Title

- Highly diverse Asgard archaea participate in organic matter degradation in coastal sediments
- Asgard archaea are capable of organic matter utilization

Authors

Mingwei Cai\textsuperscript{1,2†}, Yang Liu\textsuperscript{1†}, Xiuran Yin\textsuperscript{3,4}, Zhichao Zhou\textsuperscript{1,5}, Michael W. Friedrich\textsuperscript{3,4}, Tim Richter-Heitmann\textsuperscript{3}, Rolf Nimzyk\textsuperscript{6}, Ajinkya Kulkarni\textsuperscript{3}, Xiaowen Wang\textsuperscript{1,2}, Wenjin Li\textsuperscript{1}, Jie Pan\textsuperscript{1}, Yuchun Yang\textsuperscript{5}, Ji-Dong Gu\textsuperscript{5} and Meng Li\textsuperscript{*}

Affiliations

\textsuperscript{1}Institute for Advanced Study, Shenzhen University, Shenzhen, China.
\textsuperscript{2}Key Laboratory of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province, College of Optoelectronic Engineering, Shenzhen University, Shenzhen, China.
\textsuperscript{3}Microbial Ecophysiology Group, Faculty of Biology/Chemistry, University of Bremen, Bremen, Germany.
\textsuperscript{4}MARUM, Center for Marine Environmental Sciences, University of Bremen, Bremen, Germany.
\textsuperscript{5}Laboratory of Environmental Microbiology and Toxicology, School of Biological Sciences, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, China.
\textsuperscript{6}Department of Microbe-Plant Interactions, Faculty of Biology/Chemistry, University of Bremen, Bremen, Germany.
ABSTRACT

Asgard is an archaeal superphylum that might hold the key to understand the origin of eukaryotes, but its diversity and ecological roles remain poorly understood. Here, we reconstructed 15 metagenomic-assembled genomes (MAGs) from coastal sediments covering most known Asgard archaea and a novel group, which is proposed as a new Asgard phylum named as the “Gerdarchaeota”. Genomic analyses predict that Gerdarchaeota are facultative anaerobes in utilizing both organic and inorganic carbon. Unlike their closest relatives Heimdallarchaeota, Gerdarchaeota have genes encoding for cellulase and enzymes involving in the tetrahydromethanopterin-based Wood–Ljungdahl pathway. Transcriptomic evidence showed that all known Asgard archaea are capable of degrading organic matter, including peptides, amino acids and fatty acids, in different ecological niches in sediments. Overall, this study broadens the diversity of the mysterious Asgard archaea and provides evidence for their ecological roles in coastal sediments.
Introduction

Asgard archaea, proposed as a new archaeal superphylum, are currently composed of five phyla, i.e., Lokiarchaeota\(^1\), Thorarchaeota\(^2\), Odinarchaeota\(^3\), Heimdallarchaeota\(^3\), and Helarchaeota\(^4\), some of which encompasses lineages formerly named Marine Benthic Group B (MBG-B)\(^5\), Deep-Sea Archaeal Group (DSAG)\(^6\), Ancient Archaeal Group (AAG)\(^7\), and Marine Hydrothermal Vent Group (MHVG)\(^7, 8\). Since Asgard archaea contain abundant eukaryotic signature proteins (ESPs) and form a monophyletic group with eukaryotes in the phylogenetic tree inferred from 55 concatenated archaeo-eukaryotic ribosomal proteins, they are regarded as the closest relatives of Eukarya and have attracted increasing research interests\(^1, 3, 9\).

Metabolic potentials of several Asgard phyla have been predicted based upon their genomic inventories: hydrogen dependency\(^10\) in Lokiarchaeota (which were originally described in deep-sea sediments), mixotrophy\(^11\) (i.e., using both inorganic and organic carbon for growth) and acetogenesis\(^2\) by Thorarchaeota from estuary sediments, metabolization of halogenated organic compounds\(^12\) by Loki- and Thorarchaeota, phototrophy\(^13, 14\) in Heimdallarchaeota from coastal sediment and hydrothermal vent samples, anaerobic hydrocarbon oxidation in hydrothermal deep-sea sediment by Helarchaeota\(^4\), and nitrogen and sulfur cycling\(^15\) in all Asgard archaeal phyla. Metatranscriptomics revealed a few transcripts encoding NiFe hydrogeneases from deep-sea Lokiarchaeota\(^16\). Since no comprehensive information about the in situ activity of these Asgard archaea has been compiled so far, our understanding of these phylogenetically and evolutionally important archaea is based on prediction, and thus, severely limited.

