Describing novel mitochondrial genomes of Antarctic amphipods

Louraine Salabao\textsuperscript{a,b}, Tim Plevoets\textsuperscript{c}, Bruno Frédéric\textsuperscript{a}, Gilles Lepoint\textsuperscript{d}, Marc Kochzius\textsuperscript{e} and Isa Schön\textsuperscript{b,f}

\textsuperscript{a}Laboratory of Functional and Evolutionary Morphology, FOCUS, University of Liège, Liège, Belgium; \textsuperscript{b}Centre for Environmental Sciences, Zoology: Toxicology and Biodiversity, Diepenbeek, Belgium; \textsuperscript{c}Unit Animal Sciences - ILVO Marine Research, Flanders Research Institute for Agriculture, Fisheries and Food, Oostende, Belgium; \textsuperscript{d}Laboratory of Trophic and Isotopes Ecology, FOCUS, University of Liège, Liège, Belgium; \textsuperscript{e}Marine Biology, Vrije Universiteit Brussel (VUB), Brussels, Belgium; \textsuperscript{f}OD Nature, Freshwater Biology, Royal Belgian Institute of Natural Sciences, Brussels, Belgium

ABSTRACT

To date, only one mitogenome from an Antarctic amphipod has been published. Here, novel complete mitochondrial genomes (mitogenomes) of two morphospecies are assembled, namely, Charcotia amundseni and Eusirus giganteus. For the latter species, we have assembled two mitogenomes from different genetic clades of this species. The lengths of Eusirus and Charcotia mitogenomes range from 15,534 to 15,619 base pairs and their mitogenomes are composed of 13 protein coding genes, 22 transfer RNAs, 2 ribosomal RNAs, and 1 putative control region CR. Some tRNAs display aberrant structures suggesting that minimalization is also ongoing in amphipod mitogenomes. The novel mitogenomes of the two Antarctic species have features distinguishing them from other amphipod mitogenomes such as a lower AT-richness in the whole mitogenomes and a negative GC-skew in both strands of protein coding genes. The genetically most variable mitochondrial regions of amphipods are nad6 and atp8, while cox1 shows low nucleotide diversity among closely and more distantly related species. In comparison to the pancrustacean mitochondrial ground pattern, E. giganteus shows a translocation of the nad1 gene, while cytb and nad6 genes are translocated in C. amundseni. Phylogenetic analysis based on mitogenomes illustrates that Eusirus and Charcotia cluster together with other species belonging to the same amphipod superfamilies. In the absence of reference nuclear genomes, mitogenomes can be useful to develop markers for studying population genetics or evolutionary relationships at higher taxonomic levels.

KEYWORDS

Amphipoda; Eusirus giganteus; Charcotia amundseni; gene rearrangements; mitochondrial genome; nucleotide diversity

Introduction

Mitogenome DNA sequence data or parts of mitogenomes have been widely used to reconstruct evolutionary relationships or detect cryptic diversity (Caterino et al. 2000; Tang et al. 2020). For instance, in amphipods, sequencing mitochondrial cox1 or cytb together with nuclear genes (e.g. 18S, 28S, ITS2) has revealed cryptic species of Hyallela S.I. Smith, 1874 (Witt et al. 2006), Caprella penantis Leach, 1814 (Pilar Cabezas et al. 2013), Gammarus fossarum Koch, 1836 (Grabowski et al. 2017) and some Eusirus Krøyer, 1845 species (Baird et al. 2011). Molecular data from 13 protein coding genes of Alicella gigantea Chevreux, 1899 (Li et al. 2019b), Baikalian amphipods (Romanova et al. 2016), Gammarus roeselii Gervais, 1835 (Cormier et al. 2018), Halice sp. Boeck, 1871 (Li et al. 2019a), Metacrangonyctidae Boutin & Messouli, 1988 (Bauzá-Ribot et al. 2012), and from all mitochondrial genes (protein coding genes, rRNA, tRNA) of Gammarus pisin- rus Hou, Li & Li, 2014 and Gammarus lacustris G.O. Sars, 1863 (Sun et al. 2020) have been used to reconstruct evolutionary relationships. The broad application of molecular data from mitogenomes can be explained by several advantages, which the mitogenome has compared to the nuclear genome. These include its simpler structure, conserved gene content and limited size (Boore 1999; Li et al. 2019b; Krebes and Bastrop 2012) facilitating sequencing of mitogenomes from those species for which reference nuclear genomes are not yet available. The uniparental, usually maternal inheritance of mitogenomes furthermore simplifies analyses because recombination is either totally absent or very rare (Barr et al. 2005, Lin and Danforth 2004). The relatively high evolutionary rate of mitogenomes generating relative large genetic differences makes mitogenomic DNA sequence data furthermore suitable for studies at the genus or species level investigating population genetic or phylogeographic patterns (Ballard and Whitlock 2004; Krebes and Bastrop 2012, Tang et al. 2020; Li et al. 2019a). The inclusion of whole mitogenomes has resulted in phylogenies with better statistical supports (Haran et al. 2013; López-López and Vogler 2017) and clearer phylogeographic patterns (Keis et al. 2013). Moreover, despite the highly conserved gene content of the mitogenome, gene
order has been found to be variable and can provide additional data for reconstructing phylogenetic relationships and evolutionary histories (Cormier et al. 2018; Krebes and Bastrop 2012; Zhang et al. 2020).

