Taylor & Francis
Taylor & Francis Group

MITOGENOME ANNOUNCEMENT

🔓 OPEN ACCESS   ✔ Check for updates

Complete chloroplast genome of *Eucampia zodiacus* (Mediophyceae, Bacillariophyta)

Mengjia Zhang[a,b,c,d], Zongmei Cui[a,b,c,d], Feng Liu[a,b,d] (ID) and Nansheng Chen[a,b,d,e] (ID)

[a]CAS Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China; [b]Laboratory of Marine Ecology and Environmental Science, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China; [c]University of Chinese Academy of Sciences, Beijing, China; [d]Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao, China; [e]Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, Canada

ABSTRACT

The cosmopolitan phytoplankton species *Eucampia zodiacus* Ehrenberg 1839 is a common harmful algal bloom (HAB) species with significant negative ecological impact. However, molecular information for this HAB species is limited. In this study, the complete chloroplast genome (cpDNA) of *E. zodiacus* was constructed for the first time. The circular genome was 118,107 bp in length, containing a pair of inverted repeats (IR) (6991 bp each). The overall GC content of *E. zodiacus* cpDNA was 31.64%. It encoded 169 genes, including 131 protein-coding genes (PCGs), 30 tRNA genes, one tmRNA gene, one ncRNA gene, and six rRNA genes in IR regions. Phylogenetic analysis using concatenated PCGs of 56 diatom cpDNAs strongly supported that *E. zodiacus* was closely related to *Cerataulina daemon*, which belongs to the same family Hemiaulaceae according to AlgaeBase. Syntenic analysis revealed nearly identical gene order between the two cpDNAs, except for an inversion in the small single-copy (SSC) region.

Eucampia zodiacus Ehrenberg 1839 is a common harmful algal bloom (HAB) species with a worldwide distribution except for the polar regions. It can be detected almost all-year round in the water column, providing considerable primary production (Horner 2002; Ito et al. 2013; Nishikawa et al. 2013). *Eucampia zodiacus* can form dense blooms in coastal waters, which have been observed in many countries including Japan, Canada and China (Huo et al. 2001; Zhang et al. 2002; Nishikawa et al. 2007; Martin et al. 2008; Nishikawa et al., 2011; Liang 2012; Matsubara 2012). *E. zodiacus* blooms have been reported to cause bleaching of aquacultured nori, damages to fisheries and economic losses (Martin et al. 2008; Nishikawa et al., 2011). Despite of its important role in environment and ecology, molecular information for this HAB species is limited. In recent study, we have constructed the first mitochondrial genome (mtDNA) of *E. zodiacus*, and through the comparative analysis, we defined a high-resolution molecular marker for tracking its genetic diversity (Zhang et al. 2021). In this project, we constructed the chloroplast genome (cpDNA) of *E. zodiacus* for the first time, and submitted the annotated genomic sequence to GenBank under accession number MW412838.
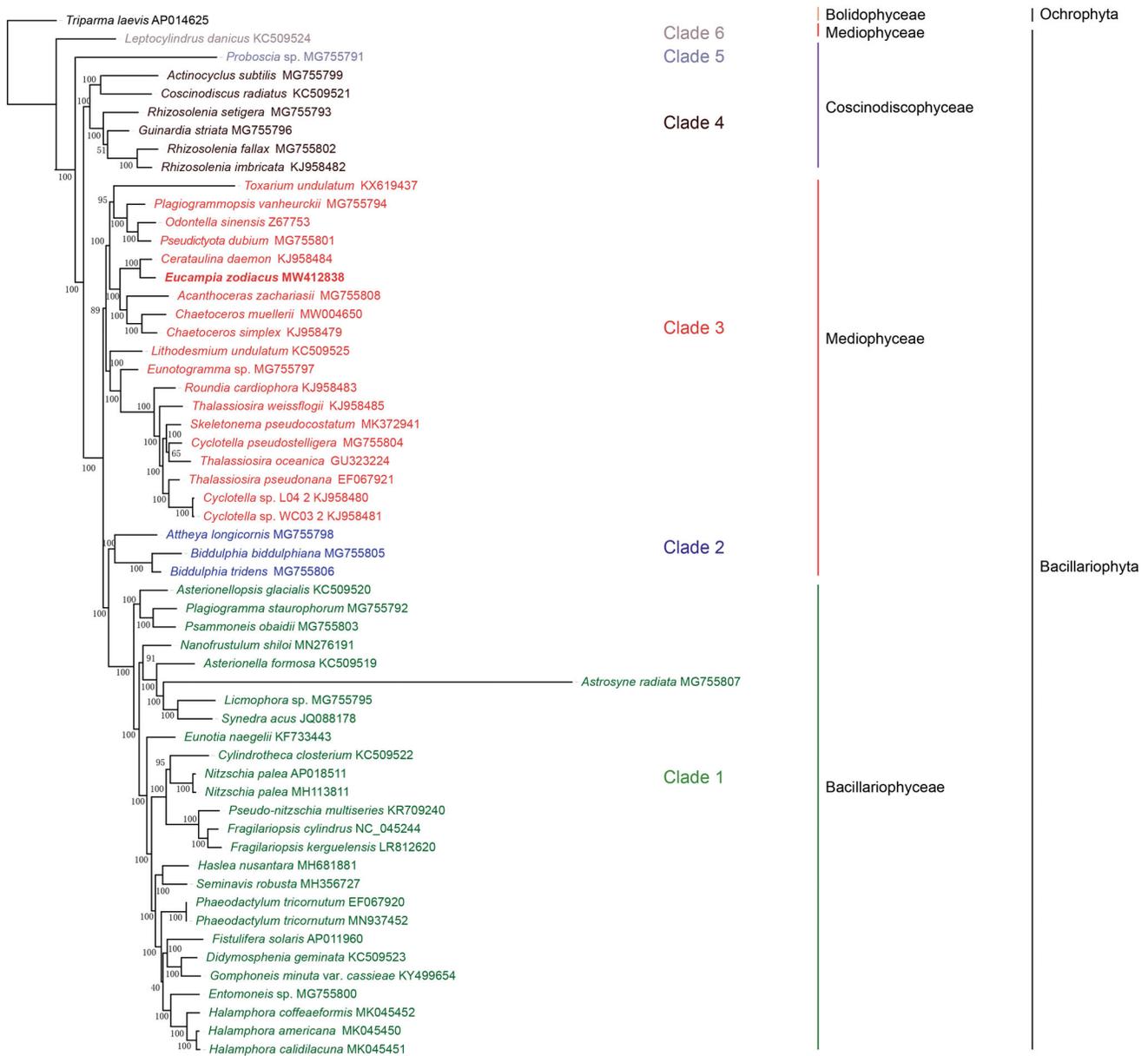
The *E. zodiacus* strains CNS00061 was isolated from seawater samples collected from the Changjiang Estuary (30.36°N, 122.86°E) (July, 2019). The specimen was deposited in the collection of marine algae in KLMEES of IOCAS

