This is an Accepted Manuscript of an article published by Elsevier in Zoologischer Anzeiger on 3 June 2021, available online: https://doi.org/10.1016/j.jcz.2021.05.011

Mismatch between similarity of mitochondrial gene order and phylogenetic distance in Podocopa (Crustacea: Ostracoda)

Keiichi Kakui[a,*], Mizuho Munakata[b], Hayato Tanaka[c], Chizue Hiruta[a]

[a]*Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan. E-mail: kakui@eis.hokudai.ac.jp (KK); chizueh@gmail.com (CH)*

[b]*Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan. E-mail: munakata.mizuho.k0@elms.hokudai.ac.jp*

[c]*Tokyo Sea Life Park, Edogawa-ku, Tokyo 134-8587, Japan. E-mail: cladocopina@gmail.com*

*Corresponding author: Keiichi Kakui; kakui@eis.hokudai.ac.jp; Faculty of Science, Hokkaido University, N10 W8 Kita-ku, Sapporo 060-0810, Japan

ABSTRACT

Ostracoda is a diverse group of tiny crustaceans. Although more than 9330 extant species have been described, whole-mitogenomic data were available for only five species. Here we present a complete mitogenomic sequence for an additional species, the parthenogenetic podocopan ostracod *Heterocypris spadix*. The mitogenome is 15,205 bp long and contains the typical animal mitogenomic complement of 13 protein coding, two ribosomal RNA, and 22 transfer RNA genes. In a mitogenome-based phylogenetic tree, Myodocopa and Podocopa were each monophyletic. Gene order in Podocopa was much more similar to the pancrustacean ground pattern than was gene order in Myodocopa. Gene order was invariant in the three myodocopan species examined, all in the family Cypridinidae. Within Podocopa, two species from different families (*Fabaeformiscandona kushiroensis*, Candonidae; *H. spadix*, Cyprididae) were identical in gene order and differed from the pancrustacean ground pattern by only a single recombination. In Cyprididae, however, *Cypridopsis vidua* differed from *H. spadix* by five rearrangements. These conflicting patterns—apparent gene-order conservation within (Cypridinidae) and between families (Cyprididae and Candonidae), but marked divergence within a family (Cyprididae)—suggest that the evolution of gene order in ostracods was more complex than that expected from the cypridinid data.

Keywords: Cypridoidea – high-throughput DNA sequencing – mitochondrion – Podocopida – phylogeny – seed shrimp

1. Introduction

Ostracoda is a group of tiny crustaceans, typically less than a few millimeters in length, with more than 9330 extant species described to date (Tran Van et al., 2021). Although most species inhabit marine environments ranging from coastal interstitial habitats to hadal bottoms (e.g., Danielopol and Hartmann, 1986; Brandão et al., 2019), ostracods also inhabit brackish, freshwater, and terrestrial environments (Pinto et al., 2005; Karanovic, 2012). Ostracoda comprises two main subgroups, Myodocopa and Podocopa (Rodriguez-Lazaro and Ruiz-Muñoz, 2012). The sister relationships of two subgroups, i.e., the monophyly of Ostracoda is still under debate (Oakley et al., 2013; Schön and Martens, 2016; Schwentner et al., 2018; Lozano-Fernandez et al., 2019).

The mitochondrial genome ("mitogenome") contains a wealth of phylogenetic information and can be sequenced from specimens in various states of preservation (e.g., ethanol- or formalin-fixed). Mitogenomic data are particularly useful for resolving taxonomic problems and reconstructing phylogeny in groups like Ostracoda, whose taxonomy is difficult due to the small size of the animals, with few morphological characters available for species delineation (Schön and Martens, 2016) and existence of species complexes (e.g., *Heterocypris salina*; Meisch, 2000). Relevant data, however, remain limited for ostracods; currently, annotated mitogenomic sequences for only five species are available in public databases (INSDC, International Nucleotide Sequence Database Collaboration; http://www.insdc.org/; last accessed on 5 March 2021) (Table 1).

