Baeza and Pirro
Revista Chilena de Historia Natural (2024) 97:9
https://doi.org/10.1186/s40693-024-00132-w

Revista Chilena de Historia Natural

RESEARCH

Open Access

Genomics resources for the Rapa Nui (Eastern Island) spiny lobster *Panulirus pascuensis* (Crustacea: Decapoda: Achelata)

J. Antonio Baeza[1,2,3*] and Stacy Pirro[4]

Abstract

Background The Easter Island spiny lobster *Panulirus pascuensis* (Reed, 1954) or '*Ura*' in the Rapa Nui language, is a little known species native to the south eastern Pacific Ocean, distributed along the coasts of Easter Island, Pitcairn Island, and the Salas y Gómez Ridge. In Easter Island, *P. pascuensis* is the target of a small and profitable and probably overexploited fishery. In this study, we profited from a series of bioinformatic analyses to mine biological insight from low-pass short-read next generation sequencing datasets; we have estimated genome size and ploidy in *P. pascuensis* using a k-mer strategy, discovered, annotated, and quantified mobile elements in the nuclear genome, assembled the 45S rRNA nuclear DNA cassette and mitochondrial chromosome, and explored the phylogenetic position of *P. pascuensis* within the genus *Panulirus* using the signal retrieved from translated mitochondrial protein coding genes.

Results K-mer analyses predicted *P. pascuensis* to be diploid with a haploid genome size ranging between 2.75 Gbp (with k-mer = 51) and 3.39 Gbp (with k-mer = 18). In *P. pascuensis*, repetitive elements comprise at least a half and a maximum of three fourths of the nuclear genome. Almost a third (64.94%) of the repetitive elements present in the studied nuclear genome were not assigned to any known family of transposable elements. Taking into consideration only annotated repetitive elements, the most abundant were classified as Long Interspersed Nuclear Elements (22.81%). Less common repetitive elements included Long Terminal Repeats (2.88%), Satellite DNA (2.66%), and DNA transposons (2.45%), among a few others. The 45S rRNA DNA cassette of *P. pascuensis* was partially assembled into two contigs. One contig, 2,226 bp long, encoded a partially assembled 5' ETS the entire ssrDNA (1,861 bp), and a partial ITS1. A second contig, 6,714 bp long, encoded a partially assembled ITS1, the entire 5.8S rDNA (158 bp), the entire ITS2, the entire lsrDNA (4,938 bp), and a partial 3' ETS (549 bp). The mitochondrial genome of *P. pascuensis* was 15,613 bp long and contained 13 protein-coding genes (PCGs), 22 transfer RNA (tRNA) genes, and two ribosomal RNA genes (12S ribosomal RNA [rrnS] and 16S ribosomal RNA [rrnL]). A phylomitogenomic analysis based on PCGs retrieved *Panulirus pascuensis* as sister to a fully supported clade comprising *P. cygnus* and *P. longipes*.

Conclusion We expect that the information generated in this study will guide the assembly of a chromosome-level nuclear genome for *P. pascuensis* in the near future. The newly assembled 45S rRNA nuclear DNA cassette and mitochondrial chromosome can support bioprospecting and biomonitoring of *P. pascuensis* using environmental DNA. The same elements can help to survey the public market place and detect mislabelling of this and other spiny lobsters. Overall, the genomic resources generated in this study will aid in supporting fisheries management and conservation strategies in this iconic spiny lobster that is likely experiencing overexploitation.

*Correspondence:
J. Antonio Baeza
baeza.antonio@gmail.com
Full list of author information is available at the end of the article

Background

In the speciose and morphologically dissimilar crustacean order Decapoda [22], spiny and slipper lobsters (infraorder Achelata) characteristically lack chelae on the first, second, and third pair of pereopods and exhibit a remarkable type of larva, the long-lived phyllosomata [41]. In this clade, many species are the target of subsistence, artisanal, and/or industrial fisheries in all subtropical and tropical oceans and in a few warm temperate coasts given that they often attain large body sizes and high population densities [41].

Among spiny lobsters (Achelata: Palinuridae), the Easter Island *Panulirus pascuensis* (Reed, 1954) or '*Ura*' in the Rapa Nui language, is a little known species native to the south eastern Pacific Ocean, distributed along the coasts of Rapa Nui or Easter Island, Pitcairn Island, and the Salas y Gómez Ridge [41]. In Easter Island, *P. pascuensis* inhabits a variety of habitats between 2 and 200 m depth and is the target of a small and profitable but probably overexploited fishery (see [10]). The life history of *P. pascuensis* is poorly known, specially when compared to that of the Caribbean spiny lobster *P. argus* ([1, 3], and references therein) and western rock lobster *P. cygnus* ([72] and references therein).

Only a few genetic and genomic resources have been developed for *P. pascuensis* [23, 57, 62]. Using two mitochondrial gene fragments (i.e., large-subunit ribosomal RNA [16S] and cytochrome oxidase subunit I (*cox1*), Ptacek et al. [62] determined that *P. pascuensis* belonged to the *P. japonicus* clade, confirming previous studies based on the analysis of morphological traits [41]. Also, a microsatellite panel ($n=9$ polymorphic SSRs) developed by Díaz-Cabrera et al. [23] was used to examine connectivity between populations in Eastern Island and the Salas y Gómez Island. *Panulirus pascuensis* exhibits high larval retention in the two studied islands as well as quite low and asymmetric larval connectivity between islands (greater from Salas y Gómez to Eastern Island - [57].

This study is part of a comprehensive program for the development of genetic and genomic resources in *P. pascuensis* and other species targeted by fisheries in the south eastern Pacific Ocean [5]. In this study, we profited from a series of bioinformatic analyses to mine biological insight from low-pass short-read next generation sequencing datasets. Using a series of bioinformatic tools tailored for retrieving information from such low coverage datasets, we have estimated genome size and ploidy in *P. pascuensis* using a k-mer strategy. We also discovered, annotated, and quantified mobile elements in the nuclear genome of the studied species. We assembled the 45S rRNA nuclear DNA cassette and mitochondrial chromosome. We have provided a detailed analysis of the latter genomic element. We explored the phylogenetic position of *P. pascuensis* within the genus *Panulirus* using the signal retrieved from translated mitochondrial protein coding genes. Lastly, we discovered a large set of microsatellites. We expect that the information generated in this study will guide the assembly of a gold-standard nuclear genome for *P. pascuensis* in the near future. The newly assembled 45S rRNA nuclear DNA cassette and mitochondrial chromosome can support bioprospecting and biomonitoring of *P. pascuensis* using environmental DNA. The same elements can help to survey the public market place and detect mislabelling of this and other spiny lobsters. Overall, the genomic resources generated in this study will aid in supporting fisheries management and conservation strategies in this iconic spiny lobster that is likely experiencing overexploitation.

Methods

Panulirus pascuensis specimen and DNA extraction, library preparation, and sequencing

The specimen used for sequencing belonging to *P. pascuensis* was deposited at the Clemson University Crustacean Collection. A small tissue sample (approx. 5 mm^3) was dissected from a pereiopod and immediately stored in sterile centrifuge tubes containing ethyl alcohol (95%) that was shipped to Iridian Genomes, Inc. (Bethesda, MD) for genomic DNA (gDNA) extraction and next generation sequencing (NGS). gDNA was extracted from the tissue sample with the DNeasy Blood and Tissue Kit (Qiagen, Germany) using the manufacturer's protocol. Library preparation was performed using the Illumina TruSeq kit following the manufacturer's instructions. NGS was conducted in a Illumina HiSeq X Ten system (Illumina, San Diego, CA, USA) using a 2×150 cycle. A total of 74,534,229 pairs (PE) reads were produced by Iridian Genomes and are available in the short read archive (SRA) repository (Bioproject: PRJNA996211, BioSample: SAMN36530443; SRA accession number: SRR25340) at NCBI's GenBank.

Cleaning and decontamination of raw reads

Illumina adapters and low quality sequences (Phred scores <20) were removed from the dataset using the program fastp v.0.20.1 with default parameters [15]. Next, the clean set of reads was 'decontaminated' from viral archaeal, bacterial, fungal, protozoan, and human reads with the program Kraken2 v2.1.2 [75] and the database kraken2-microbial-fatfree (https://lomanlab.github.io/mockcommunity/mc_databases.html).

