Systematic assessment of the Panopeidae and broader Eubrachyura (Decapoda: Brachyura) using mitochondrial genomics

Lucas A. Jennings¹, April M. H. Blakeslee², Krista A. McCoy³, Donald C. Behringer¹,⁴, Jamie Bojko⁵,⁶

¹ Fisheries and Aquatic Sciences, University of Florida, Gainesville, Florida, 32653, USA
² East Carolina University, Greenville, North Carolina, 27858, USA
³ Harbor Branch Oceanographic Institute at Florida Atlantic University, Fort Pierce, Florida, 34946
⁴ Emerging Pathogens Institute, University of Florida, Gainesville, Florida, 32611, USA
⁵ School of Health and Life Sciences, Teesside University, Middlesbrough, TS1 3BA, UK
⁶ National Horizons Centre, Teesside University, Darlington, DL1 1HG, UK

http://zoobank.org/1C8C96D7-E691-43DD-8122-4A2A7CF054DC

Corresponding author: Jamie Bojko (J.Bojko@tees.ac.uk)

Received 16 June 2021
Accepted 28 August 2021
Published 23 November 2021

Abstract

This study provides a broad phylogenetic analysis for the Eubrachyura, with the inclusion of three new Panopeidae mitochondrial genomes: *Eurypanopeus depressus* (flatback mud crab) (15,854bp), *Panopeus herbstii* (Atlantic mud crab) (15,812bp) and *Rhithropanopeus harrisii* (Harris, or ‘white-fingered’ mud crab) (15,892bp). These new mitogenomes were analyzed alongside all available brachyuran mitochondrial genomes (n = 113), comprising 80 genera from 29 families, to provide an updated phylogenetic analysis of the infra-order Brachyura (“true crabs”). Our analyses support the subsection Potamoida within the Eubrachyura as the sister group to Thoracotremata. The family Panopeidae aligns with the family Xanthidae to form the Xanthoidea branch, which is supported by current morphological and genetic taxonomy. A unique gene arrangement termed ‘XanGO’ was identified for the panopeids and varies relative to other members of the subsection Heterotremata (within the Eubrachyura) via a transposition of the *trnV* gene. This gene arrangement is novel and is shared between several Xanthoidea species, including *Etisus anaglyptus* (hairy spooner crab), *Atergatis floridus* (brown egg crab), and *Atergatis integerrimus* (red egg crab), suggesting that it is a conserved gene arrangement within the Xanthoidea superfamily. Our study further reveals a need for taxonomic revision of some brachyuran groups, particularly the Sesarmidae. The inclusion of panopeid mitogenomes into the greater brachyuran phylogeny increases our understanding of crab evolution and higher level Eubrachyuran systematics.

Keywords

Xanthidae, Panopeus, Eurypanopeus, Rhithropanopeus, mud crab, marine, genomics
1. Introduction

Brachyura (“true” crabs) is the largest subgroup of the Decapoda (Crustacea). It is a ubiquitous group, whose members thrive in terrestrial and aquatic habitats but are particularly prevalent in marine environments (Tsang 2009; Jia et al. 2018; Tan et al. 2018; Ma et al. 2019; Tan et al. 2019). Marine Brachyura boast a broad range of morphological and ecological diversity, leading to a complex taxonomy (Yong-kun et al. 2014). Historically, brachyurans were divided into two sections: the Eurbrachyura and the Podotremata, with the Eurbrachyura being further divided into the subsections: Heterotremata and Thoracotremata (Guinot 2013). The sub-sectioning of these Eurbrachyura is based almost entirely on typological morphology (particularly the genital openings) and has been subject to debate with regards to monophyly. Due to the morphological complexity of this group, genetic tools and analytical methods are typically used to resolve systematic discrepancies (Basso et al. 2017; Bai et al. 2018; Jia et al. 2018; Tan et al. 2018; Wang et al. 2018).

High throughput sequencing (HTS) has proven effective in advancing and resolving taxonomies (Tan et al. 2018). Early studies on brachyuran mitogenomics relied on long PCR and primer walking techniques to read and assemble the mitogenome (Yamauchi 2003; Miller 2005). The rapid sequencing and assembly of the mitochondrial genome (mitogenome) using HTS has proven to be a powerful tool for conducting phylogenetic studies of eukaryotes (Gan et al. 2018). The small size of mitogenomes (~14–16 kb), potential for high mutation rate, and a simple closed structure, make them an ideal marker for inferring an organism’s mitogenetic phylogeny (Boore 1999). Brachyura has been classified into 93 families with over 7000 species (Ng 2008); for 111 genera, excluding global invasions (Thoma 2014). The number of families and subfamilies within the Xanthidea has changed drastically in recent years (Lai 2011). Two common families, the Xanthidae and Panopeidae, share several morphological features that can lead to systematic confusion and difficulty in identifying them beyond the family level (Shih 2011). Both families are found in temperate and tropical shallow intertidal and subtidal zones, but xanthid crabs have a circumtropical distribution while panopeids are only found in the Americas, excluding global invasions (Thoma 2014). To date, there are only four mitogenomes available for the Xanthidae and none for the Panopeidae, whose systematics have primarily relied upon a select number of genes or morphological keys (Williams 1984; Schubart 2000). Studies using conventional PCR to amplify and sequence mitochondrial and nuclear markers revealed that the genera Eurypanopeus and Panopeus are polyphyletic (Schubart 2000). Similarly, studies on the panopeid genus Hexapanaopeus using 12S and 16S genes as markers have also suggested that this genus is polyphyletic (Thoma 2009). Later studies using three mitochondrial markers (COI, 12S and 16S) and three nuclear markers [18S, enolase (ENO) and Histone H3 (H3)] revealed that Xanthoidea is monophyletic, but its two families are not and are in need of taxonomic revision (Thoma 2014).

In this study, we enhance understanding of brachyuran systematics by adding three complete mitogenomes for the Panopeidae: Eurypanopeus depressus, Panopeus herbstii and Rhithronpanopeus harrisii from their native range along the Atlantic coast of North America. The genetic composition, genetic similarity and gene ar-
rangement of these three panopeid species are described relative to other brachyuran mitogenomes, allowing us to update the brachyuran mitogenomic phylogeny and explore brachyuran-wide classification. A new gene arrangement for the superfamily Xanthoidea is described as well as a renaming of previously reported gene arrangements suggested for other Brachyura.

2. Materials and methods

2.1. Specimen collection and dissection

Three species of panopeid mud crabs were collected for this study. First, an individual *Eurypanopeus depressus* was sampled on December 1, 2018 from Hoop Pole Creek, a polyhaline site located in Atlantic beach, North Carolina (NC), USA. The individual was hand-collected at low tide from an intertidal oyster reef and then brought back to the lab for dissection. Second, an individual *Panopeus herbstii* was sampled on August 12, 2019 from Middle Marsh (Beaufort, NC), another polyhaline site, using a passive sampler attached to a wooden stake that had been driven into the sediment. The sampler design is a small plastic milk crate (19×22×16 cm) filled with autoclaved oyster shell (Roche 2007). Third, an individual *Rhithropanopeus harrisii* was sampled on February 5, 2020 from Mallard Creek (Washington, NC), a mesohaline site, using the same passive sampling design as above, but this time attached to a small fishing dock. Crabs were brought back to the lab and anesthetized prior to dissection in a −20°C freezer. Dissections for all three species were carried out using a sterilized razor blade, and part of the hepatopancreas and gills were removed and placed into separate tubes for later DNA extraction.

2.2. DNA extraction, sequencing and assembly

The DNA extractions were conducted on the hepatopancreas and gill tissue of *E. depressus*, *P. herbstii*, and *R. harrisii* using a Zymo DNA extraction kit, according to manufacturer’s protocols. The DNA samples were shipped on dry ice to Novogene, California, who conducted library preparation using the NEBNext Ultra DNA Library Prep Kit. The library was loaded on to a NovaSeq 6000 (Illumina) system using the 150 bp NovaSeq 600 SP reagent kit (300 cycles) for paired end metagenomic sequencing for each individual sample. The resulting data were delivered to the University of Florida for bioinformatic analysis. The data were quality checked and trimmed using Trimmomatic v.0.36 (Bolger 2014) using default parameters. The paired and unpaired reads were assembled using SPAdes v.3.13.0 (Bankevich et al. 2012) with default parameters and k-mer lengths; 21, 33, 55, 77 and 99. The resulting datasets provided a series of contigs that were compared to the NCBI nr database using BLASTx. The mitochondrial genomes of *E. depressus* (574.088X coverage), *P. herbstii* (122.084X coverage) and *R. harrisii* (400.520X coverage) were each identified and circularized. Confirmation of their sequence coverage was conducted using CLC genomics workbench v.12.