In this study, we sequenced and annotated the mitogenome from the parthenogenetic podocopan ostracod *Heterocypris spadix* Munakata et al., 2021 and reconstructed a phylogeny, based on available mitogenomic sequences from ostracods on the public databases. This allowed us to examine and briefly discuss evolutionary patterns in mitochondrial gene order among known ostracod mitogenomes in a phylogenetic context.

2. Material and methods

DNA was extracted from a mass sample containing 187 ethanol-fixed *H. spadix* individuals that hatched in an aquarium and were the descendants of individuals collected from the type locality (for details, see Munakata et al., 2021), by using a Nucleospin Tissue XS kit (TaKaRa Bio, Japan). Whole-genome shotgun sequencing (2 x 200 bp) was performed on a DNBSEQ-G400 platform (MGI Tech, China) at Bioengineering Lab Co., Ltd., Japan. In total, 15,329,305 paired-end reads (6,131,722,000 bp) were assembled with NOVOPlasty 4.2 (Dierckxsens et al., 2017), using a published cytochrome c oxidase subunit I sequence from *H. spadix* as the seed (INSDC accession number LC557032; Munakata et al., 2021) and a k-mer value of 33.

Genome annotation was performed on the MITOS webserver (Bernt et al., 2013). Upstream and downstream regions detected for each protein-coding gene (PCG) were translated with the invertebrate mitochondrial code; start and stop codons were determined from reciprocal BLASTP (Altschul et al., 1990) searches of the Conserved Domain Database (Lu et al., 2020). Lengths of the 16S and 12S ribosomal RNA (rRNA) genes were determined from reciprocal BLASTn (Altschul et al., 1990) searches of the NCBI nucleotide collection (nt) database. A circular map of the mitogenome was generated with the CGView Server (Grant and Stothard, 2008) and edited in Adobe Illustrator CS6. The GC- and AT-skew plots (Supplementary material SM.01) were generated with the IMC EE ver. 7.32L (In Silico Biology, Japan) using the following parameters: Window size in ratio = 0.05; Step size in ratio = 0.001. The mitochondrial genome obtained was deposited in the INSDC through the DNA Data Bank of Japan.

A phylogenetic analysis of ostracods was conducted that included the sequences of 13 PCGs and two rRNA genes from *H. spadix*, and homologous sequences from another five

ostracod species (Table 1) and one outgroup taxon (*Belzebub intermedius* [Hansen, 1919], Decapoda, Malacostraca; MG719343; Ju et al., 2018); the alignment was 12,673 bp in length. The methods for alignment and selection of optimal substitution models were as described by Kakui and Kano (2021); aligned sequences and optimal substitution models determined for different genes and codon positions are presented as supplementary materials (SM.02 and SM.03). A partitioned maximum-likelihood (ML) analysis was conducted in RAxML-NG (Kozlov et al., 2019), with nodal support values obtained by bootstrap analysis of 500 pseudoreplicates. The ML tree was drawn with FigTree v1.4.4 (Rambaut, 2021).

3. Results

The mitogenome of *Heterocypris spadix* is 15,205 bp long (accession number LC626010). Genomic annotation on the MITOS webserver identified 13 PCGs, two rRNAs, and 22 transfer RNAs (tRNAs), as is typical for animal mitogenomes (Fig. 1A, E; Supplementary material SM.04). The putative control region (CR) was detected at positions 4810–5681, flanked by the 12S rRNA and trnI genes (Fig. 1A, E). This was the longest (872 bp) continuous non-coding region in the *H. spadix* mitogenome, characterized by high (76.5%) AT content and predicted to form multiple stem-loop structures (Supplementary material SM.05).