Amphipods are widely distributed crustaceans inhabiting a range of different habitats (Vainolä et al. 2008; Li et al. 2019b). In Antarctica, amphipods are among the most diverse components of the benthic community (Gallardo 1987) and show high levels of endemism (Knox and Lowry 1977) making them ideal model organisms to study evolutionary patterns and divergences based on mitogenomes. Currently, there is only one published complete mitogenome of an Antarctic amphipod, namely of Gondogeneia antarctica Chevreux, 1905 (Shin et al. 2012), and no mitogenomes are yet available for abundant amphipods of the genera Eusirus Krøyer, 1845 and Charcotia Chevreux, 1905.

In this paper, we have assembled and analyzed complete mitogenomes of three Antarctic amphipods from two morphospecies (Charcotia amundseni D’Udekem d’Acoz, Schön & Robert, 2018 and Eusirus giganteus Andres et al., 2002) and two genetic clades of the latter species. Our aims are to (1) provide full mitogenomic data of selected amphipod species for future research and (2) compare gene content and order with published amphipod mitogenomes to unravel shared and unique patterns of mitogenome evolution in amphipods.

**Materials and methods**

**Sample collection**

Specimens of two species of Antarctic amphipods, Charcotia amundseni d’Udekem d’Acoz, Schön & Robert, 2018 and two genetic clades of Eusirus giganteus Andres et al., 2002 (G1 and G2; which might resemble different genetic species (Verheyen and D’Udekem D’Acoz 2021) have been collected during different Antarctic expeditions (Table 1) and are curated in the collections of the Royal Belgian Institute of Natural Sciences, Brussels, Belgium.

Eusirus amphipods belong to the superfamily Eusiroidea Stebbing, 1888. Eusirus cf. giganteus has previously been confused with Eusirus perdentatus Chevreux, 1912 due to small morphological differences (Andres et al. 2002). The genetic study of Baird et al. (2011) reveals cryptic diversity of Eusirus giganteus including the so-called clades G1-G4, and the existence of a species complex is supported by Verheyen and D’Udekem D’Acoz (2021). The same authors report that potential Eusirus giganteus species that still need to be formally described showed at least minor morphological differences and different color morphs but that a thorough morphological analysis of the putative genetic species is still required. Given the possibility of multiple cryptic species, we follow here the suggestion of Greco et al. (2021) to use the name Eusirus cf. giganteus in our study. Our other target species, Charcotia amundseni, belongs to the superfamily Lysianassoida Dana, 1849. The genus Charcotia has formerly been known as Waldeckia (Chevreux 1905) but recently has undergone a change in nomenclature (D’Udekem D’Acoz et al. 2018) which we follow here.

**Mitochondrial genome sequencing, assembly, annotation, and analyses**

DNA has been extracted from a pleopod of each specimen using the DNeasy Blood & Tissue Kit (Qiagen, Germany) for both Eusirus cf. giganteus clades and the Qiamp DNA Minikit (Qiagen, Germany) for Charcotia amundseni following the manufacturer’s protocol. DNA concentration and quality have been checked with a Nanodrop ND-1000 Spectrophotometer (ThermoFisher Scientific, USA) and a Qubit 2.0 fluorometer (Life Technologies, USA).