Coastal environments, e.g., mangroves, salt marshes and seagrass beds, are known sinks of blue carbon\(^17\). Although these vegetated coastal ecosystems make up less than 0.5% of the seabed, they hold ~50% of organic carbon of the surface marine sediments globally\(^17-19\). Here, we reconstructed 15 Asgard metagenomic-assembled genomes (MAGs) from diverse coastal
sediments and analysed them together with transcriptomes from mangrove sediments to clarify
the ecological roles of the different Asgard clades in these important environments. Based on
phylogenetic analysis of this dataset, we propose a novel Asgard phylum, the "Gerdarchaeota".
Additionally, we recruited all publicly available 16S rRNA sequences to see whether or not
distinct Asgard lineages show distinct habitat preferences. Our findings substantially extend our
knowledge of the lifestyles of these mysterious archaea in coastal sediments and their ecological
roles on carbon cycling.

Results and Discussion

Mining Asgard archaea genomes lead to the proposal of a new Asgard phylum

Sediments from several coastal sites (mangrove, mudflat and seagrass bed) were collected
for deep metagenomic and metatranscriptomic sequencing (totally 2.3 Tbp, table S1). By
combining individual samples from the same site but from different depth layers for assembly and
binning, we recovered 15 Asgard MAGs with completeness of >80% through hybrid binning
strategies (MetaBAT20 in combination with Das Tool21) and manual bin refinement. Based on a
phylogenetic analysis with a concatenated set of 122 archaeal-specific marker genes, we found
that eight of the MAGs belong to known Asgard lineages, covering almost all phyla (except
Odinarchaeota), i.e., Helarchaeota (1 MAG), Lokiarchaeota (2), Thorarchaeota (3),
Heimdallarchaeota-AAG (1), and Heimdallarchaeota-MHVG (1) (Fig. 1A and table S2).
The remaining 7 MAGs clustered with an unclassified MAG (B18_G1) and formed a
monophyletic group in phylogenetic tree of concatenated 122 archaeal marker genes (Fig. 1A).
The amino acid identity (AAI, 43% to 50% compared to other Asgard archaea), pan-genome
analysis, and non-metric multidimensional scaling (NMDS) ordination agreed on assigning
phylum level identity to this new clade, (figs. S1, S2 and S3). Likewise, phylogenetic analysis of
a 16S rRNA (1152 bp) gene found in of those seven MAGs (YT_re_metabat2_2.057) showed that
it formed a new monophyletic branch (Fig. 1B) with an identity below 74% to 16S rRNA genes of other Asgard archaeal phyla (table S3). Thus, we propose this lineage as a new phylum named Gerdarchaeota, after Gerd, the Norse goddess of fertile soil, because these Asgard archaea genomes were obtained from organic-rich coastal environments\textsuperscript{22}, such as mangrove, mudflat and seagrass (table S1).

The presence of eukaryotic signature proteins (ESPs) is a characteristic feature of Asgard (Fig. 1C and table S4), and we consistently found that their phylogenies support the discrete position of Gerdarchaeota in the Asgard lineage (fig. S4). Homologs encoding eukaryotic-type topoisomerase IB and fused RNA polymerase subunit A were identified in Gerdarchaeota, while neither genes for eukarya-specific DNA polymerase epsilon and DNA-directed RNA polymerase subunit G were detected (Fig. 1C, and table S5). Gerdarchaeotal topoisomerase IB is clustered into a sister group of Thorarchaeota and the two Asgard sequences together were monophyletic with high supporting values (fig. S4A). The fused RNA polymerase A genes of Gerdarchaeota are transcriptionally expressed (table S5) and clustered into a sister group of Heimdallarchaeota, and those of Helarchaeota branch with Loki- and Thorarchaeota (fig. S4B). In both cases, the eukaryotic genes do not cluster with the Asgard homologs. Besides, Gerdarchaeota also comprise expressed homologues of ribophorin I and STT3 subunit and lack OST3/OST6 homologues in most genomes except YT_bin5.010. The phylogenetic tree of ribophorin I showed that Gerdarchaeota is monophyletic and branches with Heimdallarchaeota, and the eukaryotes branch within the cluster of Thor-, Hel- and Lokiarchaeota (fig. S4C). In addition to the reported ESPs, we identified DAD/OST2 homologues within Gerdarchaeotal MAGs, which is a component of the N-ologosaccharyl transferase for the N-linked glycosylation\textsuperscript{23}.

