará state (GU929471.1 and GU929450.1, respectively). These data show the accuracy of the in silico assembly of the mtDNA of *M. amazonicum* because the animals used for transcriptome sequencing came from offspring from nearby locations, in communicating water bodies. In this sense, the recovery of the 13 PCGs from the mitogenome of the coastal population of *M. amazonicum* (including COI, 16S rRNA, and 12S rRNA) may enable the use of alternative mtDNA genes in future population studies of the species, such as Cytb, ATP6, ATP8, COII, COIII, and ND2. These genes show conservation of their nucleotide content, in addition to consistent variation between different species [90–93], which can assist in the resolution of the incongruence and taxonomic problems of the *M. amazonicum* complex.

The cladogram showed the formation of the well-defined *Macrobrachium* clade, with high support value in the established inner branches (91–100), using maximum likelihood with a bootstrapping of 1,000

pseudoreplicates. *Macrobrachium* showed the formation of three groups, *M. amazonicum* (I), *M. rosenbergii* + *M. lanchesteri* (II), and *M. nipponense* + *M. bullatum* (III). The species *Exopalaemon carinicauda* formed a group external to the analysed genus, with 35–36% divergence from *Macrobrachium*, while *Rimicaris variabilis* was located external to the family Palaemonidae, showing 35–39% divergence from the other taxa (Fig. 3 and Table 3).

Due to the lack of standardization of the precise divergence percentages required for species separation, other aspects end up being considered, such as morphological and biological data. For the Amazon river prawn, three main populations have been established in Brazil (Amazon, North–Northeast coast, and Paraná–Paraguay watersheds), with maximum intraspecific divergences of 3.3% and 0.1% for the COI and 16S rRNA genes, respectively [6]. However, the Pantanal population (Paraná–Paraguay watershed), which presents ~3% divergence, was given a new species assignment, *Macrobrachium pantanalense* [7], taking into account mainly the variations in morphological, reproductive, and physiological aspects [8, 10, 94–97].

Three-dimensional protein structures

Most proteins had three-dimensional conformations rich in α -helices. Berg et al. (2014) [98] reported that the percentage of this type of structure can vary considerably in the final arrangement of a secondary composition. The three-dimensional spatial conformation of the mtDNA-encoded proteins showed that most of them were mainly composed of major α -helices. They show hydrophobic residues typical in those proteins inserted in the membrane (mitochondrial). In our study, some proteins were composed exclusively of this type of arrangement (COI, ATP6, ATP8, COIII, ND4,

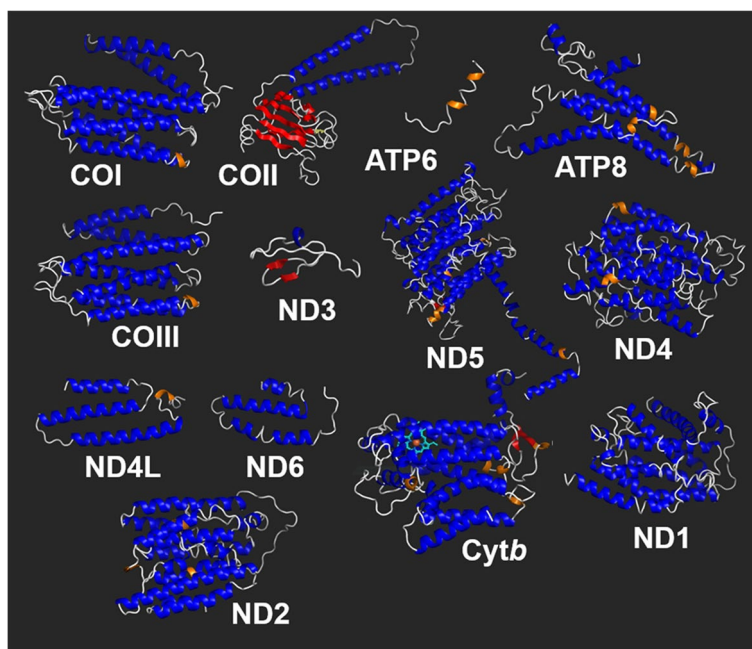


Fig. 4 Representations of the three-dimensional structures of the mitochondrial-encoded proteins. A prominent characteristic of the α -helices (blue and orange) is the folding of the polypeptide skeleton around the main chain, while the β -sheets (red) are characterized by hydrogen bonds between the polypeptide chains, presenting a flat and rigid pattern

ND4 L, ND6, ND1, and ND2), suggesting the existence of internal hydrogen bonds between the carboxyl (CO) end of an amino acid and the amine end (NH) of the other, both located in the main chain, which leads to a recognizable torsional pattern along its axis [99].