A low coverage skimming sequencing approach has been applied at the Genomics Core at the KU Leuven (Leuven, Belgium) using an Illumina HiSeq2500 sequencing platform in the 2 x 150 bp mode. Samples were indexed separately as unique libraries. Reads have been quality-checked using FASTQC (Andrews 2010) and pre-processed with Geneious primer trimming and quality trimming to a minimum quality set to 20. These pre-processed reads have then been used for de novo assemblies in MITObim v1.9.1 (Hahn et al. 2013) with the MIRA 4.0.2 (Chevreux et al. 1999) assembler with default settings (kmer size = 31) and an iteration limit of 100. The Onisimus nanseni G.O. Sars, 1900 mitogenome (GenBank accession number FJ555185.1) which belongs to the same superfamily as Charcotia and a partial 16S to COI sequence of Eusirus perdentatus have been used as seed references. The longest resulting contigs from the de novo assembly have been imported into Geneious and further assembled with the ‘map to reference’ approach with medium–low sensitivity and 50 iterations. Identity of the resulting consensus sequences have been verified with BLAST searches (Altschul et al. 1997). Automatic annotation has subsequently been conducted with the MITOS web server, versions 1 and 2 (Bernt et al. 2013). The identity of the rrnL region of both Eusirus species has been confirmed by BLAST searches only, since it has not been annotated by MITOS. The resulting annotations have been viewed and gene boundaries manually corrected in Geneious. The boundaries of the 13 protein coding genes and 2 rRNA genes have been

---

**Table 1.** Sampling details of specimens analyzed in this study, including date of sample, expedition, locality, geographical coordinates, voucher ID provided by Royal Belgian Institute of Natural Sciences (RBINS), and gear used during sampling.

| Species                      | Date collected | Expedition   | Locality, coordinates | Voucher ID       | Gear         |
|------------------------------|----------------|--------------|------------------------|-----------------|--------------|
| Eusirus cf. giganteus (G1)   | 23 February 2013| PS81, ANT-XXIX-3 | Bransfield Strait, 62°43.73'S 57°29.04'W | INV. 12279 spec. C | Agassiz trawl |
| Eusirus cf. giganteus (G2)   | 15 January 2008 | CEAMARC       | Adélie Coast, 66°10'14.3'S 139°21'11.3'E | MNHN-IU-2019-3365 | Beam trawl  |
| Charcotia amundseni          | 23 December 2008 | BELARE 08-09  | Crown Bay, 70° S 23°E  | INV.180000      | Baited trap  |

Eusirus cf. giganteus (G1) and E. cf. giganteus (G2) are taxonomically undescribed putative species that belong to E. giganteus complexes as verified genetically by Baird et al. (2011) and Verheyen and D’Udekem D’Acoz (2021).
identified by comparing alignments of the novel assemblies with mitochondrial genes of other amphipod species. Protein coding genes boundaries have been further corrected by avoiding any overlap with the subsequent tRNA gene and by noticing any partial stop codons (T or TA). Such partial stop codons are atypical features of mitochondrial protein coding genes (Cameron 2014). Transfer RNA (tRNA) genes and their secondary structures have been predicted with tRNAscan-SE (Lowe and Eddy 1997) and ARWEN 1.2.3 (Laslett and Canbäck 2008). Potential control regions (CRs) have been identified from their typical features such as high AT content, poly-T stretches, and hairpin structures (Zhang and Hewitt 1997).

Gene orders of the novel mitogenome assemblies were compared to the putative pancrustacean ground pattern which is derived from both Crustacea and Hexapoda (often referred to as pancrustacea) as they share the same ground pattern in terms of their mitochondrial gene order (Kilpert and Podsiadlowski 2006; Boore et al. 1998). Possible gene rearrangements have been analyzed with the CREx web service (Bernt et al. 2007). CREx utilizes a strong common interval tree to heuristically deduce the plausible rearrangement scenarios to change one gene order to another (Bernt et al. 2007). AT and GC skew have been calculated using the formulas of Perna and Kocher (1995): AT skew = (A − T)/(A + T) and GC skew = (G − C)/(G + C). Only other amphipod species with complete and published mitogenomes have been analyzed for their AT and GC skew (Supplementary Table 1).