Metabolic potential of the new phylum Gerdarchaeota
Gerarchaeota harbour the gene set for oxidative phosphorylation, including V/A-type ATPs, succinate dehydrogenase, NADH-quinone oxidoreductase, the key enzyme cytochrome c oxidase (subunits I, II, and III) (Fig. 2, fig. S5, and tables S6 and S7), and enzymes for the non-typical cytochrome bc1 complex (i.e., SoxL) but lacking the genes for other respiratory complex III (e.g., SoxN and CbsA), suggesting that Gerdarchaeota most likely perform aerobic respiration.

Gerdarchaeotal cytochrome c oxidases are phylogenetically separated into two lineages (fig. S6), one of which clusters closely with the facultative anaerobic Crenarchaeotal Acidianus brierleyi and the other one groups with the bacterial Synechocystis sp., which is capable of aerobic respiration in the dark. Meanwhile, Gerdarchaeotal MAGs harbour genes encoding heliorhodopsins (fig. S7), which might sense light in the top layers of sediment. A complete tricarboxylic acid cycle (e.g., citrate synthase and malate dehydrogenase, Fig. 2) further supports aerobic respiration as important dissimilatory pathway. Besides, Gerdarchaeota are equipped with enzymes for the conversion of AsO$_4^{3-}$ to (CH$_3$)$_n$-As, thus they could remove the toxic As(V) and As(III), which can disrupt oxidative phosphorylation and inhibit respiratory enzymes.

Within Gerdarchaeota MAGs, complete gene sets for acetogenesis pathways (e.g., acetyl-CoA hydrolase) are present (Fig. 2), implying the ability of these archaea to perform fermentation in the absence of oxygen. We further identified all subunits of [NiFe]-hydrogenase heterodisulfide reductase hdrABC. Since this community does not contain the key enzymes mcrABC for methanogenesis, hdrABC might function to accept electrons and reduce oxidized ferredoxin instead of a heterodisulfide anaerobically. Additionally, we identified homologues for ferric reductase but their capability to use Fe(III) as electron acceptors remained open. Notably, the canonical nitrate reductase (previously identified in Heimdallarchaeota) was not detected in Gerdarchaeota MAGs.

Gerdarchaeota appear to use diverse organic compounds (e.g., formaldehyde, amino acid, peptide, lipid and ethanol) as electron donors (Fig. 2). For genes encoding peptidase, serine
peptidases dominated (~44.1% of total peptidase, fig. S8A). We also identified abundant genes for cellulose degradation (e.g., AA3 and GT2, fig. S8B and table S8), which may be further degraded through the Embden–Meyerhof–Parnas (EMP) pathway. Different from the facultative anaerobic relatives Heimdallarchaeota, Gerarchaeota contain the complete genes for tetrahydromethanopterin Wood-Ljungdahl (THMPT_WL) pathway and the key enzyme acetyl-CoA decarbonylase/synthase (Fig. 2). The presence of the genes for groups 3b and 3c [NiFe]-hydrogenases (fig. S9) implies that these archaea may grow lithoautotrophically using H₂ as electron donor as also reported for Lokiarchaeota¹⁰. Meanwhile, we identified other CO₂ assimilation pathways in Gerarchaeota MAGs. For example, Gerarchaeota are equipped to fix carbon via the complete reductive citric acid cycle (e.g., 2-oxoglutarate oxidoreductase and isocitrate/isopropylmalate dehydrogenase) and Calvin–Benson–Bassham (CBB) cycle, and the pyruvate dehydrogenase, which is responsible for the generation of pyruvate from acetyl-CoA and CO₂, underpin the importance of inorganic carbon for biomass synthesis. Different from other Asgard phyla, we did not find type III or type IV ribulose 1,5-bisphosphate carboxylase (RuBisCO) (fig. S10), which might function in the nucleotide salvage pathway.¹¹, 2¹, 2⁶, 2⁷.