Some proteins also had β -sheets (COII, ND3, ND5, and Cytb). This arrangement is common and results from branches in the β -carbon atom of some amino acid, in which case α -helices would not be viable, as this would lead to steric collisions between amino acids with β -branching and NH-CO groups [100].

The most notable particularity was detected in Cytb, corresponding to the presence of the haem group (light blue), attached to an iron atom (orange circle) (Fig. 4). This is characteristic of Cytb and acts in the transfer of the electron of ubiquinone (coenzyme Q) to cytochrome *c* [101, 102], a vital process for cellular activity. The toxic agent cyanide has affinity to the iron binding site, where it interrupts energy production in the cell [103–105]. In addition, the response of mitochondria to stress and different types of pollutants [20, 106–108] make mtDNA a possible bioindicator of environmental quality [21]. Thus, the *M. amazonicum* mitogenome made available in the present study can be used not only for future genetic analyses but also for investigations of ecosystem data, aided by the fact that the species is widely distributed across South America [5].

Conclusions

The de novo assembly method using transcriptomic data yielded a satisfactory in silico assembly because it was possible to determine the composition, location, and distribution of the mitogenomic regions of the Amazon river prawn. The *M. amazonicum* mtDNA has features similar to that of similar species, but with significant interspecific divergences to distinguish them. The light strand encodes most regions, characterized by its high A + C content, which was one of the starkest differences from the other species of the genus, which have six or eight regions in the heavy strand. The complete recovery of proteins paves the way for using the mitochondrial genome of *M. amazonicum* as an indicator of environmental quality. The database deposited in GenBank, corresponding to the coastal population of the species, will aid in more refined studies on the distinct populations and speciation of the *M. amazonicum* complex.

Abbreviations

mtDNA	Mitochondrial DNA
tRNAs	RNA carriers
L	Light strand
H	Heavy strand
Ile	Isoleucine
Gln	Glutamine
Met	Methionine
Ala	Alanine
Leu L1	Leucine L1
Leu L2	Leucine L2

His	Histidine
Ser S1	Serine S1
Ser S2	Serine S2
ND2	NADH Dehydrogenase Subunit 2
ND3	NADH 3: NADH Dehydrogenase Subunit 3
ND4	NADH Dehydrogenase Subunit 4
ND4L	NADH Dehydrogenase Subunit 4L
ND5	NADH Dehydrogenase Subunit 5
ND6	NADH Dehydrogenase Subunit 6
COI	Cytochrome Oxidase C Subunit I
GCPs	Protein Encoding Genes
bp	Base pairs
CR	Control Region

Acknowledgements

We thank the Brazilian Agency for Support and Evaluation of Graduate Education (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES) and the Amazonia Research Support Foundation (Fundação Amazonia de Amparo à Estudos e Pesquisas – FAPESPA) for funding this study. We also thank the PRONEX Project (CNPq/Fapespa), as part of which the transcriptome used here was sequenced. As well as the support of PROPESP/UFPA (PAPQ).

Authors' contributions

Jerônimo Costa Marques-Neto participated in writing the manuscript and de novo genome assembly. Gabriel Monteiro de Lima prepared the library. Carlos Murilo Tenório Maciel participated in raw material screening and data refinement. Bruna Ramalho Maciel participated in the bioinformatic analyses as well as writing. Fernando Araujo Abrunhosa helped in capturing the animals used here, as well as in writing. Iracilda Sampaio and Cristiana Ramalho Maciel designed the study and contributed to the preparation of the manuscript, in addition to acquiring funding to fund the research. All authors read and approved the manuscript.

Funding

This work was supported by the Coordination for the Improvement of Higher Education Personnel – CAPES (financial code 001) and the Fundação Amazônia de Amparo à Estudos e Pesquisas – FAPESPA (grant number 001/2017 FAPESPA/SECTE). As well as the PRONEX Project (CNPq/Fapespa), through which the transcriptome used here was sequenced. PROPESP/UFPA (PAPQ) provided financial support regarding the manuscript publication fee (23073.080653/2023-19).

