Heimdallarchaeota harness light energy through photosynthesis

Rui Liu[1,2,4], Ruining Cai[1,2,3,4], Jing Zhang[1,2,3,4], Chaomin Sun[1,2,4*]

[1]Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China

[2]Laboratory for Marine Biology and Biotechnology, Pilot National Laboratory for Marine Science and Technology, Qingdao, China

[3]College of earth science, University of Chinese Academy of Sciences, Beijing, China

[4]Center of Ocean Mega-Science, Chinese Academy of Sciences, Qingdao, China

[*] Corresponding author

Chaomin Sun Tel.: +86 532 82898857; fax: +86 532 82898857.

E-mail address: sunchaomin@qdio.ac.cn

Keywords: Heimdallarchaeota, photosynthesis, deep-sea, metagenomics, ancestor

Abstract

Photosynthesis is an ancient process that originated after the origin of life, and has only been found in the Bacterial and Eukaryotic kingdoms, but has never been reported in any member of the domain Archaea. Heimdallarchaeota, a member of Asgard archaea, are supposed as the most probable candidates (to date) for the archaeal protoeukaryote ancestor and might exist in light-exposed habitats during their evolutionary history. Here we describe the discovery that Heimdallarchaeota genomes are enriched for proteins formerly considered specific to photosynthetic apparatus and are suggestive performing oxygenic photosynthesis. Our results provide strong support for hypotheses in which Heimdallarchaeota harvest light by bacteriochlorophyll and/or carotenoid, then transport electron from photosystems to Calvin-Benson-Bassham cycle along with CO_2 fixation and ATP biosynthesis, and release oxygen as a waste product. Given the possessing of phototrophic lifestyle together with other anaerobic and aerobic metabolic pathways, Heimdallarchaeota are firmly believed to be photomixotrophic and have a facultative aerobic metabolism. Our results expand our knowledge that archaea have played an important role in the molecular evolution of eukaryotic photosynthesis and raise the significant possibility that Heimdallarchaeota might be ancestor of eukaryotic photosynthetic organisms.

Introduction

Photosynthesis, the foundation for life, result in an enormous increase in biomass production on Earth and produce oxidized compounds serving as electron acceptors for heterotrophic metabolism[1]. In microbiology, organisms that perform photosynthesis include Cyanobacteria, Chloroflexi, Firmicutes, Chlorobi, Proteobacteria, Acidobacteria and Gemmatimonadetes[2]. Among them, only Cyanobacteria carry out oxygenic photosynthesis using electrons originating from water to generate oxygen as product, and evolve in ancestral near the time to the rise of oxygen and eventually result the Great Oxidation Event on Earth[3]. Till date, microbial photosynthesis has been only found in bacteria but not in any reported archaea except a functional bacteriochlorophyll synthase identified in an uncultivated Crenarchaeota, which provides the only clue of photosynthesis existing in Archaea[4].

Breakthroughs in environmental and metagenomic sequencing technologies are rapidly transforming the landscape for microbial evolution, especially the discovery of Asgard phylum archaea and their supposed position at the base of the eukaryotic tree of life[5]. It stands to reason that many phototrophs remain to be discovered, thus we want to ask if these metagenomic efforts help us to uncover phototrophic Archaea. Heimdallarchaeota are a member of Asgard superphylum archaea and currently represent the predicted closest archaeal relative of eukaryotes, and might exist in light-exposed habitats in their evolution history[5-8]. However, due to far less assembled genomes of Heimdallarchaeota than other Asgard members, no more light-dependent lifestyle details of this uncultured archaea have been disclosed. In the present study, nine high-quality assembled genomes of Heimdallarchaeota were obtained, and specific analyses of their reconstructed genomes provide solid proof that Heimdallarchaeota could utilize light energy through bacteriochlorophyll and/or carotenoid-based oxygenic photosynthesis. Given the closest eocytic lineage to eukaryotes, Heimdallarchaeota are proposed to be ancestor of eukaryotic photosynthetic organisms.

Results and Discussion

Photosynthetic apparatus in Heimdallarchaeota

To gain insight into the light utilizing characteristics in Heimdallarchaeota, we sampled aquatic sediments from a typical cold seep in South China Sea and two hydrothermal vents in Western pacific (Expanded Data Fig. 1 and Extended Data Table 1) with distinct chemical parameters (Expanded Data Fig. 2). Total DNA was extracted from all samples and sequenced, and nine high quality (>50% completeness, <10% contamination) metagenome-assembled genomes (MAGs) of Heimdallarchaeota were obtained by utilizing a hybrid binning strategy and performing manual inspection and data curation (Extended Data Table 2). The maximum-likelihood phylogenetic tree, based on concatenation of 37 marker genes including 13 small subunit (SSU) and 16 large subunit (LSU) ribosomal RNA genes, showed that both our and published Heimdallarchaeota MAGs clustered with other Asgard superphylum members and displayed a much closer evolutionary linkage with eukaryotes than other archaeal superphyla including DPANN, TACK and Euryarchaeota (Extended Data Fig. 3 and Supplementary Table 1). In accordance with other Asgard members, credible eukaryote-specific proteins (ESPs) were identified in our Heimdallarchaeota MAGs (Extended Data Fig. 4, Supplementary Table 2), which confirms Heimdallarchaeota as the current best candidate for the closest archaeal relatives of the eukaryotic nuclear lineage as described previously[5,6]. Different with previous report about the existence of rhodopsins in Heimdallarchaeota[1], there is no any rhodopsin homologs identified in the present nine Heimdallarchaeota MAGs. However, many typical chloroplastic proteins (including protochlorophyllide reductase, chlorophyll(ide) b reductases NOL/NYC1, NAD(P)H quinone oxidoreductase, photosystem I assembly proteins Ycf3 and phycocyanobilin lyase) were surprisingly identified in our Heimdallarchaeota MAGs (Extended Data Fig. 4 and Supplementary Table 2), indicating that Heimdallarchaeota might be a kind of unprecedented photosynthetic organism.

To carefully determine the photosynthetic position of Heimdallarchaeota, we performed various in-depth photosynthetic analyses based on our and published Heimdallarchaeota MAGs. It is well known that for the energy of sunlight to be converted and stored into biological systems, it must first be captured by the pigments present in the photosynthetic organisms. All photosynthetic organisms synthesize two

types of pigments: (a) bacteriochlorophylls and/or chlorophylls, which function in both light harvesting and photochemistry, and (b) carotenoids, which primarily act as photoprotective pigments but can also function in light harvesting[9]. Notably, almost all necessary bacteriochlorophyll synthesis components widely distribute in Heimdallarchaeota MAGs, which provides sufficient evidence that Heimdallarchaeota could synthesize bacteriochlorophyll (Fig. 1a, Supplementary Table 3). Among the key enzymes synthesizing bacteriochlorophyll, protochlorophyllide reductase (Por) could catalyze the reaction of transition between divinyl protochlorophyllide and divinyl chlorophyllide a[10]. By the phylogenetic analysis, most of Por homologs in our Heimdallarchaeota MAGs clustered in a single clade at the root of the tree (Fig. 1b, Supplementary Table 3). However, one Por homolog in H2.bin.2 was found to locate in a sister clade with typical photosynthetic organisms including Cyanobacteria, Algae and Streptophytina (Plants), and they further clustered with phototrophic bacteria containing Chlorobi, Chloroflexi and Proteobacteria (Rhodospirillales and Chromatiales), revealing the potential Heimdallarchaeota-phototrophs affiliation. The bacteriochlorophyll synthase (BCS) is capable of synthesizing bacteriochlorophyll a by esterification of bacteriochlorophyllide with phytyl diphosphate or geranylgeranyl diphosphate[4], and it is also annotated as digeranylgeranylglyceryl phosphate synthase (DGPS). We further phylogenetically analyzed the evolutionary relationship between BCS and DGPS, and the results showed that all archaeal DGPS located at the root, which separating from the clade containing BCS from phototrophic bacteria and chlorophyll synthase from photosynthetic organisms (Fig. 1c, Supplementary Table 3). And homologous proteins of BCS in our Heimdallarchaeota MAGs clustered a clade with the DGPS in Heimdallarchaeota LC2, which located at a position between DGPS and BCS branches (Fig. 1c, Supplementary Table 3). Furthermore, a previous reported functional bacteriochlorophyll synthase derived from the uncultured Crenarchaeota[4] was found to cluster a branch with the DGPS from Heimdallarchaeota LC3[5], and this cluster displayed a close evolutionary relationship with those photosynthetic bacteriochlorophyll and chlorophyll synthase branches (Fig. 1c). Although the function of these putative bacteriochlorophyll associated proteins remains to be elucidated, it is tempting to speculate that Heimdallarchaeota have a bacteriochlorophyll producing capability. It is known that bacteriochlorophylls or chlorophylls exist in only photosynthetic organisms[3,11]. The evidence of bacteriochlorophyll production in both Heimdallarchaeota and Crenarchaeota

suggests that archaea should have played an important role in the molecular evolution of bacteriochlorophyll synthase, and raises the significant possibility that the origin of photosynthesis probably predates the divergence of bacteria and archaea.