The circularized mitogenomes were annotated using MITOS (Bernt 2013). Using the MITOS output, the location of the cox1 gene was determined and the sequences were re-annotated with the cox1 gene at the start of the genome. The putative amino acid and rRNA sequences determined by MITOS were checked using BLASTn and BLASTp (Tables 1–3). The completed genomes were then annotated graphically using Circa (http://omgogenomics.com/circa). The genomes are deposited in GenBank under accession numbers MN399962 (**E. depressus**), MT024989 (*R. harrisii*), and MT024990 (*P. herbstii*).

2.3. Phylogenetic and mitochondrial gene order assessment

There were 112 brachyuran mitogenomes (see Supplementary material 1) obtained from the GenBank database (NCBI) for phylogenetic comparison using the Brachyura taxonomic ID (txid6752) and filtering results to yield DNA sequences of 10,000–25,000 bp (search date: January 2020). The amino acid and nucleotide sequences were retrieved and annotated from these genomes (13 and 15 sequences, respectively) using the Mitophast pipeline (Tan et al. 2015) which downloads each gene in the mitochondrial genome as separate files. The amino acid and nucleotide sequences files were then aligned individually using MAFFT in Geneious (v10.0.2), trimmed to the smallest sequence and concatenated using Geneious. Phylogenetic analyses were conducted in IQtree (Trifinopoulos 2016), which computed the most appropriate evolutionary model (mtMet+F+I+G4) according to BIC for both the amino acid sequences and the nucleotide sequences. A maximum likelihood tree using the amino acid sequences was created using 1000 bootstrap replicates and an SH-aLRT branch test (Guindon 2010) over 3733 positions; the tree had a log score of −157295.9511. A maximum likelihood tree using the nucleotide sequences was created using 1000 bootstrap replicates and an SH-aLRT branch test over 13790 positions; the tree had a log score of −598390.2632. The resulting trees were annotated using FigTree (http://tree.bio.ed.ac.uk/software/figtree). Both trees were rooted with the mitogenomes of *Coenobita brevimanus* (KY352233), *C. rugosus* (KY352235) and *C. perlatus* (KY352234).

All previously reported gene orders for the Brachyura were annotated according to Basso et al. (2017) and Tan et al. (2018). Pairwise comparisons of the gene orders were performed using CREx software (Bernt et al. 2007) at common intervals. The nomenclature for the gene orders follows Basso et al. (2017). MITOS was used to determine the putative location of the control region (CoRe) for *E. depressus, P. herbstii, R. harrisii, Eiusus*
anaglyptus, Leptodius sanguineus, Atergatis floridus, and A. integerrimus through manual examination of the start and stop codons of the open reading frames to look for intergenic spacers. The CoRe for the crabs in the family Potamidae were obtained from Genbank. The CoRe for Echinococcus nipponicus and Pilumnus vespertilio were determined using MITOS following the same method as for the Panopeidae.

3. Results

3.1. The mitochondrial genomes of panopeid crabs

The mitochondrial genomes of the panopeid crabs used in this study were closed circular molecules containing 13 protein coding genes, 22 tRNA genes, 2 rRNA genes, and a single control region (CoRe) (Fig. 1). The E. depressus mitogenome was 15,854 bp in length. The P. herbstii mitogenome was 15,812 bp in length. The R. harrisii mitogenome was 15,967 bp in length. As with most brachyurans, the rrnL (16S) and rrnS (12S) genes were located on the negative strand, as are the nad5, nad4, nad4L, nad1 and 8 tRNA genes (Table 1).

The nucleotide composition of the complete E. depressus mitochondrial genome was as follows: A=5442 (34.32%), T=5509 (34.75%), G=1652 (10.42%), C=3251 (20.51%). The A+T and G+C contents were 69.07% and 30.93%, respectively. The protein coding regions include 7 NADH dehydrogenases (nad1–nad6 and nad4L), three cytochrome c oxidases (cox1–cox3), 2 ATPases (atp6 and atp8) and 1 cytochrome b (cob) and account for 10,838 bp of the mitogenome. The 22 rRNA genes present in the mitogenome range in size from 62 (trnD)–71 (trnL I) bp in length, and the ribosomal RNA genes rrnL (16S) and rrnS (12S) have a length of 1393 bp and 817 bp, respectively. The 13 protein coding genes and majority of the ncRNA sequences showed similarity among the panopeid crabs used in this study (Table 1).