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the GenBank repository, <https://www.ncbi.nlm.nih.gov/nucleotide/ON513382.1>, under the accession number ON513382.1.

Declarations

Ethics approval and consent to participate

The authors declare that the preparation of the manuscript followed the ethical standards of scientific communication. The authors declare that the research that gave rise to the manuscript followed good ethical practices and that the necessary approvals from research ethics committees, when applicable, are described in the manuscript.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Laboratory of Aquaculture, Coastal Studies Institute, Federal University of Pará, Alameda Leandro Ribeiro S/N, Aldeia, Bragança, Pará CEP: 68600-000, Brazil. ²Coastal Studies Institute, School of Biological Sciences, Laboratory of Aquaculture/BioDatta, Federal University of Pará, Alameda Leandro Ribeiro S/N, Aldeia, Bragança, Pará CEP: 68600-000, Brazil. ³Coastal Studies Institute, School of Biological Sciences, Laboratory of Carcinology, Federal University of Pará, Alameda Leandro Ribeiro S/N, Aldeia, Bragança, Pará CEP: 68600-000,

Brazil. ⁴Coastal Studies Institute, Federal University of Pará, Alameda Leandro Ribeiro S/N, Aldeia, Bragança, Pará CEP: 68600-000, Brazil.

Received: 18 July 2023 Accepted: 27 October 2023

Published online: 10 November 2023

References

- Bentes B, Martinelli J, Souza L, Cavalcante D, Almeida M, Isaac V. Spatial distribution of the amazon river prawn *Macrobrachium Amazonicum* (Heller, 1862) (Decapoda, Caridea, Palaemonidae) in two perennial creeks of an estuary on the northern coast of Brazil (Guajará Bay, Belém, Pará). *Brazilian J Biol.* 2011;71:925–35.
- Costa TV da, Mattos LA de, Machado N de JB. Estrutura populacional de *Macrobrachium amazonicum* em dois lagos de várzea da Amazônia. *Bol do Inst Pesca.* 2016;42:281–93.
- Scaico MA. Fecundity and fertility of *Macrobrachium amazonicum* (Crustacea, Decapoda) at a Brazilian Northeastern dam. *Bol do Inst Pesca.* 1992;19 Único:89–96.
- Costa e Silva R, Cunha MC, Mossolin EC, Jacobucci GB. Population structure of *Macrobrachium amazonicum* (Heller, 1862) (Decapoda: Palaemonidae) in Miranda Hydroelectric Plant Reservoir, Araguari river, Minas Gerais, Brazil. *Acta Limnol Bras.* 2019;31:e14.
- Anger K. Neotropical *Macrobrachium* (Caridea: Palaemonidae): on the biology, origin, and radiation of freshwater-invading shrimp. *J Crustac Biol.* 2013;33:151–83.
- Vergamini FG, Pileggi LG, Mantelatto FL. Genetic variability of the Amazon river prawn *Macrobrachium amazonicum* (Decapoda, Caridea, Palaemonidae). *Contrib to Zool.* 2011;80:67–83.
- Dos Santos A, Hayd L, Anger K. A new species of *Macrobrachium Spence Bate*, 1868 (Decapoda, Palaemonidae), *M. pantanalense*, from the Pantanal, Brazil. *Zootaxa.* 2013;3700:534–46.
- Hayd L, Anger K. Reproductive and morphometric traits of *Macrobrachium amazonicum* (Decapoda: Palaemonidae) from the Pantanal, Brazil, suggests initial speciation. *Rev Biol Trop.* 2013;61:39–57.
- Weiss R, Anger K, Hayd L, Schubart CD. Interpreting genetic distances for species recognition: the case of *Macrobrachium amazonicum* Heller, 1862 and the recently described *M. pantanalense* Dos Santos, Hayd & Anger, 2013 (Decapoda, Palaemonidae) from Brazilian fresh waters. *Crustaceana.* 2015;88:1111–26.
- Nogueira CS, Perroca JF, Batista AC, Costa RC. Reproductive traits of the freshwater prawn *Macrobrachium amazonicum* (Decapoda: Palaemonidae) in an isolated water reservoir. *Rev Mex Biodivers.* 2020;91:e913387.
- Miller AD, Murphy NP, Burrridge CP, Austin CM. Complete mitochondrial DNA sequences of the decapod crustaceans *Pseudocarcinus gigas* (Menippidae) and *Macrobrachium rosenbergii* (Palaemonidae). *Mar Biotechnol.* 2005;7:339–49.
- Ma K, Feng J, Lin J, Li J. The complete mitochondrial genome of *Macrobrachium nipponense*. *Gene.* 2011;487:160–5.
- Li Y, Song J, Shen X, Cai Y, Cheng H, Zhang X, et al. The first mitochondrial genome of *Macrobrachium rosenbergii* from China: phylogeny and gene rearrangement within Caridea. *Mitochondrial DNA Part B.* 2019;4:134–6.
- Wallace DC. Mitochondrial DNA in evolution and disease. *Nature.* 2016;535:498–500.
- Evolution of the cetacean mitochondrial D-loop region. *Mol Biol Evol.* 1991. <https://doi.org/10.1093/oxfordjournals.molbev.a040662>.
- Marlow FL. Mitochondrial matters: Mitochondrial bottlenecks, self-assembling structures, and entrapment in the female germline. *Stem Cell Res.* 2017;21:178–86.
- Alexeyev M. Mitochondrial DNA: the common confusions. *Mitochondrial DNA Part A.* 2020;31:45–7.
- Brown TA, Cecconi C, Tkachuk AN, Bustamante C, Clayton DA. Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. *Genes Dev.* 2005;19:2466–76.
- Avise JC. Phylogeography: retrospect and prospect. *J Biogeogr.* 2009;36:3–15.
- Jayasundara N. Ecological significance of mitochondrial toxicants. *Toxicology.* 2017;391:64–74.