In addition bacteriochlorophyll, other photosynthetic pigments including carotenoid[9] and bacteriophytochrome[12] were also identified to be synthesized in Heimdallarchaeota. Particularly for carotenoid, it is a kind of ubiquitous and essential pigment in photosynthesis[9], and functions as accessory light-harvesting pigment and transfers the absorbed energy to bacteriochlorophylls, which expands the wavelength range of light that is able to drive photosynthesis[13,14]. We reconstructed the complete synthesis pathway of lycopene[15], a biologically important carotenoid, derived from acetyl-CoA according to the Heimdallarchaeota MAGs (Fig. 1d, Supplementary Table 3). It has been mentioned that chlorophylls in photosynthesis ineffectively absorbed much light in the 450-550 nm (blue-green light) region of the solar radiation spectrum, while the light within this range can be effectively absorbed by carotenoids[9]. Moreover, carotenoids protect the organisms from photodamage by quenching both singlet or triplet states of bacteriochlorophylls under strong illumination and function as photosynthetic membrane stabilizers in chloroplasts[9]. Therefore, the biosynthesis of carotenoid in Heimdallarchaeota could coordinate with bacteriochlorophyll for high-efficiency photosynthesis, which provides an opportunity for competitive advantage in any particular habitat.

Besides photosynthetic pigments, many key factors related to reaction centers (including photosystem I (PS I) assembly proteins BtpA and Ycf3[16,17]; PSI subunit VII PsaC[18]; photosystem II (PS II) stability/assembly factor related protein Ycf48[19]; antenna proteins in PSII (phycocyanobilin lyase, CpcE and bilin biosynthesis protein, CpeU)[20,21]), carbon fixation system and ATP synthase have also been identified in Heimdallarchaeota MAGs (Extended Data Fig. 5, Supplementary Table 4). Taken together, the ubiquitous identification of photosynthetic apparatus and existence of both PS I and PS II in the genomes strongly suggest that Heimdallarchaeota are potential oxygenic photosynthetic organism[9].

Electron transfer and energy production in Heimdallarchaeota

Several lines of evidence support that Heimdallarchaeota perform photosynthesis based on our study. Logically, in the primary steps of photosynthesis, solar photons are absorbed by special membrane-associated pigment-protein complexes

(light-harvesting antennas) and the electronic excitations are efficiently transferred to a reaction center[3,22]. The oxygen-evolving photosynthetic organisms have two photochemical reaction center complexes, PS I and PS II, that work together in a noncyclic electron transfer chain[22]. As a crucial electron carrier in PS I, phylloquinone was believed to be synthesized in Heimdallarchaeota (Fig. 2a, Supplementary Table 4). It is known that phylloquinone and its related compound menaquinone shared the same synthesis route for producing demethylphylloquinol[23], which is further catalyzed to phylloquinone or menaquinone by a key enzyme named demethylphylloquinol methyltransferase (MenG) or demethylmenaquinone methyltransferase (UbiE)[23]. MenG is responsible for phylloquinone (vitamin K1) synthesis in both plants and Cyanobacteria, and UbiE is only in charge of menaquinone (vitamin K2) synthesis in microorganism[23]. Interestingly, both MenG and UbiE homologs were identified in Heimdallarchaeota MAGs (Supplementary Table 4). Consistently, in the phylogenetic tree, MenGs from photosynthetic organisms were clustered in a single branch which made a long distance from UbiE branches. While some MenG and UbiE homologs from plants, archaea together with Heimdallarchaeota were located exactly between these two branches (Fig. 2b, Supplementary Table 4). In addition, all enzymes catalyzing reactions to synthesize plastoquinol, which is further oxidized to plastoquinone by oxygen to phylloquinone (a key electron transporter in PS II), could also be identified in Heimdallarchaeota MAGs (Fig. 2c, Supplementary Table 4). Collectively, Heimdallarchaeota are believed to possess both PS I and PS II. Furthermore, subsequent electron acceptors including ferredoxin (Fd) and ferredoxin $NADP^+$ reductase (FNR)[3] were also found in Heimdallarchaeota MAGs (Fig. 3, Supplementary Table 4), which confers Heimdallarchaeota oxygenic photosynthetic capability through two photosystems as that in Cyanobacteria.

Overall, we proposed a complete oxygenic photosynthetic pathway existing in Heimdallarchaeota (Fig. 3). Firstly, PS II absorbs a photon of light to generate a redox-potential by antenna proteins and reduces plastoquinone as the terminal electron acceptor[3,22]. The electrons extracted from water are further transported via a quinone and the cytochrome c complex to PS I with the presence of complex III[24] or alternative complex III (ACIII)[25] or other homologs existing in Heimdallarchaeota. Meanwhile, electrons are removed from water by PS II, oxidizing it to molecular oxygen, which is released as a waste product. Then electrons are carried by phylloquinone and transferred to Fe_4-S_4 cluster to generate NADH with electrons from

ferredoxin in PSI[3]. After the electron transfer, NAD(P)H is reduced with the electron delivering from ferredoxin[3,22], which further participates in Calvin-Benson-Bassham cycle for carbon fixation and finally to synthesize ATP[3,22].

Heimdallarchaeota are photomixotrophic

Next, to gain further insight into the lifestyle of Heimdallarchaeota, we reconstructed a complete metabolic pattern according to our and published Heimdallarchaeota MAGs. Similar with other Asgard superphylum[26], Heimdallarchaeota possess a mixotrophic lifestyle, which can simultaneously utilize the reverse tricarboxylic acid cycle (rTCA) for autotrophic CO_2 assimilation[5,6] and transport the exogenous organic matter through the metabolic circuitry for their catabolism (Fig. 4, Supplementary Table 5)[6]. For autotrophic metabolism, the Calvin-Benson-Bassham cycle was found to participate in carbon fixation through the RuBisCo and act as an intermediary between photosystems and ATP synthase for energy producing in Heimdallarchaeota (Fig. 4). Furthermore, a variety of polysaccharide-degrading enzymes, including chitinase, xylan/chitin deacetylase, diacetylchitobiose deacetylase and cellulase, were also found in Heimdallarchaeota by CAZy analysis (Extended Date Fig. 6 and Supplementary Table 5). These enzymes degraded polysaccharides like chitin, xylan and cellulose to produce oligosaccharides or monosaccharides for energy metabolism in Heimdallarchaeota. Moreover, methane metabolism pathway and Wood-Ljungdahl pathway were believed to exist in Heimdallarchaeota MAGs (Fig. 4).

Notably, aerobic metabolic pathways were also found to be ubiquitous in Heimdallarchaeota[6,27], which completely differed from metabolic characteristics of anaerobic Lokiarchaeota and Thorarchaeota[5,6,26,27]. Consistent with the results of previous studies[6], the complete tricarboxylic-acid cycle (TCA) and oxidative phosphorylation pathway were found to support the aerobic respiration in Heimdallarchaeota. And both the aerobic kynurenine pathway and aspartate pathway for NAD^+ *de novo* synthesis were also reconstructed in the present and previous published Heimdallarchaeota[6,28]. Particularly, the aerobic kynurenine pathway of NAD^+ biosynthesis was exclusively found in Heimdallarchaeota compared with other archaea, which is only present in eukaryotes and very few bacterial groups[28]. This pathway has been considered to originate from the protoeukaryote ancestor in oxygen-containing niche, which might be acquired through horizontal gene transfer in Heimdallarchaeota[6]. In addition, other molecules in oxygen-dependent metabolism,

such as aerotaxis receptor and bacterioferritin, were also identified in Heimdallarchaeota MAGs. All the above evidence indicates that oxygen must be present in the environment where Heimdallarchaeota inhabit. But this seems to contradict the strict anaerobic lifestyle existing in other Asgard archaea like Thorarchaeota and Lokiarchaeota in the same niche[5,6]. Therefore, we infer a more possible hypothesis that the oxygen presenting intracellular environment of Heimdallarchaeota might be generated by their performing photosynthesis. Consistently, Heimdallarchaeota have already evolved protection mechanisms against the formation of reactive oxygen species (ROS) via superoxide dismutase, catalase, and carotenoids (Fig. 4). Together, the metabolic reconstructions indicate that Heimdallarchaeota are photomixotrophic and have a facultative aerobic metabolism. If this is the fact, what is the ecological function or benefit for having the ability of photosynthesis in Heimdallarhaeota, which reside predominantly in marine sediments[5]? Recent study about rhodopsins identified in Heimdallarchaeota provides evidence of their light-exposed habitats, where Heimdallarchaeota could obtain enough energy from sunlight through photosynthesis[6]. The recovery of Heimdallarchaeota from deeper environments may be due to the high deposition rates characteristic for the sampling locations[6]. Nevertheless, there are plenty evidence showing that blue-green light with a wavelength range of 450-550 nm might exist in cold seep environment (~1100 m deep), and both long wavelength light (>650 nm) and (short wavelength light (<650 nm) could be detected in hydrothermal vents[29,30], which provides necessary condition for photosynthetic process. The Heimdallarchaeota capable of detecting geothermal light and phototaxis could preferentially occupy an optimum habitat, which confer them evolutionary advantages in the competition for nutrient resources. Accordingly, a kind of cyanobacterium collects light and passes excitation energy uphill to the photochemically active pigments through longer-wavelength chlorophyll f, which facilitates this bacterium to survive in dark condition[31]. Thus, we suppose that possessing a presumptive photomixotrophic lifestyle may give Heimdallarchaeota more flexibility to survive or adapt to the deep-sea harsh conditions.