Genome size of *Panulirus pascuensis*

We calculated the genome size of *P. pascuensis* using the high quality and decontaminated set of reads using

the program KMC 3 v. 3.2.1 [51] with 11 different k-mer sizes,21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51 bp long following Baeza et al. [5]. Then, the pipeline REPeat SPECTra Estimation (RESPECT) v.1.0.0 [64] was used to analyze each k-mer frequency distribution retrieved from each k-mer size and calculate nuclear genome size in *P. pascuensis*.

Ploidy in *Panulirus pascuensis*

We estimated ploidy in *P. pascuensis* in the software Smudgeplot v0.2.5 [63] specifically using the k-mer frequency distribution obtained with k-mer size = 21 in the program KMC. Immediately before the analysis of heterozygous k-mer pairs in the program Smudgeplot, we visually examined the coverage of the used k-mer frequency distribution in the web server GenomeScope (http://qb.cshl.edu/genomescope/genomescope2.0 - [63]) and selected for the analysis high coverage k-mers ranging between 20× and 140×.

Repetitive elements in the genome of *Panulirus pascuensis*

To discover, annotate, and quantify repetitive elements in the nuclear genome of *P. pascuensis*, we first mapped with the software HISAT2 v2.2.1 [50] the high quality and decontaminated set of PE reads to a newly assembled mitochondrial genome of the same species (see below). Next, we exclusively used those reads that did not map to the newly assembled mitochondrial genome (n = 72,688,900 PE reads [99.81%]) for the analysis of the nuclear 'repeatome' in *P. pascuensis* using the pipeline dnaPipeTE v1.4c [31, 32]. As a first step, DnaPipeTE assembled repetitive elements using the program Trinity [33] and subsequently annotated them based on homology with the pipeline RepeatMasker [29]. Lastly, DnaPipeTE quantified the abundance of repetitive elements by mapping a random sample of the reads onto the assembled repetitive elements. DnaPipeTE was executed using two iterations of the assembler Trinity with independent sets of read (sampled at 0.15X) each time [32]. For the analysis, we used the Protostomia-specific database of transposable elements from the consortium Dfam [43, 47]. Finally, we estimated the repetitive elements landscape of *P. pascuensis* when requesting dnaPipeTE to calculate the divergence (blastn) between transposable elements copies in the genomes (estimated from reads) and their respective assembled consensus sequences [32].

Nuclear ribosomal operon in *Panulirus pascuensis*

The nuclear ribosomal cassette or operon encodes the large (28S or lsrDNA) and small (18S or ssrDNA) nuclear rRNA genes along the 5.8S rDNA gene, two internal transcribed spacers (ITS1 and ITS2 that flank the 5.8S rDNA gene), and two external transcribed spacers (5′ ETS and 3′ ETS) [2]. This genomic element in the genome of *P. pascuensis* was assembled using the program CAP3 [42] as implemented on the platform RepeatExplorer 2.3.8 (http://repeatexplorer.org/ - [59, 60]). The exact coding positions of the small and large nuclear rDNAs along the boundaries of the 5′ and 3′ ETS were determined using RNAmmer with default parameters [53]. In turn, the boundaries of the ITS1 and ITS2 along the exact coding positions of the 5.8S nuclear rDNA were determined with the program ITSx [6].

Mitochondrial genome of *Panulirus pascuensis*

The mitochondrial genome of *P. pascuensis* was assembled 'de novo' with the software GetOrganelle v1.6.4 [45]. The complete mitochondrial genome of the congeneric Caribbean spiny lobster *P. argus* (GenBank's accession number MH068821- [1] was used as a 'seed' during the assembly process that was run with k-mer sizes of 21, 55, 85, and 115 bp. Next, the program MITOS2 [24] as implemented in the platform Galaxy [70] was employed for annotating the newly assembled mitochondrial genome. To depict the newly assembled mitochondrial genome as a circular map, we used the web server Genome Vx (http://wolfe.ucd.ie/GenomeVx/— [17]). Nucleotide composition of the complete mitochondrial genome and specific genes were estimated with the program MEGA X [52]. The Codon Usage Tool available in the web server Sequence Manipulation Suite (https://www.bioinformatics.org/sms2/codon_usage.html— [68]) was used to estimate the codon usage profile of all concatenated PCGs. We also estimated Relative Synonymous Codon Usage (RSCU) using the tool EZcodon available in the web server EZmito (http://ezmito.unisi.it/— [19]).

To analyze selective pressures in each PCG encoded in the newly assembled mitochondrial genome, the program KaKs_calculator 2.0 [73] was employed. Specifically, using the aforementioned program, we calculated for each PCG the number of nonsynonymous substitutions per nonsynonymous site (dN), the number of synonymous substitutions per synonymous site (dS), and the ratio dN/dS (= ω). The observed ω ratio is expected to be equal to 1, < 1, or > 1, if a particular PCG is exposed to neutral selection, purifying (negative), or diversifying (positive) selection, respectively. We used the congeneric Caribbean spiny lobster *P. argus* as an outgroup (GenBank accession number MH068821.1 - [1]) and the γ-MYN model to reflect mutation rate variability along the studied sequences during the analysis.

The secondary structure of each tRNA gene detected in the studied mitochondrial genome was predicted with the software MITFI [46], as implemented in MITOS2, and visualized using the on-line platform Forna (http://rna.tbi.univie.ac.at/forna/; [49]).

Lastly, the control region (CR) of the newly assembled mitochondrial genome was described in detail. We detected the presence of Simple Sequence Repeats (SSRs or microsatellites) and tandem repeats in this region using the web servers Microsatellites Repeats Finder (http://insilico.ehu.es/mini_tools/microsatellites/— [9]) and Tandem Repeats Finder (https://tandem.bu.edu/trf/trf.html— [7]), respectively. Furthermore, we detected the existence (or not) of 'stem and loops' or 'hairpins' in the CR using the the web server RNAfold (http://rna.tbi.univie.ac.at//cgi-bin/RNAWebSuite/RNAfold.cgi— [36]).

Phylogenetic position of *Panulirus pascuensis*

We examined the phylogenetic position of *P. pascuensis* in the genus *Panulirus* using the phylogenetic signal retrieved from PCGs (nucleotides) within a maximum likelihood (ML) framework. We conducted the phylogenetic analysis using the newly assembled mitochondrial genome plus those of 12 congeneric and 4 other cofamilial species whose annotated mitochondrial genomes were available in NCBI's GenBank. As outgroups, we included 7 other species belonging to family Scyllaridae, one species belonging to the infraorder Caridea, and one species belonging to the infraorder Astacidea. First, each set of PCG nucleotide sequences were aligned with the program MUSCLE [25, 67] as implemented in the software MEGA X. Next, poorly aligned regions in each PCG alignment were trimmed using the program GBlocks [12, 14]. Then, all alignments were concatenated and submitted to the web server IQ-TREE version 1.6.10 for ML analysis [56]. IQ-TREE used the program ProtTest [20] to partition the dataset and select the best fitting models of sequence evolution for each partition. The robustness of the ML tree topology was assessed using the Shimodaira–Hasegawa approximate likelihood ratio test ([SH]-aLRT) and 1,000 (ultra-fast) bootstrap iterations.

Results and discussion

Out of 74,534,229 PE raw reads produced by the sequencing facility, a total of 73,387,237 (98.46%) high quality PE clean reads remained after low quality sequences and Illumina adapters removal using the software fastp. The pipeline Kraken2 classified a total of 2.29% ($n=1,703,443$) of the clean reads as contaminants (Supplementary Materials, Fig. S1). None of the reads were classified as human and among those classified as microbial (2.24%), bacterial (0.719%), viral (0.0309%), fungal (0.328%), and protozoan (0.406%), no overrepresentation of a specific taxon (species) was observed. We considered the sequenced specimen and dataset devoid of any parasite and/or pathogen that might have biased our downstream analyses.

Genome size of *Panulirus pascuensis*

The haploid genome size (GS) estimated for *Panulirus pascuensis* using an *in-silico* k-mer approach ranged between 2,752,094,588 bp (2.75 Gbp) using a kmer size equal to 51 and 3,387,547,447 (3.39 Gbp) using a kmer size = 18. Increases in k-mer word size resulted in a decrease in the estimated genome size. The difference in estimate is roughly 19% of the maximum genome size estimate.

