ISME

ARTICLE

Metabolic potentials of archaeal lineages resolved from metagenomes of deep Costa Rica sediments

Ibrahim F. Farag[1] · Jennifer F. Biddle ◉[1] · Rui Zhao[1] · Amanda J. Martino[2] · Christopher H. House[3] · Rosa I. León-Zayas[1,4]

Abstract

Numerous archaeal lineages are known to inhabit marine subsurface sediments, although their distributions, metabolic capacities, and interspecies interactions are still not well understood. Abundant and diverse archaea were recently reported in Costa Rica (CR) margin subseafloor sediments recovered during IODP Expedition 334. Here, we recover metagenome-assembled genomes (MAGs) of archaea from the CR margin and compare them to their relatives from shallower settings. We describe 31 MAGs of six different archaeal lineages (Lokiarchaeota, Thorarchaeota, Heimdallarchaeota, Bathyarchaeota, Thermoplasmatales, and Hadesarchaea) and thoroughly analyze representative MAGs from the phyla Lokiarchaeota and Bathyarchaeota. Our analysis suggests the potential capability of Lokiarchaeota members to anaerobically degrade aliphatic and aromatic hydrocarbons. We show it is genetically possible and energetically feasible for Lokiarchaeota to degrade benzoate if they associate with organisms using nitrate, nitrite, and sulfite as electron acceptors, which suggests a possibility of syntrophic relationships between Lokiarchaeota and nitrite and sulfite reducing bacteria. The novel Bathyarchaeota lineage possesses an incomplete methanogenesis pathway lacking the methyl coenzyme M reductase complex and encodes a noncanonical acetogenic pathway potentially coupling methylotrophy to acetogenesis via the methyl branch of Wood–Ljungdahl pathway. These metabolic characteristics suggest the potential of this Bathyarchaeota lineage to be a transition between methanogenic and acetogenic Bathyarchaeota lineages. This work expands our knowledge about the metabolic functional repertoire of marine benthic archaea.

Introduction

Marine subsurface sediments are enriched with diverse groups of archaeal lineages, although their distributions, ecological roles, and adaptation strategies are understudied

✉ Jennifer F. Biddle
jfbiddle@udel.edu

[1] School of Marine Science and Policy, University of Delaware, Lewes, DE 19968, USA

[2] Department of Biology, St. Francis University, Lorretto, PA 15940, USA

[3] Department of Geosciences, Pennsylvania State University, University Park, PA 16802, USA

[4] Department of Biology, Willamette University, Salem, OR 97301, USA

[1–4]. Metagenomic sequencing and single cell genomics have enabled the discovery of novel organisms, the elucidation of new metabolisms, the expansion of known lineages, and the redefinition of portions of the tree of life [5–13]. However, there is still a paucity of genomes resolved from the deep marine subsurface, meaning niche specific adaptations in deep biosphere are not yet well understood.

Recently, the tree of life has been expanded with the discovery of the Asgard superphylum, a deeply-branching monophyletic group thought to be some of the closest relatives to the eukaryotic branch of life [10, 14]. Genome analyses of Asgard archaea noted high abundances of eukaryotic homologs and have suggested diverse metabolic functions extending from an autotrophic lifestyle, primarily dependent on carbon fixation via Wood–Ljungdahl pathway and acetogenesis, to a heterorganotrophic lifestyle consuming proteins and aliphatic hydrocarbons, using methyl-CoM reductase-like enzymes [15–17]. Asgard members have been proposed to be engaged in symbiotic partnerships involving syntrophic transfers of hydrogen and electrons following the "reverse flow model" [15].

Members of another archaeal phylum in the deep subsurface, Bathyarchaeota, are characterized by their wide metabolic range enabling heterotrophic scavenging of proteins, carbohydrates, short chain lipids, and other reduced compounds as substrates as well as their methane-metabolizing potential [18]. However, bathyarchaeotal genomes have also suggested the potential for carbon fixation and acetogenesis [19]. The evolutionary path describing the acquisition of both methanogenesis and acetogenesis pathways in Bathyarchaeota remains unresolved [19, 20].