The total length of the obtained complete mitochondrial genomes of Eusirus cf. giganteus (G1), Eusirus cf. giganteus (G2), and Charcotia amundseni is 15,558, 15,534, and 15,619 bp, respectively (Genbank accession nos. OK489458, OK489459, OK489457, respectively) which is within the range of complete mitogenomes from other amphipods (13,517–18,424 bp) (Table 2). The three newly assembled mitogenomes are each composed of 13 protein coding genes, 22 tRNAs and 2 rRNAs. For E. cf. giganteus, 23 genes are encoded on the positive (+) strand and 14 on the negative (−) strand while 17 genes are encoded on the + strand and 20 on the − strand in C. amundseni (Supplementary Figure 1a and b, Supplementary Table 2). A putative control region (CR) has also been identified in all three mitogenomes and is located between trnS2 and rrnL in Eusirus and between trnF and nad5 in C. amundseni. The mitogenome also contains 20 intergenic regions for E. cf. giganteus and 18 intergenic regions for C. amundseni. The whole mitogenomes of the two species show AT-richness of 61.9% for E. cf. giganteus and 68.7% for C. amundseni, respectively, which contributes to the positive AT skew (0.008 to 0.092) and negative GC skew (−0.317 to −0.201) values observed in the three mitogenomes (Table 2). A relatively high AT content is also observed in the complete mitogenomes of other amphipod species varying from 61.09 to 77% (Table 2).

**Results**

**Mitogenome organization**

The most frequent start codon in E. cf. giganteus and C. amundseni is ATG (Supplementary Table 2). Defining the protein coding gene boundaries following a tRNA results in a few partial or incomplete stop codons (T or TA). The AT content of the protein coding genes of the three amphipod mitogenomes is estimated as 59.6% for E. cf. giganteus (G1 and G2) and 67.3% for C. amundseni (Table 2). Mitochondrial genomes of the two species in this study have negative GC skew values in the protein coding genes encoded on both strands (Supplementary Table 4). The highest AT content is found in the third codon position of C. amundseni and the second codon position of E. cf. giganteus while the lowest AT

**Protein coding genes**

A translocation of nad1 gene in E. cf. giganteus is observed while cyt b and nad6 are translocated in C. amundseni as compared to the pancrustacean ground pattern (Supplementary Figure 2). We furthermore also find shifts in the position of tRNAs and the control region in E. cf. giganteus and C. amundseni as compared to the pancrustacean ground pattern (Supplementary Figure 2). While also trnG has been translocated in the three species investigated here, we find other tRNA gene strings consisting of trnA, trnS1, trnR, trnN, and trnE for E. cf. giganteus and trnS1, trnN, trnE, and trnF for C. amundseni (Supplementary Figure 2). Similar with the pancrustacean ground pattern, the trnV is located between the rrnL and rrnA genes in E. cf. giganteus while trnC and trnV are inserted between these genes in C. amundseni (Supplementary Figure 2).

Results of the CREx analyses indicate that E. cf. giganteus and C. amundseni have undergone multiple transpositions and rearrangements relative to the pancrustacean ground pattern (Supplementary Figure 3b).
In the three mitogenomes of this study, 22 tRNAs are present in the length ranging from 52 to 67 base pairs (Supplementary Table 2). The AT content of tRNAs of E. giganteus is 66.2% and 69.9% for C. amundseni (Table 2). In E. cf. giganteus, 14 tRNAs are encoded in the + strand and 8 in the − strand. In C. amundseni, 10 tRNAs are encoded in the + and 12 in the − strand. Typical clover leaf secondary structures are observed in most predicted tRNAs although some tRNAs show wobble base pairs, atypical pairing or the DHU and TΨC arm are missing (Supplementary Figure 4a–c). More specifically, the DHU arm is missing in tRN1, tRN2 and tRNV of E. cf. giganteus (Supplementary Figure 4a and b) and tRN1, tRN2, and tRNl of C. amundseni (Supplementary Figure 4c). We also find that the TΨC loop is absent in tRNK, tRNb, tRN and tRNb of E. cf. giganteus (Supplementary Figure 4a and b) and in tRNb, tRNA, tRNl1, tRNc, tRNA, tRNl, and tRN of C. amundseni (Supplementary Figure 4c).

### Nucleotide diversity

When estimating nucleotide diversity (π) between the two Eusirus genetic clades, we observe high values for nad6
(0.013), nad5 (0.012), and nad1 (0.011) (Supplementary Figure 5a) and low ones for nad4 (0.002), nad3 (0.003), nad2 (0.004), and cox1 (0.005) with nad4 (0.002) being the least variable. We also find high variability in nad6 (0.569), atp8 (0.566), and nad2 (0.515) between C. amundseni and E. cf. giganteus (G1) and low variability in cox1 (0.279), cytB (0.328), and cox3 (0.342) (Supplementary Figure 5b). Moreover, between C. amundseni and E. cf. giganteus (G2), high variability is observed in nad6 (0.567), atp8 (0.560), and nad2 (0.515), while the lowest variability is found in cox1 (0.28), cytB (0.327), and cox3 (0.343) (Supplementary Figure 5b).