(Nansheng Chen, chenn@qdio.ac.cn) under the voucher number CNS00061. The strain was confirmed as *E. zodiacus* both through morphological (Guo 2004; Yang and Dong 2006) and molecular identification using five common molecular markers, including full-length 18S rDNA (Sorhannus 2007), 28S rDNA D1–D2 region (Hamsher et al. 2013), ITS (Guo et al. 2015), *COI* (Guo et al. 2015) and *rbcL* (Guo et al. 2015). Genomic DNA was extracted with TIANGEN DNAsecure Plant Kit (TIANGEN, DP121221) and sequenced on the Illumina Novaseq 6000 platform at Novogene (Beijing, China), generating 150 bp paired-end reads from libraries 350 bp in length.

The filtered reads were assembled into scaffolds with Platanus-allee (v2.2.2) (Kajitani et al. 2019) and SPAdes (v3.14.0) (Bankevich et al. 2012). With the cpDNA of *Cerataulina daemon* (KJ958484) (Sabir et al. 2014) used as the reference, scaffolds corresponding to cpDNA of *E. zodiacus* were screened using BLASTN. To fill the gap, Sanger sequencing was used to extend the ends of scaffolds and the gap was covered twice to ensure the accuracy. Assembly errors were corrected and N regions were replaced by using BWA (v0.7.17–r1188) (Li and Durbin 2009), results of which were extracted with SAMtools (v1.10) (Li et al. 2009) and viewed with IGV (v2.7.2) (Robinson et al. 2011). Protein-coding genes (PCGs) and open reading frames (*orfs*) were annotated using NCBI ORF Finder and BLAST similarity searches of the non-redundant databases at NCBI (Altschul et al. 1997). tRNAs

Figure 1. Phylogenetic tree based on ML analysis of aa sequence dataset of 95 cpDNA PCGs in Bacillariophyta. *Triparma laevis* (AP014625) (Bolidophyceae, Ochrophyta) was used as out-group taxa. Numbers on the branches represent bootstrap values.

were determined by reconstructing their cloverleaf structures using the tRNAscan-SE (v1.3.1) (Lowe and Chan 2016). rRNAs were identified using RNAmmer (v1.2) (Lagesen et al. 2007), Barrnap (v0.9) and MEGA (v7.0) for homologous comparison. The gene map of the circular cpDNA of *E. zodiacus* was generated with Organellar Genome DRAW (OGDraw) (Lohse et al. 2007).