Deep sediment from the Costa Rica (CR) margin subseafloor, sampled during the International Ocean Discovery Program (IODP) Expedition 334, was recently shown to host abundant archaea [21]. Here we examine in detail the archaeal genomes recovered from metagenomes from the CR margin and compare them to their relatives recovered from shallower sites. In this study, we report 31 archaeal metagenome-assembled genomes (MAGs) belonging to six different archaeal lineages (Lokiarchaeota, Thorarchaeota, Heimdallarchaeota, Bathyarcheota, Thermoplasmatales, and Hadesarchaea). We thoroughly analyze representative MAGs of two novel archaeal lineages belonging to Lokiarchaeota and Bathyarchaeota. Our analysis suggests that Lokiarchaeota genomes encode for genes to process and degrade aliphatic and aromatic hydrocarbon anaerobically. We also describe a novel Bathyarchaeota genome, which lacks methyl coenzyme M reductase (MCR) complex and possesses a noncanonical acetogenic pathway linking methylotrophy to acetogenesis via the methyl branch of Wood–Ljungdahl pathway. Lastly, we integrate genomic and thermodynamic modeling to underline the ecological and physiological conditions that could drive the syntrophic interactions among CR_Asgard and use phylogenetic inference to elucidate the development of noncanonical acetogenesis in Bathyarchaeota in CR deep sediments.

Materials and methods

Site information and sample collection

Samples used in this study were collected under aseptic conditions from Sites U1378 and U1379 of CR margin during IODP Expedition 334. The sample depths are in the range of 2–93 m below seafloor (mbsf). Detailed site descriptions were previously reported in the IODP Proceedings for Expedition 334 [21, 22].

DNA extraction and sequencing and metagenome-assembled genome binning

DNA extraction and metagenomic sequencing have been described previously [21] and data have been deposited at NCBI GenBank SRA under project PRJEB11766. Metagenomic reads were quality trimmed using Nesoni following default parameters and applying q20 for quality score (www.vicbioinformatics.com/software.nesoni.shtml). Quality-controlled reads in individual samples were assembled separately using IDBA-UD [23] with default settings. Contigs longer than 1 kb were binned into MAGs using MaxBin V2.2.7 [24], further curated manually using VizBin [25], through filtering outlier scaffolds not falling within the same GC%, and differential coverage levels across different CR datasets. The quality and completeness of the MAGs were assessed using CheckM (v.1.0.7) [26]. MAGs from all five metagenomes were dereplicated using dRep (version v2.0.5 with ANI >99%) [27] and the most complete MAG per taxon was selected for downstream analyses. Phylogenetic placements of archaeal MAGs were further determined by Phylosift [28] analysis based on single copy marker genes and phylogenomic analysis using 16 concatenated ribosomal proteins (see description below). Assembled contigs larger than 1 kb were annotated using PROKKA [29]. Encoded proteins were predicted using Prodigal v2.6.3 with the default translation table (Table 11) was applied [30].

Concatenated ribosomal protein phylogeny

A maximum-likelihood tree was calculated based on the concatenation of 16 ribosomal proteins (L2, L3, L4, L5, L6, L14, L15, L16, L18, L22, L24, S3, S8, S10, S17, and S19) using IQ-Tree (v1.6.6) [31] (located on the CIPRES web server) [32]. References sequences used were collected [11] with the addition of more representatives of the Bathyarchaeota and Asgard lineages collected from all MAGs available in JGI_IMG and NCBI database. Evolutionary distances were calculated based on best fit substitution model (VT + F + R10), and single branch location was tested using 1000 ultrafast bootstraps and approximate Bayesian computation [31, 33], branches with bootstrap support >80% were marked by black circles.

Metabolic reconstruction and functional annotation

Predicted proteins from all MAGs were screened using HMMsearch tool against custom HMM databases representing the key genes for specific metabolic pathways [34]. The completion of the pathways was assessed through querying the predicted proteins against KEGG database using BlastKoala tool [35]. Carbohydrate-active enzymes were identified using dbCAN-fam-HMMs (v6) database [36]. Cellular localizations of predicted proteins were identified using SignalP v5.0 [37]. Proteases, peptidases, and peptidase inhibitors were identified using USEARCH-ublast tool [38] against the MEROPS database v12.1 [39].

Transporters were identified using USEARCH-ublast tool [38] against the TCDB database [40]. ESPs were detected using Interpro v75.0 [41]. Phylogenetic distributions of the predicted proteins in each bin were detected through comparing these proteins against the NCBI (nr) protein database via the DarkHorse software [42].