**Phylogenetic analysis**

Phylogenetic analysis revealed that the phylogenetic grouping follows the superfamly identity and fall under different family groups (Figure 1). The two genetic clades of Eusirus cf. giganteus G1 and G2 under family Eusiridae, cluster together. They belong to the superfamly Eusiroidae and are found to be closely related to both Epimeria frankei Beermann & Raupach, 2018 in Beerman, Westbury, Hofreiter, Higlers, Deister, Neumann & Raupach, 2018 and Epimeria cornigera Fabricius, 1779 (family Epimeriidae Boeck, 1871) from the Iphimedioidea Boeck, 1871 superfamily. Similarly, Charcotia amundseni from family Lysianassidae Dana, 1849 is clustering together with Eurythenes megallanicus H. Milne Edwards, 1848 and Eurythenes maldoror d’Udekem d’Acoz & Havermans, 2015 (family Eurytheneidae Stoddart & Lowry, 2004), Hirondella gigas Birstein & Vinogradov, 1955 (family Hirondelleidae Lowry & Stoddart, 2010), and Onisimus nansei G.O. Sars, 1900 (family Uristidae Hurley, 1963) and all belonging to the superfamily Lysiannoidea.

**Discussion**

In the current study, we have assembled and annotated novel complete mitogenomes from two Antarctic amphipod species with a low coverage skimming sequencing approach. We have obtained very low percentages of ambiguities (<0.01%) illustrating that this cost efficient approach is very successful. Besides our study, only two complete mitogenomes from amphipods of the polar regions are currently available, namely, from Gondogeneia antarctica Chevreux, 1905 from Antarctica (Shin et al. 2012) and Onisimus nansei G.O. Sars, 1900 from the Arctic (Ki et al. 2010). Our study thus provides important novel genomic data for further research and the first complete mitogenomes of the widely spread amphipod genus Eusirus and Charcotia. Our comparisons of mitogenomes between two genetic clades, possibly resembling two different genetic species of E. cf. giganteus illustrate that mitogenomic features such as length, gene order, AT content, and tRNA structure are similar at the intra-specific level (Table 2, Supplementary Figure 1, Supplementary Figure 2; Supplementary Table 3, Supplementary Figure 4a and b).

All three newly obtained mitogenomes are with 15,534 and 15,619bp at the middle range of reported lengths of published amphipod mitogenomes (14,113bp to 18,424bp) (Romanova et al. 2016; Li et al. 2019b). The observed AT-richness of the mitogenomes of the current study (61.9% and 68.7%) is slightly lower than in other studies based on complete (Table 2) and incomplete amphipod mitogenomes where AT range between 69.79% and 74.35% (Li et al. 2019a). However our data are in line with Wilson et al. (2000) reporting such an AT-rich bias as typical for arthropods.

The negative GC skews on both strands of the protein coding genes of the two species in this study differ from the so far known common Malacostraca pattern where genes encoded on the − strand usually exhibit negative and genes encoded on the + strand positive GC skews (Hassanin 2006). The strand bias in nucleotide composition of metazoan mitogenomes is attributed to varying mutational pressure during replication or transcription (Pons et al. 2014; Hassanin et al. 2005). Future research will need to test if these factors are responsible for the different GC patterns observed in the two Antarctic amphipod species of the current study.

**Gene order and rearrangements**

The translocations of trnG and a commonly derived pattern of a gene string consisting of trnA, trnS1, trnN, trnE, and trnR are presumed to be apomorphic features of certain amphipods (Kilpert and Podsiaidloewski 2010; Krebes and Bastrop 2012; Li et al. 2019a). The two studied species exhibit the translocation of trnG relative to the pancrustacean ground pattern. However, the altered tRNA gene order of the two species results in a unique tRNA string that is dissimilar to the apomorphic gene string of trnA, trnS1, trnN, trnE, and trnR. Moreover, the observed rml-trnV-rmS pattern of E. cf. giganteus is known to be common in most Malacostraca (Ki et al. 2010) and is also observed in the pancrustacean ground pattern. This is, however, not the case for C. amundseni with the trnC being present. In addition, the large-scale gene reversals that have been found in three species have also been observed in Halice sp. Boeck, 1871 (Li et al. 2019a). It may be attributed to intramitochondrial recombination allowing breaking and rejoining of the mitochondrial genome (Dowton and Austin 1999; Li et al. 2019a).