Figure 1. Annotated Circa plots for the circular mitochondrial genomes of Eurypanopeus depressus, Panopeus herbstii and Rithropanopeus harrisii. Each mitogenome is represented by a thick circular black line near the centre of the plot. Protein coding genes are on the outside of this line (negative = dark violet, positive = maroon). Non-coding RNA genes are on the inside of this line (negative = light violet, positive = light maroon). The genome sizes are written in the centre of each plot. The protein coding gene names are represented in the outer most circle (dark grey). The ncRNA gene names are listed in the second internal circle (light grey). The green rectangle labelled “CoRe” indicates the putative control region of the mitochondrial genomes. Figure 1 layout: Portrait. Associated with section 3.1
| Gene     | Strand | Gene hit          | Gene similarity (%) | Gene coverage (%) | Gene e-value | Gene accession                                      | Protein hit                          | Protein similarity | Protein coverage | Protein e-value | Protein accession          |
|----------|--------|-------------------|---------------------|------------------|--------------|---------------------------------------------------|-------------------------------------|-------------------|------------------|----------------|-----------------------------|
| cox1     | +      | *Rhithropanopeus harrisii* MCG | 88.05               | 100              | 0.0          | Present study                                      | cytochrome c oxidase subunit I [*Rhithropanopeus harrisii*] | 99.8             | 100              | 0.0           | Present study               |
| tmL2(tu) | +      | Panopeus herbstii MCG | 92.42               | 100              | 5e-23        | Present study                                      | —                                   | —                | —                | —              | —                           |
| cox2     | +      | *Rhithropanopeus harrisii* MCG | 88.99               | 100              | 0.0          | Present study                                      | cytochrome c oxidase subunit II [Panopeus herbstii] | 99.11            | 100              | 1e-170        | Present study               |
| tmK(aaa) | +      | Panopeus herbstii MCG | 95.52               | 100              | 2e-27        | Present study                                      | —                                   | —                | —                | —              | —                           |
| tmD(gac) | +      | *Rhithropanopeus harrisii* MCG | 100                 | 100              | 3e-32        | Present study                                      | —                                   | —                | —                | —              | —                           |
| atp8     | -      | —                  | —                   | —                | —            | ATP synthase F0 subunit 8 [Panopeus herbstii]      | 90.20                 | 100              | 6e-18          | Present study               |
| atp6     | +      | Panopeus herbstii MCG | 87.80               | 100              | 0.0          | Present study                                      | ATP synthase F0 subunit 6 [Panopeus herbstii] | 99.10            | 100              | 6e-158         | Present study               |
| cox3     | +      | *Rhithropanopeus harrisii* MCG | 89.83               | 100              | 0.0          | Present study                                      | cytochrome c oxidase subunit III [*Rhithropanopeus harrisii*] | 99.61            | 100              | 0.0           | Present study               |
| tmG(gga) | +      | *Rhithropanopeus harrisii* MCG | 98.41               | 100              | 1e-30        | Present study                                      | —                                   | —                | —                | —              | —                           |
| nad3     | +      | —                  | —                   | —                | —            | NADH dehydrogenase subunit 3 [Panopeus herbstii]   | 96.49                 | 99               | 1e-79          | Present study               |
| tmA(gaa) | +      | Panopeus herbstii MCG | 98.41               | 100              | 1e-28        | —                                   | —                                   | —                | —                | —              | —                           |
| tmR(gaa) | +      | *Rhithropanopeus harrisii* MCG | 98.44               | 100              | 4e-31        | Present study                                      | —                                   | —                | —                | —              | —                           |
| tmN(gac) | +      | —                  | —                   | —                | —            | —                                   | —                                   | —                | —                | —              | —                           |
| tmS1(aaa) | +   | *Rhithropanopeus harrisii* MCG | 98.51               | 100              | 9e-31        | Present study                                      | —                                   | —                | —                | —              | —                           |
| tmI(gaa) | +      | *Rhithropanopeus harrisii* MCG | 95.45               | 100              | 6e-29        | Present study                                      | —                                   | —                | —                | —              | —                           |
| tmH(gac) | —      | *Rhithropanopeus harrisii* MCG | 95.31               | 100              | 9e-26        | Present study                                      | —                                   | —                | —                | —              | —                           |
| tmF(ttc) | —      | *Rhithropanopeus harrisii* MCG | 95.31               | 100              | 9e-26        | Present study                                      | —                                   | —                | —                | —              | —                           |
| nad5     | —      | *Rhithropanopeus harrisii* MCG | 88.79               | 98               | 0.0          | Present study                                      | NADH dehydrogenase subunit 5 [*Rhithropanopeus harrisii*] | 92.75            | 98               | 0.0           | Present study               |
| nad4     | —      | *Rhithropanopeus harrisii* MCG | 87.30               | 99               | 0.0          | Present study                                      | NADH dehydrogenase subunit 4 [Panopeus herbstii] | 96.15            | 99               | 0.0           | Present study               |
| nad4L    | —      | *Rhithropanopeus harrisii* MCG | 91.21               | 98               | 1e-107       | Present study                                      | NADH dehydrogenase subunit 4L [Panopeus herbstii] | 100.00           | 100              | 8e-66         | Present study               |
| tmT(aca) | +      | *Rhithropanopeus harrisii* MCG | 98.39               | 95               | 2e-29        | Present study                                      | —                                   | —                | —                | —              | —                           |
| tmP(aca) | +      | *Rhithropanopeus harrisii* MCG | 98.46               | 100              | 1e-31        | Present study                                      | —                                   | —                | —                | —              | —                           |
| nad6     | +      | *Rhithropanopeus harrisii* MCG | 85.51               | 98               | 4e-147       | Present study                                      | NADH dehydrogenase subunit 6 [*Rhithropanopeus harrisii*] | 92.73            | 100              | 6e-90         | Present study               |
| cob      | +      | *Rhithropanopeus harrisii* MCG | 86.77               | 99               | 0.0          | Present study                                      | cytochrome b [Panopeus herbstii]     | 98.68            | 100              | 0.0           | Present study               |
| Genome            | Start  | End    | Gene    | Strand | Gene hit                          | Gene similarity (%) | Gene Coverage (%) | Gene e-value | Gene accession | Protein hit | Protein similarity | Protein cover | Protein e-value | Protein accession |
|-------------------|--------|--------|---------|--------|-----------------------------------|---------------------|------------------|--------------|----------------|-------------|-------------------|---------------|-----------------|------------------|
| Eurypanopeus depressus mitochondrial genome | 10117  | 10183  | trnS2(tca) | +      | Echinococus nipponicus voucher MABIK CR00241788 MCG | 95.85              | 97               | 2e-15        | NC_039618.1     | —           | —                 | —             | —               |
|                   | 10235  | 11134  | nad1    | —      | Rhithropanopeus harrisi MCG       | 88.95              | 99               | 0.0          | Present study   | NADH dehydrogenase subunit 1 [Rhithropanopeus harrisi] | 98.67      | 100             | 0.0              | Present study   |
|                   | 11171  | 11242  | trnL1(ctt) | —      | —                                  | —                  | —                | —             | —              | —           | —                 | —             | —               |
|                   | 11217  | 12610  | rrnL    | —      | Eurypanopeus depressus voucher USNM 16S RNA gene | 91.03              | 98               | 0.0          | KT959469.1      | —           | —                 | —             | —               |
|                   | 12705  | 13522  | rrnS    | —      | Eurypanopeus depressus voucher ULLZ 12S ribosomal RNA gene, partial sequence; mitochondrial | 99.73              | 44               | 0.0          | EU863325.2      | —           | —                 | —             | —               |
|                   | 14140  | 14204  | trnV(gta) | —      | Rhithropanopeus harrisi MCG       | 95.45              | 100              | 3e-26        | Present study   | —           | —                 | —             | —               |
|                   | 14422  | 14489  | tm(lct) | +      | Panopeus herbsti MCG              | 97.06              | 100              | 4e-31        | Present study   | —           | —                 | —             | —               |
|                   | 14487  | 14555  | trnQ(caa) | —     | Rhithropanopeus harrisi MCG       | 95.65              | 100              | 2e-30        | Present study   | —           | —                 | —             | —               |
|                   | 14579  | 14646  | tm(atg) | +      | Atergatis floridus MCG            | 98.53              | 100              | 1e-22        | NC_037201.1     | —           | —                 | —             | —               |
|                   | 14659  | 15621  | nad2    | +      | Rhithropanopeus harrisi MCG       | 82.18              | 100              | 0.0          | Present study   | NADH dehydrogenase subunit 2 [Panopeus herbsti] | 90.62      | 99              | 0.0              | Present study   |
|                   | 15656  | 15723  | trnW(tga) | +      | —                                  | —                  | —                | —             | —              | —           | —                 | —             | —               |
|                   | 15724  | 15787  | trnC(tgc) | —      | Panopeus herbsti MCG              | 96.88              | 100              | 2e-29        | NC_037201.1     | —           | —                 | —             | —               |
|                   | 15788  | 15852  | trnY(tac) | —      | Etisus anaglyptus MCG             | 90.77              | 100              | 4e-12        | NC_042208.1     | —           | —                 | —             | —               |
Table 2. Nucleotide and protein similarity data for the protein-coding and non-coding genes of the Panopeus herbstii mitochondrial genome. The data represented were acquired from BLASTn and BLASTp outputs via comparison against the complete non-redundant database. The accession number of the specific nucleotide or amino acid sequence are provided in addition to the species, if known, belonging to the sequence isolate. The similarity (%), coverage comparison (%) and e-value are all provided. MCG = mitochondrion, complete genome.