Most reported Asgard genomes contain genes for Glycerol-1-phosphate dehydrogenase (G1PDH) responsible for ether-bound phospholipids synthesis.¹², ³³. A recent study reported a co-existence of enzymes for both ether- (G1PDH) and ester-bound (bacterial/eukarya type, glycerol-3-phosphate dehydrogenase (G3PDH)) phospholipid synthesis in some Lokiarchaeotal genomes, providing a hint that Asgard archaea might produce chimeric lipids. Intriguingly, Gerarchaeota MAGs lack the key enzyme G1PDH for archaeal lipid biosynthesis, but contain the bacterial/eukarya-type G3PDH for synthesis of bona fide bacterial lipids (Fig. 2). This finding indicated that Gerarchaeota might have evolved a bacterial-like membrane predates the eukaryogenesis. Further investigation suggests that the Asgard archaeal G3PDH is synthesized by the glpA (responsible for the reverse reaction catalyzed by GpsA) instead of GpsA, which means
that G3PDH may participate in organic carbon degradation rather than lipid synthesis\textsuperscript{34, 35}, leaving the mechanism of lipid synthesis to be further explored.

Transcriptomic activities of Asgard archaea in different niches of coastal sediments

Gene expression patterns derived from metatranscriptomic analysis have been used in a number of recent studies to deduce active microbial processes in marine sediments, especially in the deep sea\textsuperscript{36-38}. This technique might be constrained by the largely unknown pool size of mRNAs maintained in endospores and dormant cells\textsuperscript{39}, although the former seems to be irrelevant for Asgard archaea. Through recruiting 16S rDNA/RNA sequences (n=10,448, read length >600bp) belonging to Asgard archaea from public databases, we found that all Asgard phyla were transcriptionally expressed and diversely distributed (with ~92% of Asgard OTUs originated from sediment samples, fig. S11 and table S9). Thus, to better uncover their activities, we used 818,479 transcripts (mRNA) belonging to Asgard archaea including Loki-, Thor-, Hel-, Heimdall- and Gerdarchaeota from mangrove sediments to elucidate their ecological roles in coastal sediments.

Organic carbons in coastal sediments are mainly composed of carbohydrates, amino acids, and lipids\textsuperscript{22, 40}; here, amino acids are the most dominant component of organics in mangrove sediments\textsuperscript{41}. Accordingly, we detected high expression levels of genes encoding extracellular peptidases, ABC transporter, and the enzyme sets for the conversion of amino acids to acetyl-CoA in both surface and subsurface coastal sediments, implying that Asgard archaea might be essential participants in the degradation of these substrates (Fig. 2, fig. S5, and tables S6, S7 and S10). This feature is supported by the high proportion of peptidases in Asgard archaea MAGs (4.1–6.3 % of the functional genes, fig. S8A). We also detected transcripts for ethanol metabolism, suggesting that ethanol might be another substrate or product. Although glucose accounts for 6–
18% of the total organic carbon in mangrove sediments\textsuperscript{41}, it might be not the first nutritional choice for Asgard archaea because no transcript for the key enzyme glucokinase was detected.