**rRNA genes**

The shortest complete rml in amphipods of 577bp is currently known from Hirondellea gigas Birstein & Vinogradov, 1955 (Lan et al. 2016), which is much shorter than the rml that has been found in the species of the current study. On the other hand, the rrmS of C. amundseni has with 529bp the same length as Alicella gigantea (Li et al. 2019b) which has so far been the shortest reported rrmS length in amphipods. Also, the shortest total length of rrml and rrmS together has been described from the amphipod Hirondellea gigas Birstein & Vinogradov, 1955 (Lan et al. 2016) with 1120bp, and we find that the total length of the two rRNAs in C. amundseni is with 1268bp rather similar. Short rRNA genes have also been observed in Gammarus duebeni Lilljeborg, 1852 (Krebes and Bastrop 2012) where they have been attributed to a minimization strategy of the mitogenome.
tRNA secondary structures

Aberrant tRNA structures as we find them in the three novel mitogenomes are common. Jühling et al. (2012) have described a loss of a D-domain in trnS1 in almost all metazoan, while the D-domain in trnS2 has only been lacking in Lophotrochozoa and Ecdysozoa. Mitogenome studies of other amphipod species also report the lack of the DHU arm in trnS1 and trnS2 in Epimeria cornigera Fabricius, 1779, Epimeria frakei Beermann & Raupach, 2018 in Beerman, Westbury, Hofreiter, Hilgers, Deister, Neumann & Raupach, 2018 (Beermann et al. 2020), Caprella scabra Templeton, 1836 (Ito et al. 2010) and ‘Metacrangonyx boveii’ (Pons et al. 2014) and in trnV in Brachyuropus growingkii Dybowski, 1874, Acanthogammarus victorii Dybowski, 1874, Eulimnogammarus cyaneus Dybowski, 1874, and Garjajewia cabanissii Dybowski, 1874 (Romanova et al. 2016), Halice sp. Boeck, 1871 (Li et al. 2019a) and ‘Metacrangonyx boveii’ (Pons et al. 2014). The absence of the TΨC loop is another aberrant and common structure in amphipod that has also been observed in trnC, trnE, and trnT of Caprella mutica Schurin, 1935 (Kilpert and Podsiadlowski 2010), trnQ and trnV of Gammarus duebeni Liljeborg, 1852 (Krebes and Bastrop 2012), and trnC, trnQ, trnK, and trnF of Onisimus nanseni G.O. Sars, 1900 (KI et al. 2010). The pressure for minimization of the mitogenome has been put forward as one of the explanations for these aberrant tRNA structures (Yamazaki et al. 1997). Other explanations could be replication slippage resulting in sequence deletions or insertions (Macey et al. 1997). Despite these aberrant structures in tRNAs, these are most likely still functional (Watanabe et al. 2014).

Nucleotide diversity

Information on nucleotide diversity can be helpful for the design of new molecular markers (Romanova et al. 2016; Zhang et al. 2018). Here, we have shown that the most variable mitogenes of Eusirus for intraspecific comparisons between genetic clades are nad6, nad5, and nad1 while for comparisons between Eusirus and Charcotia, atp8, nad6, and nad2 are most variable, which could be suitable for future phylogeographic and population genetic studies. Contrary, the least variable mitogenes for Eusirus are nad4, nad3, nad2, and cox1 and for interspecies comparisons between Eusirus and Charcotia cox1, ctb, and cox3 could be more suitable for future deep phylogeny investigations. Surprisingly, despite its wide use in DNA barcoding initiatives (Witt et al. 2006; Hebert et al. 2003), the cox1 gene appears to have relatively low nucleotide diversities between closely and distantly related amphipods. Consistent with our results, also Romanova et al. (2016) describe the mitogenes atp8, nad2, nad4l, nad5, and nad6 as most variable in Baikalian amphipods and cox genes to be less functional, with cox1 having the lowest nucleotide diversity.

Phylogenetic analysis

Our evolutionary tree (Figure 1) constructed from mitochondrial protein coding genes is well supported and shows phylogenetic clades according to amphipod superfamily identity. Moreover, our results are congruent to the current taxonomic classification where the species were categorized into their respective family and superfamily (Horton et al. 2021). Previous classification have placed Eusirus in the same Eusiroidae superfamily as Epimeria (Bousfield 1978) while the recent classification have placed Eusirus under superfamily Eusiroidae and Epimerida under Iphimoedae (Lowry and Myers 2017). Phylogenetic evidence using 18S rDNA have shown that Eusirus has a close relationship with Epimeria which showed a well-supported clade of Eusiroidae, Calliopiidae, Astyridae, Iphimoedae, Epimeridae, and Pleustidae families (Englisch 2001). Phylogenetic evidence using 13-protein coding genes further corroborates these close relationships (Figure 1).

