



News & Views

Important role of Marine Group II *Euryarchaeota* in organic carbon cycling in the South China Sea [☆]

Ye Lu ¹, Qian Li ¹, Duo Zhao, Yao Zhang ^{*}

State Key Laboratory of Marine Environmental Science & College of Ocean and Earth Sciences, Xiamen University, Xiamen 361102, China

Planktonic archaea are among the most abundant members of marine microbes and play a crucial role in driving biogeochemical cycles. It is now recognized that there are four major phylogenetic groups of planktonic archaea (Marine Group I–IV, MGI–MIV), which include the chemoautotrophic ammonia-oxidizing *Thaumarchaeota* (MGI), as well as the heterotrophic *Euryarchaeota* (MGII and MGIII) and *Haloarchaea*-affiliated archaea (MGIV) [1]. The MGII *Euryarchaeota*, the most predominant planktonic heterotrophic archaea in ocean waters, exhibit a global distribution and can at times account for over 30% of the marine microbial community across at least seventeen candidate genera [1,2]. Previous studies have indicated that the distribution of MGII often coincides with that of phytoplankton because they tend to degrade phytoplankton-derived organic carbon, playing a significant role in the cycling of organic carbon in the ocean [3].

The South China Sea (SCS) is a typical oligotrophic semi-enclosed marginal sea in the Western Pacific. A widespread presence of MGII in the SCS is evident, with high abundance reported in the euphotic zone across different regions: the Pearl River Estuary (up to 10⁸ 16S rRNA gene copies/L [4]), continental shelf and slope (frequently over 40% of the archaeal community [5]), northern offshore and central basin (over 60% of the archaeal community [6,7]), and southwest offshore area (reaching approximately 10⁸ cells/L [8]). MGII *Euryarchaeota* in the SCS, similar to findings from studies in other marine regions worldwide, were phylogenetically classified into two main lineages, II.a and II.b, which have recently been placed within the new families *Ca. Poseidoniaceae* and *Ca. Thalassarchaeaceae*, respectively (Fig. 1a) [1,9]. MGII.a typically exhibits higher abundance than MGII.b in nearshore environments, such as the Pearl River Estuary, accounting for ~48%–99% of the MGII abundance [4]. However, MGII.a could be surpassed by MGII.b in the open ocean and is restricted to the photic zone [7,10]. In contrast, MGII.b shows a dominant abundance throughout the subsurface water column in the open ocean [7]. Recent studies have further distinguished the genus-level clades within both MGII.a and MGII.b in the SCS (Fig. 1a) [10,11]. Furthermore, multiple genus-level clades within the MGII.a and MGII.b lineages

also exhibit different niche preferences. For instance, in the Pearl River plume and nearshore niches of the SCS, clades K1, L1, and L2 of MGII.a (Fig. 1a), which represent globally widespread clades, are abundant [9–11]. However, the N1 and O3 clades of MGII.b, which exhibit a preference for tropical and subtropical oligotrophic waters [9], are abundant in the open basin of the SCS [10]. By analyzing the MGII 16S rRNA amplicon dataset from the SCS in Ref. [12], we further demonstrate distinct vertical profiles of the families *Poseidoniaceae* (MGII.a) and *Thalassarchaeaceae* (MGII.b) (Fig. 1b). Moreover, when comparing the predominant clades L2 (within MGII.a) with O3 and O4 (both within MGII.b), three distinct ecotypes emerge: L2 is dominant in surface waters but decreases with depth, while O3 is most abundant at the DCM layer, and O4 prevails at greater depths (Fig. 1b). It has been inferred that the niche partitioning between MGII.a and MGII.b may be attributed to their distinct adaptations to the temperature, salinity, nutrient, and light conditions [13]. In the Pearl River Estuary, MGII.a was found to be positively correlated with light intensity and nitrite concentration, whereas MGII.b was positively correlated with salinity [4]. In contrast, in the surface water of the open SCS, MGII.a L1 and MGII.b N1 showed a positive correlation with temperature and a moderate negative correlation with nitrate and phosphate concentrations [10]. In addition, both MGII.a and MGII.b exhibited multiple interactions with Cyanobacteria in the co-occurrence networks [4,10], suggesting the diverse physiological and ecological characteristics of MGII involved in utilizing phytoplankton-derived organic matter [3].