Genome size has been calculated only in 5 representatives of the family Paniluridae using flow cytometry or static cell fluorometry (Animal Genome Size Database (https://www.genomesize.com/) – [34] [consulted on 03 03 2024]) and varies between 3.08 Gb in the Pink spiny lobster *Palinurus mauritanicus* (estimated using flow cytometry - [21]) and 5.43 Gbp in the Caribbean spiny lobster *Panulirus argus* (estimated with static cell fluorometry - [44]). Our estimate of genome size using a k-mer approach in *Panulirus pascuensis* is within the range (when using k-mer size = 18) or somewhat lower (when using k-mer size = 21–51) than that reported for cofamilial species. The relatively small genome size estimated for the studied species might be explained by the relatively small number of reads used during the analysis or due to the relatively large portion of repetitive elements in the nuclear genome of this species (see below). Genome size estimations are biased downwards when using small datasets and/or when transposable elements account for a large portion of a genome [2].

Ploidy in *Panulirus pascuensis*

Panulirus pascuensis was determined to be diploid after analyzing the abundance of heterozygous k-mer pairs with the program Smudgeplot (Fig. 1). Decapod crustaceans, including spiny lobsters in the Infra-Order Achelata are assumed to be diploid even though research on ploidy is rare in this clade [48, 58]. Studies focusing on ploidy estimation are common in other eumetazoan clades and ancient polyploidization events appear to have driven the reshaping of genomes and spur diversification and evolutionary innovations in other pan-arthropod clades [48, 58]. A recent study has demonstrated that a crayfish, the marmorkrebs *Procambarus virginalis*, belonging to the closely related infraorder Achelata, is a triploid [61]. Originally, it was thought that a triploid and clonal marmorkrebs population descended as recently as 25 years ago from a single specimen of the slough crayfish *P. fallax* in the laboratory [35]. However, recent studies have shown that the two parental haplotypes of *P. virginalis* were inherited from natural populations of *P. fallax* and that *P. fallax* triploids are relatively common in nature [35]. A hybrid origin has recently been reported for a few subspecies of spiny lobsters (i.e., *P. homarus*

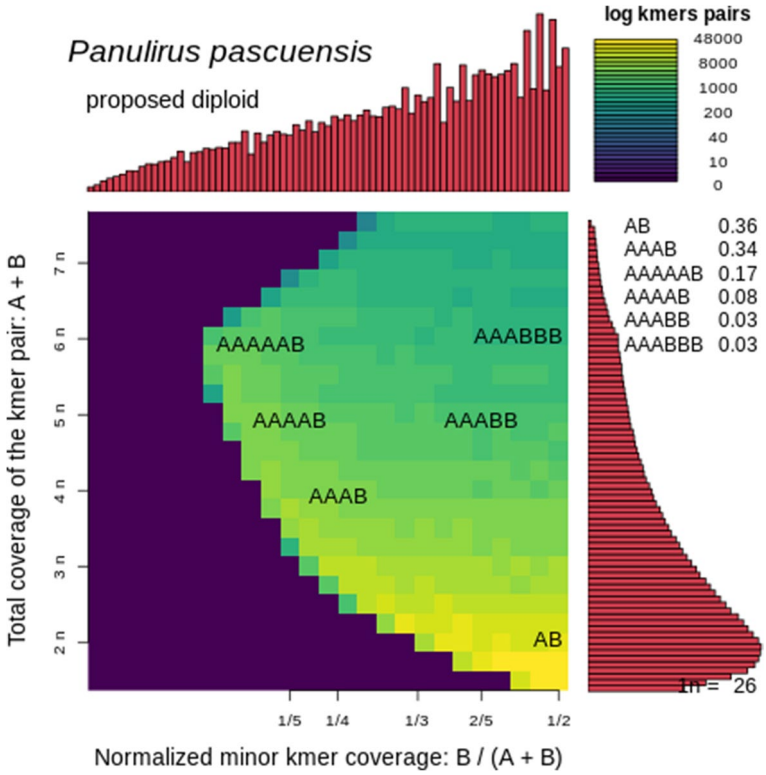

Fig. 1 Relationship between coverage of heterozygous k-mer pairs and normalized minor k-mer coverage in *Panulirus pascuensis*

rubellus) but nothing is known about the ploidy of hybrid specimens [27]. Research examining ploidy in the genus *Panulirus* and other decapod crustaceans, as already conducted in other arthropod clades (e.g., in Trichoptera - [39]), is needed to understand how often polyploidy occurs in nature and the genomic consequences of such events in the Achelata and beyond.

Nuclear repetitive elements of *Panulirus pascuensis*

The nuclear repetitive genome content in *P. pascuensis* ranged between 44% using a kmer size equal to 51 and 74% (with kmer = 18) as estimated by the program RESPECT. Increases in k-mer word size resulted in decreases (30% difference) in the nuclear repetitive genome content in our analysis. In turn, the pipeline dnaPipeTE determined that 57.64% of the genome in *P. pascuensis* comprised repetitive elements (Fig. 2), a value within the range estimated by the program RESPECT. Overall, repetitive elements comprise at least a half and a maximum of three fourths of the nuclear genome of *Panulirus pascuensis*. Repetitive content in the long-legged spiny lobster *P. longipes* and the Caribbean spiny lobster *P. argus* is 56.28% [2] and 69.02% [4], respectively. No estimation of repetitive content is available for *P. ornatus*, a third species of spiny lobster with an assembled genome available [71].

DnaPipeTE didn't annotate almost a third (64.94%) of the repetitive elements present in the nuclear genome of *P. pascuensis*; these 'unknown' repetitive elements were not assigned to any known family from the database of transposable elements specific to the Protostomia developed by the Dfam consortium (Fig. 2). Taking into consideration only annotated repetitive elements, the most abundant were classified as LINEs (Long Interspersed Nuclear Elements, 22.81%). Less common repetitive elements included LTRs (Long Terminal Repeats, 2.88%), Satellite DNA (2.66%), and DNA transposons (2.45%), among a few others (Fig. 3). The 'repeatome' has been characterized only in two other spiny lobsters, *P. argus* and *P. longipes*, and the portion of unannotated repetitive elements as well as repetitive element content in those species is similar to that reported for *P. pascuensis* in this study [2, 4]. Both in *P. argus* and *P. longipes*, LINEs and Satellite DNA were the most abundant annotated repetitive elements [2, 4]. Also, LINEs and LTRs comprise a large proportion of the genome in the few other decapods in which mobile elements have been quantified (e.g., in the Chinese mitten crab *Eriocheir japonica sinensis* - [69].

Lastly, the repetitive elements 'landscape' of *P. pascuensis* estimated with the program dnaPipeTE exhibited a right-skewed distribution suggesting an expansion (burst) of repetitive elements in the recent past (Fig. 2).

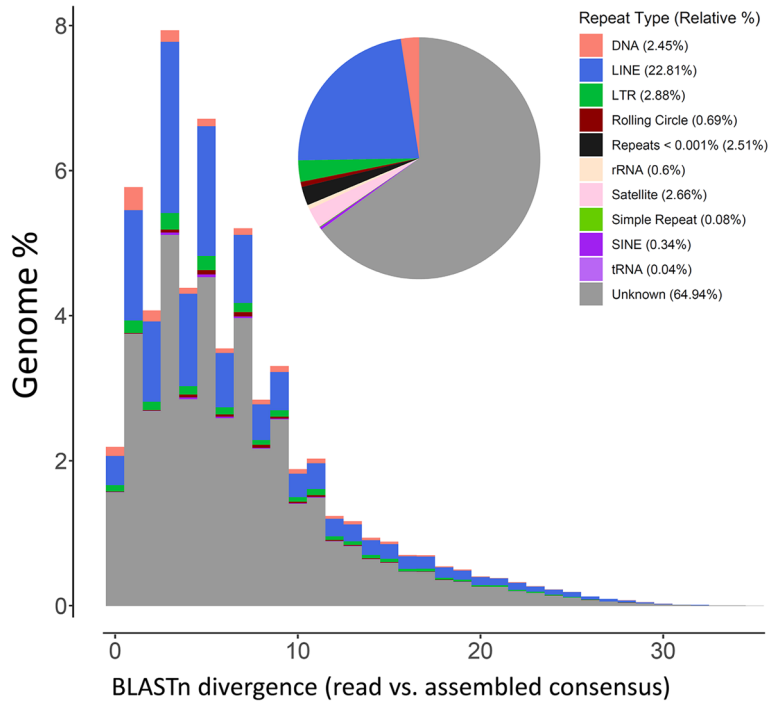

Fig. 2 Transposable elements genome composition (circle graph) and landscape in the genome of *Panulirus pascuensis*

Other than in *P. pascuensis*, no previous studies have examined the repetitive elements landscape in the Achelata and other decapod crustaceans.