Our grouping of Charcotia amundseni with other species from the superfamily Lysianassoidea (Figure 1) is supported with the morphological phylogeny of Lowry and Myers (2017), which characterized this superfamily as often having a type 3 lysianassoid calceolus and a cleft telson. Molecular phylogenetic analyses using concatenated 16S-COX1-18S
data in Ritchie et al. (2015) show clustering of families and superfamilies similar to our study which further backs up our results. The phylogenetic grouping of the two morphospecies invested here based on the three novel mitogenomes thus follow the expected patterns according to taxonomic relationships.

Conclusions

The current study provides three additional novel complete mitogenomes of Antarctic amphipod species and the first complete mitogenomes of the abundant amphipod genera Eusirus and Charcotia. In comparison to other published amphipod mitogenomes, the novel mitogenomes show distinct features such as a lower AT-richness in their whole mitogenomes, negative GC skews on both strands of the protein coding genes, and unique gene rearrangements. The novel mitogenomes also share characteristics with other amphipod mitogenomes including aberrant tRNA and short rRNA genes, which could be linked to minimalization of mitogenomes. Moreover, the estimation of the nucleotide diversity (π) provides information to choose mitogenes as most suitable markers for future phylogenetic studies of amphipods. The novel mitogenomes are certainly useful for future phylogenetic analyses as put the investigated species into phylogenetic positions matching superfamily and family identity.

Acknowledgements

The authors are grateful to the Royal Belgian Institute of Natural Sciences, Belgium, and the National Museum of Natural History, Paris, for providing the samples used in this study. Also, the authors are also greatly indebted to Marie Verheye (University of Liege, Belgium) and Cédric d’Udekem d’Acoz (Royal Belgian Institute of Natural Sciences, Belgium) for providing suggestions for this manuscript. Lastly, the Genomics Core in KU Leuven is acknowledged for conducting the high throughput sequencing.

Ethics statement

This study did not require any ethical approval from the University of Hasselt, Belgium.

Author contributions

Molecular analyses (LS and TP), analysis and interpretation of the data (LS, IS, and TP), and writing of this manuscript (LS, TP, BF, GL, MK, and IS). All authors agree to be accountable for all aspects of the work.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the Bijzonder Onderzoeksfonds (BOF) of the Universities Hasselt and Liège with grant number [BOF18DOC101] and by Belgian Science Policy (Belspo) as part of the project ‘Refugia and Ecosystem Tolerance in the Southern Ocean’ (RECTO) [BR/154/A1/RECTO]. This is RECTO contribution 028.

Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at https://www.ncbi.nlm.nih.gov/ under the accession numbers Eusirus cf. giganteus (G1) (OK489458), Eusirus cf. giganteus (G2) (OK489459), and Charcotia amundseni (OK489457). The associated Bioproject is PRJNA769065 and the Biosample numbers are Eusirus cf. giganteus (G1) (SAM22086850), Eusirus cf. giganteus (G2) (SAM22087742), and Charcotia amundseni (SAM22087745). The SRA accession numbers are Eusirus cf. giganteus (G1) (SRX13936485), Eusirus cf. giganteus (G2) (SRX13936486), and Charcotia amundseni (SRX13936484).

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. 366Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25(17):810–3402.

Andres HG, Lörz AN, Brandt A. 2002. A common but undescribed huge species of Eusirus kroyer, 1845 (Crustacea, Amphipoda, Eusiridae) from Antarctica. Mitteilungen Aus Dem Hamburgischen Zoologischen Museum Und Institut. 99:109–126.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data [Online]. [cited 2019 July 28] http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Baird HP, Miller KJ, Stark JS. 2011. Evidence of hidden biodiversity, ongoing speciation and diverse patterns of genetic structure in giant Antarctic amphipods. Mol Ecol. 20(16):3439–3454.

Ballard JW, Whitlock MC. 2004. The incomplete natural history of mitochondria. Mol Ecol. 13(4):729–744.

Barr CM, Neiman M, Taylor DR. 2005. Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. New Phytol. 168(1):39–50.

Bauzá-Ribot MM, Juan C, Nardi F, Oromí P, Pons J, Jaume D. 2012. Mitogenomic phylogenetic analysis supports continental-scale vicariance in subterranean thalassoid crustaceans. Curr Biol. 22(1):2069–2074.