| Genome | Start | End | Gene | Strand | Gene hit | Gene similarity (%) | Gene Coverage (%) | Gene e-value | Protein hit | Protein similarity | Protein cover | Protein e-value | Protein accession |
|--------|-------|-----|------|--------|----------|---------------------|-----------------|-------------|-------------|------------------|---------------|----------------|------------------|
| 1      | 1515  |     | cox1 | +      | Rhithropanopeus harrisi MCG | 87.52             | 100              | 0.0         | Present study | cytochrome c oxidase subunit I | 100.00        | 100            | 0.0              | Present study |
| 1535   | 1600  |     | tmL2(tu) | +     | Eurypanopeus depressus MCG | 92.42             | 100              | 6e-23       | Present study | —              | —             | —               | —               |
| 1607   | 2278  |     | cox2 | +      | Rhithropanopeus harrisi MCG | 88.24             | 100              | 0.0         | Present study | cytochrome c oxidase subunit II | 99.11         | 100            | 1e-170           | Present study |
| 2292   | 2358  |     | tmK(aaa) | +    | Eurypanopeus depressus MCG | 95.52             | 100              | 2e-29       | Present study | —              | —             | —               | —               |
| 2359   | 2421  |     | tmD(gac) | +     | Rhithropanopeus harrisi MCG | 95.24             | 100              | 3e-27       | Present study | —              | —             | —               | —               |
| 2422   | 2574  |     | atp8 | +      | —         | —                  | —                | —           | —           | ATP synthase F0 subunit 8 | 90.20         | 100            | 6-e18            | Present study |
| 2577   | 3239  |     | atp6 | +      | Eurypanopeus depressus MCG | 87.80             | 100              | 0.0         | Present study | ATP synthase F0 subunit 6 | 99.10         | 100            | 6e-158           | Present study |
| 3257   | 4033  |     | cox3 | +      | Rhithropanopeus harrisi MCG | 89.32             | 100              | 0.0         | Present study | cytochrome c oxidase subunit III | 98.46         | 100            | 0.0              | Present study |
| 4039   | 4102  |     | tmG(gga) | +    | Rhithropanopeus harrisi MCG | 96.88             | 100              | 6e-29       | Present study | —              | —             | —               | —               |
| 4109   | 4450  |     | nad3 | +      | —         | —                  | —                | —           | —           | NADH dehydrogenase subunit 3 | 96.49         | 99             | 1e-79            | Present study |
| 4458   | 4520  |     | tmA(gca) | +    | Eurypanopeus depressus MCG | 98.41             | 100              | 1e-28       | Present study | —              | —             | —               | —               |
| 4521   | 4583  |     | tmR(gga) | +     | Rhithropanopeus harrisi MCG | 96.88             | 100              | 6e-29       | Present study | —              | —             | —               | —               |
| 4584   | 4651  |     | tmN(gac) | +     | —         | —                  | —                | —           | —           | —              | —             | —             | —               | —               |
| 4653   | 4719  |     | tmS1(gaa) | +   | Eurypanopeus depressus MCG | 97.01             | 100              | 4e-29       | Present study | —              | —             | —               | —               |
| 4722   | 4785  |     | trnE(gaa) | +    | Eurypanopeus depressus MCG | 95.45             | 100              | 2e-28       | Present study | —              | —             | —               | —               |
| 4805   | 4868  |     | tmH(gac) | —     | Rhithropanopeus harrisi MCG | 96.88             | 100              | 2e-27       | Present study | —              | —             | —               | —               |
| 4869   | 4935  |     | tmF(tsc) | —     | —         | —                  | —                | —           | —           | —              | —             | —             | —               | —               |
| 4943   | 6550  |     | nad5 | —      | Rhithropanopeus harrisi MCG | 87.93             | 99               | 0.0         | Present study | NADH dehydrogenase subunit 5 | 93.64         | 99             | 0.0              | Present study |
| 6725   | 8047  |     | nad4 | —      | Rhithropanopeus harrisi MCG | 85.54             | 99               | 0.0         | Present study | NADH dehydrogenase subunit 4 | 96.15         | 99             | 0.0              | Present study |
| 8044   | 8319  |     | nad4L | —     | Rhithropanopeus harrisi MCG | 90.84             | 98               | 6e-106      | Present study | NADH dehydrogenase subunit 4L | 100.00        | 100            | 8e-66            | Present study |
| 8346   | 8409  |     | trnT(aac) | +   | —         | —                  | —                | —           | —           | —              | —             | —             | —               | —               |
| 8410   | 8474  |     | tmP(gca) | —     | Rhithropanopeus harrisi MCG | 98.46             | 100              | 1e-31       | Present study | —              | —             | —               | —               |
| 8477   | 8974  |     | nad6 | —      | Eurypanopeus depressus MCG | 83.54             | 97               | 2e-130      | Present study | NADH dehydrogenase subunit 6 | 91.52         | 99             | 7e-89            | Present study |
| Genome      | Start | End   | Gene | Strand | Gene hit | Gene similarity (%) | Gene Coverage (%) | Gene e-value | Gene accession | Protein hit | Protein similarity (%) | Protein cover | Protein e-value | Protein accession |
|-------------|-------|-------|------|--------|----------|---------------------|------------------|--------------|----------------|-------------|-----------------------|---------------|----------------|-------------------|
| Panopeus herbstii Mitochondrial Genome | 8983  | 10119 | cob  | +      | Rhithropanopeus harrisii MCG | 87.13             | 99              | 0.0          | Present study | cytochrome b [Eurypanopeus depressus] | 98.68       | 100            | 0.0               | Present study |
| 10118       | 10184 | trnS2(tca) | +    | Rhithropanopeus harrisii MCG | 92.65             | 100             | 4e-26         | Present study | —             | —          | —                    | —             | —             | —                 |
| 10230       | 11135 | nad1 | —    | Rhithropanopeus harrisii MCG | 89.40             | 98              | 0.0          | Present study | NADH dehydrogenase subunit 1 [Eurypanopeus depressus] | 97.00       | 99            | 0.0               | Present study |
| 11171       | 11239 | trnL1(cca) | —    | —      | —        | 100.00            | 37              | 0.0          | KT959516.1 | —           | —                    | —             | —             | —                 |
| 11194       | 12584 | rNL  | —    | Panopeus herbstii voucher USNM: 16S RNA gene, mitochondrial | 99.46             | 44              | 0.0          | EU863296 | —           | —                    | —             | —             | —                 |
| 12683       | 13502 | rNS  | —    | Panopeus herbstii voucher ULLZ 8457 12S ribosomal RNA gene, partial sequence; mitochondrial | 99.46             | 44              | 0.0          | NC_042208 | —           | —                    | —             | —             | —                 |
| 14124       | 14190 | tmV(gta) | —    | —      | —        | 97.06             | 100             | 4e-31         | Present study | —           | —                    | —             | —             | —                 |
| 14357       | 14423 | tmI(ac) | +    | Eurypanopeus depressus MCG | 95.65             | 100             | 2e-30         | Present study | —           | —                    | —             | —             | —                 |
| 14421       | 14489 | tmQ(caa) | —    | Eurypanopeus depressus MCG | 98.55             | 100             | 9e-25         | NC_042208 | —           | —                    | —             | —             | —                 |
| 14541       | 14609 | tmM(atg) | +    | Etisus anglyptus MCG | 98.55             | 100             | 9e-25         | NC_042208 | —           | —                    | —             | —             | —                 |
| 14622       | 15581 | nad2 | +    | —      | —        | NADH dehydrogenase subunit 2 [Eurypanopeus depressus] | 90.62       | 100            | 0.0          | Present study | —           | —                    | —             | —             | —                 |
| 15619       | 15685 | tmW(tga) | +    | —      | —        | —                | —               | —             | —           | —                    | —             | —             | —                 |
| 15685       | 15748 | tmC(tgc) | —    | Rhithropanopeus harrisii MCG | 98.44             | 100             | 4e-31         | Present study | —           | —                    | —             | —             | —                 |
| 15749       | 15812 | tmY(tac) | —    | Rhithropanopeus harrisii MCG | 92.31             | 100             | 2e-24         | Present study | —           | —                    | —             | —             | —                 |
Table 3. Nucleotide and protein similarity data for the protein-coding and non-coding genes of the Rhithropanopeus harrisii mitochondrial genome. The data represented were acquired from BLASTn and BLASTp outputs via comparison against the complete non-redundant database. The accession number of the specific nucleotide or amino acid sequence are provided in addition to the species, if known, belonging to the sequence isolate. The similarity (%), coverage comparison (%) and e-value are all provided. MCG = mitochondrion, complete genome.