21. Roubicek DA, de Souza-Pinto NC. Mitochondria and mitochondrial DNA as relevant targets for environmental contaminants. *Toxicology*. 2017;391:100–8.
22. Wowor D, Muthu V, Meier R, Balke M, Cai Y, Ng PKL. Evolution of life history traits in Asian freshwater prawns of the genus *Macrobrachium* (Crustacea: Decapoda: Palaemonidae) based on multilocus molecular phylogenetic analysis. *Mol Phylogenet Evol*. 2009;52:340–50.
23. Liu M-Y, Cai Y-X, Tzeng C-S. Molecular systematics of the freshwater prawn genus *Macrobrachium* Bate, 1868 (Crustacea: Decapoda: Palaemonidae) inferred from mtDNA sequences, with emphasis on East Asian species. *Zool Stud*. 2007;46:272.
24. Pileggi LG, Mantelatto FL. Molecular phylogeny of the freshwater prawn genus *Macrobrachium* (Decapoda, Palaemonidae), with emphasis on the relationships among selected American species. *Invertebr Syst*. 2010;24:194.
25. Jose D, Harikrishnan M. Evolutionary history of genus *Macrobrachium* inferred from mitochondrial markers: a molecular clock approach. *Mitochondrial DNA Part A*. 2019;30:92–100.
26. Jia X, Xu S, Bai J, Wang Y, Nie Z, Zhu C, et al. The complete mitochondrial genome of *Somanniathelphusa boyangensis* and phylogenetic analysis of Genus *Somanniathelphusa* (Crustacea: Decapoda: Parathelphusidae). *PLoS ONE*. 2018;13: e0192601.
27. Soroka M, Rymaszewska A, Sańko T, Przyłucka A, Lubośny M, Śmietanka B, et al. Next-generation sequencing of *Dreissena polymorpha* transcriptome sheds light on its mitochondrial DNA. *Hydrobiologia*. 2018;810:255–63.
28. Choi B-S, Lee YH, Kim H-J, Hagiwara A, Lee J-S. Complete mitochondrial DNA of the marine water flea *Diaphanosoma celebensis* (Cladocera, Sidaidae). *Mitochondrial DNA Part B*. 2020;5:2254–5.
29. Kumar A, Chordia N. In silico PCR primer designing and validation. *PCR Prim Des*. 2015;1275:143–51.
30. Bacalhau M, Pratas J, Simões M, Mendes C, Ribeiro C, Santos MJ, et al. In silico analysis for predicting pathogenicity of five unclassified mitochondrial DNA mutations associated with mitochondrial cytopathies' phenotypes. *Eur J Med Genet*. 2017;60:172–7.
31. Forni G, Puccio G, Bourguignon T, Evans T, Mantovani B, Rota-Stabelli O, et al. Complete mitochondrial genomes from transcriptomes: assessing pros and cons of data mining for assembling new mitogenomes. *Sci Rep*. 2019;9:14806.
32. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29:644–52.
33. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed 12 Feb 2019.
34. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20.
35. Notepad++ (Version 7.8.6) [<https://notepad-plus-plus.org/downloads/>].
36. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser*. 1999;41:95–8.
37. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*. 1997;25:4876–82.
38. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol*. 1980;16:111–20.
39. Kumar S, Stecher G, Li M, Kryaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol*. 2018;35:1547–9.
40. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis*. 1997;18:2714–23.
41. Gouet P. ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res*. 2003;31:3320–3.
42. Janson G, Zhang C, Prado MG, Paiardini A. PyMod 2.0: improvements in protein sequence-structure analysis and homology modeling within PyMOL. *Bioinformatics*. 2017;33:444–6.
43. Gillett CPDT, Crampton-Platt A, Timmermans MJTN, Jordal BH, Emerson BC, Vogler AP. Bulk De Novo Mitogenome Assembly from Pooled Total DNA Elucidates the Phylogeny of Weevils (Coleoptera: Curculionidae). *Mol Biol Evol*. 2014;31:2223–37.
44. Sharma S, Singhal A. De-Novo Assembly of Short Reads in Minimal Overlap Model. In: *Proceedings of the International Conference on Bioinformatics Models, Methods and Algorithms*. Lisbon-Portugal: SCITEPRESS - Science and Technology Publications; 2015. p. 44–54.