Due to its higher energetic efficiency compared to fermentative processes\textsuperscript{43}, aerobic respiration contributes to 50% or more of the total organic matter decomposition in the offshore marine sediments. The expressed gene set for aerobic respiration in Gerdarchaeota, Heimdallarchaeota-AAG and Heimdallarchaeota-MHV\G, including the key transcript of cytochrome $c$ oxidase (belonging to Gerdarchaeota) indicates that these Asgard archaea might participate aerobically in organic matter degradation in surface sediments (Fig. 3). Although co-existing within the same depth layers, Asgard archaea might play phylum specific ecological roles. For example, unlike Heimdallarchaeota-AAG and Heimdallarchaeota-MHV\G, Gerdarchaeota contain and expressed genes for autotrophy and cellulose degradation, but lack the complete gene set for glucose degradation (Fig. 3). Like Gerdarchaeota, other Asgard archaea have the potential to perform anaerobic metabolisms (e.g., acetogenesis) under anoxic conditions (subsurface layers)\textsuperscript{42}. Notably, Helarchaeota-like $mcrA$ gene transcripts found in unbinned scaffolds (e.g., SZ\_4\_scaffold\_203331\_2, fig. S12) highlight the involvement of Helarchaeota in alkane oxidation in coastal sediments, in which ethane and butane might originate from oil-gas seepage or human activities\textsuperscript{43}, are preferentially used as revealed by molecular modelling and dynamics studies (fig. S13 and Supplementary Results).

Previous studies suggested that Asgard archaea could account for up to 50% of the total prokaryotes in some marine sediments\textsuperscript{44}; and attribute to 40% of the total archaeal sequences in coastal sediments\textsuperscript{44, 45}. The discovery of novel lineages, as well as the discovery of co-occurring diverse lineages in one vertical biosphere in this study, may elevate their relative abundance and ecological roles in natural environments. Therefore, we propose that Asgard archaea might be essential archaeal lineages for organic carbon degradation in coastal sediments, similar to previously proposed roles for Bathyarchaeota\textsuperscript{45} and Thermoprofundales\textsuperscript{46}, which also support the
notion that Asgard archaea are critical participants for organic matter utilization in coastal sediments\textsuperscript{44, 47}.

\section*{Conclusion}

Currently, the research of the diversity and ecology of Asgard archaea is still in its infancy. Contrasting previous studies that Asgard archaeal MAGs covered not more than three phyla\textsuperscript{3, 4, 14}, our study features a wider taxonomic breadth, as we obtained almost all known Asgard phyla in the coastal sediments, and additionally found a new Asgard archaea phylum. Metabolic comparison and transcriptomic evidence suggest divergent ecological roles and niches for different Asgard phyla but they shared key transcripts involved in the degradation of specific compartments of organic matter (e.g., peptides and amino acids). Thus, considering their high relative abundance\textsuperscript{44, 45} and ubiquitous distribution in the coastal area, we infer that Asgard archaea are important players for organic matter utilization\textsuperscript{45, 48}. However, their contribution to the coastal sediment carbon budget remains to be further examined. Overall, the metabolic features, transcript evidence, and their global distribution imply that Asgard archaea are essential players in carbon cycling of coastal sediments.

\section*{Materials and Methods}

\textbf{Sediment sample collection and processing}

Samples for metagenome analysis were collected from the coastal sediment (i.e., mangrove, mudflat and seagrass sediments) of China and Helgoland coastal mud area during the RV HEINCKE cruise HE443 (table S1). They were sampled using custom corers, sealed in plastic bags in duplicates, stored in sampling box with ice bags, and transported to the lab within 4 hours. The physiochemical parameters of the samples were determined as previously described\textsuperscript{49}. Samples for RNA extraction were preserved in RNALater (Ambion, Life Technologies). For each
sample, 10 g of sediment each was used for DNA and RNA isolation with the PowerSoil DNA Isolation Kit (MO BIO) and RNA Powersoil™ Total RNA Isolation Kit (QIAGEN), respectively. The rRNA genes were removed from the total RNA using the Ribo-Zero rRNA removal kit (Illumina, Inc., San Diego, CA, USA) and the remaining mRNA was reverse-transcribed. DNA and cDNA were sequenced using an Illumina HiSeq sequencer (Illumina) with 150-bp paired-end reads at BerryGenomics (Beijing, China). Metatranscriptomic reads were quality-trimmed using Sickle (version 1.33)\(^{50}\) with quality score ≥25, and the potential rRNA reads were removed using SortMeRNA (version 2.0)\(^{51}\) against both the SILVA 132 database and the default databases (E-value cutoff ≤1e-5).