Functional protein-based trees

All functional protein-based trees were built by aligning the query protein sequences to the reference sequences belonging to the same protein family using Muscle v3.8.31 [43]. Reference sequences were collected from AnnoTree [44] using the corresponding KEGG entry as search keyword. Aligned sequences were manually curated using Geneious v9.0.5 (https://www.geneious.com). The phylogenetic trees were computed using IQ-TREE (v1.6.6) [31], through the CIPRES web server [45] and the evolutionary relationships were described using the best fit model. Branch locations were tested using 1000 ultrafast bootstraps and approximate Bayesian computation [31, 33].

Thermodynamic calculations

Initial Gibbs free energy (ΔrG$'$) calculations were performed for the redox reactions proposed in the CR Lokiarchaeota and their potential partners using eQuilibrator [46] applying pH 8, similar to the approach reported in [47], and reactant concentrations of 1 mM each.

We further confirmed if the coupled redox reactions proposed for the novel CR Lokiarchaeota genomes were feasible, by calculating the Gibbs free energy of five redox reactions (Table S1) under the near in situ conditions of CR sites U1378 and U1379, following the method described in ref. [48]. Gibbs free energy was calculated using the equation:

$$\Delta G_r = \Delta G_r^0 + RT \ln Q_r,$$

where ΔG_r^0 and Q_r refer to the standard molar Gibbs energy and the reaction quotient of the indicated reaction, respectively, R represents the gas constant, and T denotes temperature in Kelvin. In this study, ΔG_r^0 was calculated using the thermodynamic data of standard Gibbs free energy of formation of each species and corrected to near in situ pressure and temperature (4 °C), using the R package $CHNOSZ$ [49]. Q_r stands for the reaction quotient, which can be calculated with the relation

$$Q_r = \Pi\left(a_i^{\nu i}\right),$$

where a_i is the activity of species i and ν_i is its stoichiometric coefficient. a_i is the product of chemical species concentration [i] and its activity coefficient γ_i, which was computed as a function of temperature and ionic strength by using an extended version of the Debye–Huckel equation [50]. Because most of the reactant concentrations were hard to measure or were below detection limits, we assumed 0.1 µM for the concentration of NO_3^-, NO_2^-, SO_3^{2-}, and H_2S. We benzoate ($C_7H_5O_2^-$) concentrations in the range of 0.0001–100 µM to explore the feasibility of these reactions over a wide range of substrate concentrations.

Results

Diversity of the dominant archaeal lineages based on 16S rRNA gene analyses

High abundances of archaea across sediment samples collected from five depths of the CR Margin (2, 32, and 93 mbsf from Site 1378 and 22 and 45 mbsf from Site 1379 of IODP Expedition 334), were previously reported [21]. Metagenomic data from these samples were assembled and examined for small subunit ribosomal genes. A total of 126 16S rRNA gene sequences were recovered from the metagenome assemblies. Sequences affiliated to archaea (31 sequences representing 25% of the total) represented a wide diversity of lineages, including Bathyarchaeota (14 sequences representing 11% of the total 16S rRNA gene sequences), Thermoplasma (5 sequences representing 4%) and Lokiarchaeota (4 sequences representing 3%) (Fig. S1).

General genomic features of the abundant archaeal lineages

Across the five depths analyzed, 31 different draft archaeal MAGs were recovered (Table 1). Genomic analyses were only performed on the 11 MAGs showing high completeness and low contamination (above 60% completeness and below 5% contamination). Overall, completeness varied from 32 to 99% with an average of 50% and contamination varied from 0 to 10% with an average of 8% (Table 1). Genome qualities were further assessed by comparing their predicted proteins against the NCBI (nr) database to evaluate the extent of phylogenetic consensus within the binned genomes (Fig. 2a). Overall, the taxonomic affiliations of the majority of the predicted proteins (60–75%) in each MAG agreed with their respective phylogenetic group, except CR_12 in which only 15% of the encoded proteins assigned to Heimdallarchaeota and was excluded from further analysis.

Phylogenetic placement of the draft genomes was determined using 16 ribosomal proteins (Fig. 1) [11]. The recovered MAGs were affiliated to six different phylogenetic lineages, namely, Lokiarchaeota, Thorarchaeota,

Table 1 Details of the MAGs constructed from the Costa Rica margin.