The complete cpDNA of *E. zodiacus* (strain CNS00061) was a circular molecule that was 118,107 bp in size. The whole cpDNA contained a pair of inverted repeats (IR) regions of 6,991 bp in length. The IR regions divided the genome into two single-copy regions, small single copy (SSC) and large single copy (LSC) with 38,079 and 66,046 bp, respectively. The GC contents of the LSC, SSC, IR regions and the cpDNA as a whole, are 30.37%, 30.57%, 40.58%, and 31.64%, respectively. It possessed 169 genes, out of which, there were 131 PCGs, 30 tRNA genes, one tmRNA, one ncRNA, and six rRNA genes (three rRNA species). Among the 131 PCGs, 129 genes started with the canonical ATG start codons, *rpl23* with GTG and *rbcS* with ATA. No introns were found in the *E. zodiacus* cpDNA.

To confirm the phylogenetic position of *E. zodiacus* within the phylum Bacillariophyta, the amino acid (aa) sequence dataset of 95 concatenated PCGs of 56 diatom species was obtained to construct phylogenetic tree. *Triparma laevis* (AP014625) (Bolidophyceae, Ochrophyta) was used as out-group taxa. Maximum likelihood (ML) phylogenetic tree (Figure 1) constructed using IQ-TREE (v1.6.12) (Trifinopoulos et al. 2016) with 1000 bootstrap replicates demonstrated that the 56 species in Bacillariophyta formed six clades in three classes, Coscinodiscophyceae, Mediophyceae, and Bacillariophyceae. Bacillariophyceae ('Clade 1') was recovered as monophyletic, within which *Astrosyne radiata* was recovered on an extremely long branch. Mediophyceae, which

consisted of three clades ('Clade 2', 'Clade 3' and 'Clade 6'), was not monophyletic. Among these three clades, *Attheya longicornis* and two *Biddulphia* species formed Clade 2, and were sister to the Bacillariophyceae diatoms ('Clade 1') with 100% bootstrap support. Furthermore, the species *Leptocylindrus danicus*, which as annotated at AlgaeBase as a species in the class Mediophyceae, did not cluster together with other species of this class. Coscinodiscophyceae was not monophyletic, including two clades ('Clade 4' and 'Clade 5'). The species *Proboscia* sp. ('Clade 4'), which was annotated as a species in the class Coscinodiscophyceae, did not cluster together with the species of this class. This result was consistent to findings reported previously (Yu et al. 2018).

E. zodiacus was grouped together with *C. daemon*, which was consistent to the current classification in AlgaeBase. Furthermore, syntenic analysis between the cpDNAs of *E. zodiacus* and *C. daemon* revealed nearly identical gene order, except for an inversion in the SSC region, which is a well-accepted feature of cpDNAs (Doyle, 1999). The orientation of the SSC region may be different between these two species, due to evolutionary changes. However, we could not rule out the possibility that the two species have heteroplasmy (Walker et al. 2015), which remains to be confirmed by further analyses. Compared to *C. daemon*, cpDNA of *E. zodiacus* also lacked *syfB* gene, which is absent from all species in another order Thalassiosirales (Mediophyceae) (Sabir et al. 2014). The cpDNAs of more species in the order Hemiaulales are needed to track gene loss and clarify the evolutionary relationships between two orders.