45. Timmermans MJTN, Viberg C, Martin G, Hopkins K, Vogler AP. Rapid assembly of taxonomically validated mitochondrial genomes from historical insect collections. *Biol J Linn Soc*. 2016;117:83–95.
46. Moreira DA, Furtado C, Parente TE. The use of transcriptomic next-generation sequencing data to assemble mitochondrial genomes of *Ancistrus* spp. (Loricariidae). *Gene*. 2015;573:171–5.
47. Cheng Y-Z, Xu T-J, Jin X-X, Tang D, Wei T, Sun Y-Y, et al. Universal primers for amplification of the complete mitochondrial control region in marine fish species. *Mol Biol*. 2012;46:727–30.
48. Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. *Nature*. 1981;290:457–65.
49. Tanaka M, Ozawa T. Strand Asymmetry in Human Mitochondrial DNA Mutations. *Genomics*. 1994;22:327–35.
50. Reyes A, Gissi C, Pesole G, Saccone C. Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. *Mol Biol Evol*. 1998;15:957–66.
51. Li P-W, Wang X-Q, Chen S-C, Peng P. The complete mitochondrial genome of the tea lace bug, *Stephanitis chinensis* (Hemiptera: Tingidae). *Mitochondrial DNA Part B*. 2017;2:607–8.
52. Li Y, Song P, Feng J, Zhang N, Zhang R, Lin L. Complete mitochondrial genome sequence and phylogenetic analysis of *Myoxocephalus scorpius* (Linnaeus, 1758). *Mitochondrial DNA Part B*. 2019;4:862–3.
53. Song S-L, Yong H-S, Suana IW, Lim P-E, Eamsobhana P. Complete mitochondrial genome of *Dacus conopsoides* (Insecta: Tephritidae) with tRNA gene duplication and molecular phylogeny of Dacini tribe. *J Asia Pac Entomol*. 2019;22:997–1003.
54. González VL, Kayal E, Halloran M, Shrestha Y, Harasewych MG. The complete mitochondrial genome of the land snail *Cerion incanum* (Gastropoda: Stylommatophora) and the phylogenetic relationships of Cerionidae within Panpulmonata. *J Molluscan Stud*. 2016;82:525–33.
55. Wang Q, Feng R, Li L, Wang C, Zhu C. Characterization of the complete mitochondrial genome for the freshwater shrimp *Exopalaemon modestus*. *Conserv Genet Resour*. 2018;10:805–8.
56. Liu Y-Y, Zhou Z-C, Chen X-S. Characterization of the Complete Mitochondrial Genome of *Epicauta impressicornis* (Coleoptera: Meloidae) and Its Phylogenetic Implications for the Infraorder Cucujiformia. *J Insect Sci*. 2020;20(2):1–10.
57. Schon EA. The mitochondrial genome. In: *Rosenberg's Molecular and Genetic Basis of Neurological and Psychiatric Disease*. Elsevier; University of Texas Southwestern-Texas-US. 2020. p. 389–400.
58. Liu Y, Landick R, Raman S. A Regulatory NADH/NAD⁺ Redox Biosensor for Bacteria. *ACS Synth Biol*. 2019;8:264–73.
59. Ma H, Ma C, Li C, Lu J, Zou X, Gong Y, et al. First mitochondrial genome for the red crab (*Charybdis feriata*) with implication of phylogenomics and population genetics. *Sci Rep*. 2015;5:11524.
60. Shen K-N, Yen T-C, Chen C-H, Li H-Y, Chen P-L, Hsiao C-D. Next generation sequencing yields the complete mitochondrial genome of the flathead mullet, *Mugil cephalus* cryptic species NWP2 (Teleostei: Mugilidae). *Mitochondrial DNA*. 2014;27(5):1–2.
61. Tokishita S, Shibuya H, Kobayashi T, Sakamoto M, Ha J-Y, Yokobori S, et al. Diversification of mitochondrial genome of *Daphnia galeata* (Cladocera, Crustacea): Comparison with phylogenetic consideration of the complete sequences of clones isolated from five lakes in Japan. *Gene*. 2017;611:38–46.
62. Andriyono S, Alam MJ, Pramono H, Abdillah AA, Kim HW. Next-generation sequencing yields the complete mitochondrial genome of mud spiny lobster, *Panulirus polyphagus* (Crustacea: Decapoda) from Madura water. *IOP Conf Ser Earth Environ Sci*. 2019;348: 012020.
63. Kilpert F, Podsiadlowski L. 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*. 2006;7:241.