Bin_ID	Dataset	Phylogenetic affiliation	Completeness	Redundancy	Strain heterogeneity	Genome size (Mbp)	No. of scaffolds	GC	Coverage
CR_01	1378_2mbsf	Bathyarchaeota	38.71	10.09	59.09	0.62	335	39	155.6
CR_02	1378_2mbsf	Bathyarchaeota	45.64	5.61	16.67	1.05	225	38.7	29.9
CR_03	1378_2mbsf	Hadesarchaea	74.86	10.67	50	0.98	333	40.2	19.3
CR_04	1378_2mbsf	Bathyarchaeota	60.62	10.08	5.88	1.09	368	42.7	16.05
CR_05	1378_2mbsf	Bathyarchaeota (RPL2)	65.2	10.26	0	1.92	900	40.9	11.4
CR_06	1378_32mbsf	Asgard (Lokiarchaeota)	86.78	1.87	33.33	2.41	1110	42.7	216.64
CR_07	1378_32mbsf	Asgard (Lokiarchaeota)	47.66	5.43	63.64	1.525	829	41.9	72.54
CR_08	1378_32mbsf	Asgard (Lokiarchaeota)	61.06	9.65	20	3.54	693	42.9	71.06
CR_09	1378_32mbsf	Bathyarchaeota	41.64	0	0	3.6	89	39.3	30.63
CR_10	1378_32mbsf	Asgard (Lokiarchaeota)	32.81	10.6	23.33	2.91	1865	43.3	24.53
CR_11	1378_32mbsf	Asgard (Heimdallarchaeota)	94.39	9.06	0	3.47	167	38.6	23.11
CR_12	1378_32mbsf	Asgard (Thorarchaeota)	32.05	1.08	0	1.49	1019	38.8	10.48
CR_13	1379_22mbsf	Bathyarchaeota	35.05	0	0	0.3	118	39.2	116.39
CR_14	1379_22mbsf	Bathyarchaeota	95.64	4.15	0	2.49	140	39.4	77.12
CR_15	1379_22mbsf	Bathyarchaeota	92.61	3.12	0	2.11	130	39	47.44
CR_16	1379_22mbsf	Bathyarchaeota	36.14	0	0	2.99	59	39.4	47.78
CR_17	1379_22mbsf	Bathyarchaeota	71.03	6.54	0	0.88	28	39.4	26.16
CR_18	1379_22mbsf	Bathyarchaeota (RPL2)	42.3	5.61	0	1.13	103	40.8	24.29
CR_19	1379_22mbsf	Bathyarchaeota	35.84	5.99	0	1.03	340	41	13.83
CR_20	1379_22mbsf	Bathyarchaeota	58.35	2.8	0	1.01	184	41	18.65
CR_21	1379_22mbsf	Bathyarchaeota	38.51	1.94	0	0.58	249	41.8	17.31
CR_22	1379_22mbsf	Bathyarchaeota	42.88	9.88	15.38	1.31	344	38.6	19.48
CR_23	1379_22mbsf	Thermoplasmata	62.44	2.4	0	1.96	284	43.1	19.88
CR_24	1379_22mbsf	Bathyarchaeota	53.74	6.31	0	2.36	556	41.8	17.98
CR_25	1379_22mbsf	Bathyarchaeota (RPS3)	56.65	1.94	0	1.68	531	39.6	13.92
CR_26	1379_22mbsf	Hadesarchaea	50.96	7.79	0	0.58	329	39.6	14.9
CR_27	1379_45mbsf	Bathyarchaeota	79.77	5.49	33.33	1.24	404	39.8	63.85
CR_28	1379_45mbsf	Bathyarchaeota	32.78	6.15	0	1.25	296	39.1	45.38
CR_29	1379_45mbsf	Bathyarchaeota	56.96	0.97	100	0.7	154	37.7	40.92
CR_30	1379_45mbsf	Bathyarchaeota	49.54	3.4	28.57	4	191	40	21.23
CR_31	1379_45mbsf	Hadesarchaea	31.55	0.64	33.33	3.75	176	40.2	25.74

Heimdallarchaeota, Bathyarcheota, Thermoplasmatales, and Hadesarchaea.