In the ML tree (Fig. 1B), Myodocopa and Podocopa were each monophyletic, with 100% bootstrap support. Within Podocopa, *H. spadix* was the sister taxon to the confamilial species *Cypridopsis vidua* (O. F. Müller, 1776) with 100% bootstrap support, and together they were sister to *Fabaeformiscandona kushiroensis* Hiruta and Hiruta, 2015. Mitochondrial gene order was incongruent with these relationships: the order in *H. spadix* (Cyprididae) differed from that in *C. vidua* (Cyprididae) but was same as that in *F. kushiroensis* (Candonidae) (Fig. 1E, F). The order in *H. spadix* and *F. kushiroensis* differs from the

hypothetical ancestral pancrustacean (‘pancrustacean ground pattern’; Boore et al., 1998), with translocation of trnK (Fig. 1D, E).

4. Discussion

We determined a complete mitogenomic sequence for *H. spadix* by using a mass DNA sample, thus eliminating the need for a long PCR amplicon. This approach may be an effective option for determining the mitogenomic sequences in other tiny parthenogenetic animals when long PCR amplifications fail.

The relationships in our mitogenome-based phylogenetic tree were consistent with those inferred in previous molecular phylogenetic studies using one or two molecular markers (Hiruta et al., 2016; Pham et al., 2020). Congruent with an analysis based on 16S rRNA (Pham et al., 2020), *Vargula* was not monophyletic, as *Cypridina dentata* (G. W. Müller, 1906) was nested within *Vargula* (Fig. 1B); this conclusion should be further tested by including more species. Hiruta et al. (2016) concluded that 18S and 28S rRNA data alone are not sufficient for resolving relationships among podocopan families or those within Cyprididae. Whole-mitogenomic data may contribute to investigating these relationships, as our phylogeny provided a well-resolved phylogeny, with 100% bootstrap support values over all nodes.

The order of mitochondrial genes differs considerably between Myodocopa and Podocopa (Fig. 1C–F), as pointed out by Hiruta and Hiruta (2015). The myodocopan mitogenomes currently available contain all the genes typically found in animal mitogenomes, but differ greatly in gene order from the pancrustacean ground pattern, with many positional changes (with or without inversions) and subdivision of the control region (Fig. 1C, D). Podocopan mitogenomes differ less from the pancrustacean ground pattern; *F. kushiroensis* and *H. spadix* show a positional change of only one tRNA gene (trnK; Fig. 1D, E). Five

positional changes were observed between *H. spadix* and *C. vidua*, including trnQ, trnC, trnK, trnT, and nad6+cob+trnS2 (Fig. 1E, F). The two species in Cyprididae, *H. spadix* and *C. vidua*, thus differ in mitochondrial gene order despite their sister-group relationship in our tree.

Our results indicate that mitochondrial gene order evolved independently in Podocopa and Myodocopa (Fig. 1B: α and β), as the two groups show no positional changes in common relative to the pancrustacean ground pattern. Gene order in the common ancestor of Podocopa and Myodocopa was likely identical to (or nearly so) the pancrustacean ground pattern; Myodocopa diverged much more from the ground pattern than Podocopa. Gene order appears to be more conservative in Myodocopa than in Podocopa, but this conclusion is premature, because the three myodocopan species examined are confamilial and more closely related to one another than the representative species included for Podocopa. While gene order was conserved across two families in Podocopa (the *Fabaeformiscandona–Heterocypris* pattern in Candonidae and Cyprididae), several positional changes occurred in *Cypridopsis* after divergence of this genus from *Heterocypris* (Fig. 1B: γ) within Cyprididae. These conflicting patterns—apparent gene-order conservation within and between families, but marked divergence within a family—suggest that the evolution of gene order in ostracods was more complex than that expected from the cypridinid data.

Declaration of Competing Interest

The authors declare no competing interests.

Appendix A. Supplementary materials

SM.01 The GC- and AT-skew plots and gene order for *Heterocypris spadix* mitogenome. Window size in ratio = 0.05; Step size in ratio = 0.001. The labels for transfer RNA genes are

omitted. CR, putative control region.

SM.02 Alignment lengths and optimal substitution models for the two rRNA and 13 protein coding genes (PCGs) used for the phylogenetic analysis.