Beermann J, Westbury MV, Hofreiter M, Hilgers L, Deister F, Neumann H, Raupach MJ. 2018. Cryptic species in a well-known habitat: applying taxonomics to the amphipod genus Epimeria (Crustacea, Peracarida). Sci Rep. 8(1):1–26.

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 69(2):313–319.

Bernt M, Merkle D, Ramsch K, Fritzsch G, Perseke M, Bernhard D, Schlegel M, Stadler PF, Middendorf M. 2007. CReX: inferring genomic rearrangements based on common intervals. Bioinformatics. 23(21):2957–2958.

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

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

Bousfield EL. 1978. A revised classification and phylogeny of amphipod crustaceans. Ottawa: Royal Society of Canada.

Cameron SL. 2014. How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research. Syst Entomol. 39(3):400–411.

Caterino MS, Cho S, Sperling FA. 2000. The current state of insect molecular systematics: a thriving Tower of Babel. Annu Rev Entomol. 45(1):1–54.

Cheveux B, Wetter T, Suahi S. 1999. Genome sequence assembly using trace signals and additional sequence information. Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB 1999), Vol. 99, p. 45–56.

Cheveux E. 1905. Diagnoses d’amphipodes nouveaux provenant de l’expédition antarctique du Français. I. Lysianassidae. Bull Soc Zool Fr. 30:159–165.
Evolution of mitochondrial genomes in Baikalian amphipods. BMC Genomics. 17(Suppl 14):1016.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19(12):1572–1574.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol. 34(12):3299–3302.

Shin SC, Cho J, Lee JK, Ahn DH, Lee H, Park H. 2012. Complete mitochondrial genome of the Antarctic amphipod Gondogeneia antarctica (Crustacea, amphipod). Mitochondrial DNA. 23(1):25–27.

Sun S, Wu Y, Ge X, Jakovlić I, Zhu J, Mahboob S, Al-Ghanim KA, Al-Misned F, Fu H. 2020. Disentangling the interplay of positive and negative selection forces that shaped mitochondrial genomes of Gammarus pisinus and Gammarus lacustris. R Soc Open Sci. 7(1):190669.

Tang Y, Zheng X, Liu H, Sunxie F. 2020. Population genetics and comparative mitogenomic analyses reveal cryptic diversity of Amphioctopus neglectus (Cephalopoda: Octopodidae). Genomics. 112(6):3893–3902.

Väinölä R, Witt JDS, Grabowski M, Bradbury JH, Jazdzewski K, Sket B. 2008. Global diversity of amphipods (Amphipoda; Crustacea) in freshwater. Hydrobiologia 595:241–255.

Verheye ML, D’Udekem D’Acoz C. 2021. Integrative taxonomy of giant crested Eusirus in the Southern Ocean, including the description of a new species (Crustacea: Amphipoda: Eusiridae. Zool J Linn Soc. 193(1):31–77.

Watanabe YI, Suematsu T, Ohtsuki T. 2014. Losing the stem-loop structure from metazoan mitochondrial tRNAs and co-evolution of interacting factors. Front Genet. 5:109.

Wilson K, Cahill V, Ballment E, Benzie J. 2000. The complete sequence of the mitochondrial genome of the crustacean Penaeus monodon: are malacostracan crustaceans more closely related to insects than to branchiopods? Mol Biol Evol. 17(6):863–874.

Witt JD, Threlfall DL, Hebert PD. 2006. DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation. Mol Ecol. 15(10):3073–3082.

Yamazaki N, Ueshima R, Terrett JA, Yokobori S, Kaifu M, Segawa R, Kobayashi T, Numachi K, Ueda T, Nishikawa K. 1997. Evolution of pulmonate gastropod mitochondrial genomes: comparisons of gene organizations of Euhadra, Cepaea and Albinaria and implications of unusual tRNA secondary structures. Genetics. 145(3):749–758.

Zhang D, Li WX, Zou H, Wu SG, Li M, Jakovlić I, Zhang J, Chen R, Wang GT. 2018. Mitochondrial genomes of two diplanctanids (Platyhelminthes: Monogenea) expose paraphyly of the order Dactylogyridae and extensive tRNA gene rearrangements. Parasit Vectors. 11(1):1–13.

Zhang DX, Hewitt GM. 1997. Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. Biochem Syst Ecol. 25(2):99–120.

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). Mol Phylogenet Evol. 143:106669.