We argue in favor of future studies characterizing the 'repeatome' in spiny lobsters. Such studies likely will result in the discovery of numerous new mobile genetic elements, given the relatively high number of unannotated repetitive elements observed in this and other spiny lobster genomes [2, 4]. The same studies will advance the understanding of the role that these elements have in driving genome size (e.g., [37]), genomic architecture (e.g., [65]) and evolutionary innovations (e.g., [74]) in the Achelata and other decapod crustaceans. Furthermore, Casacuberta and González [13] suggested that repetitive elements can affect the ability of their hosts to react to environmental insults. To what extent the 'repeatome' plays a role in determining the ability of spiny lobsters, decapod crustaceans, and other marine invertebrates to acclimate and adapt to pervasive global change needs to be examined by future studies.

Nuclear ribosomal operon in *Panulirus pascuensis*

The 45S rRNA DNA operon of *P. pascuensis* was partially assembled by the pipeline RepeatExplorer2. One contig, 2,226 bp long, encoded a partially assembled 5′ ETS (length = 105 bp), the entire ssrDNA (1,861 bp, GenBank accession number: PP417726), and a partial ITS1 (260 bp). A second contig, 6,714 bp long, encoded a partially assembled ITS1 (172 bp), the entire 5.8S rDNA (158 bp, GenBank accession number: PP416763), ITS2 (897 bp, fully assembled), the entire lsrDNA (4,938 bp, GenBank accession number: PP417725), and a partial 3′ ETS (549 bp). The two contigs matched nuclear ribosomal sequences available in GenBank that belonged to the genus *Panulirus* with E-values < $1 \times 10 - 6$.

During the last decades, fragments of the 45S rRNA DNA cassette have been employed to reveal phylogenetic relationships at multiple taxonomic levels in crustaceans, including representatives of the infraorder Achelata [11]. Yet, research focusing on the organization and genomic localization of the nuclear ribosomal RNA gene is rare in this and other clades of decapod crustacean. The study of Yu et al. [77] in the Chinese mitten crab *Eriocheir sinensis* is one of a few exceptions. Considering that the 45S rRNA DNA cassette can be assembled (either partially, nearly completely, or on its entirety) using low-coverage sequencing using the strategy used in this or other studies (see [1, 4]), we argue that this element, coupled with mitochondrial genomes (also assembled from low coverages sequencing data, see below), can be used to explore phylogenetic relationships among closely related species at a fraction of a cost, especially when compared to strategies like ultraconserved elements [28], anchored hybrid enrichment [54], or single nucleotide variants (SNVs) mining from assembled nuclear genomes [38].

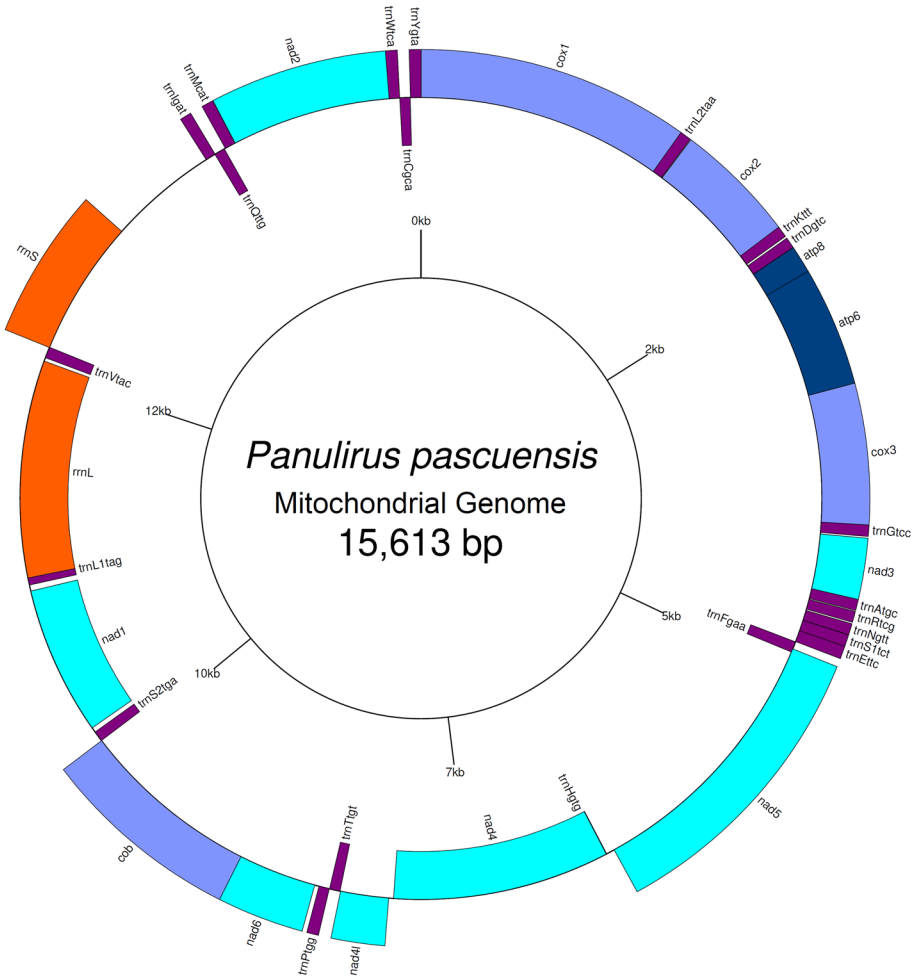

Fig. 3 Circular map of the mitochondrial genome of *Panulirus pascuensis*

Mitochondrial genome of *Panulirus pascuensis*

The program GetOrganelle assembled a complete (circularized) mitochondrial genome for the Eastern Island spiny lobster *P. pascuensis* with an average coverage of 129.4×and 539.3×per k-mer and base pair, respectively. The mitochondrial genome of *P. pascuensis* (GeneBank accession number OR612316) was 15,613 bp long, AT-rich (A+T content=63.02%), and encoded for 13 protein-coding genes (PCGs), 22 transfer RNA (tRNA) genes, and two ribosomal RNA genes (12S ribosomal RNA [rrnS] and 16S ribosomal RNA [rrnL]). Nine PCGs and 14 tRNA genes were located on one of the strands while 4 PCGs, the two ribosomal RNA genes, and 8 tRNA genes were encoded in the opposite strand (Table 1, Fig. 3). A single long intergenic region 688 bp long in the studied mitochondrial genome was assumed to be the Control Region (CR) or D-loop (Fig. 3, Table 1). The length of the mitochondrial genome assembled for *P. pascuensis* is within the range previously reported for congeneric species. In the genus *Panulirus*, the length of the mitochondrial genome varies between 15,739 bp in *P. argus* [1] and 15,665 bp in *P. homarus* [76]. Also, mitochondrial synteny in P. *pascuensis* is identical to that previously reported for the genus *Panulirus* and the infraorder Achelata [1] and references therein). Indeed, gene composition in the mitochondrial genome of *P. pascuensis* corresponds to the assumed Pancrustacean (Hexapoda+Crustacea) ground pattern [40].