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/nuccore/MW412838, under the accession no. MW412838. The associated BioProject, SRA and Bio-Sample numbers are PRJNA682714 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA682714), SRR13201560 (https://www.ncbi.nlm.nih.gov/sra/SRR13201560) and SAMN17005956 (https://www.ncbi.nlm.nih.gov/biosample/SAMN17005956), respectively.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25(17):3389–3402.
Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19(5):455–477.
Doyle JJ. 1999. Nuclear protein-coding genes in phylogeny reconstruction and homology assessment: some examples from Leguminosae. Syst Assoc. 57:229–254.
Guo H. 2004. Illustrations of planktons responsible for the blooms in Chinese coastal waters. Beijing: China Ocean Press.
Guo L, Sui Z, Zhang S, Ren Y, Liu Y. 2015. Comparison of potential diatom 'barcode' genes (the 18S rRNA gene and ITS, COI, rbcL) and their effectiveness in discriminating and determining species taxonomy in the Bacillariophyta. Int J Syst Evol Microbiol. 65(Pt 4):1369–1380.
Hamsher SE, LeGresley MM, Martin JL, Saunders GW. 2013. A comparison of morphological and molecular-based surveys to estimate the species richness of Chaetoceros and Thalassiosira (Bacillariophyta), in the Bay of Fundy. PLOS One. 8(10):e73521.
Horner R. 2002. A taxonomic guide to some common marine phytoplankton. London: Biopress; p. 195.
Huo W, Yu Z, Zou J, Han X. 2001. Analysis of dynamic process and the cause of Eucampia zoodiacus red tide in Jiaozhou Bay. J Fish China. 25(03):222–226.
Ito Y, Katano T, Fujii N, Koriyama M, Yoshino K, Hayami Y. 2013. Decreases in turbidity during neap tides initiate late winter blooms of Eucampia zodiacus in a macrotidal embayment. J Oceanogr. 69(4):467–479.
Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35(9):3100–3108.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 25(14):1754–1760.
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics. 25(16):2078–2079.
Liang Y. 2012. Investigation and evaluation of red Tide disaster in China（1933–2009). Beijing: China Ocean Press.
Lohse M, Drechsel O, Bock R. 2007. Organellar Genome DRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Curr Genet. 52(5–6):267–274.
Lowe TM, Chan PP. 2016. tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res. 44(W1):W54–W57.
Martin JL, Hastey CD, Legresley MM, Page FH, Canada O, Station B. 2008. Temporal and spatial characteristics of the diatom Eucampia zodiacus in the Western Isles region of the Bay of Fundy. Acta Biologica et Medica Germanica. 41:805–811.
Matsubara T. 2012. Current status and issues of the Ariake Bay, a potentially highly productive coastal sea. Oshima Y, editor. Kouseisha-kouseikaku; p. 9–24. Tokyo.
Nishikawa T, Hori Y, Harada K, Imai I. 2013. Annual regularity of reduction and restoration of cell size in the harmful diatom Eucampia zodiacus, and its application to the occurrence prediction of nori bleaching. Plankton Benthos Res. 8(4):166–170.

Nishikawa T, Hori Y, Nagai S, Miyahara K, Nakamura Y, Harada K, et al. (2011). Long time-series observations in population dynamics of the harmful diatom *Eucampia zodiacus* and environmental factors in Harima-Nada, eastern Seto Inland Sea, Japan during 1974–2008. Plankton Benthos Res. 6(1):26–34.

Nishikawa T, Hori Y, Tanida K, Imai I. 2007. Population dynamics of the harmful diatom *Eucampia zodiacus* Ehrenberg causing bleachings of Porphyra thalli in aquaculture in Harima-Nada, the Seto Inland Sea, Japan. Harmful Algae. 6(6):763–773.

Kajitani R, Yoshimura D, Okuno M, Minakuchi Y, Kagoshima H, Fujiyama A, Kubokawa K, Kohara Y, Toyoda A, Itoh T, et al. 2019. Platanus-allee is a *de novo* haplotype assembler enabling a comprehensive access to divergent heterozygous regions. Nat Commun. 10(1):1702.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. Nat Biotechnol. 29(1):24–26.

Sabir JSM, Yu M, Ashworth MP, Baeshen NA, Baeshen MN, Bahieldin A, Theriot EC, Jansen RK. 2014. Conserved gene order and expanded inverted repeats characterize plastid genomes of Thalassiosirales. PLoS One. 9(9):e107854.

Sorhannus U. 2007. A nuclear-encoded small-subunit ribosomal RNA timescale for diatom evolution. Mar Micropaleontol. 65(1–2):1–12.

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.

Walker JF, Jansen RK, Zanis MJ, Emery NC. 2015. Sources of inversion variation in the small single copy (SSC) region of chloroplast genomes. Am J Bot. 102(11):1751–1752.

Yang S, Dong S. 2006. Illustrations of common planktons of diatoms in Chinese waters. China: Ocean University Press.

Yu MJ, Ashworth MP, Hajrah NH, Khiyami MA, Sabir MJ, Alhebshi AM, et al. 2018. Evolution of the Plastid Genomes in Diatoms. Plastid Genome Evol. 85:129–155.

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.

Zhang Y, Wu Y, Zou J, Yu Z, Pu X. 2002. A red tide caused by diatom *Eucampia zoodiacus* in the Jiaozhou Bay. Oceanol Limnol Sin. 33(01):55–61.