**Metagenomic assembly, genome binning and gene annotation**

Raw metagenomic DNA reads of the coastal sediments were dereplicated (identical reads) and trimmed using Sickle (version 1.33)\(^{50}\) with the option “-q 25”. Paired-end Illumina reads for each sample were assembled *de novo* using IDBA-UD (version 1.1.1)\(^{52}\) with the parameters “-mink 65, -maxk 145, -steps 10”. Scaffolds were binned into genomic bins using a combination of MetaBAT\(^{20}\) and Das Tool\(^{21}\). Briefly, twelve sets of parameters were set for MetaBAT binning, and Das Tool was further applied to obtain an optimized, non-redundant set of bins. To improve the quality of the bins (e.g., scaffold length and bin completeness), each Asgard-related bin was remapped with the short-read mapper BWA and re-assembled using SPAdes (version 3.0.0)\(^{53}\) or IDBA-UD (version 1.1.1)\(^{52}\), followed by MetaBAT and Das Tool binning. Asgard MAGs with high contamination were further refined with Anvi’o software (version 2.2.2)\(^{54}\). The completeness, contamination and strain heterogeneity of the genomic bins were estimated by CheckM (version 1.0.7) software\(^{55}\). Anvi’o software (version 2.2.2)\(^{54}\) was applied for pan-genome analysis of Asgard MAGs with the option “--min-occurrence 3”.
Protein-coding regions were predicted using Prodigal (version 2.6.3) with the “-p meta” option\(^5\). The KEGG server (BlastKOALA)\(^57\), eggNOG-mapper\(^58\), InterProScan tool (V60)\(^59\), and BLASTp vs. NCBI-nr database searched on December 2017 (E-value cutoff ≤1e-5) were used to annotate the protein-coding regions. Archaeal peptidases were predicted against the MEROPS database\(^60\), and the extracellular peptidases were further identified using PRED-SIGNAL\(^61\) and PSORTb\(^62\) (table S11).

Phylogenetic analyses of Asgard MAGs

The 16S rRNA gene sequences and a concatenated set of 122 archaeal-specific conserved marker genes\(^63\) were used for phylogenetic analyses of Asgard archaea. Ribosomal RNA genes in the Asgard MAGs were extracted by Barrnap (version 0.3, http://www.vicbioinformatics.com/software.barrnap.html). An updated 16S rRNA gene sequence dataset from reference papers\(^64, 65\) with genome-based 16S rRNA genes were aligned using SINA (version 1.2.11)\(^66\). The 16S rRNA gene sequences maximum-likelihood tree was built with IQ-TREE (version 1.6.1)\(^67\) using the GTR+I+G4 mixture model (recommended by the “TESTONLY” model), with option “-bb 1000”. Marker genes for protein tree were identified using hidden Markov models (HMMs) and were aligned separately using hmmalign from HMMER3\(^68\) with default parameters. The 122 archaeal marker genes were identified using hidden Markov models. Each protein was individually aligned using hmmalign\(^69\). The concatenated alignment was trimmed by BMGE with flags “-t AA -m BLOSUM30”\(^70\). Then, maximum-likelihood trees were built using IQ-TREE with the best-fit model of “LG+F+R10” followed by extended model selection with FreeRate heterogeneity and 1000 times ultrafast bootstrapping. The final tree was rooted with the DPANN superphylum and Euryarchaeota.

Metabolic pathway construction
Potential metabolic pathways were reconstructed based on the predicted annotations and the reference pathways depicted in KEGG and MetaCyc. Metatranscriptome data from mangrove and mudflat sediments of Shenzhen Bay (table S1) were analyzed to clarify the transcriptomic activity of Asgard archaea. The abundance of transcripts for each gene was determined by mapping all non-rRNA transcripts to predicted genes using BWA with default setting. Normalized expression was expressed in transcript per million units (TPM), followed by normalization by genome number of each phylum.

**ESP identification**

As predicted by prodigal (v2.6.3) with default parameters, genes of Gerdarchaeota were searched against InterPro and eggNOG databases to gain the IPRs and arCOGs. The list of those annotations in Zaremba-Niedzwiedzka et al (2017) was searched for in the Gerdarchaeota bins. We also manually inspected the IPRs and arCOGs only present in Gerdarchaeota. MAFFT-linsi and trimAl (-gappyout) were used to align and trim the protein sequences. IQ-TREE (version 1.6.1) was used to infer phylogeny of under best-fit models with 1000 ultrafast bootstraps with SH-aLRT test values.
References and Notes


34. Yokobori, S.-i., Nakajima, Y., Akanuma, S. & Yamagishi, A. Birth of archaeal cells: molecular phylogenetic analyses of G1P dehydrogenase, G3P dehydrogenases, and glycerol


Acknowledgments

**General:** We thank Dr. Nidhi Singh for her suggestions in molecular modeling. We thank the captain, crew and scientists of R/V HEINCKE expeditions HE443.