In the mitochondrial genome of *P. pascuensis*, eight out of the 13 PCGs started and terminated with canonical mitochondrial crustacean codons (Table 1). *Cox1* exhibited the alternative (putative) start codon ACG while *nad1*, *nad3*, and *nad6* exhibited the alternative (putative) start codons ATT (Table 1). Non-canonical mitochondrial start codons have been reported before for other decapod crustaceans, including spiny lobsters [1] and references therein). One gene (*cox3*) terminated with TGA while *cob* ended with the incomplete stop codon T. Incomplete stop codons are commonly observed in the

Table 1 Mitochondrial genome of *Panulirus pascuensis*. Arrangement and annotation

Name	Type	Start	Stop	Strand	Length (bp)	Start	Stop	Continuity
cox1	PCG	1	1539	(+)	1539	ACG	TAA	-4
trnL2(taa)	tRNA	1535	1599	(+)	65			3
cox2	PCG	1603	2307	(+)	705	ATG	TAA	-16
trnK(ttt)	tRNA	2291	2355	(+)	65			11
trnD(gtc)	tRNA	2367	2429	(+)	63			0
atp8	PCG	2430	2588	(+)	159	ATG	TAA	-4
atp6	PCG	2585	3259	(+)	675	ATA	TAA	-1
cox3	PCG	3259	4050	(+)	780	ATG	TGA	-1
trnG(tcc)	tRNA	4050	4114	(+)	65			9
nad3	PCG	4124	4468	(+)	345	ATT	TAG	-2
trnA(tgc)	tRNA	4467	4530	(+)	64			4
trnR(tcg)	tRNA	4535	4598	(+)	64			7
trnN(gtt)	tRNA	4606	4670	(+)	65			0
trnS1(tct)	tRNA	4671	4738	(+)	68			-1
trnE(ttc)	tRNA	4738	4810	(+)	73			1
trnF(gaa)	tRNA	4812	4880	(+)	69			22
nad5	PCG	4858	6558	(-)	1729	ATG	TAG	51
trnH(gtg)	tRNA	6610	6674	(+)	65			-62
nad4	PCG	6613	7977	(-)	1387	ATG	TAA	40
nad4l	PCG	8007	8309	(-)	308	ATG	TAA	2
trnT(tgt)	tRNA	8312	8379	(+)	68			0
trnP(tgg)	tRNA	8380	8446	(-)	67			26
nad6	PCG	8473	8964	(+)	500	ATT	TAA	-1
cob	PCG	8964	10,098	(+)	1153	ATG	T	0
trnS2(tga)	tRNA	10,099	10,165	(+)	60			30
nad1	PCG	10,196	11,131	(-)	951	ATT	TAA	37
trnL1(tag)	tRNA	11,169	11,238	(-)	70			-23
rrnL	rRNA	11,214	12,577	(-)	1386			21
trnV(tac)	tRNA	12,599	12,669	(-)	71			-3
rrnS	rRNA	12,667	13,519	(-)	867			0
CR		13,520	14,207	n	689			0
trnI(gat)	tRNA	14,208	14,273	(+)	66			-3
trnQ(ttg)	tRNA	14,271	14,339	(-)	69			6
trnM(cat)	tRNA	14,346	14,414	(+)	69			0
nad2	PCG	14,415	15,416	(+)	1018	ATG	TAA	-2
trnW(tca)	tRNA	15,415	15,481	(+)	67			-2
trnC(gca)	tRNA	15,481	15,546	(-)	66			0
trnY(gta)	tRNA	15,547	15,613	(+)	67			0

mitochondrial genome of decapod crustaceans, including spiny lobster [1] and references therein) and these truncated termination codons are assumed to be completed via post-transcriptional poly-adenylation [66].

In the mitochondrial PCGs of *P. pascuensis*, codons were not used evenly. The most frequently used codons were AT-rich, and included TTT (Phe, $n=206$ times used), ATT (Ile, $n=204$), TTA (Leu, $n=180$), GTT, (Val, $n=113$), ATA (Met, $n=113$), TTG (Leu, $n=112$), TCT, (Ser, $n=111$), and TTC (Phe, $n=105$). In turn, other than stop codons, codons least frequently used were GC-rich and included CGC (Arg, $n=9$), ACG (Thr, $n=12$), CGG (Arg, $n=12$), TGC, (Cys, $n=13$), AGC (Ser, $n=14$) (Supplementary Materials, Table S1). The RSCU analysis also indicated that among synonymous codons, most frequently used codons were AT-rich (Fig. 4). Our results are in line with those reported for *P. argus*, the only species of spiny lobster for which codon usage has been estimated [1].

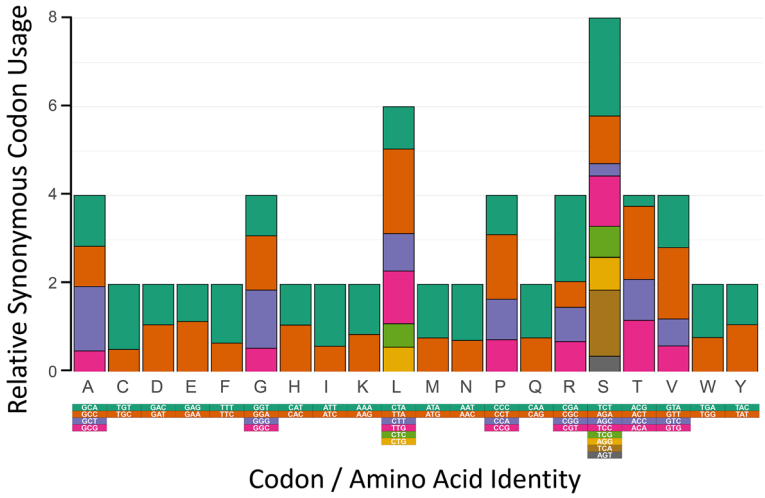

Fig. 4 Relative synonymous usage in the 13 protein coding genes encoded in the mitochondrial genome of *Panulirus pascuensis*

The analysis of selective pressures in the mitochondrial PCGs of *P. pascuensis* indicated that all the studied genes are evolving under purifying selection; all estimated KA/KS ratios exhibited values < 1 (Table 2). The KA/KS ratios estimated for *cob*, *cox1*, and *cox2* (KA/KS = 0.0042, 0.0015, and 0.0056, respectively) were one or two orders of magnitude lower than those estimated for the rest of the PCGs (Table 2). The aforementioned differences in KA/KS ratios suggest strong evolutionary constraints in the *cob*, *cox1*, and *cox2* genes (Table 2). In spiny lobsters, selective pressure in mitochondrial PCGs has rarely been explored [1]. Nonetheless, widespread negative selection in mitochondrial PCGs s] is well documented in eumetazoans, including other decapod crustaceans [1].

In the studied mitochondrial genome, tRNA genes ranged in length between 63 pb (tRNA-D) and 73 bp (tRNA-E) and all of them featured a standard 'cloverleaf' secondary structure with the exception of tRNA-Ser1 that was missing the D-arm stem and featured a short loop compared to other tRNA genes (Supplementary Materials, Fig. S2). In eumetazoans, including decapod crustaceans, either tRNA-Ser1 or tRNA-Ser2 are often reported as truncated and whether or not truncated tRNA genes are functional remains to be addressed in most eumetazoans, including decapod crustaceans and spiny lobsters [8].

Table 2 Selective pressure analysis of the protein coding genes in the mitochondrial genome of *Panulirus pascuensis*

PCG	KA	KS	KA/KS	*P*-value
atp6	0.0185549	1.50396	0.0123373	7.16E-71
atp8	0.121534	3.02484	0.0401787	4.29E-13
cob	0.00492923	1.19756	0.00411606	6.76E-123
cox1	0.00182467	1.20659	0.00151225	4.51E-169
cox2	0.00775123	1.38463	0.00559806	2.46E-74
cox3	0.0188861	1.56849	0.0120409	2.53E-83
nad1	0.0187797	0.894983	0.0209833	3.68E-66
nad2	0.0913878	1.785000	0.0511977	5.09E-83
nad3	0.0382029	1.88172	0.0203021	3.50E-37
nad4	0.0872272	0.131893	0.661348	4.58E-36
nad4l	0.0143833	1.31851	0.0109087	1.87E-32
nad5	0.0487075	1.4077	0.0346008	1.24E-144
nad6	0.0572806	1.29943	0.0440814	2.35E-41

In the mitochondrial genome of *P. pascuensis*, the 688 bp long CR is located between tRNA-I and the 12 S ribosomal RNA (Fig. 3). The region was AT-rich (70.64%) having an overall base composition equal to A = 36.19%, T = 34.45%, C = 18.75%, and G = 10.61%. The analysis conducted in the online tool Microsatellite repeats finder revealed the presence of four SRRs in the studied CR, all of them were dinucleotide SRRs repeated between 3 and 5 times and 4 out of the 5 SRRs were AT-rich (Supplementary Materials, Table S2). In turn, the online tool Tandem Repeat Finder failed to detect tandem repeats in the studied CR, in line to that observed in the Caribbean spiny lobster [1] but in disagreement to that reported for the Chinese spiny lobster *Panulirus stimpsoni* [55]. Lastly, prediction (using minimum free energy [MFE] and Centroid optimization) of the secondary structure in RNAFold revealed stem-loop structures along the entirety of the CR (Supplementary Materials, Fig. S2), similarly to that observed in other spiny lobster mitochondrial genomes in which a detailed analysis of this region has been conducted [1]. Additional detailed studies focusing on the organization of the CR are needed to understand its function during mitochondrial transcription and replication.