| Genome Start | End | Gene | Strand | Gene hit | Gene similarity (%) | Gene Coverage (%) | Gene e-value | Protein hit | Protein similarity | Protein cover | Protein e-value | Protein accession |
|--------------|-----|------|--------|----------|---------------------|------------------|--------------|-------------|------------------|---------------|-----------------|------------------|
| 1            | 1515| cox1 | +      | Rhithropanopeus harrisii mitochondrial partial COI gene for cytochrome oxidase subunit I, isolate R617-8 | 99.39 | 65 | 0.0 | LN810615 | cytochrome c oxidase subunit I [Panopeus herbstii] | 100.00 | 100 | 0.0 | Present study |
| 1535         | 1599| trnL2(tta) | + | - | 88.99 | 100 | 0.0 | Present study | cytochrome c oxidase subunit II [Panopeus herbstii] | 99.11 | 100 | 8e-170 | Present study |
| 1607         | 2278| cox2 | + | Eurypanopeus depressus MCG | 95.52 | 100 | 7e-29 | Present study | — | — | — | — |
| 2292         | 2357| tmK(aaa) | + | Panopeus herbstii MCG | 100.00 | 100 | 3e-32 | Present study | — | — | — | — |
| 2358         | 2420| tmD(gac) | + | Eurypanopeus depressus MCG | 98.41 | 100 | 1e-30 | Present study | — | — | — | — |
| 2421         | 2573| aps8 | + | - | 88.54 | 100 | 0.0 | Present study | ATP synthase F0 subunit 8 [Panopeus herbstii] | 84.31 | 100 | 2e-16 | Present study |
| 2576         | 3238| ap6 | + | Eurypanopeus depressus MCG | 98.41 | 100 | 0.0 | Present study | — | — | — | — |
| 3256         | 4032| cox3 | + | Eurypanopeus depressus MCG | 98.41 | 100 | 0.0 | Present study | — | — | — | — |
| 4038         | 4100| trnG(gga) | + | Eurypanopeus depressus MCG | 98.41 | 100 | 0.0 | Present study | — | — | — | — |
| 4107         | 4448| nad3 | + | - | 98.41 | 100 | 1e-28 | Present study | — | — | — | — |
| 4455         | 4517| tmA(gca) | + | Eurypanopeus depressus MCG | 98.41 | 100 | 4e-31 | Present study | — | — | — | — |
| 4518         | 4581| tmR(cga) | + | Eurypanopeus depressus MCG | 98.44 | 100 | 4e-31 | Present study | — | — | — | — |
| 4582         | 4648| tmN(aac) | + | - | 98.51 | 100 | 4e-31 | Present study | — | — | — | — |
| 4651         | 4717| tmS1(aaa) | + | Eurypanopeus depressus MCG | 98.51 | 100 | 4e-31 | Present study | — | — | — | — |
| 4721         | 4786| tmE(gaa) | + | Eurypanopeus depressus MCG | 95.45 | 100 | 6e-29 | Present study | — | — | — | — |
| 4803         | 4866| tmH(cac) | — | Panopeus herbstii MCG | 96.88 | 100 | 2e-17 | Present study | — | — | — | — |
| 4867         | 4930| trnT(ttc) | — | Eurypanopeus depressus MCG | 95.31 | 100 | 8e-28 | Present study | — | — | — | — |
| 4941         | 6554| nad5 | — | Eurypanopeus depressus MCG | 88.79 | 99 | 0.0 | Present study | NADH dehydrogenase subunit 5 [Eurypanopeus depressus] | 92.75 | 98 | 0.0 | Present study |
| 6712         | 8037| nad4 | — | Eurypanopeus depressus MCG | 87.30 | 99 | 0.0 | Present study | NADH dehydrogenase subunit 4 [Eurypanopeus depressus] | 94.80 | 100 | 0.0 | Present study |
| 8034         | 8309| nad4L | — | Eurypanopeus depressus MCG | 91.21 | 98 | 1e-107 | Present study | NADH dehydrogenase subunit 4L [Eurypanopeus depressus] | 96.74 | 100 | 2e-64 | Present study |
| 8336         | 8400| trnT(aaa) | — | Eurypanopeus depressus MCG | 88.39 | 95 | 2e-29 | Present study | — | — | — | — |
| 8401         | 8465| trnP(cca) | — | Panopeus herbstii MCG | 89.46 | 100 | 1e-31 | Present study | — | — | — | — |
| Accession | Tissue | Mitochondrial Genome Length (Kb) | P. harrisii Maternal | P. harrisii Paternal | E. depressus Maternal | E. depressus Paternal | B. simillimis Maternal | B. simillimis Paternal | H. serrata Maternal | H. serrata Paternal | Y. alaskensis Maternal | Y. alaskensis Paternal |
|-----------|--------|-------------------------------|---------------------|---------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| KT959486.1 | USNM 12S ribosomal RNA gene, partial sequence | 100.00 | 98 | 0.0 | | | | | | | | |
| EU863280 | ULLZ 3995 12S ribosomal RNA gene, partial sequence | 98.90 | 44 | 0.0 | | | | | | | | |
| MCG | | | | | | | | | | | | |
The nucleotide composition of the complete _P. herbstii_ mitochondrial genome was as follows: A=5520 (34.91%), T=5687 (35.97%), G=1627 (10.29%), C=2980 (18.85%). The A+T and the G+C contents were 70.87% and 29.13%, respectively. The protein coding region contains 7 NADH dehydrogenases (nad1–nad6 and nad4L), three cytochrome c oxidases (cox1–cox3), 2 ATPases (atp6 and atp8) and 1 cytochrome b (cob) and accounts for 10947 bp of the mitogenome of _P. herbstii_.

The 22 rRNAs present in the mitogenome range in size from 63 (trnD, trnA, trnR) – 69 (trnL1, trnQ, trnM) bp in length, and the ribosomal RNA genes rrnL (16S) and rrnS (12S) have a length of 1392 bp and 820 bp, respectively. All 13 protein coding genes showed high similarity to the panopeid crabs used in this study. The ncRNAs all showed similarity to decapod crustaceans with the majority having high similarity with the Panopeidae (Table 2).

The nucleotide composition of the complete _R. har risi i_ mitochondrial genome was as follows: A=5595 (34.21%), T=5873 (37.00%), G=1556 (9.82%), C=2866 (17.99%). The A+T and the G+C contents were 72.20% and 27.080%, respectively. The protein coding region contains 7 NADH dehydrogenases (nad1–nad6 and nad4L), three cytochrome c oxidases (cox1–cox3), 2 ATPases (atp6 and atp8) and 1 cytochrome b (cob) and accounts for 10,848 bp of the mitogenome. The 22 rRNAs present in the mitogenome range in size from 63 (trnD, trnG, trnR) – 69 (trnL1, trnQ, trnM) bp in length, and the ribosomal RNA genes rrnL (16S) and rrnS (12S) have a length of 1392 bp and 820 bp, respectively. All 13 protein coding genes showed high similarity to the panopeid crabs used in this study (Table 2).

### 3.2. Phylogenetics

To establish where the panopeid crabs align within the Eubrachyura, amino acid and nucleotide sequences from 112 mitogenomes comprising 77 genera from 28 families of brachyuran crabs were used along with the three new mitogenomes (Fig. 2). Two sequences that are publicly available for brachyurans were not included in our analysis due to inconsistencies with the sequences. (1) The protein sequences for _Gecarcoidea natalis_ contained ambiguous amino acid identifications, resulting in poor alignment with other members within the superfamily Grapsoidea. (2) The protein sequences for _Pyrhila pisum_ aligned poorly with other members of the Brachyura; however, there were no missing protein codes. When tested in BLASTp, the proteins for _P. pisum_ yielded low identity with other brachyurans; < 60% identity in most cases.

Four distinct clades were identified (Fig. 2). One clade belongs to crabs in the subsection Heterotremata (n=40), a second belongs to crabs in the subsection Thoracotremata (n=44) and a third belongs to crabs in the section Podotremata (n=7). The fourth clade belongs to the ‘Old World’ freshwater crabs in the superfamilies Potamoidea and Gecarcinucoidea (n=20). This fourth clade forms a subsection termed Potamoidea, a sister group to Thoracotremata. The split between the Heterotremata and the Potamoida/Thoracotremata clades is well supported using both amino acid sequences (Sh-aLRT/UFBoot: 100/100) as well as nucleotide sequences (Sh-aLRT/UF-Boot:100/100). The Potamoida and Thoracotremata split is also well supported using both sequence types (amino acids and nucleotides).
The panopeid crab species *E. depressus*, *P. herbstii* and *R. harrisii* formed a branch for the family Panopeidae (Fig. 2 and Fig. 3; “β”) aligned alongside the xanthid branch to form the superfamily Xanthoidea (amino acids- Sh-aLRT/UFBoot: 100/100; nucleotides- Sh-aLRT/UFBoot: 100/100). The xanthid branch contains members of the Xanthidae family: *E. anaglyptus*, *A. floridus* and *A. integerrimus* (Fig. 2 and Fig. 3; “α”). When considering its amino acid sequences, the crab species *Epixanthus frontalis* from the family Oziidae aligns with the Xanthoidea superfamily with moderate support (Sh-aLRT/UFBoot: 89.4/94) (Fig. 2). The nucleotide sequences for *E. frontalis* show a similar pattern; however, *Leptodius sanguineus* is part of the branch with middling support (Sh-aLRT/UFBoot: 66.7/93) (Fig. 3). Based on amino acid comparison, *L. sanguineus* (considered a member of the Xanthidae) aligns between *E. frontalis* and members of the Pilumnidae, on a branch separate from other xanthid crabs (Sh-aLRT/UFBoot: 66.7/93) (Fig. 3). On the other hand, the new gene arrangement XanGO shares 870 common intervals with PanGO and 988 common intervals with BraGO (Fig. 3), suggesting it to be a low-level rearrangement relative to the common gene arrangements. The new XanGO is most different to the MaVaGO, sharing only 80 common intervals.

The gene arrangements for the panopeid crabs *E. depressus*, *P. herbstii* and *R. harrisii* (Fig. 3) correspond in synteny to other sequenced xanthid species: *E. anaglyptus*, *A. floridus* and *A. integerrimus* (Fig. 2 and Fig. 3; “α”). When considering its amino acid sequences, the crab species *Epixanthus frontalis* from the family Oziidae aligns with the Xanthoidea superfamily with moderate support (Sh-aLRT/UFBoot: 89.4/94) (Fig. 2). The nucleotide sequences for *E. frontalis* show a similar pattern; however, *Leptodius sanguineus* is part of the branch with middling support (Sh-aLRT/UFBoot: 66.7/93) (Fig. 3). Based on amino acid comparison, *L. sanguineus* (considered a member of the Xanthidae) aligns between *E. frontalis* and members of the Pilumnidae, on a branch separate from other xanthid crabs (Sh-aLRT/UFBoot: 16/66) (Fig. 2). The amino acid phylogeny suggests that the hydrothermal vent crabs in the family Xenograpsidea align with the terrestrial crabs in the family Ocypodidae (Sh-aLRT/UFBoot: 83.2/72), yet the nucleotide sequences suggest that the xenograpids form their own branch alongside of the sesarmid crabs (Sh-alRT/UFBoot: 100/99).