Recent studies have recovered metagenome-assembled genomes (MAGs) of MGII from metagenomes [2,9,14], revealing the metabolic pathways and networks of MGII *Euryarchaeota*, partly bridging the gap caused by the absence of any cultivated representatives. Near all of the MGII MAGs, including both MGII.a and MGII.b lineages, encode suites of enzymes involved in central carbon metabolisms, such as glycolysis, the citric acid cycle, and gluconeogenesis (Fig. 1c). Additionally, they have been identified with the capacity to degrade high-molecular-weight organic matter, including polysaccharides, proteins, and lipids, encoding carbohydrate-active enzymes (CAZymes), peptidases, and fatty-acid-degrading dehydrogenases [2,9,14]. Several of these MGII peptidases and CAZymes are extracellular enzymes, which can facilitate the degradation of organic matter in the surrounding environment (Fig. 1c). In addition, a variety of ATP-binding cassette transporters (ABC

[☆] SPECIAL TOPIC: The South China Sea Deep Program.

^{*} Corresponding author.

E-mail address: yaozhang@xmu.edu.cn (Y. Zhang).

¹ These authors contributed equally to this work.

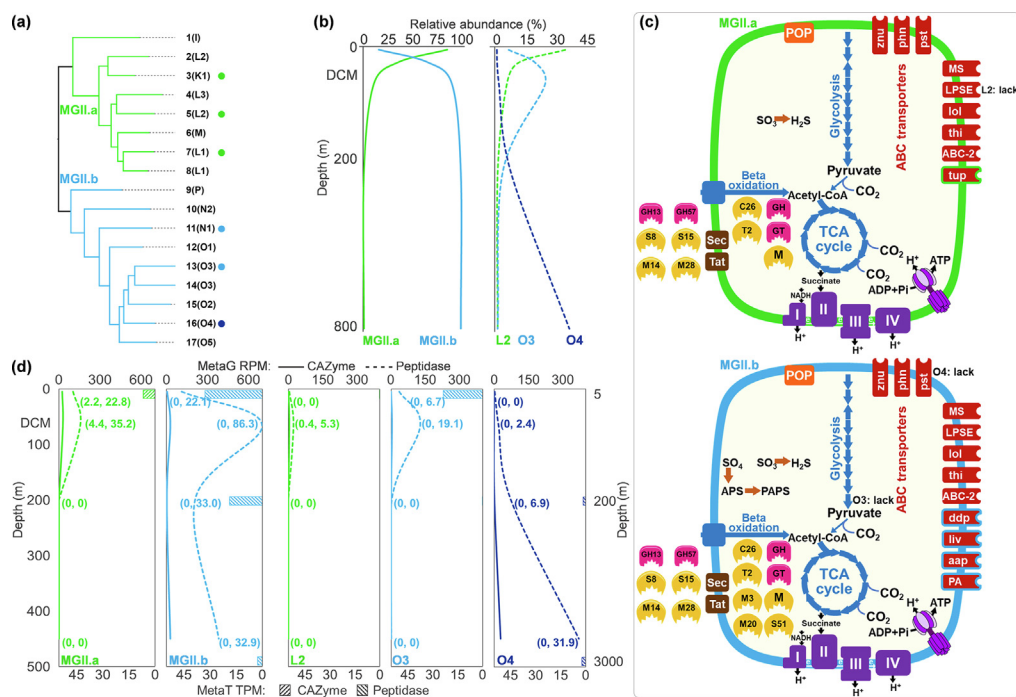


Fig. 1. A summary diagram illustrating the distribution and metabolic characteristics of MGII *Euryarchaeota*. (a) Phylogenetic tree of MGII MAGs constructed with FastTree based on the concatenated 53 archaeal marker genes derived from their GTDB-Tk outputs. These MAGs (completeness > 70%) were retrieved from two available SCS metagenome datasets (accession number PRJNA880762 and PRJNA521300) from the NCBI Sequence Read Archive. Dots indicate the abundant clades of MGII in the SCS. (b) Vertical distribution of the families Poseidoniaceae (MGII.a) and Thalassarchaeaceae (MGII.b), along with the genus-level clades L2, O3, and O4 in the SCS. (c) Metabolic characteristics of MGII.a and MGII.b, focusing on the degradation of high-molecular-weight organic matter as well as central carbon metabolism and energy metabolism, based on the contents of multiple MGII genomes from the literature [2,3,9]. ATP-binding cassette transporters (ABC transporters, red icons) include znu (zinc), pst (phosphate), phn (phosphonate), MS (multiple sugar), LPSE (lipopolysaccharide), lol (lipoprotein), thi (thiamine), tup (tungstate), ddp (peptide/nickel), liv (branched-chain amino acid), aap (general L-amino acid), and PA (polar amino acid) transporters. GH (in pink icons), carbohydrate-active enzyme families GH76, GH101, GH103, and GH130; GT (in pink icons), carbohydrate-active enzyme families GT2, GT4, GT29, and GT66. M (in yellow icons), peptidase families M14, M17, M23, M24, M32, M38, and M50; POP, proteorhodopsin. (d) Vertical distribution of the relative abundance of CAZyme and peptidase genes (left Y-axis and upper X-axis) and transcripts (right Y-axis and lower X-axis) in MGII.a and MGII.b, as well as the L2, O3, and O4 clades. The relative abundance of genes and transcripts was determined through the following procedure: (1) the predicted genes from the SCS metagenomes were mapped to annotated MGII MAGs; (2) the identified MGII CAZyme and peptidase genes were mapped to the SCS metagenomes and metatranscriptomes to calculate their coverages [12,15]; (3) the gene and transcript abundances were calculated using normalized read counts with the reads per kilobase million (RPM) and transcripts per kilobase million (TPM) methods, respectively. The two values in brackets indicate the relative abundance of extracellular CAZyme and peptidase genes, respectively. MetaG, metagenomes; MetaT, metatranscriptomes.

