Taylor & Francis
Taylor & Francis Group

MITOGENOME ANNOUNCEMENT

🔓 OPEN ACCESS   Check for updates

The complete mitochondrial genome of *Gruberia lanceolata* (Gruber, 1884) Kahl, 1932 (Ciliophora: Heterotrichea)

Mi-Hyun Park 🄳 and Gi-Sik Min 🄳

Department of Biological Sciences, Inha University, Incheon, The Republic of Korea

ABSTRACT

The ciliate *Gruberia lanceolata* (Gruber, 1884) Kahl, 1932 belonging to the class Heterotrichea was sampled from the coastal waters of South Korea. The complete mitogenome in its linear form and large size (∼40 kb) was obtained. It consisted of 27 protein-coding genes (PCGs), two ribosomal subunit RNA (rRNA) genes, four transfer RNAs (tRNAs), and ten unclassified open reading frames (ORFs). Their telomeric structures were capped, with repeat regions at both ends. We analyzed its phylogenetic tree using the data of its respiratory chain complex I genes. It can be suggested that the complete mitochondrial genome of *G. lanceolata* can be recorded as a new class of the mitogenome.

To date, mitogenomes of only 12 species have been identified in members belonging to the phylum Ciliophora (Pritchard et al. 1990; Burger et al. 2000; de Graaf et al. 2009, 2011; Swart et al. 2012; Li et al. 2018; Park et al. 2019). The genus *Gruberia* is not well-known and comprises only of seven species (Chen et al. 2018) in the class Heterotrichea. It was sampled from the coastal waters of the Seonnyeobawi Beach, Yellow Sea, Korea (37°26′N, 126°22′E; salinity: 29 psu), on 24 October 2014. A slide of protargol-impregnated specimens was deposited at the National Institute of Biological Resources, Korea (NIBRPR0000106616) (Park and Min 2016). The mitochondrial DNA was sequenced using the Illumina Miseq platform (Macrogen, Seoul, Korea) and the mitogenome from partial *cox1* was assembled using NOVOplasty 2.6.3 (Dierckxsens et al. 2016). Furthermore, we determined complete mitogenome sequence by comparing it with the contigs available in SOAPdenovo v2.01 (Luo et al. 2012). The annotation was performed using MFANNOT webserver (http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl) and RNAweasal webserver (http://megasun.bch.umontreal.ca/cgi-bin/RNAweasel/RNAweaselInterface.pl) based on ERPIN version 5.2.1. (Gautheret and Lambert 2001), with assistance from the mold mitochondrial genetic code (i.e. genetic code 4). The transfer RNA-encoding genes were identified using the tRNAscan-SE (Lowe and Eddy 1997; Lowe and Chan 2016). The maximum-likelihood (ML) phylogenetic tree for concatenated mitochondrial complex I genes (*nad1*, *nad2*, *nad3*, *nad4*, *nad5*, *nad7*, *nad9*, and *nad10*; except *nad4L* and *nad 6*) was constructed using RAxML (Stamatakis 2014) with 1000 bootstrap replicates based on the amino acids aligned using MAFFT (Katoh and Standley 2013). All gaps representing many contiguous non-conserved positions were removed using the Gblocks Server Version 0.91 b (Talavera and Castresana 2007).

The complete mitochondrial genome of *Gruberia lanceolata* consisting of 39,989 bp was deposited in the GenBank (accession number MK301177) for identification. The mitogenome is a linear structure comprising repeat regions in the telomeres at both ends, and its transcription occurs in single, as well as reverse directions. The nucleotide composition of *G. lanceolata* is as follows: A: 44.2%, C: 12.3%, G: 8.1%, T: 35.4%. The GC content of the complete genome was found to be 20.4%. The genome contained 27 protein-coding genes (PCGs), 2 ribosomal subunit RNA (rRNA) genes, 4 tRNAs, and 10 unclassified open reading frames (ORFs). The 27 PCGs included genes encoding the following molecules: respiratory chain complex I (*nad1*, *nad2*, *nad3*, *nad4*, *nad4_i*, *nad4L*, *nad5*, *nad6*, *nad7*, *nad9*, and *nad10*), complex III (*cob*), complex IV (*cox1* and *cox2*), and complex V (*atp9*); cytochrome c-related genes (*ccmf_i* and *ccmf_ii*); small and large ribosomal subunit genes (*rps3*, *rps4*, *rps8*, *rps12*, *rps14*, *rps19*; *rpl2*, *rpl6*, *rpl14*, and *rpl16*).