The family Sesarmidae (10 mitogenomes) appears to be polyphyletic. Rather than grouping together, the genus *Chiromantes* is split, where *C. dehaani* aligns with *Sesarma neglectum* (amino acids- Sh-alRT/UFBoot: 99.5/100; nucleotides- Sh-alRT/UFBoot: 100/100), and *C. haematocheir* aligns with *Sesarmops sinensis* (amino acids- Sh-alRT/UFBoot: 99.7/100; nucleotides- Sh-alRT/UFBoot: 100/100) (Fig. 2 and Fig. 3).
The mitogenomes of the crabs in the family Panopeidae all shared a ~600 bp long intergenic spacer between the \textit{rrnS} and \textit{trnV} ncRNA genes (\textit{E. depressus}, 618 bp; \textit{P. herbstii}, 622 bp; \textit{R. harrisii}, 644 bp) representing the control region (CoRe) (Fig. 1). The CoRe in the panopeid mitogenomes are A + T skewed (78.40–80.22%) and contain the repeated motifs TA (125–107), AT (112–104), TAA (47–39), TTA (30–40), ATA (43–35) and TAT (41–37).

The mitogenomes of the xanthid crabs used in this study also have similar sized intergenic spacers in this region, suggesting that this is the putative location of the CoRe for members of the superfamily Xanthoidea. The CoRe

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Gene orders (\textit{\textit{GO}}) found among brachyuran crabs. Red boxes indicate protein coding genes. Blue boxes indicate tRNA’s. Green boxes indicate rRNAs. Purple boxes indicate the control region (CoRe). The red lines along the bottom of the gene orders represents areas within the gene order that are located on the negative strand. Not shown are the 9 unique gene orders for the freshwater crabs (see Zhang et al. 2020). The CREx results are listed for the different gene orders. In the associated table, gene orders with high similarity (> 1000) have red boxes while those with low similarity (< 200) have blue boxes. Intermediate similarity remains white.}
\end{figure}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|}
\hline
 & PanGo & BraGo & XanGo & MajGO & MaVA Go & SesGO & XenGo & SinGo & GeoGo & DamGo & DynGo & HuaGo & PotGo & PliGo & SomGo \\
\hline
PanGo & 1326 & & & & & & & & & & & & & & \\
BraGo & 1326 & 1326 & & & & & & & & & & & & & \\
XanGo & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
MajGO & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
MaVA Go & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
SesGO & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
XenGo & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
SinGo & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
GeoGo & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
DamGo & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
DynGo & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
HuaGo & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
PotGo & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
PliGo & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
SomGo & 1326 & 1326 & 1326 & & & & & & & & & & & & \\
\hline
\end{tabular}
\end{table}
nucleotide sequence for all species within Xanthoidea were isolated and run through BLASTn, resulting in a lack of any significant similarity, suggesting high mutability.

4. Discussion

This study provides the first mitochondrial genomes for three members of the Panopeidae and an updated concatenated mito-phylogenetic analysis for the Eubrachyura (excluding nuclear genetic data), informing upon the systematics of multiple families and higher taxonomic rankings. In addition, the mitochondrial genomes for members of the Panopeidae are identified with a consensus gene arrangement shared with other Xanthoidea (XanGO). These results advance our systematic understanding of the brachyurans through the exploration of mitochondrial genomics and gene synteny rearrangement events.

4.1. Xanthid systematics considering panopeid mitogenomic data

The mitogenomes of the panopeid crabs *E. depressus*, *P. herbstii* and *R. harrisi* support the position of the Panopeidae within the Heterotremata, helping to build/support the branch belonging to the superfAMILY Xanthoidea (Ng 2008). Along this branch, the Xanthidae and Panopeidae form sister groups, additionally supported by previous genetic data using five or less mitochondrial and nuclear genes (Thoma 2009; Lai 2011; Thoma 2014). The genera within these families have been historically identified as polyphyletic (Thoma 2009) and the limited number of mitogenomes available makes it difficult to determine their validity. We acknowledge that the families Xanthidae and Panopeidae both occur in two forms: sensu stricto and sensu lato (Ng 2008). There are 4 publicly available mitogenomes for the Xanthidae (GenBank) and we provide 3 additional mitogenomes for the Panopeidae. We have treated these families in their simple form due to the lack of genetic information to split them further. As more mitogenomes become available, the validity of the two forms should be revisited.

Several taxonomic conflicts appear when considering mitogenetics surrounding the Xanthoidea. First, based on morphology and limited mitogenome availability, the genus *Leptodius* is considered a member of the family Xanthidae. However, despite this genus having 12 separate species, only one mitogenome (the species *L. sanguineus*) is available for analysis. Previous studies showed that *L. sanguineus* aligns closely with other members of the Xanthidae (Sung 2016; Karagozlu 2018; Xie 2018; Ma 2019), but these studies use fewer brachyuran mitogenomes in their analysis prior to our study. When considering all mitogenomes available for the Brachyura in our investigation, *L. sanguineus* aligns more closely with the members of the family Oziidae, rather than the Xanthidae. This interesting observation merits further exploration.

4.2. Mitogenomic gene arrangements across the Brachyura

Gene arrangement changes were once thought to be rare (Boore 2000) but with greater availability of mitogenome sequencing, it appears that changes in gene arrangements can be common across groups. For example, gene order is conserved within Ostechithyes and some subgroups of Mammalia, while it varies strongly in e.g. Ctenophora (Arafat 2018), Mollusca (Guerra 2018), Hymenoptera (Dowton 1999) and Anomura (Tan et al. 2018). For Crustacea, some species within the Stomatopoda, Amphipoda and Dendrobranchiata stay within the PanGO ground pattern of Pancrustacea (Shen 2011), while no sequenced species within the Brachyura have retained this gene order. Studies on gene order rearrangement are ongoing with some hypothesizing that the evolution to living within harsh environments, such as the deep sea or hydrothermal vents, can lead to new gene synteny (Nakajima 2016; Gan 2018; Tan et al. 2019).

The brachyurans include several families found in the deep sea. Two of them are represented herein: Bythograeidae and Xenograpsidae. Bythograeidae possess the BraGO arrangement plesiomorphic for Brachyura, while Xenograpsidae have their own gene arrangement (XenGO). In contrast, the freshwater crab family Potamidae has 9 different gene arrangements (Zhang 2020). Brachyuran crabs represent both cases: the adaptation from a marine to a freshwater environment was likely harsh and may have resulted in several new gene arrangements, while in contrast, the evolution of crabs to the deep-sea benthos resulted in some retaining the ancestral gene order in the face of a new environmental extreme. Therefore, when considering crabs, living within harsh environments does not seem to be the only answer to gene arrangement plasticity, but perhaps requires consideration at the finer scale of environmental adaptation. Similar findings have been reported by Tan et al. (2019) who found little evidence for linking gene order rearrangements with adaptations to extreme environments, concluding that these cues are poorly understood and merit a more detailed approach.

A comparison of the eubrachyuran subsections shows that Heterotremata has a higher diversity of gene arrangements than Thoracotremata. Both subsections share species whose gene arrangement follows the basic BraGO pattern. Aside from the BraGO, Thoracotremata only has 3 unique gene arrangements while Heterotremata has 8 unique gene arrangements (including the herein newly established XanGO). This does not include the gene arrangements for the freshwater crabs in the superfamilies Potamoidea and Gecarcinucoidea. The freshwater crabs have more unique gene arrangements than the known Heterotremata.

The panopeid crabs *E. depressus*, *P. herbstii* and *R. harrisi* all have the *rrnV* gene transposed from between the *rrnL* and *rrnS* genes to a location past the CoRe.
This differs from the PanGO, BraGO, SesGO, XenGO, DamGO, MajGO and DynGO, which all have the trnV gene located between the rrnL and rrnS genes, with the CoRe following the rrnS gene. The xanthid crabs *E. anglyptus, L. sanguineus, A. floridus* and *A. integerrimus* all share the latter gene arrangement, suggesting that it might be a conserved arrangement within Xanthoidea and thus support our interpretation of the new Xanthoidea gene arrangement (XanGO). The intergenic spacer found between the rrnS and trnV genes in panopeids appears to be the putative location of the CoRe for these species and is shared with xanthid species, *E. anglyptus, A. floridus* and *A. integerrimus*. All have similarly sized intergenic spacers (600–750 bp long) at this location, suggesting that this may be the location of the CoRe across Xanthoidea. Apart from *L. sanguineus*, the Xanthidae all follow the new gene arrangement XanGO. *Leptodius sanguineus* follows the plesiomorphic brachyuran gene arrangement BraGO and based on its amino acid sequences, it groups more closely with the family Pilummiidae than the members of the Xanthidae or the Panopeidae presented here; however, nodal support is low, meriting further study and sequencing of closer relatives. Higher nodal support is offered with the nucleotide tree, where *L. sanguineus* groups with *Epixanthus frontalis* from Oziidae rather than with the xanthids. Based on the molecular taxonomy and its gene arrangement, the placement of *L. sanguineus* within Xanthidae appears to be invalid and in need of revision, adding to our explanation above.