A potential ancestor of eukaryotic photosynthetic organisms

Notably, given the close match of the emission spectra of geothermal light and the absorption spectra of bacteriochlorophylls *a* and *b*, photosynthesis was proposed to

arise from bacteriochlorophyll *a*- or *b*-containing organisms near oceanic hydrothermal vents where weak infrared radiation could be detected[32]. In the evolution history of photosynthesis, the organisms were thought to initially use the bacteriochlorophyll pigments to sense infrared light, and started making use of the near-infrared part of sunlight when moving to shallow water through further adaptation of a primitive photosystem, and chlorophylls would be eventually developed to make use of higher energy (visible) light to split water[32,33]. Meanwhile, during the evolution process of oxygenic photosynthesis, a bioinorganic water-oxidizing complex (WOC) was thought to serves as a redox capacitor to accomplish the oxidation of two water molecules to produce O_2 in PS II, which is comprised with a Mn_3CaO_4 distorted cubane structure bound to a fourth Mn by oxo-bridges[3,34]. And photoassembly of the WOC requires only Mn^{2+} and light to form the high-valent WOC[3]. Interestingly, an extremely high concentration of manganese element was detected in our sampling hydrothermal vent (Extended Data Fig. 2), which might promote WOC formation and play a key role in the acquiring and developing of photosynthesis in Heimdallarchaeota. Consistently, Heimdallarchaeota MAGs H2.bin.2 and H2.bin.81 derived from hydrothermal vent possess much more photosynthetic characteristics than those from cold seep in our present work. Therefore, in combination with the facts that Heimdallarchaeota are the most probable candidates for the archaeal protoeukaryote ancestor in previous report[6] and the first identified photosynthetic archaea in this study, we propose that Heimdallarchaeota might be ancestor of eukaryotic photosynthetic organism.

The presence of oxygen-dependent pathways in Heimdallarchaea raises the possibility that the archaeal eukaryotic phtosynthetic ancestor could have also been a facultative aerobe, and the archaeal-photobacterial endosymbiosist gave birth to the eukaryotic phtosynthetic ancestor took place after the Great Oxidation Event[35]. Horizontal gene transfer is postulated to play a major role in the evolution of microbial phototrophs and that many of the essential components of photosynthesis have conducted horizontal gene transfer[22]. And Heimdallarchaea might obtain the photosynthetic apparatus through lateral gene transfer from their cyanobacterial endosymbionts[6,28] because most of key factors associated with photosynthesis in Heimdallarchaea have high similarity with those from cyanobacteria (Supplementary Table 5). Our prediction is in agreement with the recent hypothesis that both the archaeal and bacterial eukaryotic ancestors have an oxygen-dependent metabolism[6],

in which the primordial function of the bacterial counterpart performing oxidative phosphorylation would not be detrimental to the existence of the archaeal host who exposed in oxygen environment. Even though the present results provide an updated perspective on the photosynthetic lifestyle of Heimdallarchaeota, further studies will be needed to elucidate the light utilizing strategies and evolutionary histories with their enrichment or even pure culture.

References

1. Martin, W. F., Bryant, D. A. & Beatty, J. T. A physiological perspective on the origin and evolution of photosynthesis. *FEMS Microbiol. Rev.* **42**, 205-231, doi:10.1093/femsre/fux056 (2018).
2. Cardona, T. Reconstructing the origin of oxygenic photosynthesis: do assembly and photoactivation recapitulate evolution? *Front. Plant Sci.* **7**, doi:10.3389/Fpls.2016.00257 (2016).
3. Fischer, W. W., Hemp, J. & Johnson, J. E. Evolution of oxygenic photosynthesis. *Annu. Rev. Earth Pl. Sc.* **44**, 647-683, doi:10.1146/annurev-earth-060313-054810 (2016).
4. Meng, J. *et al.* An uncultivated crenarchaeota contains functional bacteriochlorophyll a synthase. *ISME J.* **3**, 106-116, doi:10.1038/ismej.2008.85 (2009).
5. Zaremba-Niedzwiedzka, K. *et al.* Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541**, 353-+, doi:10.1038/nature21031 (2017).
6. Bulzu, P. A. *et al.* Casting light on Asgardarchaeota metabolism in a sunlit microoxic niche. *Nat. Microbiol.* **4**, 1129-1137, doi:10.1038/s41564-019-0404-y (2019).
7. Beja, O. *et al.* Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea. *Science* **289**, 1902-1906, doi:10.1126/science.289.5486.1902 (2000).
8. Spang, A. *et al.* Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**, 173-+, doi:10.1038/nature14447 (2015).
9. Hashimoto, H., Uragami, C. & Cogdell, R. J. Carotenoids and photosynthesis. *Subcell. Biochem.* **79**, 111-139, doi:10.1007/978-3-319-39126-7_4 (2016).
10. Chew, A. G. & Bryant, D. A. Chlorophyll biosynthesis in bacteria: the origins of structural and functional diversity. *Annu. Rev. Microbiol.* **61**, 113-129, doi:10.1146/annurev.micro.61.080706.093242 (2007).
11. Bryant, D. A. & Frigaard, N. U. Prokaryotic photosynthesis and phototrophy illuminated. *Trends Microbiol.* **14**, 488-496, doi:10.1016/j.tim.2006.09.001 (2006).

12. Moran, A. M. Elucidation of primary events in bacteriophytochrome photoreceptors. *Biophys J.* **111**, 2075-2076, doi:10.1016/j.bpj.2016.10.018 (2016).
13. Kirilovsky, D. Photosynthesis: Dissipating energy by carotenoids. *Nat. Chem. Biol.* **11**, 242-243, doi:10.1038/nchembio.1771 (2015).
14. Papagiannakis, E., Kennis, J. T. M., van Stokkum, I. H. M., Cogdell, R. J. & van Grondelle, R. An alternative carotenoid-to-bacteriochlorophyll energy transfer pathway in photosynthetic light harvesting. *Proc. Natl. Acad. Sci. U S A* **99**, 6017-6022, doi:10.1073/pnas.092626599 (2002).
15. Hernandez-Almanza, A. *et al.* Lycopene: progress in microbial production. *Trends Food Sci. Tech.* **56**, 142-148, doi:10.1016/j.tifs.2016.08.013 (2016).
16. Zak, E. & Pakrasi, H. B. The BtpA protein stabilizes the reaction center proteins of photosystem I in the cyanobacterium *Synechocystis* sp. PCC 6803 at low temperature. *Plant Physiol.* **123**, 215-222, doi:10.1104/pp.123.1.215 (2000).
17. Nellaepalli, S., Ozawa, S. I., Kuroda, H. & Takahashi, Y. The photosystem I assembly apparatus consisting of Ycf3-Y3IP1 and Ycf4 modules. *Nat. Commun.* **9**, 2439, doi:10.1038/s41467-018-04823-3 (2018).
18. Pi, X. *et al.* The pigment-protein network of a diatom photosystem II-light-harvesting antenna supercomplex. *Science* **365**, doi:10.1126/science.aax4406 (2019).
19. Yu, J. *et al.* Ycf48 involved in the biogenesis of the oxygen-evolving photosystem II complex is a seven-bladed beta-propeller protein. *Proc. Natl. Acad. Sci. U S A* **115**, E7824-E7833, doi:10.1073/pnas.1800609115 (2018).
20. Alvey, R. M., Biswas, A., Schluchter, W. M. & Bryant, D. A. Attachment of noncognate chromophores to CpcA of *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002 by heterologous expression in *Escherichia coli*. *Biochemistry* **50**, 4890-4902, doi:10.1021/bi200307s (2011).
21. Wittkopp, T. M. *et al.* Bilin-dependent photoacclimation in chlamydomonas reinhardtii. *Plant Cell* **29**, 2711-2726, doi:10.1105/tpc.17.00149 (2017).
22. Johnson, M. P. Photosynthesis. *Essays Biochem.* **60**, 255-273, doi:10.1042/EBC20160016 (2016).
23. Lohmann, A. *et al.* Deficiency in phylloquinone (vitamin K1) methylation affects prenyl quinone distribution, photosystem I abundance, and anthocyanin accumulation in the Arabidopsis AtmenG mutant. *J. Biol. Chem.* **281**, 40461-40472, doi:10.1074/jbc.M609412200 (2006).
24. Xia, D. *et al.* Structural analysis of cytochrome bc1 complexes: implications to the mechanism of function. *Biochim. Biophys. Acta* **1827**, 1278-1294, doi:10.1016/j.bbabio.2012.11.008 (2013).
25. Refojo, P. N., Ribeiro, M. A., Calisto, F., Teixeira, M. & Pereira, M. M. Structural composition of alternative complex III: variations on the same theme. *Biochim. Biophys. Acta* **1827**, 1378-1382, doi:10.1016/j.bbabio.2013.01.001 (2013).