SM.03 Aligned sequences used for phylogeny reconstruction, with the optimal substitution model for each partition.

SM.04 Annotations and characteristic features of genes in the mitochondrial genome of *Heterocypris spadix*. Negative values for IGR (intergenic region) reflect overlap between genes. CR, putative control region.

SM.05 Secondary structure of the putative control region in the *Heterocypris spadix* mitogenome, predicted with the RNAfold WebServer (Gruber et al., 2008; Lorenz et al., 2011); colors indicate the base-pairing probability. A, Structure from a minimum free energy prediction with a partition function. B, Structure from a thermodynamic ensemble prediction. The scale indicates probability levels between 0 and 1 for the various colors.

Acknowledgments

We thank Akira Tsukagoshi for literature; Matthew H. Dick for reviewing and editing the manuscript; and two anonymous reviewers for improving this manuscript with their comments. This study was supported by the Japan Society for the Promotion of Science (JSPS) under KAKENHI grant JP20J40122 to CH.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendofr, M., Stadler, P.F., 2013. MITOS: improved de novo metazoan mitochondrial

genome annotation. Mol. Phylogenet. Evol. 69, 313–319.

Boore, J.L., Lavrov. D.V., Brown, W.M., 1998. Gene translocation links insects and crustaceans. Nature 392, 667–668.

Brandão, S.N., Hoppema, M., Kamenev, G.M., Karanovic, I., Riehl, T., Tanaka, H., Vital, H., Yoo, H., Brandt, A., 2019. Review of Ostracoda (Crustacea) living below the Carbonate Compensation Depth and the deepest record of a calcified ostracod. Prog. Oceanogr. 178, 102144.

Danielopol, D.L., Hartmann, G., 1986. Ostracoda, in: Botosaneanu, L., (Ed.), Stygofauna Mundi. Brill, Leiden, pp. 265–294.

Dierckxsens, N., Mardulyn, P., Smits, G., 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45, e18.

Goodheart, J.A., Minsky, G., Brynjegard-Bialik, M.N., Crummond, M.S., Munoz, J.D., Fallon, T.R., Schultz, D.T., Weng, J.-K., Torres, E., Oakley, T.H., 2020. Laboratory culture of the California Sea Firefly *Vargula tsujii* (Ostracoda: Cypridinidae): developing a model system for the evolution of marine bioluminescence. Sci. Rep. 10, 10443.

Grant, J.R., Stothard, P., 2008. The CGView Server: a comparative genomics tool for circular genomes. Nucleic Acids Res. 36, W181–W184.

Hansen, H.J., 1919. The Sergestidæ of the Siboga-Expedition. Siboga-Expeditie 38, 1–65.

Hiruta, S.F., Hiruta, S., 2015. Description of a species of *Fabaeformiscandona* (Ostracoda, Crustacea) from Kushiro Marsh, Hokkaido, Japan, with the nearly complete mitochondrial genomic sequence. Biodivers. Data J. 3, e7074.

Hiruta, S.F., Kobayashi, N., Katoh, T., Kajihara, H., 2016. Molecular phylogeny of cypridoid freshwater ostracods (Crustacea: Ostracoda), inferred from 18S and 28S rDNA sequences. Zool. Sci. 33, 179–185.

Ju, S.-Y., Song, J.-H., Lee, S.-M., Min, G.-S., 2018. The complete mitochondrial genome of

Belzebub intermedius (Decapoda, Dendrobranchiata, Luciferidae). Mitochondrial DNA B 3, 594–595.

Kakui, K., Kano, Y., 2021. First complete mitochondrial genome for a tanaidacean crustacean (*Arctotanais alascensis*). Zool. Sci., in press.

Karanovic, I., 2012. Recent Freshwater Ostracods of the World. Springer, Berlin.

Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., Stamatakis, A., 2019. RAxML-NG: A fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35, 4453–4455.

