

## The complete mitochondrial genome and phylogenetic analysis of *Coscinodiscus wailesii* (Coscinodiscophyceae, Bacillariophyta)

Hailong Huang<sup>a,b,d</sup>, Yichao Wang<sup>a,b,c,d</sup>, Huiyin Song<sup>a,b,d</sup>, Jing Wang<sup>a,b,d</sup>,  
Yang Chen<sup>a,b,c,d</sup>, Yongfang Zhao<sup>b,d,e</sup>, Feng Liu<sup>a,b,d</sup>  and Nansheng Chen<sup>a,b,d,f</sup> 

<sup>a</sup>CAS Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China; <sup>b</sup>Laboratory of Marine Ecology and Environmental Science, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China; <sup>c</sup>University of Chinese Academy of Sciences, Beijing, China; <sup>d</sup>Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao, China; <sup>e</sup>Jiaozhou Bay National Marine Ecosystem Research Station, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China; <sup>f</sup>Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada

### ABSTRACT

*Coscinodiscus* is a species-rich genus with about 400 species described, many of which are harmful algal bloom species with significant negative ecological impact. Despite of their importance in primary production and as harmful algal bloom species, genome data for species in this genus is limited. No mitochondrial genome (mtDNA) of any species in this genus has been reported. Here, the complete mtDNA sequence of the *Coscinodiscus wailesii* Gran & Angst 1931 was constructed and analyzed. The circular mtDNA was 36,071 bp in length, encoding 64 genes, including 34 protein coding genes (PCGs), 24 transfer RNA (tRNA) genes, 2 ribosomal RNA (rRNA) genes and 4 conserved open reading frames (orfs). The overall AT content of *C. wailesii* mtDNA was 75.00%, which was slightly lower than that of *Melosira undulate* (78.40%). Maximum likelihood (ML) phylogenetic analysis using 29 shared protein-coding genes revealed that *C. wailesii* clustered well with *M. undulata*, which was the only species of class Coscinodiscophyceae whose mtDNA has been fully constructed. The complete mtDNAs of more *Coscinodiscus* species will be valuable for studying the evolutionary relationships among species in the genus *Coscinodiscus* and in the Class of Coscinodiscophyceae.

### ARTICLE HISTORY

Received 21 January 2021  
Accepted 15 May 2021

### KEYWORDS



Diatoms; mitochondrial genome; *Coscinodiscus wailesii*; Coscinodiscophyceae

Diatoms are a diverse phytoplankton group estimated to contribute up to 20% of global primary production (Mann 1999). *Coscinodiscus wailesii* belongs to the genus *Coscinodiscus*, the class Coscinodiscophyceae, and the phylum Bacillariophyta with a worldwide distribution (Rick and Dürselen 1995). The genus *Coscinodiscus* is species-rich with about 400 species identified globally (Li 2009) and 49 species identified in China (Xie 2006), many of which have been found to form blooms with severe negative impact directly on environment (Hasle and Lange 1992). The cell size of *Coscinodiscus* species is large, about 70–470 µm, with a high carbon content, which poses a great impact on the whole carbon pool and plays an important role in the marine ecosystem (Sun and Liu 2005). *Coscinodiscus wailesii* is a harmful algal bloom (HAB) species that can cause serious damage to *Nori* (*Pyropia*) cultivation through competitive utilization of nutrients during its bloom (Tetsuya et al. 2000; Nishikawa et al. 2010), and forms dense HABs that cause extensive clogging of fishing nets and other equipment, such as cages (Laing and Gollasch 2002).