Distribution of eukaryotic signature protein (ESP) homologs in Costa Rica archaea MAGs

All archaeal MAGs recovered from CR sediments contained eukaryotic signature proteins (ESPs). These ESPs include protein homologs dedicated for information processing, trafficking machinery, ubiquitin system, cell division, and cytoskeleton formation. The Asgard archaea genomes recovered from CR, e.g., Heimdallarchaeota and Lokiarchaeota MAGs (CR_06, CR_11), showed high numbers of eukaryotic homologs and covered broad classes of ESPs (Fig. 2b). However, ESPs were also detected in the Bathyarchaeota, Hadesarchaeota, and Thermoplasmata MAGs recovered from CR (Fig. 2b). Yet, only the Asgard genomes have homologs for cell division and cytoskeleton.

Genomic evidence of CR_Asgard hydrocarbon utilization

We screened all of the MAGs, assembled metagenomes, and raw reads for the presence of genes encoding for MCR complex to assess the role of the archaeal members in mediating the degradation of hydrocarbon compounds (Fig. 3a) that are abundant in CR margin [51]. MCR complex genes (*mcrABCDG*) were completely absent from both the MAGs and the entire metagenomes, which suggests that short chain alkanes are not oxidized using MCR complex in CR sediments. All metagenomic reads and MAGs were screened for possible alternative hydrocarbon degradation pathways using custom HMM searches that specifically targeted key metabolic genes for aliphatic and aromatic hydrocarbon degradation pathways (Fig. 3a). Multiple pathways were successfully identified including glycyl-radical enzymes related genes coupled with n-alkane succinate synthase (AssA) and benzylsuccinate synthase

Fig. 1 Phylogenetic placement of the Costa Rica archaeal draft genomes (genomes of this study are highlighted in red). The maximum-likelihood phylogenetic tree was calculated based on the concatenation of 16 ribosomal proteins (L2, L3, L4, L5, L6, L14, L15, L16, L18, L22, L24, S3, S8, S10, S17, and S19) retrieved from the Costa Rica archaeal genomes and 231 reference archaeal genomes representing 13 different archaeal phyla. The relationships were inferred using the best fit substitution model (VT + F + R10) and nodes with bootstrap support >80% were marked by black circles. Scale bar indicates substitutions per site. The tree is available with full bootstrap values in Newick format in the Supplementary Data.

(BssA), which activates *n*-alkanes and mono-aromatic compounds, respectively, by forming C–C bond between these compounds and fumarate to form hydrocarbon adducts [45, 52]. Benzylsuccinate synthesis is the initial step for aromatic hydrocarbon mineralization, in which benzylsuccinate is converted to benzoyl-CoA [52]. Interestingly, the potential for ATP-dependent Benzoyl-CoA reductase (BCR) complex utilization was identified in some CR Asgard members (Lokiarchaeota, Thorarchaeota, and Heimdallarchaeota) (Fig. 3). This reaction can dearomatize
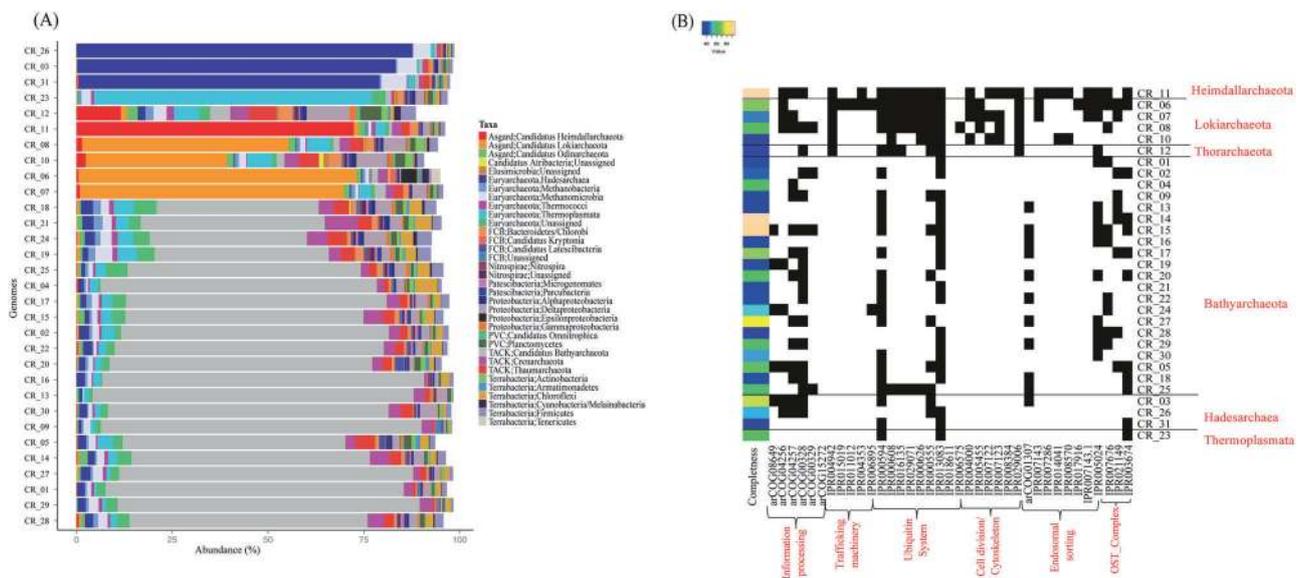
Fig. 2 Phylogenetic distribution and the occurrence of eukaryotic-specific proteins in archaeal MAGs. a Phylogenetic distributions of the predicted proteins encoded by the archaeal MAGs, the phylogenetic assignments were performed through comparing the proteins to the NCBI (nr) proteins database via the DarkHorse software [42].