Lozano-Fernandez, J., Giacomelli, M., Fleming, J.F., Chen, A., Vinther, J., Thomsen, P.F., Glenner, H., Palero, F., Legg, D.A., Iliffe, T.M., Pisani, D., Olesen, J., 2019. Pancrustacean evolution illuminated by taxon-rich genomic-scale data sets with an expanded remipede sampling. Genome Biol. Evol. 11, 2055–2070.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D., Marchler, G.H., Song, J.S., Thanki, N., Yamashita, R.A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C.J., Marchler-Bauer, A., 2020. CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Res. 48, D265–D268.

Meisch, C. 2000. Freshwater Ostracoda of Western and Central Europe. Spektrum Akademischer Verlag, Berlin.

Müller, O.F., 1776. Zoologiæ Danicæ Prodromus, seu Animalium Daniæ et Norvegiæ Indigenarum, Characteres, Nomina, et Synonyma Imprimis Popularium, Typis Hallageriis, Havniæ.

Müller, G.W., 1906. Die Ostracoden der Siboga-Expedition. Siboga-Expeditie 30, 1–40.

Munakata, M., Tanaka, H., Kakui, K., 2021. *Heterocypris spadix* sp. nov. (Crustacea: Ostracoda: Cypridoidea) from Japan, with information on its reproductive mode. Zool. Sci., in press.

Oakley, T.H., Wolfe, J.M., Lindgren, A.R., Zaharoff, A.K., 2013. Phylotranscriptomics to bring the understudied into the fold: monophyletic Ostracoda, fossil placement, and pancrustacean phylogeny. Mol. Biol. Evol. 30, 215–233.

Ogoh, K., Ohmiya, Y., 2004. Complete mitochondrial DNA sequence of the sea-firefly, *Vargula hilgendorfii* (Crustacea, Ostracoda) with duplicate control regions. Gene 327, 131–139.

Ogoh, K., Futahashi, R., Ohmiya, Y., 2021. Intraspecific nucleotide polymorphisms in seven complete sequences of mitochondrial DNA of the luminous ostracod, *Vargula hilgendorfii* (Crustacea, Ostracoda). Gene Rep. 23, 101074.

Pham, H.T.M., Tanaka, H., Karanovic, I., 2020. Molecular and morphological diversity of *Heterodesmus* Brady and its phylogenetic position within Cypridinidae (Ostracoda). Zool. Sci. 37, 240–254.

Pinto, R.L., Rocha, C.E.F., Martens, K., 2005. On new terrestrial ostracods (Crustacea, Ostracoda) from Brazil, primarily from São Paulo State. Zool. J. Linn. Soc. 145, 145–173.

Rambaut, A., 2021. FigTree v1.4.4. http://tree.bio.ed.ac.uk/software/figtree/ (accessed 5 March 2021).

Rodriguez-Lazaro, J., Ruiz-Muñoz, F., 2012. A general introduction to ostracods: morphology, distribution, fossil record and applications. Dev. Quat. Sci. 17, 1–14.

Schön, I., Martens, K., 2016. Ostracod (Ostracoda, Crustacea) genomics—promises and challenges. Mar. Genomics 29, 19–25.

Schwentner, M., Richter, S., Rogers, D.C., Giribet, G., 2018. Tetraconatan phylogeny with special focus on Malacostraca and Branchiopoda: highlighting the strength of taxon-specific matrices in phylogenomics. Proc. R. Soc. B 285, 20181524.

Tran Van, P., Anselmetti, Y., Bast, J., Dumas, Z., Galtier, N., Jaron, K.S., Martens, K., Parker,

D.J., Robinson-Rechavi, M., Schwander, T., Simion, P., Schön, I., 2021. First annotated draft genomes of non-marine ostracods (Ostracoda, Crustacea) with different reproductive modes. G3 Genes Genom. Genet., jkab043.

Wang, X., Xu, Q., Xia, J., Miao, X., Liu, P., Wang, Z., 2019. First record of the complete mitochondrial genome of *Cypridina dentata* (Myodocopida: Cypridinidae). Mitochondrial DNA B 4, 1607–1608.