64. Ahn D-H, Min G-S, Park J-K, Kim S. The complete mitochondrial genome of the red-banded lobster *Metanephrops thomsoni* (Crustacea, Astacidea, Nephropidae): a novel gene order. *Mitochondrial DNA Part A*. 2016;27:2663–4.
65. Yu Y-Q, Liu X-L, Li H-W, Lu B, Fan Y-P, Yang J-S. The complete mitogenome of the Atlantic hydrothermal vent shrimp *Rimicaris exoculata* Williams & Rona 1986 (Crustacea: Decapoda: Alvinocarididae). *Mitochondrial DNA Part A*. 2016;27:3115–7.
66. Cheng J, Chan T, Zhang N, Sun S, Sha Z. Mitochondrial phylogenomics reveals insights into taxonomy and evolution of Penaeoidea (Crustacea: Decapoda). *Zool Scr*. 2018;47:582–94.
67. Hua CJ, Li WX, Zhang D, Zou H, Li M, Jakovčić I, et al. Basal position of two new complete mitochondrial genomes of parasitic Cymothoidea (Crustacea: Isopoda) challenges the monophyly of the suborder and phylogeny of the entire order. *Parasit Vectors*. 2018;11:628.
68. Yang M, Gao T, Yan B, Chen X, Liu W. Complete mitochondrial genome and the phylogenetic position of a wood-boring Isopod *Sphaeroma terebrans* (Crustacea, Isopod, Sphaeromatidae). *Mitochondrial DNA Part B*. 2019;4:1920–1.
69. Kearse MG, Wilusz JE. Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev*. 2017;31:1717–31.
70. Brown A, Shao S, Murray J, Hegde RS, Ramakrishnan V. Structural basis for stop codon recognition in eukaryotes. *Nature*. 2015;524:493–6.
71. Rossi N, Magalhães C, Mesquita ER, Mantelatto FL. Uncovering a hidden diversity: a new species of freshwater shrimp *Macrobrachium* (Decapoda: Caridea: Palaemonidae) from Neotropical region (Brazil) revealed by morphological review and mitochondrial genes analyses. *Zootaxa*. 2020;4732:177–95.
72. de Bruyn M, Wilson JC, Mather PB. Reconciling geography and genealogy: phylogeography of giant freshwater prawns from the Lake Carpentaria region. *Mol Ecol*. 2004;13:3515–26.
73. Nguyen Thanh H, Liu Q, Zhao L, Zhang H, Liu J, Nguyen HD. Genetic diversity of cultured populations of giant freshwater prawn (*Macrobrachium rosenbergii*) in China using mtDNA COI and 16S rDNA markers. *Biochem Syst Ecol*. 2015;62:261–9.
74. Elsheikh MO, Begham Mustafa F, Ibrahim Eid I, Lutas A, Bhasu S. COI gene sequence analysis for testing cyclical mating in securing genetic diversity of *Macrobrachium rosenbergii*. *Biochem Syst Ecol*. 2015;62:178–85.
75. Maidin MSR, Anton A, Yong ASK, Chin G. Mitochondrial COI gene Sequence of giant freshwater prawn, *Macrobrachium rosenbergii*: An assessment of a community-based stock enhancement programme in Petagas River, Sabah Malaysia. *Int J Fish Aquat Stud*. 2017;5:518–26.
76. Yang P, Zhang H, Chen L, Ye J, Yu N, Gu Z, et al. Genetic structure of the oriental river prawn (*Macrobrachium nipponense*) from the Yangtze and Lancang rivers, inferred from COI gene sequence. *Zool Res*. 2007;28:113.
77. Cook BD, Bunn SE, Hughes JM. Genetic structure and dispersal of *Macrobrachium australiense* (Decapoda: Palaemonidae) in western Queensland Australia. *Freshw Biol*. 2002;47:2098–112.