transporters) have been identified in the MGII MAGs, which are responsible for transporting multiple sugars, lipopolysaccharides, lipoproteins, phosphate, phosphonate, zinc, and thiamine into the cells for growth (Fig. 1c) [2,9]. These genomic characteristics demonstrate its heterotrophic lifestyle and ecological roles in the remineralization of organic matter.

Taking into account the niche partitioning of MGII.a and MGII.b and the distinct physicochemical properties of their preferred environments, we summarized the diverse heterotrophic strategies between the MGII lineages through comparative genomic analysis (Fig. 1c). MGII.a, represented by the L2 clade, is characterized by: (1) a higher proportion of total alpha-amylases in the CAZyme categories and extracellular alpha-amylases from the GH13 and GH57 families [2], (2) prolyl aminopeptidase (S33) and Ste24 peptidase (M48) [9], and (3) the tungstate ABC transporter along with a higher proportion of the thiamine (vitamin B1) ABC transporter [9] (Fig. 1c). Since starch (amylose and amylopectin) is a typical food storage polysaccharide in a variety of phytoplankton, which are more abundant in the coastal and shallow waters, MGII.a seems to dominate these niches by fully leveraging organic matter derived from phytoplankton, such as polysaccharides (starch), amino acids (like proline), and specific vitamins.

MGII.b is characterized by: (1) Pz-peptidase A (M3), carboxypeptidase Ss1 (M20), and dipeptidase E (S51), as well as (2) ABC transporters for branched-chain amino acids, polar amino acids, L-type amino acids, and peptide/nickel (Fig. 1c). By equipping these additional peptidases and ABC transporters, MGII.b has the capability

to utilize a wider range of peptides and amino acids compared to MGII.a. Therefore, MGII.b can be well adapted to waters with low irradiance or greater depths, where there is a higher proportion of available protein-associated organic matter, coupled with a decrease in high-molecular-weight, high-carbon compounds such as polysaccharides derived from phytoplankton [3,13]. Notably, among the genera of MGII.b, the O4 clade, as the exclusive dominant genus in the deep sea of the SCS (Fig. 1b), exhibits depth niche-related genomic characteristics distinct from those of the shallow-water genera like O3 (Fig. 1c) [9]. For instance, most MAGs of the O4 clade, in comparison to the O3 clade, lack genes encoding proteorhodopsin for harvesting light energy, as well as DNA photolyase and deoxyribodipyrimidine photolyase genes associated with DNA repair (Fig. 1c). Additionally, the O4 clade is characterized by the absence of the L-aspartate oxidase gene, the PrsW peptidase (M82) gene, the flagellin gene, and the genes encoding the phosphate ABC transporter [2,9]. The O4 clade seems to abandon the essential capabilities required for adapting to the photic environment and instead adopts a free-living lifestyle strategy. In summary, heterotrophic MGII *Euryarchaeota* are composed of at least three genus-level ecotypes in the ocean, represented by MGII.a L2 and MGII.b O3 and O4, each potentially carrying out distinct organic metabolisms and contributing unique roles in organic carbon cycling.