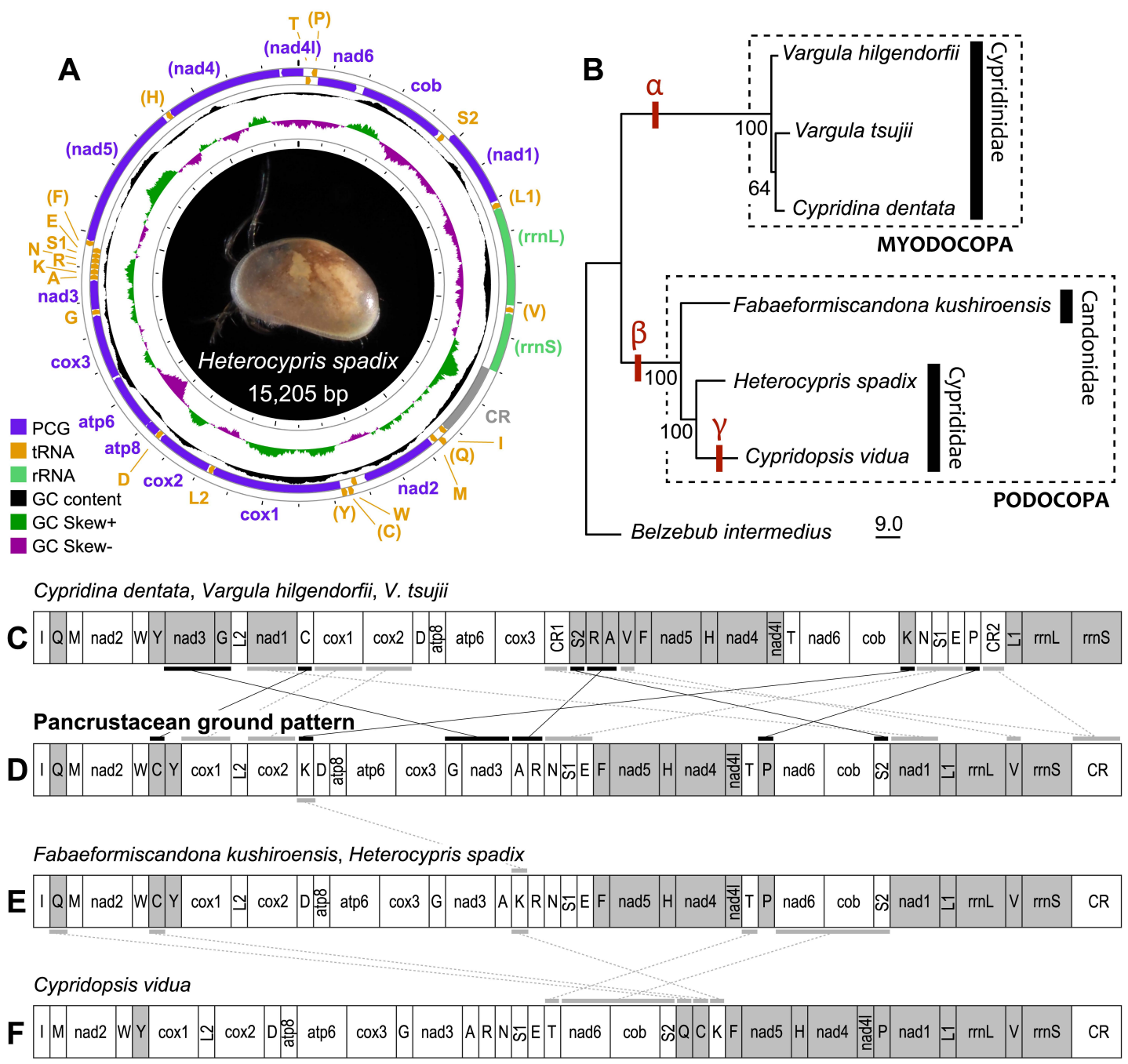
Figure caption

Fig. 1. Mitochondrial gene order and maximum-likelihood (ML) phylogenetic tree for six ostracod species. **A,** Map of the mitogenome of *Heterocypris spadix*. **B,** ML phylogeny based on mitogenomic sequences (12,673 bp from 13 PCGs and two rRNA genes). Numbers near nodes are bootstrap values. The scale indicates branch length in number of substitutions per site. α, β, γ, hypothetical gene-order rearrangement events. **C–F,** Diagrammatic linearized mitochondrial chromosomes showing gene order for myodocopans (C), the hypothetical ancestral pancrustacean (Boore *et al*., 1998) (D), *F. kushiroensis* and *H. spadix* (E), and *C. vidua* (F). Bars above or below genes indicate genes or gene clusters that have undergone rearrangements, with (black bars) or without (gray bars) inversions; fine solid/dashed lines between chromosomes indicate the locations of positional shifts. Transfer RNA genes are labeled with their one-letter amino acid codes. Genes on the reverse (–) strand are in parentheses in **A** or shaded gray in **C–F**. CR, putative control region; PCG, protein coding gene.

Table 1. Ostracod species for which a complete, annotated mitogenome sequence is currently available.

Subclass	Family	Species name	Accession no.	Source
Myodocopa	Cypridinidae	*Cypridina dentata*	MK482395	Wang et al. (2019)
		Vargula hilgendorfii	NC_005306[a], AP007266–007272	Ogoh and Ohmiya (2004); Ogoh et al. (2021)
