

Characterization of the complete mitochondrial genome of *Notospermus geniculatus*

Jin-qing Jiang^a and Rui-guang Deng^b

^aCollege of Animal Science and Veterinary Medicine, Henan Institute of Science and Technology, Xinxiang, Henan Province, China;

^bKey Laboratory of Animal Immunology, Henan Academy of Agricultural Sciences, Zhengzhou, China

ABSTRACT

In this study, the complete mitochondrial genome of *Notospermus geniculatus*, was recovered through Illumina sequencing data. This complete mitochondrial genome of *N. geniculatus* is 15,180 bp in length and has a base composition of A (14.4%), T (41.3%), C (14.6%), G (29.6%), demonstrating a bias of higher AT content (55.7%) than GC content (44.2%). The mitochondrial genome contains a typically conserved structure among Lineidae mitogenomes, encoding 13 protein-coding genes (PCGs), 23 transfer RNA genes (tRNA), 2 ribosomal RNA genes (12S rRNA and 16S rRNA), and a control region (D-loop region). All PCGs were located on the H-strand. ND4L gene and ND4 gene were overlapped by 6 bp. The whole mt genome of *N. geniculatus* and other Protostomia mitogenomes (17 species, in total) were used for phylogenetic analysis. The result indicated *N. geniculatus* has the closest relationship with *Lineus viridis* (FJ839919.1) and clustered within Heteronemertea Clade, which representing a distinct order.

ARTICLE HISTORY

Received 3 August 2018

Accepted 31 August 2018

KEYWORDS

Notospermus geniculatus;
mitochondrial genome;
assembly; phylogeny



Notospermus geniculatus is one of the largest nemerteans present in Mediterranean waters, reaching 1 m in length and 15 mm in width. This species was originally described from Naples and is also widely distributed (Kajihara 2007). Here, we provide a report of the complete mitochondrial genome of *Notospermus geniculatus*.

In this study, a specimen of *N. geniculatus* was collected from Balearic Islands (38°73'N, 1°41'E) and fixed in absolute ethanol. Genomic DNA was extracted from muscle tissues using DNEasy Extraction Kit (supplied by QIAGEN) following the manufactures instructions. The isolated DNA was stored in the sequencing company (HuiTong Tech, Shenzhen, China). Purified DNA was fragmented and used to construct the sequencing libraries following the instructions of NEBNext[®] Ultra[™] II DNA Library Prep Kit (NEB, BJ, CN). Whole genomic sequencing was performed by the Illumina HiSeq 2500 Sequencing Platform (Illumina, San Diego, CA). Adapters and low-quality reads were removed using the NGS QC Toolkit (Patel et al. 2012). Then assembly as implemented by SPAdes 3.9.0 (Bankevich et al. 2012). Circularization of this mt genome was confirmed using MITObim V1.9 (Hahn et al. 2013). The complete sequence was primarily annotated by ORF prediction in Unipro UGENE (Okonechnikov et al. 2012) combined with manual correction. All tRNAs were confirmed using the tRNAscan-SE search server (Lowe et al. 1997). Other protein-coding genes were verified by BLAST search on the

NCBI website (<http://blast.ncbi.nlm.nih.gov/>), and manual correction for start and stop codons were conducted. This complete mitochondrial genome sequence together with gene annotations were submitted to GenBank under the accession numbers of MH714705.

The complete mitochondrial genome of *Notospermus geniculatus* was 15,180 bp in length and has a base composition of (14.4%), T (41.3%), C (14.6%), G (29.6%), demonstrating a bias of higher AT content (55.7%) than GC content (44.2%). The mitochondrial genome contains a typically conserved structure among Lineidae mitogenomes, encoding 13 protein-coding genes (PCGs), 23 transfer RNA genes (tRNA), 2 ribosomal RNA genes (12S rRNA and 16S rRNA), and a control region (D-loop region). All PCGs were located on the H-strand. ND4L gene and ND4 gene were overlapped by 6 bp.

For phylogenetic analysis assessing the relationship of this mitogenome, we selected other 17 Protostomia mitogenomes from Lineidae (4 taxa), Monostilifera (4 taxa), Palaeonemertea (1 taxon), Molluscae (4 taxa), and Ecdysozoa (2 taxa) to construct genome-wide alignment. The genome-wide alignment of all mt genomes was done by HomBlocks (Bi et al. 2018) under trimAl method, which containing all phylogenetic informative gaps in alignments, resulting in 4248 positions in total, including three conserved fragments shared by all mitogenomes. The whole genome alignment was analyzed by IQ-TREE version 1.6.6 (Nguyen et al. 2014)

CONTACT Jin-qing Jiang  jjq5678@126.com  College of Animal Science and Veterinary Medicine, Henan Institute of Science and Technology, Xinxiang, Henan Province, 453000, China