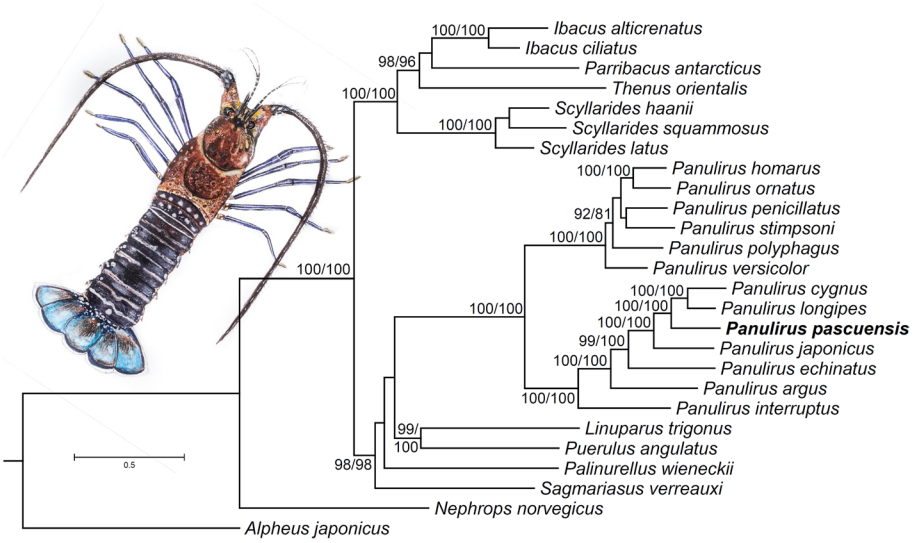

Fig. 5 Maximum likelihood phylogenetic hypothesis for the family Palinuridae and phylogenetic placement of *Panulirus pascuensis*. The phylogenetic tree was retrieved using the phylogenetic signal provided by (translated) protein coding genes. Numbers above branches near nodes represent bootstrap pseudo-replicates ($N = 1{,}000$) of the tree search. Depiction of the studied species *Panulirus pascuensis* by Brooke Fitzwater (used with permission)

Phylogenetic position of *Panulirus pascuensis*

The ML phylogenetic analysis was based on 26 terminals, 11,103 characters, and 5,974 parsimony-informative sites. The analysis recovered the infraorder Achelata and the families Scyllaridae and Paniluridae as monophyletic (in all three cases, bootstrap support [bs] values: $98 < bs > 100$). In the latter family, the genus *Panulirus* was recovered as monophyletic (bv = 100) and comprised two fully supported (bv = 100) sister clades, Lineage 1 and Lineage 2 sensu George [30]. In our analysis, species belonging to Lineage 1 included *P. argus*, *P. cygnus*, *P. echinatus*, *P. interruptus*, *P. japonicus*, *P. longipes*, and *P. pascuensis*, while Lineage 2 was represented by *P. homarus*, *P. ornatus*, *P. penicillatus*, *P. polyphagus*, *P. stimpson*, and *P. versicolor*. In Lineage 1, the Caribbean spiny lobster *P. argus* had an early branching position; it was sister to all other species in this clade that clustered together into a single fully supported clade. In turn, *P. echinatus* was arranged sister to a fully supported clade (bv = 100) containing the remaining species of Lineage 1 except *P. argus*. *Panulirus pascuensis*, the focus of this study, was sister to a fully supported clade comprising *P. cygnus* and *P. longipes*. In Lineage 2, *P. versicolor* was sister to a moderately supported clade containing all other species in this lineage. The relationships among the species comprising the latter clade were poorly resolved other than a fully supported sister relationship between *P. homarus* and *P. ornatus* (Fig. 5).

Overall, the recovered phylogenetic relationships within the Achelata support findings by previous studies using a larger set of species but with a smaller number of markers [62] or a smaller set of species using entire mitochondrial genomes [1]. Also, in line with previous studies in other infraorders of decapod crustaceans (e.g., Caridea - [26]; Dendrobranchiata - [18]; Anomura - [16] our study suggests that mitochondrial genomes can reliably recover phylogenetic relationships at and below the family level in the Achelata. We argue in favor of additional studies sequencing mitochondrial genomes from low-pass coverage datasets in spiny and slipper lobsters to advance our understanding of the evolutionary history of this remarkable clade of decapod crustaceans.

Conclusion

In this study, we have generated genomic resources for *P. pascuensis*, a spiny lobster for which little is known but that likely plays an important ecological role and that is probably overfished in Rapa Nui. We used low-coverage short-read sequencing to determine ploidy and genome size of *P. pascuensis*. Also, we identified, classified, and quantified mobile elements in the nuclear genome of the studied species. Importantly, the large size and number of mobile elements in the nuclear genome of *P. pascuensis* implies that long-reads (i.e., Oxford Nanopore Technology and/or Pacific Biosciences) plus chromosome conformation capture techniques (i.e., Hi-C) will be needed to assemble a high-quality (i.e., chromosome-level) genome in this spiny lobster. We have also assembled the ribosomal RNA cassette and mitochondrial genome of the studied species. Overall, the genomic resources generated in this study will aid in supporting fisheries management and conservation strategies in this iconic spiny lobster that is likely experiencing overexploitation.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40693-024-00132-w.

Supplementary Material 1: Table S1. Codon usage analysis in the mitochondrial protein coding genes of *Panulirus pascuensis*. Table S2. Microsatellites in the Control Region of the spiny lobster *Panulirus pascuensis* detected by the online tool Microsatellite repeats finder. Figure S1. Secondary structure prediction of tRNA-Ser1 gene in the mitochondrial genome of *Panulirus pascuensis*. Figure S2. Secondary structure prediction (minimum free energy and Centroid optimization, left and right, respectively) of the control region in the mitochondrial genome of *Panulirus pascuensis*.

Acknowledgements
JAB thanks Dr. Vince P. Richards for bioinformatics support. We thank two anonymous reviewers that provided comments that improved previous versions of this manuscript. We thanks Brooke Fitzwater for allowing us to use her painted artistic depiction of the studied species.

Field study permissions
NA.

Authors' contributions
J. Antonio Baeza conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Funding
Funding was provided by Iridian Genomes, grant# IRGEN_RG_2021-1345 Genomic Studies of Eukaryotic Taxa.

Availability of data and materials
Sequences described here are accessible via GenBank: MW252173 and MW251820, BioProject ID: PRJNA453553.

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Biological Sciences, Clemson University, Clemson, SC, USA. [2]Departamento de Biología Marina, Universidad Catolica del Norte, Coquimbo, Chile. [3]Smithsonian Marine Station at Fort Pierce, Smithsonian Institution, Fort Pierce, FL, USA. [4]Iridian Genomes, Silver Spring, MD, USA.