		Vargula tsujii	MG767172	Goodheart et al. (2020)
Podocopa	Candonidae	*Fabaeformiscandona kushiroensis*	AP014656	Hiruta and Hiruta (2015)
	Cyprididae	*Cypridopsis vidua*	KP063117	Unpublished
		Heterocypris spadix	LC626010	This study

[a], the sequence used for our phylogeny reconstruction.


A
T (P)
(nad4l)
nad6
(nad4)
cob
(H)
S2
(nad5)
(nad1)
(F)
(L1)
E
S1
N
R
K
A
(rrnL)
nad3
G
(V)
(rrnS)
cox3
Heterocypris spadix
15,205 bp
CR
atp6
I
atp8
(Q)
D
cox2
M
L2
nad2
cox1
(Y)
W
(C)
PCG
tRNA
rRNA
GC content
GC Skew+
GC Skew-
B
α
Vargula hilgendorfii
Cypridinidae
100
Vargula tsujii
64
Cypridina dentata
MYODOCOPA
Fabaeformiscandona kushiroensis
Candonidae
β
100
Heterocypris spadix
Cypridinidae
100
γ
Cypridopsis vidua
PODOCOPA
Belzebub intermedius
9.0
Cypridina dentata, Vargula hilgendorfii, V. tsujii
C
I Q M nad2 W Y nad3 G L2 nad1 C cox1 cox2 D atp8 atp6 cox3 CR1 S2 R A V F nad5 H nad4 nad4l T nad6 cob K N S1 E P CR2 L1 rrnL rrnS
Pancrustacean ground pattern
D
I Q M nad2 W C Y cox1 L2 cox2 K D atp8 atp6 cox3 G nad3 A R N S1 E F nad5 H nad4 nad4l T P nad6 cob S2 nad1 L1 rrnL V rrnS CR
Fabaeformiscandona kushiroensis, Heterocypris spadix
E
I Q M nad2 W C Y cox1 L2 cox2 D atp8 atp6 cox3 G nad3 A K R N S1 E F nad5 H nad4 nad4l T P nad6 cob S2 nad1 L1 rrnL V rrnS CR
Cypridopsis vidua
F
I M nad2 W Y cox1 L2 cox2 D atp8 atp6 cox3 G nad3 A R N S1 E T nad6 cob S2 Q C K F nad5 H nad4 nad4l P nad1 L1 rrnL V rrnS CR