78. Guerra AL, Lima AVB, Lucato Júnior R V, Chiachio MC, Taddei FG, Castiglioni L. Genetic variability and phylogenetic aspects in species of the genus *Macrobrachium*. *Genet Mol Res*. 2014;13:3646–55.
79. de Carvalho FL, Pileggi LG, Mantelatto FL. Molecular data raise the possibility of cryptic species in the Brazilian endemic prawn *Macrobrachium potiuna* (Decapoda, Palaemonidae). *Lat Am J Aquat Res*. 2013;41:707–17.
80. Mantelatto FL, Carvalho FL, Vera-Silva AL. Distribution and genetic differentiation of *Macrobrachium jelskii* (Miers, 1877) (Natantia: Palaemonidae) in Brazil reveal evidence of non-natural introduction and cryptic allopatric speciation. *J Crustac Biol*. 2016;36:373–83.
81. Castelin M, Mazancourt V De, Marquet G, Zimmerman G, Keith P. Genetic and morphological evidence for cryptic species in *Macrobrachium australe* and resurrection of *M. ustulatum* (Crustacea, Palaemonidae). *Eur J Taxon*. 2017. <https://doi.org/10.5852/ejt.2017.289>.
82. Mantelatto FL, Terossi M, Negri M, Buranelli RC, Robles R, Magalhães T, et al. DNA sequence database as a tool to identify decapod crustaceans on the São Paulo coastline. *Mitochondrial DNA Part A*. 2018;29:805–15.
83. Pileggi LG, Mantelatto FL. Taxonomic revision of doubtful Brazilian freshwater shrimp species of genus *Macrobrachium* (Decapoda, Palaemonidae). *Iheringia Série Zool*. 2013;102:426–37.
84. García-Velazco H, Maeda-Martínez AM, Obregón-Barboza H, Campos-Torres O, Murugan G. The systematics of the Mexican populations of *Macrobrachium digueti* (Bouvier, 1895) (Decapoda: Caridea: Palaemonidae). *J Crustac Biol*. 2017;37:168–86.
85. Iketani G, Pimentel L, Torres E dos S, Rêgo PS do, Sampaio I. Mitochondrial heteroplasmy and pseudogenes in the freshwater prawn, *Macrobrachium amazonicum* (Heller, 1862): DNA barcoding and phylogeographic implications. *Mitochondrial DNA Part A*. 2021; 32:1–11.
86. Lopez JV, Yuhki N, Masuda R, Modi W, O'Brien SJ. Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. *J Mol Evol*. 1994;39:174–90.
87. Kmiec B, Woloszyńska M, Janska H. Heteroplasmy as a common state of mitochondrial genetic information in plants and animals. *Curr Genet*. 2006;50:149–59.
88. Williams ST, Knowlton N. Mitochondrial Pseudogenes Are Pervasive and Often Insidious in the Snapping Shrimp Genus *Alpheus*. *Mol Biol Evol*. 2001;18:1484–93.
89. Kang AR, Kim MJ, Park IA, Kim KY, Kim I. Extent and divergence of heteroplasmy of the DNA barcoding region in *Anapodisma miramae* (Orthoptera: Acrididae). *Mitochondrial DNA Part A*. 2016;27:3405–14.
90. Waldbieser GC, Lelania Bilodeau A, Nonnemanf DJ. Complete Sequence and Characterization of the Channel Catfish Mitochondrial Genome. *DNA Seq*. 2003;14:265–77.
91. Hassanin A, Léger N, Deutsch J. Evidence for Multiple Reversals of Asymmetric Mutational Constraints during the Evolution of the Mitochondrial Genome of Metazoa, and Consequences for Phylogenetic Inferences. *Syst Biol*. 2005;54:277–98.
92. Kartavtsev YP, Lee J-S. Analysis of nucleotide diversity at the cytochrome b and cytochrome oxidase 1 genes at the population, species, and genus levels. *Russ J Genet*. 2006;42:341–62.