Received: 28 March 2024 Accepted: 19 August 2024
Published online: 30 August 2024

References

1. Baeza JA. The complete mitochondrial genome of the Caribbean spiny lobster *Panulirus argus*. Sci Rep. 2018;8(1):17690. https://doi.org/10.1038/s41598-018-36132-6.
2. Baeza JA. Genome survey sequencing of the Caribbean spiny lobster *Panulirus argus*: Genome size, nuclear rRNA operon, repetitive elements, and microsatellite discovery. PeerJ. 2020;8:e10554.
3. Baeza JA, Simpson L, Ambrosio LJ, Mora N, Childress M. Active parental care, reproductive performance, and a novel egg predator affecting reproductive investment in the Caribbean spiny lobster *Panulirus argus*. BMC Zool. 2016;1(1):6. https://doi.org/10.1186/s40850-016-0006-6.
4. Baeza JA, Baker AM, Liu H. Genome survey sequencing of the long-legged spiny lobster *Panulirus longipes* (A. Milne-Edwards, 1868) (Decapoda: Achelata: Palinuridae): improved mitochondrial genome annotation, nuclear repetitive elements classification, and SSR marker discovery. J Crustacean Biol. 2022;42(1):ruac006.
5. Baeza JA, González MT, Sigwart JD, et al. Insights into the genome of the 'Loco' *Concholepas concholepas* (Gastropoda: Muricidae) from low-coverage short-read sequencing: genome size, ploidy, transposable elements, nuclear RNA gene operon, mitochondrial genome, and phylogenetic placement in the family Muricidae. BMC Genomics. 2024;25:77.
6. Bengtsson-Palme J, Ryberg M, Hartmann M, Branco S, Wang Z, Godhe A, De Wit P, Sanchez-Garcia M, Ebersberger I, De Sousa F, Amend AS. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. Methods Ecol Evol. 2013;4:914e919.
7. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27:573–80. https://doi.org/10.1093/nar/27.2.573.
8. Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 2013;69:313–9. https://doi.org/10.1016/j.ympev.2012.08.023.
9. Bikandi J, San Millán R, Rementeria A, Garaizar J. In silico analysis of complete bacterial genomes: PCR, AFLP-PCR, and endonuclease restriction. Bioinformatics. 2004;20:798–9. https://doi.org/10.1093/bioinformatics/btg491.
10. Boyko CB. The endemic marine invertebrates of Easter Island: how many species and for how long? In: Easter Island: scientific exploration into the world's environmental problems in microcosm. 2003. p. 155–175.
11. Bracken-Grissom HD, Ahyong ST, Wilkinson RD, Feldmann RM, Schweitzer CE, Breinholt JW, Bendall M, Palero F, Chan TY, Felder DL, Robles R. The emergence of lobsters: phylogenetic relationships, morphological evolution and divergence time comparisons of an ancient group (Decapoda: Achelata, Astacidea, Glypheidea, Polychelida). Syst Biol. 2014;63(4):457–79.
12. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3. https://doi.org/10.1093/bioinformatics/btp348.
13. Casacuberta E, González J. The impact of transposable elements in environmental adaptation. Mol Ecol. 2013;22(6):1503–17. https://doi.org/10.1111/mec.12170.
14. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 2000;17:540–52.
15. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90. https://doi.org/10.1093/bioinformatics/bty560.
16. Colín A, Galván-Tirado C, Carreón-Palau L, Bracken-Grissom HD, Baeza JA. Mitochondrial genomes of the land hermit crab *Coenobita clypeatus* (Anomura: Paguroidea) and the mole crab *Emerita talpoida* (Anomura: Hippoidea) with insights into phylogenetic relationships in the Anomura (Crustacea: Decapoda). Gene. 2023;849:146896.
17. Conant GC, Wolfe KH. GenomeVx: simple web-based creation of editable circular chromosome maps. Bioinformatics. 2008;24:861–2. https://doi.org/10.1093/bioinformatics/btm598.
18. Cronin TJ, Conrad I, Kerkhove TR, Hellemans B, De Troch M, Volckaert FA, Baeza JA. Characterization of the complete mitochondrial genome of the Atlantic seabob shrimp *Xiphopenaeus kroyeri* Heller, 1862 (Decapoda: Dendrobranchiata: Penaeidae), with insights into the phylogeny of Penaeidae. J Crustacean Biol. 2022;42(1):ruac004.
19. Cucini C, Leo C, Iannotti N, Boschi S, Brunetti C, Pons J, Fanciulli PP, Frati F, Carapelli A, Nardi F. EZmito: a simple and fast tool for multiple mitogenome analyses. Mitochondrial DNA B Resour. 2021;6:1101–9. https://doi.org/10.1080/23802359.2021.1899865.
20. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics. 2011;27:1164–5. https://doi.org/10.1093/bioinformatics/btr088.
21. Deiana AM, Cau A, Coluccia E, Cannas R, Milia A, Salvadori S, Libertini A. Genome size and AT-DNA content in thirteen species of decapoda. In: Schram FR, von Vaupel Klein JC, editors. Crustaceans and the biodiversity crisis. Leiden: Koninklijke Brill NV; 1999. p. 981–5.
22. De Grave S, Pentcheff ND, Ahyong ST, Chan T-Y, Crandall KA, Dworschak PC, Felder DL, Feldmann RM, Fransen CHJM, Goulding LYD, Lemaitre R, Low MEY, Martin JW, Ng PKL, Schweitzer CE, Tan SH, Tshudy D, Wetzer RA. Classification of living and fossil genera of decapod crustaceans. Raffles Bull Zool. 2009;21:1–109.