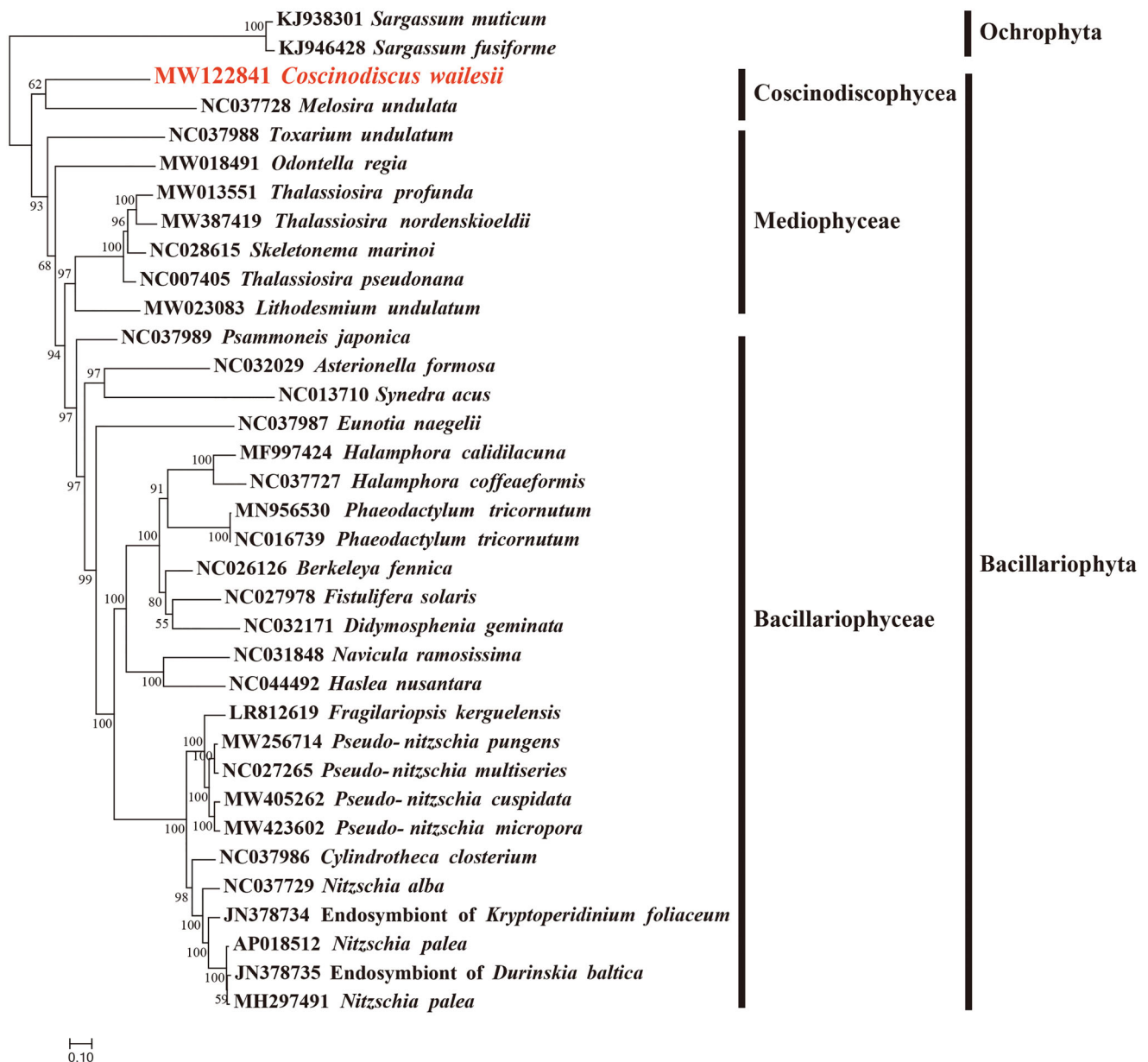
Mitochondrial genome is an ideal handle for genetic and phylogenetic analyses because of its uniparental inheritance

and high evolutionary rates (Wang et al., 2016). Although the mtDNAs of many species in phylum Bacillariophyta have been sequenced, most of these mtDNAs were constructed for species of two of the three classes including Mediophyceae and Bacillariophyceae. By now, mtDNA of only a single species (e.g., *M. undulata*) in the third class Coscinodiscophyceae has been constructed. The first aim of this study was to construct mtDNA of the species *C. wailesii*, which represents the second mtDNA of species of the class Coscinodiscophyceae. The second aim was to compare the two mtDNAs of these two species *C. wailesii* and *M. undulate* in the class Coscinodiscophyceae. Here, we assemble and annotate the complete mtDNA of *C. wailesii* for the first time, as part of our effort to gain better understanding of its genetic characteristics at the genome level.

The strain *C. wailesii* CNS00372 used in this study was isolated in water samples collected during an expedition to the Jiaozhou Bay, China (36°08.364'N, 120°11.359'E) in September 2019 onboard the research vehicle 'Innovation'. Its specimen was deposited in the collection of marine algae in KLMEES of IOCAS (Nansheng Chen, [chenn@qdio.ac.cn](mailto:chenn@qdio.ac.cn)) under the voucher number CNS00372. The annotation of

**CONTACT** Nansheng Chen  [chenn@qdio.ac.cn](mailto:chenn@qdio.ac.cn)  Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, 266071, China