26. Spang, A. *et al.* Proposal of the reverse flow model for the origin of the eukaryotic cell based on comparative analyses of Asgard archaeal metabolism. *Nat. Microbiol.* **4**, 1138-1148, doi:10.1038/s41564-019-0406-9 (2019).
27. Liu, Y. *et al.* Comparative genomic inference suggests mixotrophic lifestyle for Thorarchaeota. *ISME J.* **12**, 1021-1031, doi:10.1038/s41396-018-0060-x (2018).
28. Ternes, C. M. & Schonknecht, G. Gene transfers shaped the evolution of de novo NAD^+ biosynthesis in eukaryotes. *Genome Biol. Evol.* **6**, 2335-2349, doi:10.1093/gbe/evu185 (2014).
29. VanDover, C. L., Reynolds, G. T., Chave, A. D. & Tyson, J. A. Light at deep-sea hydrothermal vents. *Geophys. Res. Lett.* **23**, 2049-2052, doi:10.1029/96gl02151 (1996).
30. White, S. N., Chave, A. D. & Reynolds, G. T. Investigations of ambient light emission at deep-sea hydrothermal vents. *J. Geophys. Res-Sol. Ea.* **107**, doi:10.1029/2000jb000015 (2002).
31. Nurnberg, D. J. *et al.* Photochemistry beyond the red limit in chlorophyll f-containing photosystems. *Science* **360**, 1210-1213, doi:10.1126/science.aar8313 (2018).
32. Nisbet, E. G., Cann, J. R. & VanDover, C. L. Origins of photosynthesis. *Nature* **373**, 479-480 (1995).
33. Xiong, J. & Bauer, C. E. Complex evolution of photosynthesis. *Annu. Rev. Plant Biol.* **53**, 503-521, doi:10.1146/annurev.arplant.53.100301.135212 (2002).
34. Yano, J. & Yachandra, V. Mn_4Ca cluster in photosynthesis: where and how water is oxidized to dioxygen. *Chem. Rev.* **114**, 4175-4205, doi:10.1021/cr4004874 (2014).
35. Betts, H. C. *et al.* Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nat. Ecol. Evol.* **2**, 1556-1562, doi:10.1038/s41559-018-0644-x (2018).

Methods

Sample collection and processing. Samples were collected from the cold seep in South China Sea and hydrothermal vent field in Okinawa trough (Expended Data Table 1) during the cruise of the R/V *Kexue* on July of 2017-2018. The sediment samples (C1, C4, C2 and C5) were collected from cold seep area in South China Sea at depth intervals of 0-20, 20-40, 40-60 and 280 cm. The other subsurface sediment (0-20 cm) samples (H1 and H2) were taken at the outside of the “black chimney” of the hydrothermal vent. Among the specimens, samples C1, H1 and H2 were collected through the Discovery remotely operated vehicle (ROV), sample C4 was obtained by the TV grab, while samples C2 and C5 were taken from the gravity sampler. Sediments were sealed into sterile sampling bags immediately after collection, and

446 stored in -80 °C. DNA for metagenomics analysis was isolated from 20 g (wet
447 weight) sediment per sample with the PowerSoil DNA Isolation Kit (Qiagen)
448 following the manufacturer's instructions.
449 **Analyses of environmental and chemical parameters of sampling sites.** The
450 temperature, salinity, underwater depth of sampling sites were recorded in real-time
451 by SBE 25plus Sealogger CTD (SBE, USA), and concentrations of CO_2 and CH_4 of
452 surface sediments were in situ measured with the CONTROS®HydroCO_2
453 (CONTROS, Norway) and Hydro®CH_4 (CONTROS, Norway), respectively. All
454 these sensors were mounted on the Discovery ROV. For chemical element analyses,
455 sediment samples from cold seep (C1, C2, C4 and C5) in South China Sea and
456 hydrothermal vent field (H1 and H2) in Okinawa trough were dehydrated in an oven
457 at 80 °C, respectively, until completely dry. After grinded, powder of samples were
458 filtered through the 200-mesh screen. The obtained filtrate was further used to analyze
459 the contents of different chemical elements, including Na, Mg, Fe, Cl, S, P, Mn, Zn,
460 Ni and Co, by an S8 Tiger X-ray fluorescence spectrometry (BRUKER, Germany).
461 **Library construction and sequencing.** DNA extracts were treated with DNase-free
462 RNase to eliminate RNA contamination. Then the DNA concentration was measured
463 by Qubit 3.0 fluorimeter (Thermo Fisher Scientific, USA). DNA integrity was
464 evaluated by gel electrophoresis and 0.5 μg of each sample was used to prepare
465 libraries. The DNA was sheared into fragments between 50-800 bp using Covaris
466 E220 ultrasonicator (Covaris, UK). DNA fragments between 150 bp and 250 bp were
467 secreted using AMPure XP beads (Agencourt, USA) and then were repaired using T4
468 DNA polymerase (ENZYMATICS, USA). These DNA fragments were ligated at both
469 ends to T-tailed adapters and amplified for eight cycles. Finally, the amplification
470 products were subjected to single-strand circular DNA libraries. All NGS libraries
471 were sequenced on BGISEQ-500 platform (BGI, China) to obtain 100 bp paired-end
472 raw reads. Quality control was performed by SOAPnuke (v1.5.6) (setting: -l 20 -q 0.2
473 -n 0.05 -Q 2 -d -c 0 -5 0 -7 1)[36].
474 **Genomes assembly, binning and annotation.** The raw shotgun sequencing
475 metagenomic reads were dereplicated and trimmed by the BGI-Qingdao (BGI, China).
476 The clean data were assembled using MEGAHIT (v1.1.3, setting: --min-count 2
477 --k-min 33 --k-max 83 --k-step 10)[37]. Thereafter, metaBAT2[38], Maxbin2[39] and
478 Concoct[40] were used to automatically bin from assemblies. Finally, MetaWRAP[41] was
479 used to purify and generate data to get the final bins. Manual curation was adapted for

reducing the genome contamination based on differential coverage, GC content, and the presence of duplicate genes. The completeness and contamination of the genomes within bins were then estimated by using CheckM[42]. Gene prediction for individual genomes was performed using Glimmer (v 3.02)[43]. The KEGG (Kyoto Encyclopedia of Genes and Genomes, Release 87.0), NR (Non-Redundant Protein Database databases, 20180814), Swiss-Prot (release-2017_07) and EggNOG (2015-10_4.5v) databases were used to annotate protein functions by default, and the best hits were chosen. Additionally, database of CAZy (Carbohydrate-Active enZYmes Database)[44] was downloaded to search for carbohydrate active enzymes from genomic bins.

Phylogenetic analyses. To reveal the phylum composition of assembled genomes in the archaea kingdom, the genomic sequences of Archaea were downloaded from NCBI ref genomes using Aspera (v3.9.8). Then, extract 37 marker genes in genomes (Supplemental Table 1) were chosen by Phylosift (v1.0.1)[45] with automated setting. The concatenated sequences were trimmed using TrimAl (version 1.2)[46] using gappyout function. Finally, maximum likelihood tree was calculated by using IQ-TREE (v1.6.12)[47] with GTR+F+I+G4 model (-bb 1000) and shown by iTOL (v5)[48]. For phylogenetic analyses of protochlorophyllide reductase (Por), bacteriochlorophyll synthase (BCS) and demethylphylloquinol/demethylmenaquinone methyltransferase (MenG/ubiE) in Heimdallarchaeota bins, the related sequences were selected from archaea, bacteria and eukaryotes in NCBI and Swiss-Prot database. The Maximum-Likelihood phylogeny trees were constructed with WAG+G4, LG+F+I+G4 and LG+G4 model (-bb 1000) by using IQ-TREE, respectively, and showing with iTOL.

Data availability. The Heimdallarchaeota genomic bins (C2.bin.3, C4.bin.14, C4.bin.22, C5.bin.12 and H2.bin.2) supporting the results of this study are available in NCBI Genbank under the accession numbers: SAMN13483368, SAMN13483369, SAMN13483392, SAMN13483370 and SAMN13483372 in BioProject PRJNA593668, respectively.

Acknowledgements

This work was funded by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA22050301), China Ocean Mineral Resources R&D Association Grant (Grant No. DY135-B2-14), National Key R and D Program

of China (Grant No. 2018YFC0310800), the Taishan Young Scholar Program of Shandong Province (tsqn20161051), and Qingdao Innovation Leadership Program (Grant No. 18-1-2-7-zhc) for Chaomin Sun.

Author Contributions

CS and RL conceived and designed the study. RL and JZ collected samples and supported the information of genomes. RL and RC analyzed the data. RL and CS wrote the manuscript with the input from all authors. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests

References

36. Chen, Y. *et al.* SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *GigaScience* **7**, 1-6, doi:10.1093/gigascience/gix120 (2018).
37. Li, D., Liu, C. M., Luo, R., Sadakane, K. & Lam, T. W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674-1676, doi:10.1093/bioinformatics/btv033 (2015).
38. Kang, D. D. *et al.* MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**, e7359, doi:10.7717/peerj.7359 (2019).
39. Wu, Y. W., Simmons, B. A. & Singer, S. W. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**, 605-607, doi:10.1093/bioinformatics/btv638 (2016).
40. Alneberg, J. *et al.* Binning metagenomic contigs by coverage and composition. *Nat. Methods* **11**, 1144-1146, doi:10.1038/nmeth.3103 (2014).
41. Uritskiy, G. V., DiRuggiero, J. & Taylor, J. MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* **6**, 158, doi:10.1186/s40168-018-0541-1 (2018).
42. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043-1055, doi:10.1101/gr.186072.114 (2015).