By analyzing the MGII CAZymes and peptidases in the SCS metagenome (0–450 m) and metatranscriptome (5, 200, and 3000 m) datasets from the literature [12,15], we observed distinct vertical patterns of organic matter degradation potentials exhib-

ited by MGII.a and MGII.b, as well as among the three typical genus-level ecotypes (Fig. 1d). The relative abundance of total and extracellular CAZyme genes was higher in MGII.a than in MGII.b in both the surface and deep chlorophyll maximum (DCM) layers. The relative abundance of total and extracellular peptidase genes was similar between MGII.a and MGII.b in the surface layer, but lower for MGII.a compared to MGII.b in the subsurface water (Fig. 1d). In the upper water, the main CAZymes of MGII.a were identified as belonging to the GT2 family, while the predominant extracellular type was alpha-amylase (GH13 family). The primary peptidases, both intracellular and extracellular, were identified as subtilisin (serine peptidase S8 family). The transcriptional activities of these enzymes in MGII.a were mainly attributed to glycosyltransferase (GT2 family of CAZymes) in the surface layer and less than 1% of the transcripts were assigned to the serine peptidase S8 family (Fig. 1d). The GT2 family is involved in polysaccharide metabolism, while the S8 family catalyzes the nonspecific hydrolytic cleavage of peptides [2] (according to information from the MEROPS database, https://www.ebi.ac.uk/merops/cgi-bin/family_index?type=P). These results suggest that MGII.a has a preference for labile dissolved organic carbon (DOC), including polysaccharides and proteins freshly derived from phytoplankton, in the upper water of the SCS.

MGII.b exhibited its highest relative abundance of CAZyme (no extracellular retrieved) and peptidase (total and extracellular) genes in the DCM layer in the vertical profile (Fig. 1d). The predominant CAZyme and peptidase families in MGII.b were found to be glycosyltransferase (GT4 family) and the S8 family, respectively. Furthermore, the primary extracellular peptidases of MGII.b were identified to be from the S8 family and carboxypeptidase (M14 family). The M14 family comprises a variety of carboxypeptidases capable of hydrolyzing the C-terminal amino acids of polypeptides based on the MEROPS database. In the metatranscriptomes, the peptidase S8 genes exhibited higher transcriptional activity compared to other types of peptidase genes in MGII.b, while no transcripts were retrieved for CAZyme genes. These characteristics indicate that MGII.b is versatile in handling complex proteins and peptides derived from phytoplankton, thus playing an important role in the degradation of proteins and peptides from the DCM layer down to the mesopelagic water of the SCS.

Among the three typical genus-level ecotypes, the surface L2 clade displayed a profile similar to that described for MGII.a above, while the DCM O3 clade and the deep O4 clade exhibited their highest relative abundances of CAZyme and peptidase genes in the DCM and dark ocean, respectively (Fig. 1d). In the metatranscriptomes, peptidase transcripts were most abundant for O3 in the upper ocean and for O4 in the deep ocean, with no transcripts for CAZyme genes (Fig. 1d). The predominant CAZyme family in O4 was the beta-oligosaccharyltransferase (GT66), in contrast to the glycosyltransferases (GT4) involved in polysaccharide metabolism found in the O3 in the DCM layer. The predominant peptidase family in O4 was the S8 family, but the predominant extracellular peptidases were identified as belonging to the M14 family (specifically hydrolyzing the C-terminal amino acids of polypeptides), rather than the S8 family (catalyzing the nonspecific hydrolytic cleavage of peptides) found in the O3 clade in the DCM layer. These characteristics suggest that O4 can specifically degrade extracellular proteins and peptides, thus making it better adapted to the deep ocean of the SCS, where most of the labile DOC has been depleted, leaving behind (semi-)recalcitrant DOC.