93. Kumar R, Gopalakrishnan A, Divya PR, Basheer VS, Singh RK, Mohindra V, et al. Population genetic structure of *Macrobrachium rosenbergii* (Palaemonidae) from Indian waters using mitochondrial ATPase 6/8 gene. *Mitochondrial DNA Part A*. 2017;28:602–5.
94. Anger K, Hayd L. Feeding and growth in early larval shrimp *Macrobrachium amazonicum* from the Pantanal, southwestern Brazil. *Aquat Biol*. 2010;9:251–61.
95. Charmantier G, Anger K. Ontogeny of osmoregulatory patterns in the South American shrimp *Macrobrachium amazonicum*: Loss of hypo-regulation in a land-locked population indicates phylogenetic separation from estuarine ancestors. *J Exp Mar Bio Ecol*. 2011;396:89–98.
96. Meireles AL, Valenti WC, Mantelatto FL. Reproductive variability of the Amazon River prawn, *Macrobrachium amazonicum* (Caridea, Palaemonidae): influence of life cycle on egg production. *Lat Am J Aquat Res*. 2013;41:718–31.
97. Boudour-Boucheher N, Boulo V, Lorin-Nebel C, Elguero C, Grousset E, Anger K, et al. Adaptation to freshwater in the palaemonid shrimp *Macrobrachium amazonicum*: comparative ontogeny of osmoregulatory organs. *Cell Tissue Res*. 2013;353:87–98.
98. Berg JM, Tymoczko JL, Stryer L, Gatto GJ. *Bioquímica*. 7th ed. Rio de Janeiro: Guanabara Koogan; 2014.
99. Wolny M, Batchelor M, Bartlett GJ, Baker EG, Kurzawa M, Knight PJ, et al. Characterization of long and stable de novo single alpha-helix domains provides novel insight into their stability. *Sci Rep*. 2017;7:44341.
100. Cohen N, Eisenbach CD. Molecular Mechanics of Beta-Sheets. *ACS Biomater Sci Eng*. 2020;6:1940–9.
101. Esposti MD, De Vries S, Crimi M, Ghelli A, Patarnello T, Meyer A. Mitochondrial cytochrome b: evolution and structure of the protein. *Biochim Biophys Acta - Bioenerg*. 1993;1143:243–71.
102. Crofts AR, Hong S, Ugulava N, Barquera B, Gennis R, Guergova-Kuras M, et al. Pathways for proton release during ubiquinone oxidation by the bc₁ complex. *Proc Natl Acad Sci*. 1999;96:10021–6.
103. Hatefi Y. The mitochondrial electron transport and oxidative phosphorylation system. *Annu Rev Biochem*. 1985;54:1015–69.
104. Parashar A, Venkatachalam A, Gideon DA, Manoj KM. Cyanide does more to inhibit heme enzymes, than merely serving as an active-site ligand. *Biochem Biophys Res Commun*. 2014;455:190–3.
105. Dennerlein S, Wang C, Rehling P. Plasticity of Mitochondrial Translation. *Trends Cell Biol*. 2017;27:712–21.

106. Meyer JN, Leung MCK, Rooney JP, Sandoel A, Hengartner MO, Kisby GE, et al. Mitochondria as a Target of Environmental Toxicants. *Toxicol Sci.* 2013;134:1–17.
107. Shaughnessy DT, McAllister K, Worth L, Haugen AC, Meyer JN, Domann FE, et al. Mitochondria, Energetics, Epigenetics, and Cellular Responses to Stress. *Environ Health Perspect.* 2014;122:1271–8.
108. de Quadros T, Schramm H, Zeni EC, Simioni C, Allodi S, Müller YMR, et al. Developmental effects of exposure to ultraviolet B radiation on the freshwater prawn *Macrobrachium olfersi*: Mitochondria as a target of environmental UVB radiation. *Ecotoxicol Environ Saf.* 2016;132:279–87.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