549 43. Delcher, A. L., Bratke, K. A., Powers, E. C. & Salzberg, S. L. Identifying
550 bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **23**,
551 673-679, doi:10.1093/bioinformatics/btm009 (2007).
552 44. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B.
553 The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.*
554 **42**, D490-495, doi:10.1093/nar/gkt1178 (2014).
555 45. Darling, A. E. *et al.* PhyloSift: phylogenetic analysis of genomes and
556 metagenomes. *PeerJ* **2**, e243, doi:10.7717/peerj.243 (2014).
557 46. Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for
558 automated alignment trimming in large-scale phylogenetic analyses.
559 *Bioinformatics* **25**, 1972-1973, doi: 10.1093/bioinformatics/btp348 (2009).
560 47. Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast
561 and effective stochastic algorithm for estimating maximum-likelihood
562 phylogenies. *Mol. Biol. Evol.* **32**, 268-274, doi:10.1093/molbev/msu300 (2015).
563 48. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the
564 display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**,
565 W242-245, doi:10.1093/nar/gkw290 (2016).
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582

Figure Legends

Fig. 1 Phylogenomic analysis of photosynthetic pigments biosynthesis in Heimdallarchaeota. **a**, Analysis of bacteriochlorophyll biosynthesis in different Heimdallarchaeota and other Asgard archaeal genomes. **b**, **c**, Phylogenetic analyses of protochlorophyllide reductase (Por) and bacteriochlorophyll synthetase (BCS). A rooted maximum-likelihood tree of Por (b) or BCS (c) homologs derived from different photosynthetic organisms identified in this work. The bootstrap support values 1,000. All proteins and species detailed information used for phylogenetic analyses are listed in Supplementary Table 3. **d**, Analysis of lycopene biosynthesis in different Heimdallarchaeota and other Asgard archaeal genomes. For panels a and d, the solid arrows indicate the enzymes associated with bacteriochlorophyll or lycopene biosynthesis identified in Heimdallarchaeota MAGs. Dotted arrows indicate the enzymes associated with bacteriochlorophyll or lycopene biosynthesis not identified in Heimdallarchaeota MAGs. The light blue box highlights Heimdallarchaeota MAGs from cold seeps. The light pink box highlights Heimdallarchaeota MAGs from hydrothermal environment. The light purple box highlights assembled genomes of other Asgard archaea. The parallelogram box highlights Heimdallarchaeota MAGs obtained in this study. The detail information of key enzymes involved in bacteriochlorophyll and lycopene biosynthesis is listed in Supplementary Tables 3 and 5.

Fig. 2 Phylogenomic analysis of phylloquinone and plastoquinone biosynthesis in Heimdallarchaeota. **a**, Analysis of phylloquinone biosynthesis in Heimdallarchaeota. The detail information of key enzymes involved in phylloquinone biosynthesis is listed in Supplementary Table 4. SEHCC, 2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate; SHCHC, (1R, 6R)-2-Succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate. **b**, Phylogenetic analysis of MenG/UbiE. An unrooted maximum-likelihood tree of MenG/UbiE homologs derived from different photosynthetic organisms identified in this work. The bootstrap support values 1,000. The green box highlights proteins annotated as MenG. The pink box highlights proteins annotated as UbiE. The blue box highlights proteins ambiguously annotated as MenG or UbiE. **c**, Analysis of plastoquinone biosynthesis in Heimdallarchaeota. ARO8, 2-aminoadipate transaminase. HPD,

616 4-hydroxyphenylpyruvate dioxygenase. SPS, All-trans-nonaprenyl-diphosphate
617 synthase. HST, Homogentisate solanesyltransferase. VTE3, MPBQ/MSBQ
618 methyltransferase. For panels a and c, the solid arrows indicate the enzymes
619 associated with phylloquinone or plastoquinone biosynthesis identified in
620 Heimdallarchaeota MAGs. Dotted arrows indicate the enzymes associated with
621 phylloquinone or plastoquinone biosynthesis not identified in Heimdallarchaeota
622 MAGs. The detail information of key enzymes involved in phylloquinone or
623 plastoquinone biosynthesis is listed in Supplementary Tables 4 and 5.
624
625 **Fig. 3** The inferred photosynthetic model of Heimdallarchaeota. Solid lines and
626 arrows indicate the elements associated with photosystem and electron transport
627 identified in Heimdallarchaeota MAGs, respectively. Dotted lines and arrows indicate
628 the elements associated with photosystem and electron transport not identified in
629 Heimdallarchaeota MAGs, respectively. The detail information of key enzymes
630 involved in photosystem and electron transport is listed in Supplementary Table 4.
631
632 **Fig. 4** Reconstruction of photomixotrophic lifestyle in Heimdallarchaeota. Solid
633 arrows indicate the enzymes associated with corresponding metabolic pathway
634 identified in Heimdallarchaeota MAGs. Dotted arrows indicate the enzymes
635 associated with corresponding metabolic pathway not identified in Heimdallarchaeota
636 MAGs. The grey box highlights the photosynthetic pathway present in
637 Heimdallarchaeota. The detail information of key enzymes mentioned in this figure is
638 listed in Supplementary Table 5.