In conclusion, MGII *Euryarchaeota* are an important heterotrophic archaeal group with a wide distribution and transcriptional activity throughout the water column in the SCS. They are primarily composed of the families Poseidoniaceae (MGII.a) and Thalassarchaeaceae (MGII.b), within which three typical genus-level ecotypes can be identified. These ecotypes are equipped with

depth-related metabolic genes, actively participating in the degradation of organic carbon throughout the water column. The majority of MGII.a is primarily distributed in surface habitats and is capable of promoting the degradation of phytoplankton-released carbohydrates, such as polysaccharides. Among these, the L2 clade is a typical (dominant) genus-level surface ecotype in the SCS. MGII.b contributes to the degradation of organic matter, primarily proteins, in the subsurface water column, with the majority being the shallow clades. Among these, the O3 and O4 clades are typical DCM and deep-sea ecotypes, respectively, in the SCS. Notably, the O4 clade occupying a unique deep-sea niche plays a significant role in the specific degradation of proteins and oligopeptides in the deep ocean. These potential distinct preferences of genus-level ecotypes of MGII for different substrates along the depth profile suggest that niche partitioning among them may contribute to the depth-related transition from labile to recalcitrant DOC. Future research on MGII *Euryarchaeota* should focus on uncovering their contributions to the deep-sea recalcitrant DOC pool and their impact on the global cycling of organic carbon through the integration of (meta)genomics, metatranscriptomics, metaproteomics, and experiments involving isotopic labeling substrates.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China (42125603 and 92251303) and the Global Ocean Negative Carbon Emissions (Global ONCE) Program.

References

- [1] Santoro AE, Richter RA, Dupont CL. Planktonic marine archaea. *Ann Rev Mar Sci* 2019;11:131–58.
- [2] Tully BJ. Metabolic diversity within the globally abundant Marine Group II euryarchaea offers insight into ecological patterns. *Nat Commun* 2019;10:271.
- [3] Orellana LH, Francis TB, Krüger K, et al. Niche differentiation among annually recurrent coastal Marine Group II Euryarchaeota. *ISME J* 2019;13:3024–36.
- [4] Xie W, Luo H, Murugapiran SK, et al. Localized high abundance of Marine Group II archaea in the subtropical Pearl River Estuary: implications for their niche adaptation. *Environ Microbiol* 2018;20:734–54.
- [5] Zhang Y, Lu Y, Wang J, et al. Diurnal variations of the microbial community in mesopelagic fish habitats of the northern slope of the South China Sea. *Deep Sea Res PT* 2019;167:55–61.
- [6] Li J, Gu L, Bai S, et al. Characterization of particle-associated and free-living bacterial and archaeal communities along the water columns of the South China Sea. *Biogeosciences* 2021;18:113–33.
- [7] Liu H, Zhang CL, Yang C, et al. Marine Group II dominates planktonic archaea in water column of the northeastern South China Sea. *Front Microbiol* 2017;8:1098.
- [8] Zhang Y, Sintez E, Chen J, et al. Role of mesoscale cyclonic eddies in the distribution and activity of *Archaea* and *Bacteria* in the South China Sea. *Aquat Microb Ecol* 2009;56:65–79.
- [9] Rinke C, Rubino F, Messer LF, et al. A phylogenomic and ecological analysis of the globally abundant Marine Group II archaea (*Ca. Poseidoniales* ord. nov.). *ISME J* 2018;13:663–75.
- [10] Chen S, Tao J, Chen Y, et al. Interactions between Marine Group II archaea and phytoplankton revealed by population correlations in the northern coast of South China Sea. *Front Microbiol* 2022;12:785532.
- [11] Fan L, Xu B, Chen S, et al. Gene inversion led to the emergence of brackish archaeal heterotrophs in the aftermath of the Cryogenian Snowball Earth. *PNAS Nexus* 2024;3:057.
- [12] Lu Y, Lv Y, Zhang Y, et al. Metatranscriptomes reveal the diverse responses of *Thaumarchaeota* ecotypes to environmental variations in the northern slope of the South China Sea. *Environ Microbiol* 2023;25:410–27.
- [13] Zhang CL, Xie W, Martin-Cuadrado AB, et al. Marine Group II Archaea, potentially important players in the global ocean carbon cycle. *Front Microbiol* 2015;6:160251.
- [14] Iverson V, Morris RM, Frazar CD, et al. Untangling genomes from metagenomes: revealing an uncultured class of marine Euryarchaeota. *Science* 2012;335:587–90.
- [15] Zhang Y, Qin W, Hou L, et al. Nitrifier adaptation to low energy flux controls inventory of reduced nitrogen in the dark ocean. *Proc Natl Acad Sci USA* 2020;117:4823–30.