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

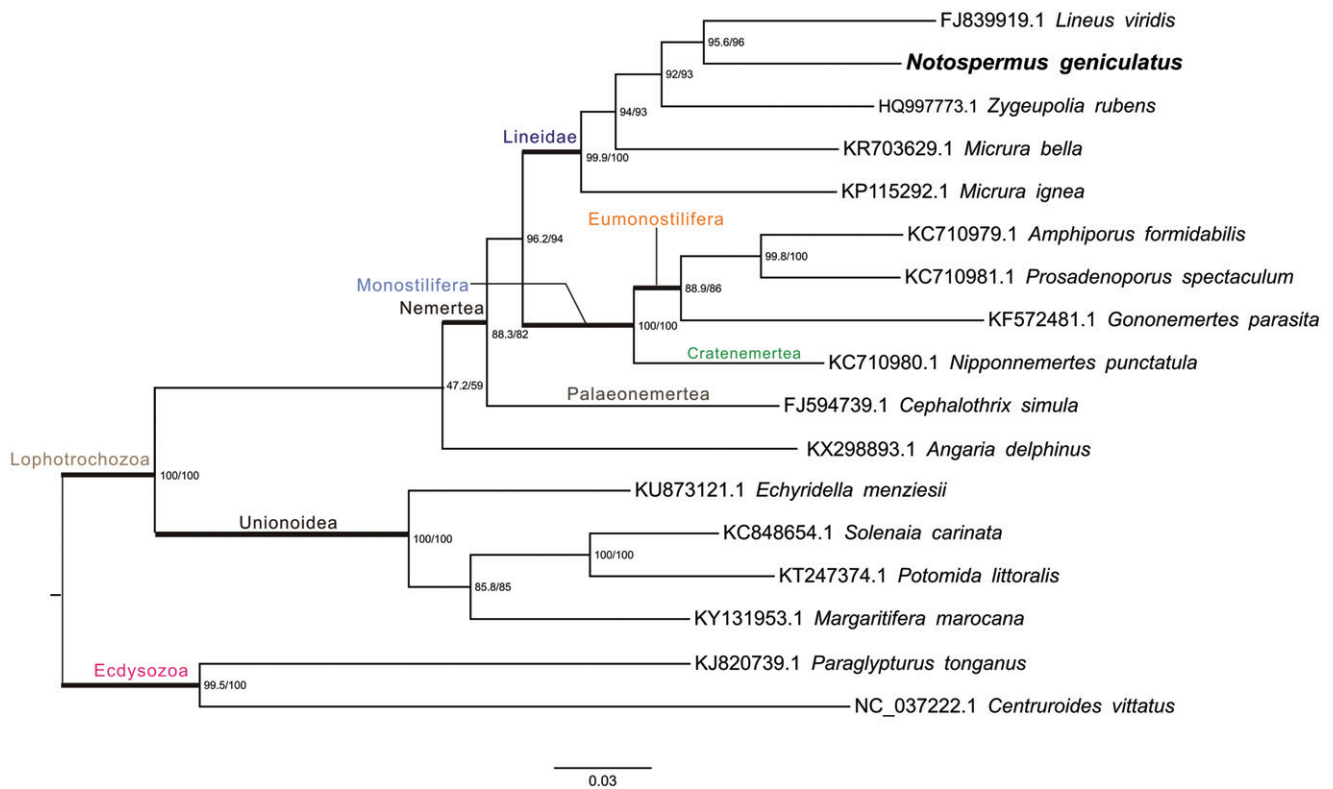


Figure 1. Phylogenetic tree yielded by IQ-TREE of 17 Protostomia mitogenomes. Consensus tree is shown with support indicated by numbers at branches, representing percentages of SH-aLRT test and bootstraps.

under the TIM3 + F + R3 model. The tree topology was verified under both 1000 bootstrap and 1000 replicates of SH-aLRT test. The resulting tree was represented and edited using FigTree v1.4.1. As shown in Figure 1, the phylogenetic positions of these 17 mt genomes were successfully resolved with high bootstrap and SH-aLRT supports except with one node. The result indicated *N. geniculatus* has the closest relationship with *Lineus viridis* (FJ839919.1) and clustered within Heteronemertea Clade, which represents a distinct order.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was financially supported by the Open Project of Henan Academy of Agricultural Sciences [PKLAI20170601].

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Bi GQ, Mao YX, Cao M. 2018. HomBlocks: a multiple-alignment construction pipeline for organelle phylogenomics based on locally collinear block searching. *Genomics.* 110(1):18–22.
- Hahn C, Bachmann L, Chevreaux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads – a baiting and iterative mapping approach. *Nucleic Acids Res.* 41(13): e129–e129.
- Kajihara, H., 2007. A Taxonomic Catalogue of Japanese Nemerteans (Phylum Nemertea). *Zoo Sci.* 24(4):287–326.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25(5): 955–964.
- Nguyen, L-T, Schmidt HA, von Haeseler A, Minh BQ. 2014. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32(1):268–274.
- Okonechnikov K, Golosova O, Fursov M, the UGENE team. 2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics.* 28(8): 1166–1167.
- Patel RK, Jain M. 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One.* 7(2):e30619.