Figures

Figure 1

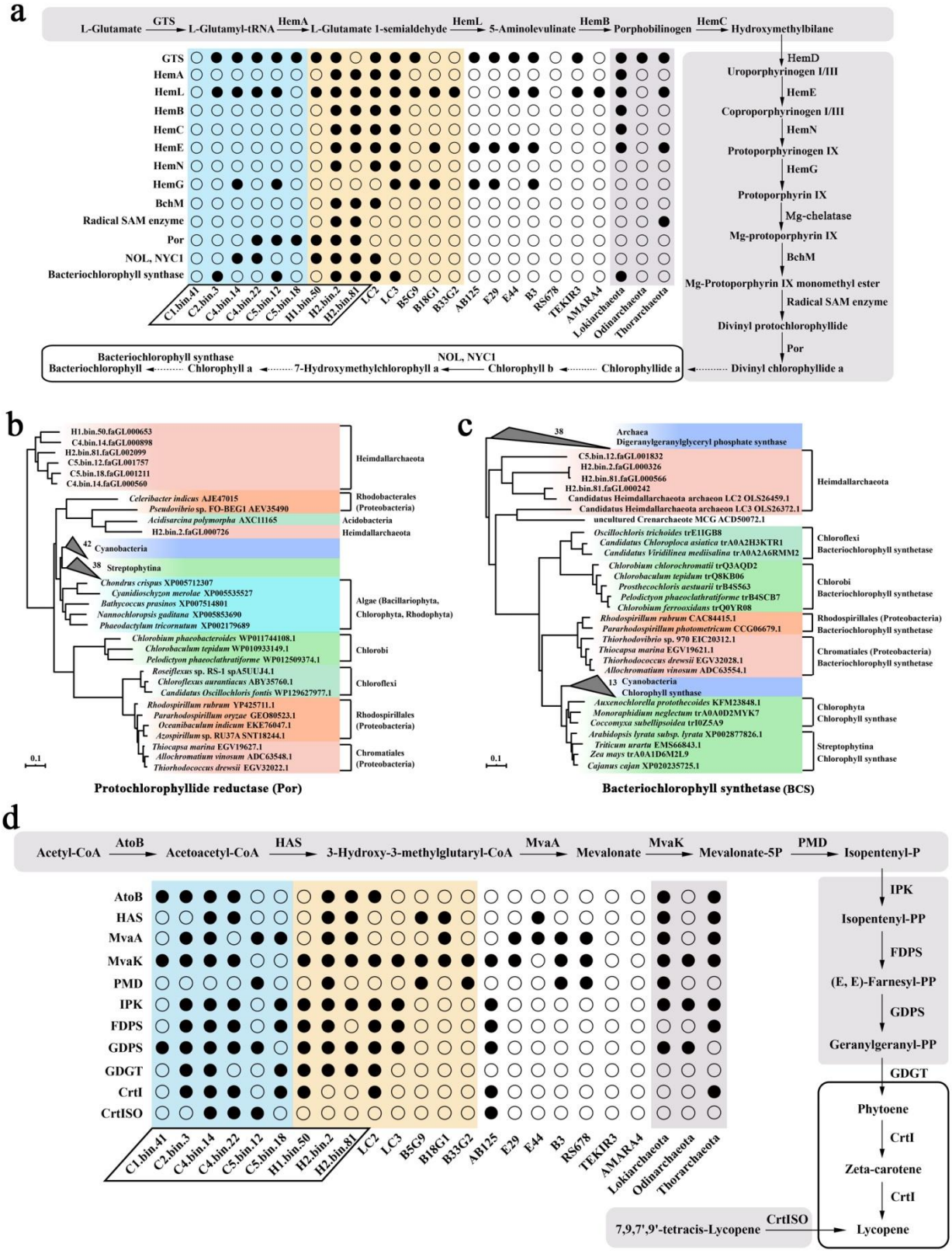


Figure 2

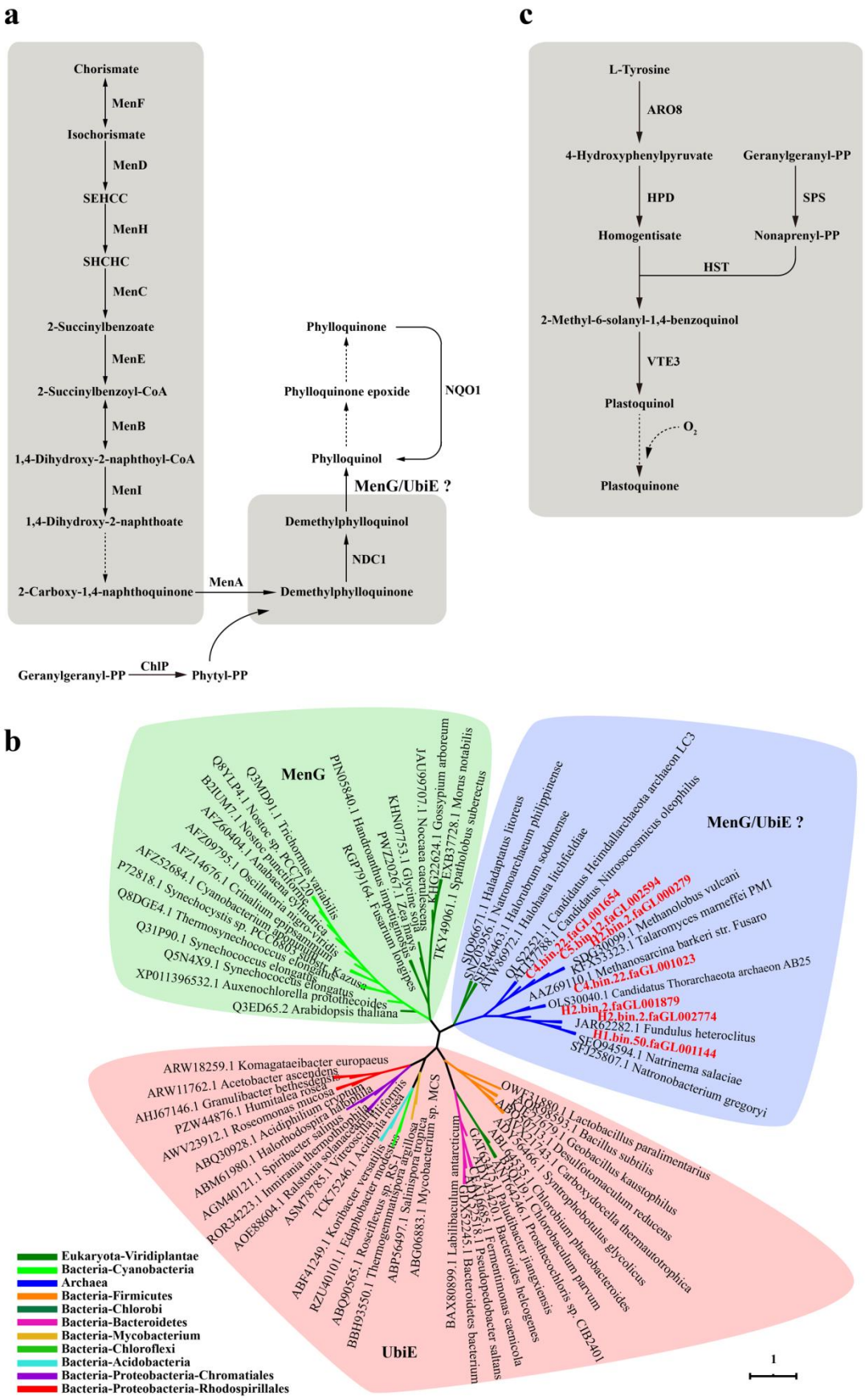
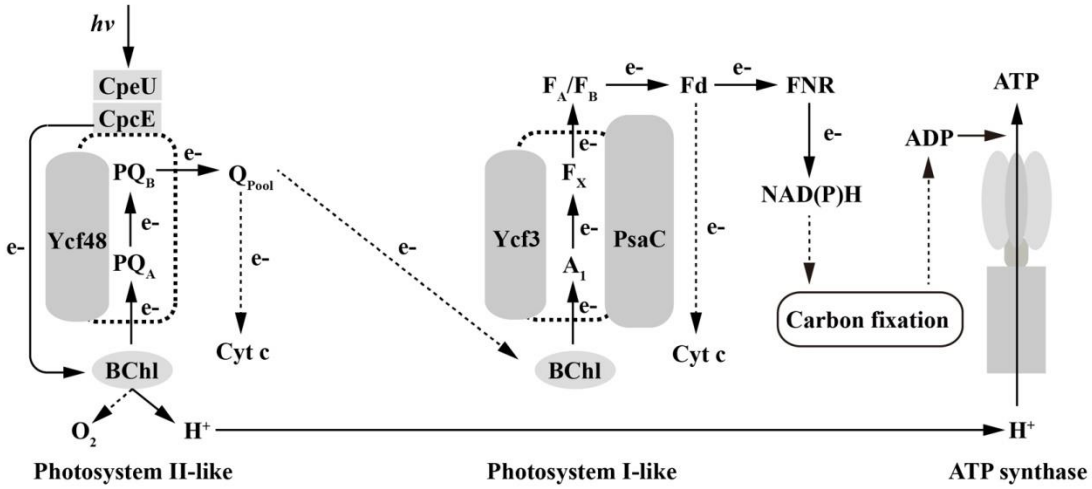
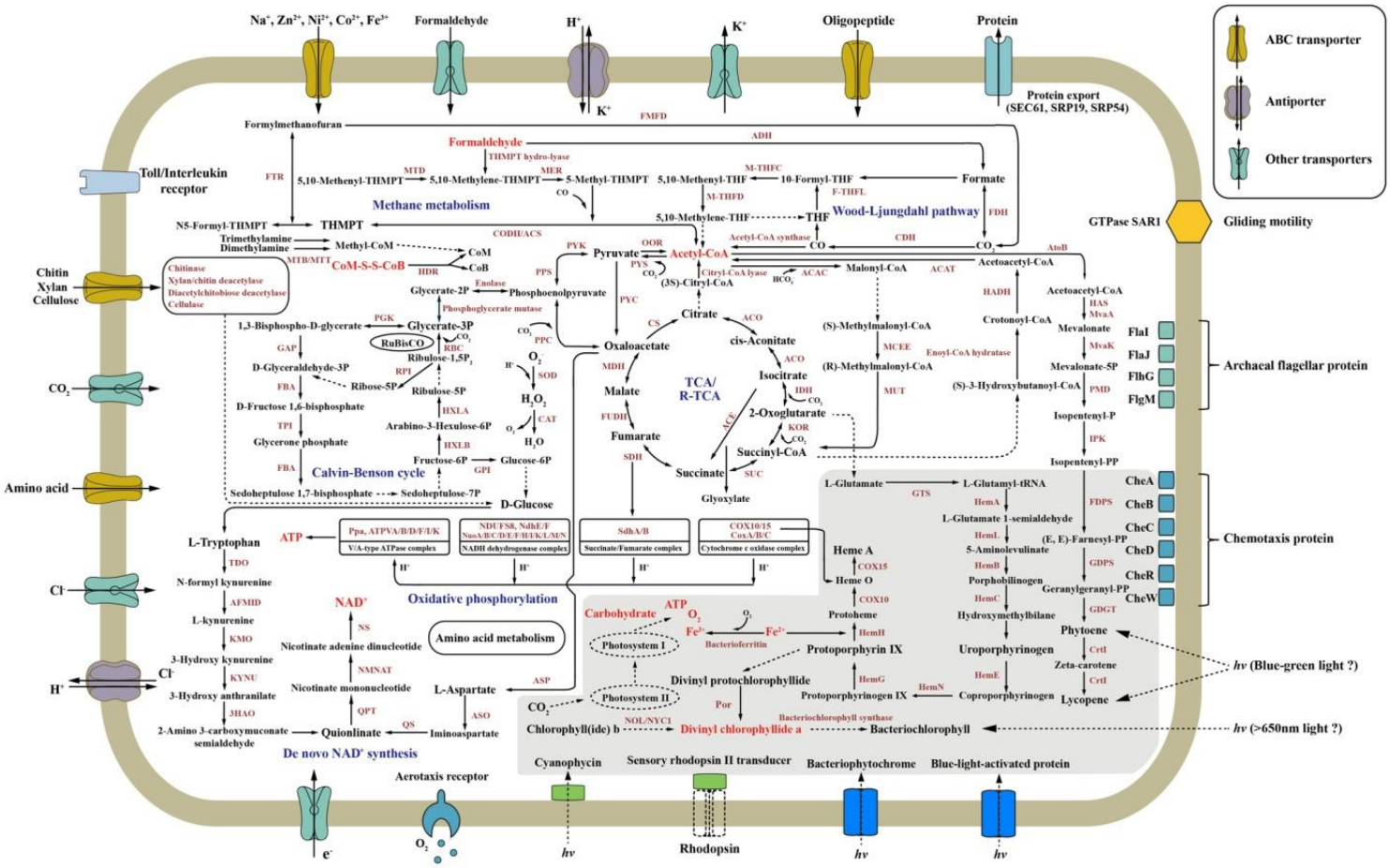


Figure 3



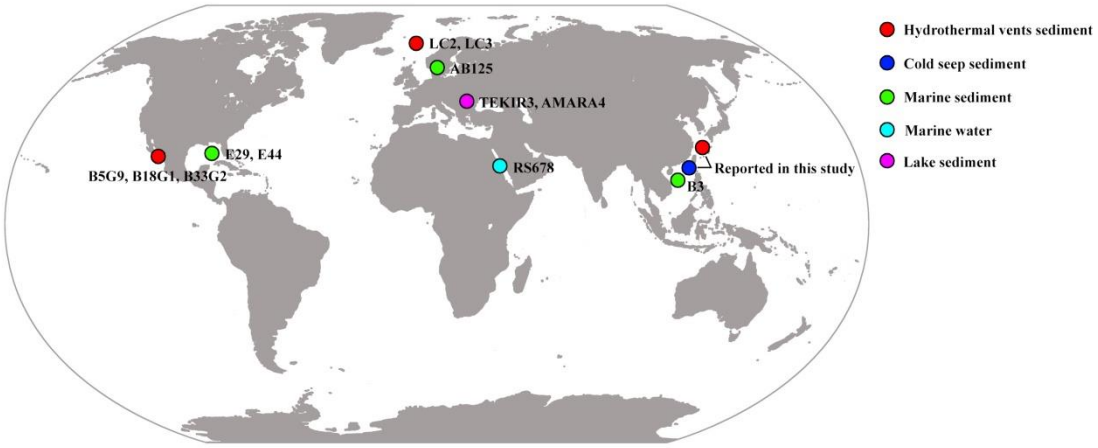
BChl: Bacteriochlorophyll
PQ: Plastoquinone
CpcE and CpeU: Antenna proteins
Ycf48: Photosystem II stability/assembly factor
A_1: Phylloquinone
Cyt c: Cytochrome c
PsaC: Photosystem I subunit VII
Ycf3: Photosystem I assembly protein
F_X: Fe_4-S_4
F_A/F_B: $2Fe_4$-S_4
Fd: Ferredoxin
FNR: Ferredoxin $NADP^+$ reductase