23. Díaz-Cabrera E, Meerhoff E, Rojas-Hernandez N, Vega-Retter C, Veliz D. Development and characterization of the first 16 microsatellites loci for *Panulirus pascuensis* (Decapoda: Palinuridae) from Easter Island using Next Generation Sequencing. Rev Biol Mar Oceanogr. 2017;52(2):395–8.
24. Donath A, Jühling F, Al-Arab M, Bernhart SH, Reinhardt F, Stadler PF, Middendorf M, Bernt M. Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. Nucleic Acids Res. 2019;47:10543–52. https://doi.org/10.1093/nar/gkz833.
25. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7. https://doi.org/10.1093/nar/gkh340.
26. Ennis CC, Haeffner NN, Keyser CD, Leonard ST, Macdonald-Shedd AC, Savoie AM, Cronin TJ, et al. Comparative mitochondrial genomics of sponge-dwelling snapping shrimps in the genus *Synalpheus*: exploring differences between eusocial and non-eusocial species and insights into phylogenetic relationships in caridean shrimps. Gene. 2021;786:145624.
27. Farhadi A, Jeffs AG, Lavery SD. Genome-wide SNPs in the spiny lobster *Panulirus homarus* reveal a hybrid origin for its subspecies. BMC Genomics. 2022;23(1):750.
28. Faircloth BC, Branstetter MG, White ND, Brady SG. Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. Mol Ecol Resour. 2015;15(3):489–501. https://doi.org/10.1111/1755-0998.12328.
29. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci. 2020;117(17):9451–7.
30. George RW. Tethys sea fragmentation and speciation of *Panulirus* spiny lobsters. Crustaceana. 2006;78(11):1281–309.
31. Goubert C, Modolo L, Vieira C, ValienteMoro C, Mavingui P, Boulesteix M. De novo assembly and annotation of the Asian tiger mosquito (*Aedes albopictus*) repeatome with dnaPipeTE from raw genomic reads and comparative analysis with the Yellow Fever mosquito (*Aedes aegypti*). Genome Biol Evol. 2015;7:1192–205. https://doi.org/10.1093/gbe/evv050.
32. Goubert C, Craig RJ, Bilat AF, Peona V, Vogan AA, Protasio AV. A beginner's guide to manual curation of transposable elements. Mob DNA. 2022;13:7. https://doi.org/10.1186/s13100-021-00259-7.
33. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat Biotechnol. 2011;29(7):644.
34. Gregory TR. Animal genome size database. 2020 .http://www.genomesize.com.
35. Gutekunst J, Maiakovska O, Hanna K, Provataris P, Horn H, Wolf S, Skelton CE, Dorn NJ, Lyko F. Phylogeographic reconstruction of the marbled crayfish origin. Commun Biol. 2021;4(1):1096.
36. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA websuite. Nucleic Acids Res. 2008;36(2):W70-74. https://doi.org/10.1093/nar/gkn188.
37. Helmkampf M, Bellinger MR, Geib SM, Sim SB, Takabayashi M. Draft genome of the rice coral *Montipora capitata* obtained from linked-read sequencing. Genome Biol Evol. 2019;11:2045–54.
38. Helyar SJ, Hemmer-Hansen J, Bekkevold D, Taylor MI, Ogden R, Limborg MT, Cariani A, Maes GE, Diopere E, Carvalho GR, Nielsen EE. Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Mol Ecol Resour. 2011;11:123–36.
39. Heckenhauer J, Frandsen PB, Sproul JS, Li Z, Paule J, Larracuente AM, Maughan PJ, Barker MS, Schneider JV, Stewart RJ, Pauls SU. Genome size evolution in the diverse insect order Trichoptera. GigaScience. 2022;11:giac011.
40. Hickerson MJ, Cunningham CW. Dramatic mitochondrial gene rearrangements in the hermit crab *Pagurus longicarpus* (Crustacea, Anomura). Mol Biol Evol. 2000;17:639–44.
41. Holthuis L. FAO species catalogue. Vol 13. Marine lobsters of the world. An annotated and illustrated catalogue of species of interest to fisheries known to date in FAO fisheries synopsis. Rome: FAO; 1991.
42. Huang X, Madan A. CAP3: a DNA sequence assembly program. Genome Res. 1999;9(9):868–77. https://doi.org/10.1101/gr.9.9.868.
43. Hubley R, Finn RD, Clements J, Eddy SR, Jones TA, Bao W, Smit AF, Wheeler TJ. The Dfam database of repetitive DNA families. Nucleic Acids Res. 2016;44(D1):D81–9. https://doi.org/10.1093/nar/gkv1272.
44. Jimenez AG, Kinsey ST, Dillaman RM, Kapraun DF. Nuclear DNA content variation associated with muscle fiber hypertrophic growth in decapod crustaceans. Genome. 2010;53(3):161–71. https://doi.org/10.1139/G09-095.
45. Jin JJ, Yu WB, Yang JB, Yu S, dePamphilis CW, Yi TS, Li DZ. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 2020;21:1–31. https://doi.org/10.1186/s13059-020-02154-5.
46. Jühling F, Pütz J, Bernt M, Donath A, Middendorf M, Florentz C, Stadler PF. Improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. Nucleic Acids Res. 2012;40(7):2833–45.
47. Kalvari I, Argasinska J, Quinones-Olvera N, Nawrocki EP, Rivas E, Eddy SR, Bateman A, Finn RD, Petrov AI. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. Nucleic Acids Res. 2018;46(D1):D335–42. https://doi.org/10.1093/nar/gkx1038.
48. Kenny NJ, Chan KW, Nong W, Qu Z, Maeso I, Yip HY, Chan TF, Kwan HS, Holland PW, Chu KH, et al. Ancestral whole-genome duplication in the marine chelicerate horseshoe crabs. Heredity. 2016;116(2):190–9.
49. Kerpedjiev P, Hammer S, Hofacker IL. Forna (force-directed RNA): simple and effective online RNA secondary structure diagrams. Bioinformatics. 2015;31(20):3377–9.
50. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37:907–15. https://doi.org/10.1038/s41587-019-0201-4.
51. Kokot M, Długosz M, Deorowicz S. KMC 3: counting and manipulating k-mer statistics. Bioinformatics. 2017;33(17):2759–61.
52. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35:1547–9. https://doi.org/10.1093/molbev/msy096.
53. Lagesen K, Hallin PF, Rødland E, Stærfeldt HH, Rognes T, Ussery DW. RNammer: consistent annotation of rRNA genes in genomic sequences. Nucleic Acids Res. 2007;35:3100–8.
54. Lemmon AR, Emme SA, Lemmon EM. Anchored hybrid enrichment for massively high-throughput phylogenomics. Syst Biol. 2012;61(5):727–44. https://doi.org/10.1093/sysbio/sys049.
55. Liu Y, Cui Z. Complete mitochondrial genome of the Chinese spiny lobster *Panulirus stimpsoni* (Crustacea: Decapoda): genome characterization and phylogenetic considerations. Mol Biol Rep. 2011;38:403–10.
56. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–74. https://doi.org/10.1093/molbev/msu300.
57. Meerhoff E, Yannicelli B, Dewitte B, Díaz-Cabrera E, Vega-Retter C, Ramos M, Bravo L, Concha E, Hernández-Vaca F, Véliz D. Asymmetric connectivity of the lobster *Panulirus pascuensis* in remote islands of the Southern Pacific: importance for its management and conservation. Bull Mar Sci. 2018;94(3):753–74.
58. Nossa CW, Havlak P, Yue JX, Lv J, Vincent KY, Brockmann HJ, Putnam NH. Joint assembly and genetic mapping of the Atlantic horseshoe crab genome reveals ancient whole genome duplication. GigaScience. 2014;3:9.
59. Novak P, Neumann P, Pech J, Steinhaisl J, Macas J. RepeatExplorer: a galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next generation sequence reads. Bioinformatics. 2013;29(6):792–3. https://doi.org/10.1093/bioinformatics/btt054.
60. Novak P, Neumann P, Macas J. Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. Nat Protoc. 2020;15:1–32.
61. Martin P, Thonagel S, Scholtz G. The parthenogenetic Marmorkrebs (Malacostraca: Decapoda: Cambaridae) is a triploid organism. J Zool Syst Evol Res. 2016;54(1):13–21.
62. Ptacek MB, Sarver SK, Childress MJ, Herrnkind WF. Molecular phylogeny of the spiny lobster genus *Panulirus* (Decapoda: Palinuridae). Mar Freshw Res. 2001;52(8):1037–47.
63. Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and smudgeplot for reference-free profiling of polyploid genomes. Nat Commun. 2020;11(1):1–10. https://doi.org/10.1038/s41467-020-14998-3.
64. Sarmashghi S, Balaban M, Rachtman E, Touri B, Mirarab S, Bafna V. Estimating repeat spectra and genome length from low-coverage genome

skims with RESPECT. PLoS Comput Biol. 2021;17:e1009449. https://doi.org/10.1371/journal.pcbi.1009449.
65. Shapiro JA. A 21st century view of evolution: genome system architecture, repetitive DNA, and natural genetic engineering. Gene. 2005;345(1):91–100.
66. Slomovic S, Laufer D, Geiger D, Schuster G. Polyadenylation and degradation of human mitochondrial RNA: the prokaryotic past leaves its mark. Mol Cell Biol. 2005;25(15):6427–35.
67. Sievers F, Higgins DG. Clustal Omega, accurate alignment of very large numbers of sequences. In: Multiple sequence alignment methods 1079. Totowa: Humana Press; 2014. p. 105–116. https://doi.org/10.1007/978-1-62703-646-7_6.
68. Stothard P. The sequence Manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques. 2000;28:1102–4. https://doi.org/10.2144/00286ir01.
69. Tang B, Wang Z, Liu Q, Zhang H, Jiang S, Li X, Wang Z, Sun Y, Sha Z, Jiang H, Wu X, Ren Y, Li H, Xuan F, Ge B, Jiang W, She S, Sun H, Qiu Q, Wang W, Wang Q, Qiu G, Zhang D, Li Y. High-quality genome assembly of *Eriocheir japonica sinensis* reveals its unique genome evolution. Front Genet. 2020;10:1340. https://doi.org/10.3389/fgene.2019.01340.
70. The Galaxy Community. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. Nucleic Acids Res. 2022;50(W1):W345-351. https://doi.org/10.1093/nar/gkac247.
71. Veldsman WP, Ma KY, Hui JHL, Chan TF, Baeza JA, Qin J, Chu KH. Comparative genomics of the coconut crab and other decapod crustaceans: exploring the molecular basis of terrestrial adaptation. BMC Genomics. 2021;22(1):313.
72. Wahle RA, Linnane AJ, Harrington AM, Lovrich G, Thiel M. Lobster fisheries. Nat Hist Crustaceans. 2020;9:55–90.
73. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. KaKs_calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genome Proteome Bioinformatics. 2010;8:77–80. https://doi.org/10.1016/S1672-0229(10)60008-3.
74. Werren JH. Selfish genetic elements, genetic conflict, and evolutionary innovation. Proc Natl Acad Sci. 2011;108(2):10863–70.
75. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 2014;15:1–12.
76. Xiao BH, Zhang W, Yao W, Liu CW, Liu L. Analysis of the complete mitochondrial genome sequence of *Palinura homarus*. Mitochondrial DNA B. 2017;2(1):60–1.
77. Yu J-H, Li H-X, Ge X-P, Li J-L, Tang Y-K. Characteristics analysis of complete ribosomal DNA sequence in Chinese mitten crab (*Eriocheir sinensis*). J Fish China. 2010;34(5):696–703. https://doi.org/10.3724/SP.J.1231.2010.06659.

Publisher's Note