**Funding:** This research was financed by the National Natural Science Foundation of China (No. 91851105, 31622002, 31970105, 31600093, and 31700430), the Science and Technology Innovation Committee of Shenzhen (Grant No. JCYJ20170818091727570), the Key Project of Department of Education of Guangdong Province (No. 2017KZDXM071), the China Postdoctoral Science Foundation (No. 2018M633111), the DFG (Deutsche Forschungsgemeinschaft) Cluster of Excellence EXC 309 “The Ocean in the Earth System - MARUM - Center for Marine Environmental Sciences” (project ID 49926684) and the University of Bremen.

**Author contributions:** M.L., M.C. and Y.L. conceived this study. M.C. analyzed the 16S rRNA data, metagenomic data and metatranscriptomic data. Y.L. collected samples and analyzed the metagenomic data. X.Y., M.W.F., T.R.H., R.N., and A.K. provided metagenomic data. Z.Z. provided support for diversity analysis. M.C., Y.C.Y., J.P. and Z.Z. prepared the DNA and cDNA for sequencing. W.L. and X.W. analyzed MCR complex protein structure and simulated the binding substrates. M.C., Y.L., and M.L. wrote, and all authors edited and approved the manuscript.

**Competing interests:** The authors declare no conflicts of interest.
Data and materials availability: Archaeal 16S rRNA gene sequences were retrieved from NCBI database, SILVA SSU r132 database, and a reference paper as described in Supplementary Materials and Methods. Public Asgard MAGs were from NCBI database and MG-RAST. The newly obtained Asgard MAGs and metatranscriptomic data are available in NCBI database under the project PRJNA495098 and PRJNA360036.
Figures and Tables

Fig. 1. Phylogenetic positions and ESPs of Asgard archaea reconstructed from coastal sediments. (A), Maximum likelihood tree of Gerdarchaeota MAGs built on a concatenated alignment of 122 archaeal marker genes. The tree was inferred with LG+F+R10 mixture mode in IQ-TREE and rooted with DPANN and Euryarchaeota. Asgard archaea MAGs obtained in this study are marked in bold face. (B), Phylogenetic position of Asgard archaeal 16S rRNA genes. Red colour represents 16S rRNA genes from newly discovered Asgard MAGs, and blue colour represents sequences from reference ones. The tree was built using the IQ-TREE software with GTR+I+G4 mixture mode and rooted with Crenarchaeota. (C), ESPs identified in Gerdarchaeota and other Asgard archaea. Asgard archaea MAGs obtained in this study are marked bold. Colours indicate phylum-level assignment (see 3A).
**Fig. 2. Key potential metabolic pathways of Gerdarchaeota.** Solid lines represent pathway steps present in 8 MAGs assigned to Gerdarchaeota. Dashed lines represent pathway steps not identified. Circles represent transcripts from surface sediments and rectangles represent transcripts from subsurface samples. The relative abundance of the transcripts (Transcripts Per Million reads, TPM) for each gene is marked with different colours. Pathway abbreviations: PPP, pentose phosphate pathway (archaea); CBB, Calvin–Benson–Bassham cycle; EMP, Embden–Meyerhof–Parnas pathway, THMPT WL, tetrahydromethanopterin-dependent Wood–Ljungdahl pathway; TCA cycle, tricarboxylic acid cycle. Detailed metabolic information for the MAGs is available in fig. S5, tables S6, and S7.
**Fig. 3. Ecological niches of Asgard archaea in coastal sediments.** Dashed lines represent pathways with no transcript for the key genes. Detailed information is available in fig. S5, tables S6, and S7.

**Supplementary Materials**

Supplementary methods and results

figures S1 to S16

tables S1 to S12