Figure 4



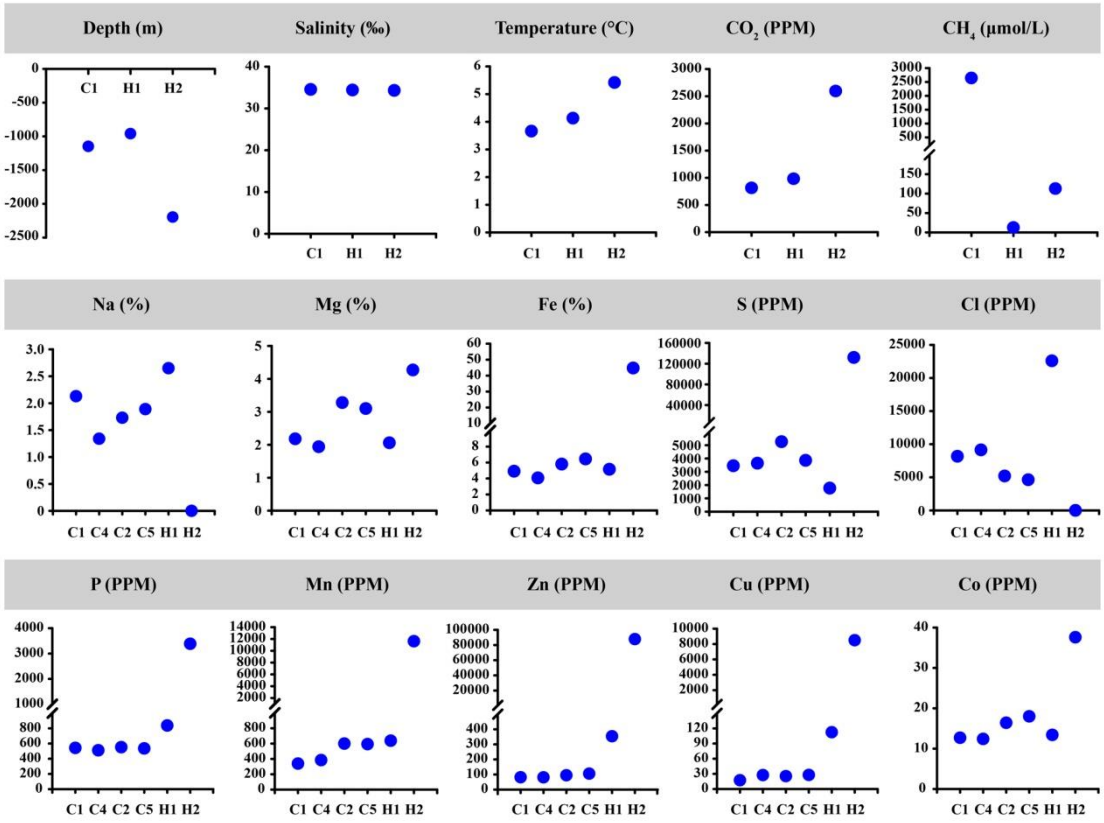
Expanded data Figures

Expanded data Figure 1



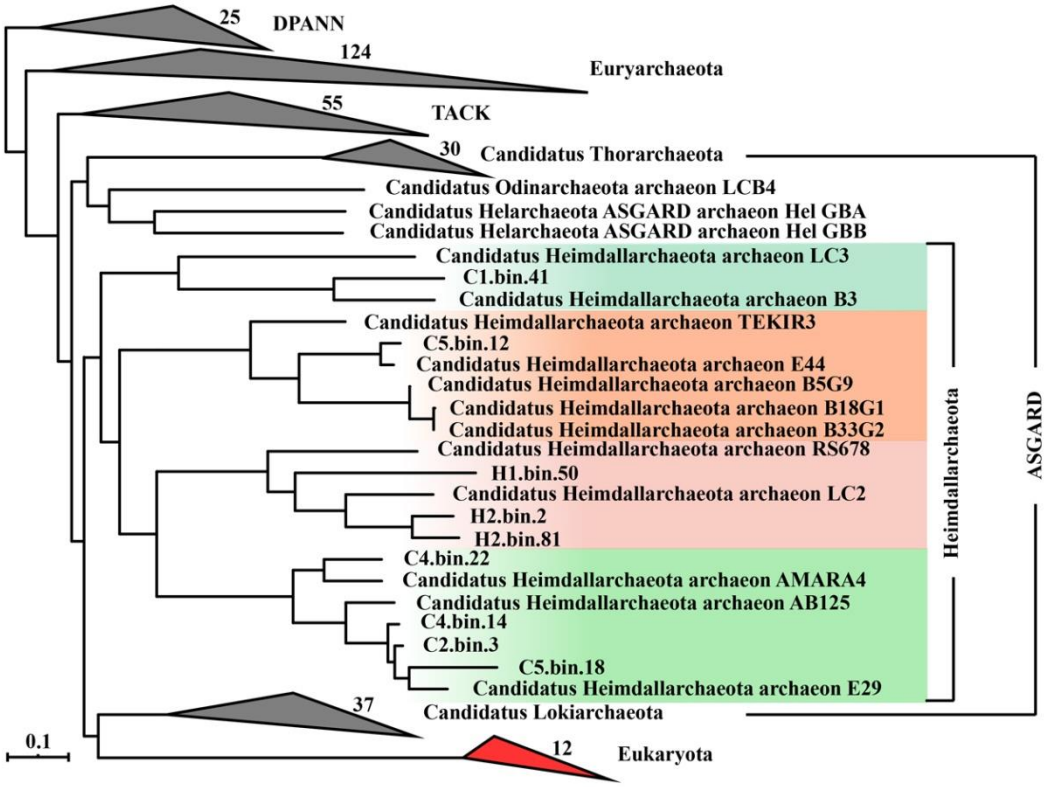
Extended Data Fig. 1 Global distribution of Heimdallarchaeota MAGs reported in previous and present studies. MAGs LC2 and LC3 are derived from Loki's Castle. MAG AB125 is derived from Aarhus Bay (Denmark). MAG TEKIR3 is derived from Tekirghiol (Romania). MAG AMARA4 is derived from Amara (Romania). MAG RS678 is derived from Red Sea (Saudi Arabia). MAGs B5G9, B18G1 and B33G2 are derived from Guaymas Basin, Gulf of California (Mexico). MAGs E29 and E44 are derived from Atlantic Ocean. MAG B3 is derived from the north of the South China Sea (China).

Expanded data Figure 2



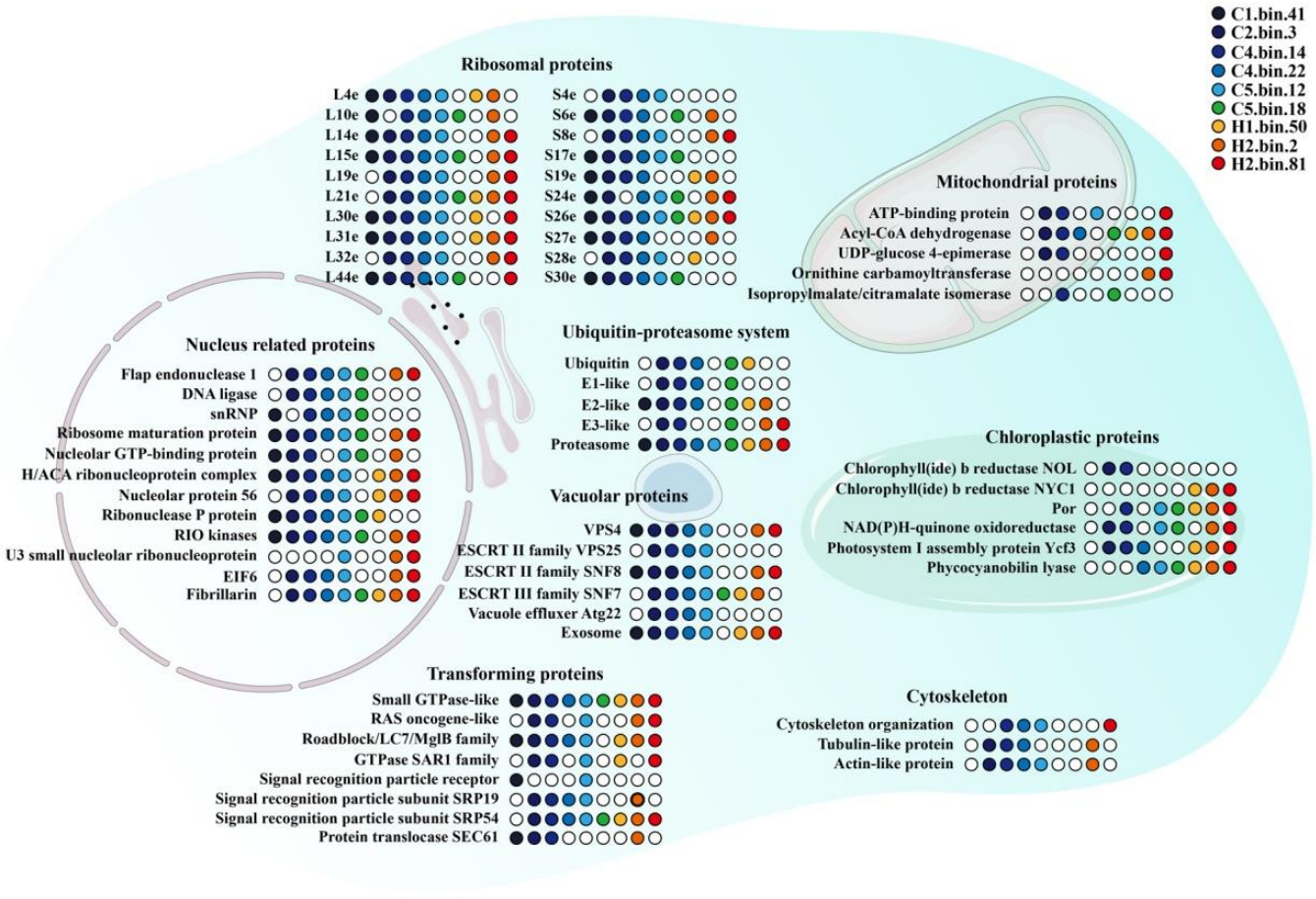
Extended Data Fig. 2 Analyses of environmental and chemical parameters of sampling sites in the deep-sea cold seep and hydrothermal vents. The temperature, salinity, underwater depth were recorded in real-time by SBE 25plus Sealogger CTD, and concentrations of CO_2 and CH_4 of surface sediments were in situ measured with the CONTROS®HydroCO_2 and Hydro®CH_4. Contents of different elements including Na, Mg, Fe, Cl, S, P, Mn, Zn, Ni and Co were measured by an S8 Tiger X-ray fluorescence spectrometry.