© 2021 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.** Maximum likelihood phylogenetic tree using concatenated amino acid sequences of 29 shared protein-coding genes (*atp* 8, 9; *cob*; *cox1*, 2, 3; *nad1-7*, 4L, 9; *rpl2*, 5, 6, 14, 16; *rps3*, 4, 8, 10, 11, 13, 14, 19; and *tatC*) from 33 publicly available diatom mtDNAs, and *Sargassum fusiforme* (KJ946428) and *Sargassum muticum* (KJ938301) of Ochrophyta were used as out-group taxa. The numbers beside branch nodes are the percentage of 1000 bootstrap values.

protein-coding genes (PCGs), transfer RNA (tRNA) genes, ribosomal RNA (rRNA) genes and open reading frames (*orfs*) were conducted using Open Reading Frame Finder (ORF finder) (<https://www.ncbi.nlm.nih.gov/orffinder>), tRNAscan-SE 2.0 (Chan and Lowe 2019) and MFannot (<https://megasun.bch.umontreal.ca/RNAweasel>).

The circular mtDNA of *C. walesii* was 36,071 bp in size (GenBank accession number: MW122841). The mtDNA encoded a set of 64 genes found in most diatom mtDNAs (Wang et al. 2021; Zhang et al. 2021), including 34 protein-coding genes (PCGs), 2 ribosomal RNA (rRNA) genes, 24 transfer RNA (tRNA) genes, and 4 open reading frames (*orfs*). The heavy strand (+) encoded 56 genes while the light strand (-) had 8 genes. The percentage of A+T of the mtDNA of *C. walesii* was 75.00%, which was slightly lower than that of *M. undulata*. The 34 PCGs included *atp6,8,9*; *cob*;

*cox1-3*; *nad1-7,4L,9,11*; *rpl2,5,6,14,16*; *rps3,4,7,8,10-14,19*; and *tatA,C*. All PCGs began with ATG except *atp8*, which began with TTA. The termination codons of PCGs were either TAA (29 of 34 genes) or TAG (5 of 34 genes). The 24 transfer RNA (tRNA) genes ranged in size from 71 bp to 89 bp. Five pairs of overlapping genes were identified in the *C. walesii* mtDNA, including *rps10-trnF* (10 bp), *rps12-rps7* (65 bp), *nad1-tatC* (48 bp), *orf143-trnP* (9 bp), and *rrs-trnS* (53 bp).

Maximum likelihood (ML) phylogenetic tree was constructed using concatenated amino acid sequences of 29 protein-coding genes (*atp* 8, 9; *cob*; *cox1*, 2, 3; *nad1-7*, 4L, 9; *rpl2*, 5, 6, 14, 16; *rps3*, 4, 8, 10, 11, 13, 14, 19; and *tatC*) shared by 33 publicly available diatom mtDNAs (Figure 1), using IQtree v1.6.12 (Trifinopoulos et al. 2016) with 1000 bootstrap replications. Two species *Sargassum fusiforme* (KJ946428) and *Sargassum muticum* (KJ938301) of Ochrophyta were used as

out-group taxa. The results demonstrated that species fell well into three clades corresponding to three classes of the phylum Bacillariophyta, including Coscinodiscophyceae, Mediophyceae and Bacillariophyceae. In a separate study, Crowell et al. (2019) showed that *Durinskia baltica* and the two *Nitzschia palea* plastomes grouped on a strongly supported branch (bootstrap 100) within a larger clade that included *Kryptoperidinium foliaceum*. The phylogenetic positions of Endosymbiont of *K. foliaceum* and *D. baltica* are consistent with our results described in this manuscript. Recent studies have also proposed that the species *Thalassiosira pseudonana* belongs to the genus *Cyclotella* (Alverson et al. 2011). As expected, mtDNA of *C. walesii* (MW122841) showed a close relationship with that of *M. undulata* (NC037728) (Figure 1), which was the only species in the class Coscinodiscophyceae whose mtDNA had been fully constructed (Pogoda et al. 2019). Comparative analysis of mtDNAs of *C. walesii* and *M. undulata* revealed extensive genome rearrangement events including translocation and inversion events. Nevertheless, many conserved gene blocks could be identified, such as *atp6-rps10-trnF(gaa)-rps8-rpl6*. These results suggested that mtDNAs of many more species of the class Coscinodiscophyceae are needed for in depth understanding of their evolutionary dynamics.

## Acknowledgements

We are thankful to all staffs of marine ecological environment genomics research group in Institute of Oceanology, Chinese Academy of Sciences. We are grateful to colleagues from the Jiaozhou Bay Marine Ecosystem Research Station for their help in field sampling.