The mitogenome analysis we performed also supports the renaming of two gene arrangements and confirms the correct gene sequence for another. Two mitogenomes were available for the pilummid crabs, *Echinoecus nipponicus* and *Pilumnus vespertilio*. They follow the gene arrangement reported by Tan et al. (2018) and differ from BraGO in having the trnL gene transposed from its location between the cox1 and cox2 genes to a location between the second trnL and rrnL genes. This gene arrangement was reported by Tan et al. (2018) as number 12, but we propose Pilummiidae gene order (PilGO) to follow the original gene nomenclature determined by Basso et al. (2017). Similarly, the gene arrangement reported as number 5 by Tan et al. (2018) we rename to the *Somanathelphusa* gene order (SomGO). Basso et al. (2017) report the gene arrangement GeoGO as having the trnL gene between the cox1 and cox2 genes, but based on the gene arrangement listed in Genbank, this is nonconcurrent. The correct gene arrangement was reported by Tan et al. (2019) and is supported here with the addition of the mitochondrial genome for *Geothelphusa* sp. (MG674171), where the trnL gene is located between nad1 and the second trnL gene. This corrected nomenclature should be incorporated into further taxonomic assessments.

### 4.3. Conclusions

This study provides an updated mitophylogeny for the Brachyura, utilizing all available mitogenomes, along with the first mitogenomes for the Panopeidae, a highly abundant group of ecologically important estuarine crabs with a limited phylogenetic understanding. Our data support the subsection, Potamoida, within the Eubrachyura. The addition of *E. depressus, P. herbstdii* and *R. harrisii* mitogenomes provides a greater phylogenetic understanding of a group that has been taxonomically challenging in the past. Moreover, the addition of mitogenomes from the Panopeidae further supports the split of the Xanthoidea into multiple families. The novel gene arrangement we describe within the Heterotremata, increases the total number of unique gene arrangements within this subsection to eight. Whilst our results clarify some phylogenetic relationships, they also highlight the need for further study of the genus *Leptodius* which appears to be incorrectly placed within the subfamily Xanthoidea. Greater sequencing efforts will provide more comparative data for these underrepresented crab groups, and should include the incorporation of nuclear genetic data where possible.

### 5. Author contributions

AMHB collected the crabs used in the study. JB performed the extraction and bioinformatic processing/assemblage of the mitogenomes. LAJ and JB performed the phylogenetics and gene similarity assessments. Gene order analysis and annotation was performed by LAJ and JB. LAJ, AMHB, KAM, DCB and JB contributed to the writing of the manuscript.

### 6. Competing interests

The authors declare no competing interests.

### 7. Acknowledgements

Thanks to Mr. Christopher Moore, who helped to collect the crabs used in the study. Funding for the research and staff time was attained from East Carolina University (AMHB and KAM), the University of Florida (LAJ and DCB) and National Horizon Centre, Teesside University (JB).

### 8. References

Arafat H, Alambrar A, Giss C, Huchon D (2018) Extensive mitochondrial gene rearrangements in Ctenophora: insights from benthic Platycetenida. BMC Ecology and Evolution 18(65): 1–11. https://doi.org/10.1186/s12862-018-1186-1

Babucci M, Basso A, Scupola A, Patarinello T, Negrisolo E (2014) Is it an ant or a butterfly? Convergent evolution in the mitochondrial gene order of Hymenoptera and Lepidoptera. Genome Biology and Evolution 6(12): 3326–3343.

Bai J, Xu S, Nie Z, Wang Y, Zhu C, Wang Y, Min W, Cai Y, Zou J, Zhou X (2018) The complete mitochondrial genome of *Huananpotamon lichuanense* (Decapoda: Brachyura) with phylogenetic implications for freshwater crabs. Gene 646: 217–226.
Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prijibelski AD, Pyshkin AV (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19(5): 455–477.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prijibelski AD, Pyshkin AV (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19(5): 455–477.

Basso A, Babucci M, Pauleto M, Riginella E, Patarnello T, Negrisolo E (2017) The highly rearranged mitochondrial genomes of the crabs Maja crispata and Maja squinado (Majidae) and gene order evolution in Brachyura. Scientific Reports 7: 1–13. https://doi.org/10.1038/s41598-017-04168-9

Bernt M, Merkle D, Ramsch M, Fritzsch G, Perseke M, Bernhard D, Schlegel M, Stadler PF, Middendorf M (2007) CREx: inferring genome rearrangements based on common intervals. Bioinformatics 23(21): 2957–2965.

Boller AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics 30(15): 2114–2120.

Boore JL, Brown WM (1998) Big trees from little genomes: mitochondrial DNA sequences of the decapod crustaceans Pseudocarcinus gigas (Menippidae) and Macrobrachium rosenbergii (Palaemonidae). Marine Biotechnology 7(4): 339–349.

Boore JL, Lavin DV, Brown WM (1998) Gene translocation links insects and crustaceans. Nature 392(6677): 667–668.

Boore JL (1999) Animal mitochondrial genomes. Nucleic Acids Research 27(8): 1767–1780.

Boore JL (2000) The duplication/random loss model for gene rearrangements exemplified by mitochondrial genomes of deuterostome animals. In: Sankoff D, Nadeau JH (eds) Comparative Genomics. Springer, Dordrecht, 133–147.

Bernt M, Donath A, Jüthing F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF (2013) MITOS: improved de novo metazoan mitochondrial genome annotation. Molecular Phylogenetics and Evolution 69(2): 313–319.

Dowton M, Austin AD (1999) Evolutionary dynamics of a mitochondrial rearrangement” hot spot” in the Hymenoptera. Molecular Biology and Evolution 16(2): 298–309.

Gan H, Tan MH, Lee YP, Schultz MB, Horwitz P, Burnham Q, Austin CM (2018) More evolution underground: accelerated mitochondrial substitution rate in Australian burrowing freshwater crayfishes (Decapoda: Parastacidae). Molecular Phylogenetics and Evolution 118: 88–98.

Guerra D, Bouvet K, Breton S (2018) Mitochondrial gene order evolution in Mollusca: inference of the ancestral state from the mtDNA of Chaetopleura apiculata (Polyplacophora, Chaetopleuridae). Molecular Phylogenetics and Evolution 120: 233–239.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic Biology 59(3): 307–321.

Guinot D, Tavares M, Castro P (2013) Significance of the sexual open-ranked podotreme taxa. Zootaxa 3665(1): 1–414.

Jia XN, Xu SX, Bai J, Wang YF, Nie ZH, Zhu CC, Wang Y, Cai YX, Zou JX, Zhou XM (2018) The complete mitochondrial genome of Somanniathelphusa boyangensis and phylogenetic analysis of genus Somanniathelphusa (Crustacea: Decapoda: Parathelphusiidae). PLoS ONE 13(2): e0192601. https://doi.org/10.1371/journal.pone.0192601

Karagözlu MZ, Dinh TD, Nguyen VQ, Kim CB (2018) Analysis of complete mitochondrial genome of Eitusana anguifrons (Arthropoda, Decapoda, Xanthidae) with phylogenetic consideration. Mitochondrial DNA Part B 3(1): 278–279.

Karasawa H, Schweitzer, CE (2006) A new classification of the Xanthidea sensu lato (Crustacea: Decapoda: Brachyura) based on phylogenetic analysis and traditional systematics and evaluation of all fossil Xanthidea sensu lato. Contributions to Zoology 75(01–02): 23–73.

Lai JC, Mendoza JCE, Guinot D, Clark PF, Ng PK (2011) Xanthidae MacLeay, 1838 (Decapoda: Brachyura: Xanthidea) systematics: a multi-genre approach with support from adult and zoeal morphology. Zoologischer Anzeiger – A Journal of Comparative Zoology 250(4): 407–448.