Expanded data Figure 3



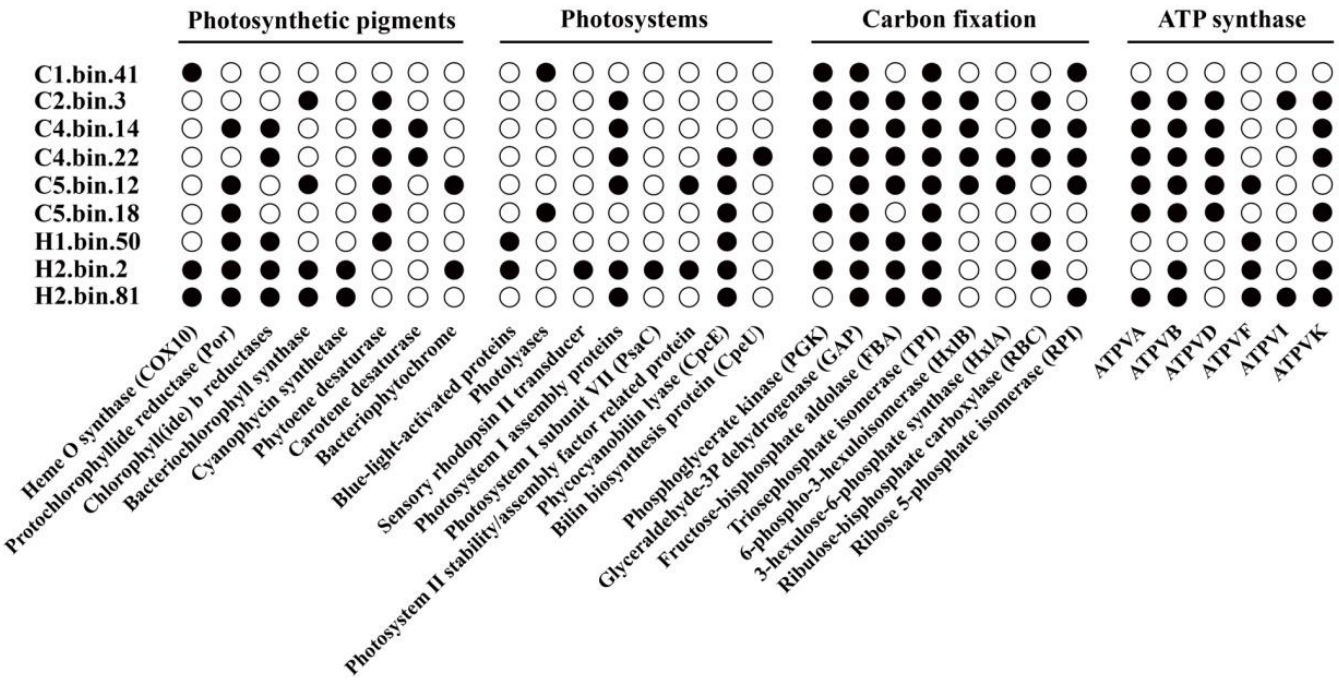
Extended Data Fig. 3 Maximum-likelihood phylogeny of superphyla Asgard, TACK, Euryarchaeota, DPANN and Eukaryota. Total 37 marker genes chosen by Phylosift, including 13 small subunit (SSU) and 16 large subunit (LSU) ribosomal RNA genes. The bootstrap support values 1000. All detailed sequence information of different species in compressed clades is listed in Supplementary Table 1.

Expanded data Figure 4



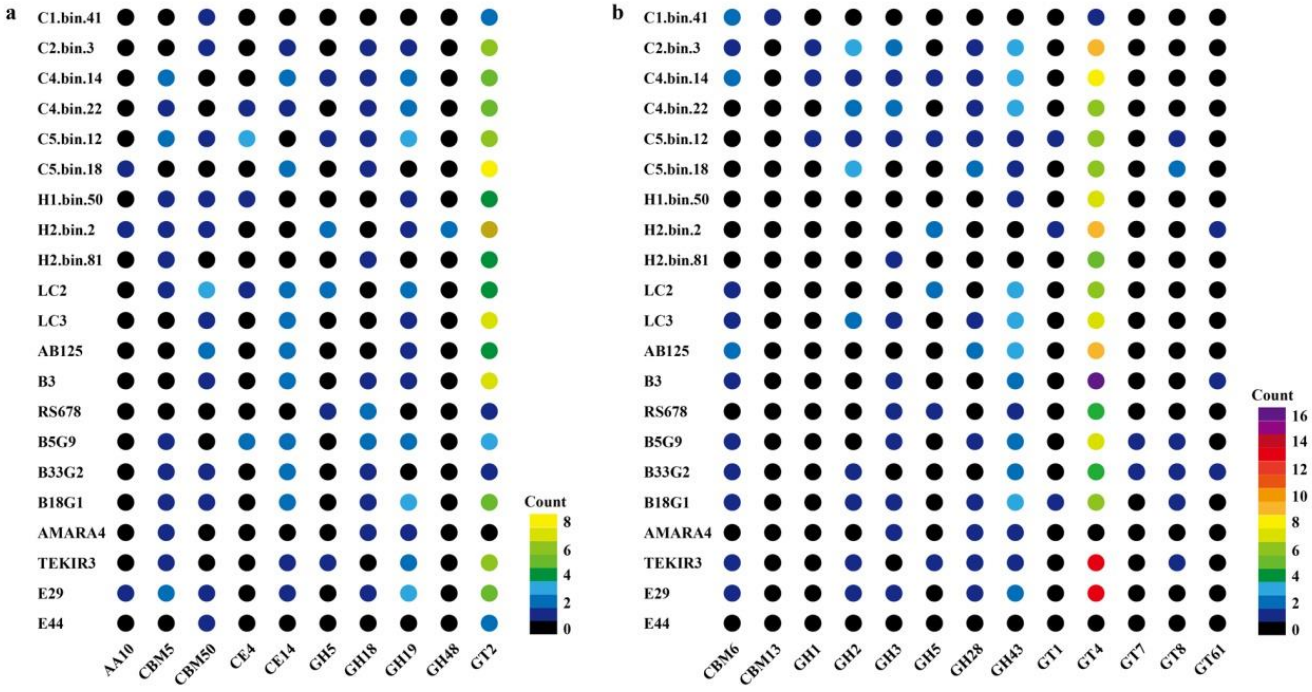
Extended Data Fig. 4 Eukaryotic signatures in Heimdallarchaeota. Schematic representation of a eukaryotic-like cell in which ESPs that have been identified in Heimdallarchaeota are highlighted. The overall illustration indicates that Heimdallarchaeota contain both reported eukaryotic signatures and unprecedented chloroplastic proteins. All detailed protein information mentioned in this figure is listed in Supplementary Table 2.

Expanded data Figure 5



Extended Data Fig. 5 Photosynthetic apparatus identified in Heimdallarchaeota MAGs. All detailed protein information mentioned in this figure is listed in Supplementary Table 5. ATPVA~ATPVK, V/A-type H^+/Na^+-transporting ATPase subunits A~K.

Expanded data Figure 6



Extended Data Fig. 6 Chitin and xylan metabolic associated enzymes identified in Heimdallarchaeota MAGs by CAZy analysis. **a**, Chitin metabolic related enzymes identified in Heimdallarchaeota MAGs. **b**, Xylan metabolic associated enzymes identified in Heimdallarchaeota MAGs. AAs, auxiliary activities. CBMs, carbohydrate-binding modules. CEs, carbohydrate esterases. GHs, glycoside hydrolases. GTs, glycosyltransferases. All detailed protein information mentioned in this figure is listed in Supplementary Table 5.