## Disclosure statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Funding

This research was supported by the Strategic Priority Research Program of Chinese Academy of Sciences [Grant No. XDB42000000], the Error! Hyperlink reference not valid. (to Nansheng Chen), the Taishan Scholar Project Special Fund (to Nansheng Chen), the Qingdao Innovation and Creation Plan (Talent Development Program – 5th Annual Pioneer and Innovator Leadership Award to Nansheng Chen, 19-3-2-16-zhc), the Key Research Program of Frontier Sciences, Chinese Academy of Sciences [No. QYZDB-SSW-DQC023] (to Feng Liu), and the Major Scientific and Technological Innovation Project of Shandong Province [No. 2019JZZY020706] (to Feng Liu), the Post-doctoral Program of Qingdao 2020 [Grant No. EOKY061] (to Hailong Huang) and the Natural Science Foundation of Shandong Province, China [Grant No. ZR2020QD095] (to Hailong Huang).

## ORCID

Feng Liu  <http://orcid.org/0000-0002-5774-900X>

Nansheng Chen  <http://orcid.org/0000-0002-0615-7215>

## 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 number MW122841. The associated BioProject, SRA and Bio-Sample numbers are PRJNA686182, SRR13269809 and SAMN17108303, respectively.

## References

- Alverson AJ, Beszteri B, Julius ML, Theirot EC. 2011. The model marine diatom *Thalassiosira pseudonana* likely descended from a freshwater ancestor in the genus *Cyclotella*. *BMC Ecol Evol.* 11:125–125.
- Chan PP, Lowe TM. 2019. tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods Mol Biol.* 1962:1–14.
- Crowell RM, Nienow JA, Cahoon AB. 2018. The complete chloroplast and mitochondrial genomes of the diatom *Nitzschia palea* (Bacillariophyceae) demonstrate high sequence similarity to the endosymbiont organelles of the dinoflagellate *Durinskia baltica*. *J Phycol.* 55(2): 352–364.
- Hasle GR, Lange CB. 1992. Morphology and distribution of *Coscinodiscus* species from the Oslofjord, Norway, and the Skagerrak. *North Atlantic Diatom Res.* 7(1):37–68.
- Laing I, Gollasch S. 2002. *Coscinodiscus walesii* – a nuisance diatom in European waters. Invasive aquatic species of Europe. Distribution, impacts and management. Amsterdam (The Netherlands): Springer.
- Li Y. 2009. Morphological characteristics comparisons between *Thalassiosira* and *Coscinodiscus*. *Bull Bot Res.* 29(3):282–288.
- Mann DG. 1999. The species concept in diatoms. *Phycologia.* 38(6): 437–495.
- Nishikawa T, Tarutani R, Yamamoto R. 2010. Nitrate and phosphate uptake kinetics of the harmful diatom *coscinodiscus walesii*, a causative organism in the bleaching of aquacultured *porphyra thalli*. *Harmful Algae.* 9(6):563–567.
- Pogoda CS, Keepers KG, Hamsher SE, et al. 2019. Comparative analysis of the mtDNAs of six newly sequenced diatoms reveals group II introns in the barcoding region of *cox1*. *Mitochondrial DNA Part A.* 30(1): 43–51.
- Rick HJ, Dürselen CD. 1995. Importance and abundance of the recently established species *Coscinodiscus walesii* Gran & Angst in the German Bight. *Helgol Mar Res.* 49(1):355–374.
- Sun J, Liu D. 2005. Net-phytoplankton community of the Bohai Sea in the autumn of 2000. *Acta Oceanolog Sin.* 27(3):124–132.
- Tetsuya N, Kazutaka M, Satoshi N. 2000. Effects of temperature and salinity on the growth of the giant diatom *Coscinodiscus walesii* isolated from Harima-Nada, Seto Inland Sea, Japan. *Nippon Suisan Gakkaishi.* 66(6):993–998.
- Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44(W1):W232–W235.
- Wang S, Fan X, Guan Z, Xu D, Zhang X, Wang D, Miao Y, Ye N. 2016. Sequencing of complete mitochondrial genome of *Saccharina latisima* ye-C14. *Mitochondrial DNA Part A.* 27(6):4037–4038.
- Wang Y, Chen Y, Wang J, Liu F, Chen N. 2021. Mitochondrial genome of the harmful algal bloom species *Odontella regia* (Mediophyceae, Bacillariophyta). *J Appl Phycol.* 33(2):855–868.
- Xie W. 2006. Community structure and dynamics of planktonic diatoms in typical areas of East China Sea. Xiamen University, Xiamen, China.
- Zhang M, Cui Z, Liu F, Chen N. 2021. Definition of a high-resolution molecular marker for tracking the genetic diversity of the harmful algal species *Eucampia zodiacus* through comparative analysis of mitochondrial genomes. *Front Microbiol.* 12..