Based on the data of the complex I genes, it was determined that the species were clustered within their own classes, and *G. lanceolata* formed a separate clade at the class level in all gene trees (Figure 1). It can be expected that the
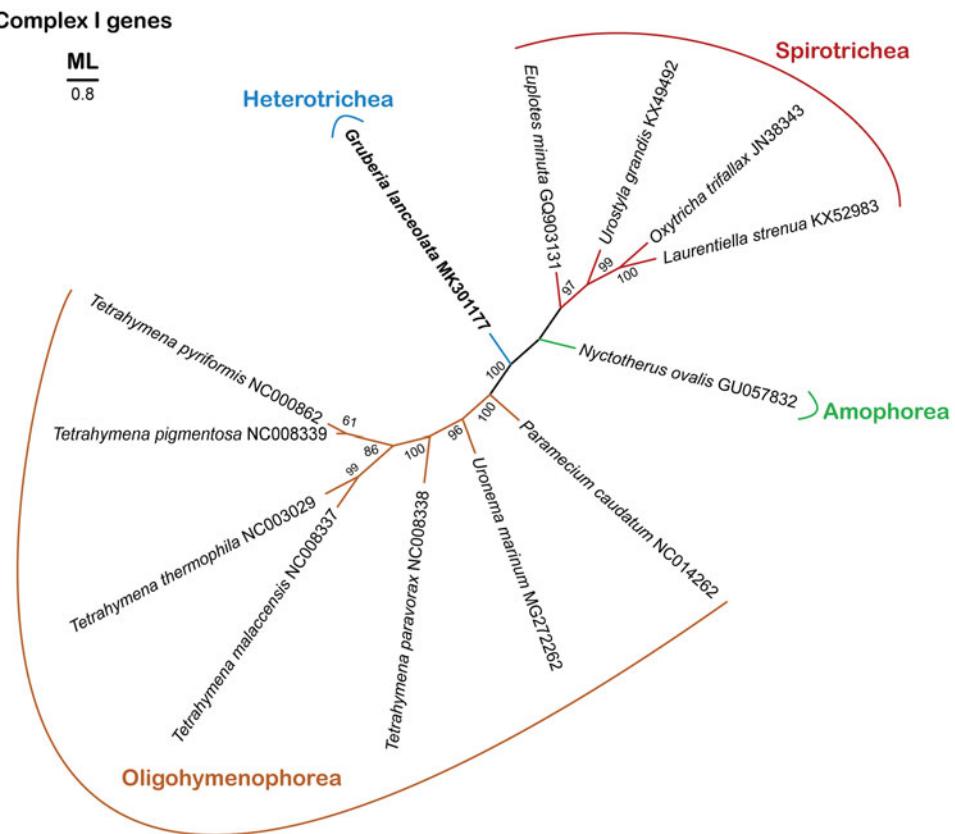
Figure 1. The maximum-likelihood (ML) tree for mitochondrial respiratory chain complex I (*nad1, nad2, nad3, nad4, nad5, nad7, nad9,* and *nad10*) genes. The ML tree of complex I genes is unrooted.

phylogenetic analysis using the complete mitogenome well represents the phylogenetic relationship of ciliates.

Disclosure statement

Funding

ORCID

Mi-Hyun Park http://orcid.org/0000-0002-2844-3546
Gi-Sik Min http://orcid.org/0000-0003-2739-3978

References

Burger G, Zhu Y, Littlejohn TG, Greenwood SJ, Schnare MN, Lang BF, Gray MW. 2000. Complete sequence of the mitochondrial genome of *Tetrahymena pyriformis* and comparison with *Paramecium aurelia* mitochondrial DNA. J Mol Biol. 297:365–380.

Chen X, Shazib SUA, Kim JH, Kim MS, Shin MK. 2018. New contributions to *Gruberia lanceolata* (Gruber, 1884) Kahl, 1932 based on analyses of multiple populations and genes (Ciliophora, Heterotrichea, Gruberiidae). Eur J Protistol. 65:16–30.

de Graaf RM, Ricard G, van Alen TA, Duarte I, Dutilh BE, Burgtorf C, Kuiper JWP, van der Staay GWM, Tielens AGM, Huynen MA, et al. 2011. The organellar genome and metabolic potential of the

hydrogen-producing mitochondrion of *Nyctotherus ovalis*. Mol Biol Evol. 28:2379–2391.

de Graaf RM, van Alen TA, Dutilh BE, Kuiper JW, van Zoggel HJ, Huynh MB, Gortz HD, Huynen MA, Hackstein JH. 2009. The mitochondrial genomes of the ciliates *Euplotes minuta* and *Euplotes crassus*. BMC Genomics. 10:514.

Dierckxsens N, Mardulyn P, Smits G. 2016. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45:e18–e18.

Gautheret D, Lambert A. 2001. Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. J Mol Biol. 313:1003–1011.

Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. Mol Biol Evol. 30:772–780.

Li R, Gao Y, Hou Y, Ye S, Wang L, Sun J, Li Q. 2018. Mitochondrial genome sequencing and analysis of scuticociliates (*Uronema marinum*) isolated from *Takifugu rubripes*. Mitochondrial DNA Part B Resour. 3: 736–737.

Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res. 44: W54–W57.

Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25: 955–964.

Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaSci. 1:18.

Park K-M, Min G-S, Kim S. 2019. The mitochondrial genome of the ciliate *Pseudourostyla cristata* (Ciliophora, Urostylida). Mitochondrial DNA Part B Resour. 4:66–67.

Park M-H, Min G-S. 2016. Morphological reports of four ciliates (Ciliophora) from coastal marine and brackish water habitats in Korea. J Species Res. 5:514–516.

Pritchard AE, Seilhamer JJ, Mahalingam R, Sable CL, Venuti SE, Cummings DJ. 1990. Nucleotide sequence of the mitochondrial genome of *Paramecium*. Nucl Acids Res. 18:173–180.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30:1312–1313.

Swart EC, Nowacki M, Shum J, Stiles H, Higgins BP, Doak TG, Schotanus K, Magrini VJ, Minx P, Mardis ER, et al. 2012. The *Oxytricha trifallax* mitochondrial genome. Gen Biol Evol. 4:136–154.

Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 56:564–577.