Klaus S, Magallhães C, Salas-Gismondi R, Gross M, Antoine PO (2017) Palaeogene and Neogene brachyuran crabs of the Amazon basin: a re-visited first appearance date for primary freshwater crabs (Brachyura, Trichodactyliidae). Crustacea 90(7–10): 953–967.

Ma KY, Qin J, Lin CW, Chan TY, Ng PK, Chu KH, Tsang LM (2019) Phylogenomic analyses of brachyuran crabs support early divergence of primary freshwater crabs. Molecular Phylogenetics and Evolution 135: 62–66.

Mindell DP, Sorenson MD, Dimche DE (1998) Multiple independent origins of mitochondrial gene order in birds. Proceedings of the National Academy of Sciences 95(18): 10693–10697.

Miller AD, Murphy NP, Burridge CP, Austin CM (2005) Complete mitochondrial DNA sequences of the decapod crustaceans Pseudocarcinus gigas (Menippidae) and Macrobrachium rosenbergii (Palaemonidae). Marine Biotechnology 7(4): 339–349.

Moret BM, Wang LS, Warnow T, Wyman SK (2001) New approaches for reconstructing phylogenies from gene order data. Bioinformatics 17(suppl_1): S165–S173.

Nakajima Y, Shinzato C, Khaliturna M, Nakamura M, Watanabe H, Satoh N, Mitarai S (2016) The mitochondrial genome sequence of a deep-sea, hydrothermal vent limpet, Lepetodrilus nasus, presents a novel vetigastropod gene arrangement. Marine Genomics 28: 121–126.

Perseke MM, Fritzsch G, Ramsch K, Bernt M, Merkle D, Middendorf M, Bernhard D, Stadler PF, Schlegel M (2008) Evolution of mitochondrial gene orders in echinoderms. Molecular Phylogenetics and Evolution 47(2): 855–864.

Roche DG, Torchin ME (2007) Established population of the North American Harris mud crab, Rhithropanopeus harrisi (Gould 1841) (Crustacea: Brachyura: Xanthidae) in the Panama Canal. Aquatic Invasions 2: 155–161.

Schubart CD, Neigel JE, Felder DL (2000) Molecular phylogeny of mud crabs (Brachyura: Panopoidae) from the northwestern Atlantic and the role of morphological stasis and convergence. Marine Biology 137(1): 11–18.

Shen X, Wang H, Wang M, Liu B (2010) The complete mitochondrial genome sequence of Euphausia pacifica (Malaconotra: Euphausiacea) reveals a novel gene order and unusual tandem repeats. Genome 54(11): 911–922.

Shih HT, Ng PK (2011) Diversity and biogeography of freshwater crabs (Brachyura: Crangonidae) from the northwestern Atlantic and the role of morphological stasis and convergence. Marine Biology 158(1): 1–16.

Sternberg RV (1997) Cladistics of the freshwater crab family Trichodactylidae (Crustacea: Decapoda: Brachyura, Xanthidea) based on phylogenetic analysis and traditional systematics and evaluation of all fossil Xanthidea sensu lato. Journal of Comparative Biology 2(1): 49–62.
Sung JM, Lee J, Kim SG, Karagozlu MZ, Kim CB (2016) Complete mitochondrial genome of Leptodius sanguineus (Decapoda, Xanthidae). Mitochondrial DNA Part B 1: 500–501.

Tan MH, Gan HM, Schultz MB, Austin CM (2015) MitoPhAST, a new automated mitogenomic phylogeny tool in the post-genomic era with a case study of 89 decapod mitogenomes including eight new freshwater crayfish mitogenomes. Molecular Phylogenetics and Evolution 85: 180–188.

Tan MH, Gan HM, Lee YP, Linton S, Grandjean F, Bartholomei-Santos ML, Miller AD, Austin CM (2018) ORDER within the chaos: insights into phylogenetic relationships within the Anomura (Crustacea: Decapoda) from mitochondrial sequences and gene order rearrangements. Molecular Phylogenetics and Evolution 127: 320–331.

Tan MH, Gan HM, Lee YP, Bracken-Grisson H, Chan TY, Miller AD, Austin CM (2019) Comparative mitogenomics of the Decapoda reveals evolutionary heterogeneity in architecture and composition. Scientific Reports 9: 1–16. https://doi.org/10.1038/s41598-019-47145-0

Tan BP, Xin ZZ, Liu Y, Zhang DZ, Wang ZF, Zhang HB, Chai XY, Zhou CL, Liu QN (2017) The complete mitochondrial genome of Sesarmops sinensis reveals gene rearrangements and phylogenetic relationships in Brachyura. PLoS ONE 12(6): 1–16. https://doi.org/10.1371/journal.pone.0179800

Tan BP, Liu Y, Xin ZZ, Zhang DZ, Wang ZF, Zhu XY, Wang Y, Zhang HB, Zhou CL, Chai XY, Liu QN (2018) Characterisation of the complete mitochondrial genome of Helice wuana (Grapsidae: Varunidae) and comparison with other brachyuran crabs. Genomics 110(4): 221–230.

Thoma BP, Schubart CD, Felder DL (2009) Molecular phylogeny of western Atlantic representatives of the genus Hexaspanopeus (Decapoda: Brachyura: Panopeidae). In: Decapod Crustacean Phylogenetics. Crustacean Issues, CRC Press, Boca Raton, 551–565.

Thoma BP, Guinot D, Felder DL (2014) Evolutionary relationships among American mud crabs (Crustacea: Decapoda: Brachyura: Xanthoidea) inferred from nuclear and mitochondrial markers, with comments on adult morphology. Zoological Journal of the Linnean Society 170(1): 86–109.

Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Research 44(W1): 232–235.

Tsang LM, Schubart CD, Ahyong ST, Lai JC, Au EY, Chan TY, Ng PK, Chu KH (2014) Evolutionary history of true crabs (Crustacea: Decapoda: Brachyura) and the origin of freshwater crabs. Molecular Biology and Evolution 31(5): 1173–1187.

Wang Q, Tan D, Guo H, Wang J, Xu X, Wang Z (2020) Comparative mitochondrial genomic analysis of Macrobrachium pacificus and insights into the phylogeny of the Ocypodoidea & Grapsoidea. Genomics 112(1): 82–91.

Wang Z, Wang Z, Shi X, Wu Q, Tao Y, Guo H, Ji C, Bai Y (2018) Complete mitochondrial genome of Parasecarma affine (Brachyura: Sesarmidae): Gene rearrangements in Sesarmidae and phylogenetic analysis of the Brachyura. International Journal of Biological Macromolecules 118: 31–40.

Wang Z, Shi X, Guo H, Tan D, Bai Y, Wang Z (2020) Characterization of the complete mitochondrial genome of Uca lactea and comparison with other brachyuran crabs. Genomics 112(1): 10–19.

Williams AB (1984) Shrimps, lobsters, and crabs of the Atlantic coast of the eastern United States, Maine to Florida. Washington, DC: Smithsonian Institution Press 550.

Xie Z, Tan H, Lin F, Guan M, Waiho K, Fang S, Ikhwanuddin M, Fazhan H, Ma H (2018) Characterization of the complete mitochondrial genome of xanthid crab, Atergatis integrerimus from China (Decapoda: Brachyura) and its phylogenetic analysis. Mitochondrial DNA Part B 3(1): 397–398.

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

Yong-Kun J, Wang A, Lu XL, Song DH, Jin YH, Lu J, Sun HY (2014) Mitochondrial genomes of two Brachyuran crabs (Crustacea: Decapoda) and phylogenetic analysis. Journal of Crustacean Biology 34(4): 494–503.

Zhang Z, Xing Y, Cheng J, Pan D, Lv L, Cumberlidge N, Sun H (2020) Phylogenetic implications of mitogenome rearrangements in East Asian potamiscine freshwater crabs (Brachyura: Potamidae). Molecular Phylogenetics and Evolution 143: 106669.

Zhuang X, Cheng CHC (2010) ND6 Gene “Lost” and Found: Evolution of the complete mitochondrial genome of Sesarmidae and comparison with other brachyuran crabs. Genomics 112(1): 82–91.

Zhuang X, Cheng CHC (2010) ND6 Gene “Lost” and Found: Evolution of mitochondrial gene rearrangement in Antarctic notothenioids. Molecular Biology and Evolution 27: 1391–1403.

**Supplementary material 1**

**Table S1**

**Authors:** Jennings et al. (2021)

**Data type:** .docx

**Explanation note:** NCBI accession numbers for species used to conduct phylogenetic analysis.

**Copyright notice:** This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** https://doi.org/10.3897/asp.79.e70234.suppl1