b Heat map shows the presence (black)/absence (white) patterns of eukaryotic homologs in the recovered archaea MAGs. The color along the left side shows the completeness of the genome bin, as less complete bins would be expected to contain fewer homologs. Only the Asgard archaea contain cell division/cytoskeleton homologs.

the benzoyl-CoA to dienoyl-CoAs as the first step in aromatic hydrocarbon degradation and then couple this reaction with a beta-oxidation pathway to ultimately produce acetyl-CoA, similar to the mechanism previously reported in the denitrifying bacteria *Thauera aromatica* [53]. Since Lokiarchaeota CR_06 had the least contamination levels <2%, we used this MAG to verify the presence of aromatic hydrocarbon degradation function mediated by BCR complex in Asgard phyla. Phylogenetic analysis of the BCR subunit B recovered from Lokiarchaeota *CR_06* showed their affiliation to class Bzd, which is composed of four subunits (BzdONPQ) (Fig. 3c). This BCR type was originally discovered in *Betaproteobacteria, Azoarcus evansii*, and the ones detected in CR_06 are closely related to BCRs detected in different *β, δ proteobacteria*, and other archaeal lineages (e.g., Lokiarchaeota, Bathyarchaeota, and Archaeoglobus) (Fig. 3c) [54]. Subunits P and Q have ATP binding/ATPase functional domains and are linked together by a [4Fe–4S] cluster. Reduced ferrodoxins transfer electrons to the CoA ester-binding domains (subunits O and N), which catalyzes the cleavage of benzoyl-CoA aromatic ring and yields dienoyl CoA product (Fig. 3a) [54]. In CR_06, the four subunits of BCR complex were located in one contiguous operon (CR_06-contig-100_3495), however due to the high fragmentation of the genomes no reliable phylogenetic marker genes were found in the adjoining genomic neighborhoods. A plausible explanation for the high fragmentation levels of the CR_06 MAG (number of scaffolds = 1110), even though the genome exhibited high coverage levels >200×, is that there are high levels of

intralineage strain heterogeneity. This is confirmed by the ANI values (ANI = 70–80%) between MAGs of the same phylogenetic group and disparities in their coverage levels (Tables 1 and S3).

The affiliations of the BCR complex related proteins, BCRA and BCRB in Lokiarchaeota CR_06 were confirmed by the following observations: (1) contigs encoding for the four subunits of BCR complex fall within the same coverage and GC% range of the rest of the contigs in the same MAG (Fig. S2); (2) phylogenetic trees of the BzdO and Q protein sequences detected in the CR_06 were placed as siblings to BCR sequences belonging to Lokiarchaeota MAGs recovered from deep sediment from south China Sea (GenBank accession numbers TET61664 and TKJ21612) (Figs. 3c and S3) [55].

Fate of the degraded aromatic hydrocarbons and syntrophic interactions

CR_ Lokiarchaeota (CR_06) as well as other Asgard archaea in the CR margin could potentially grow heterotrophically on aromatic hydrocarbons. Though, the absence of genes encoding for different types of cytochrome oxidase and anaerobic respiration from the genome content of the Asgard MAGs (CR_06, 07, 08, 11, and 12) indicate their inability to completely mineralize these hydrocarbons to CO_2 and H_2O, which would allow a high energy yield. They encode for the genes mediating the fermentation of these organic macromolecules to acetate and other reduced products. However, this is thermodynamically unfavorable
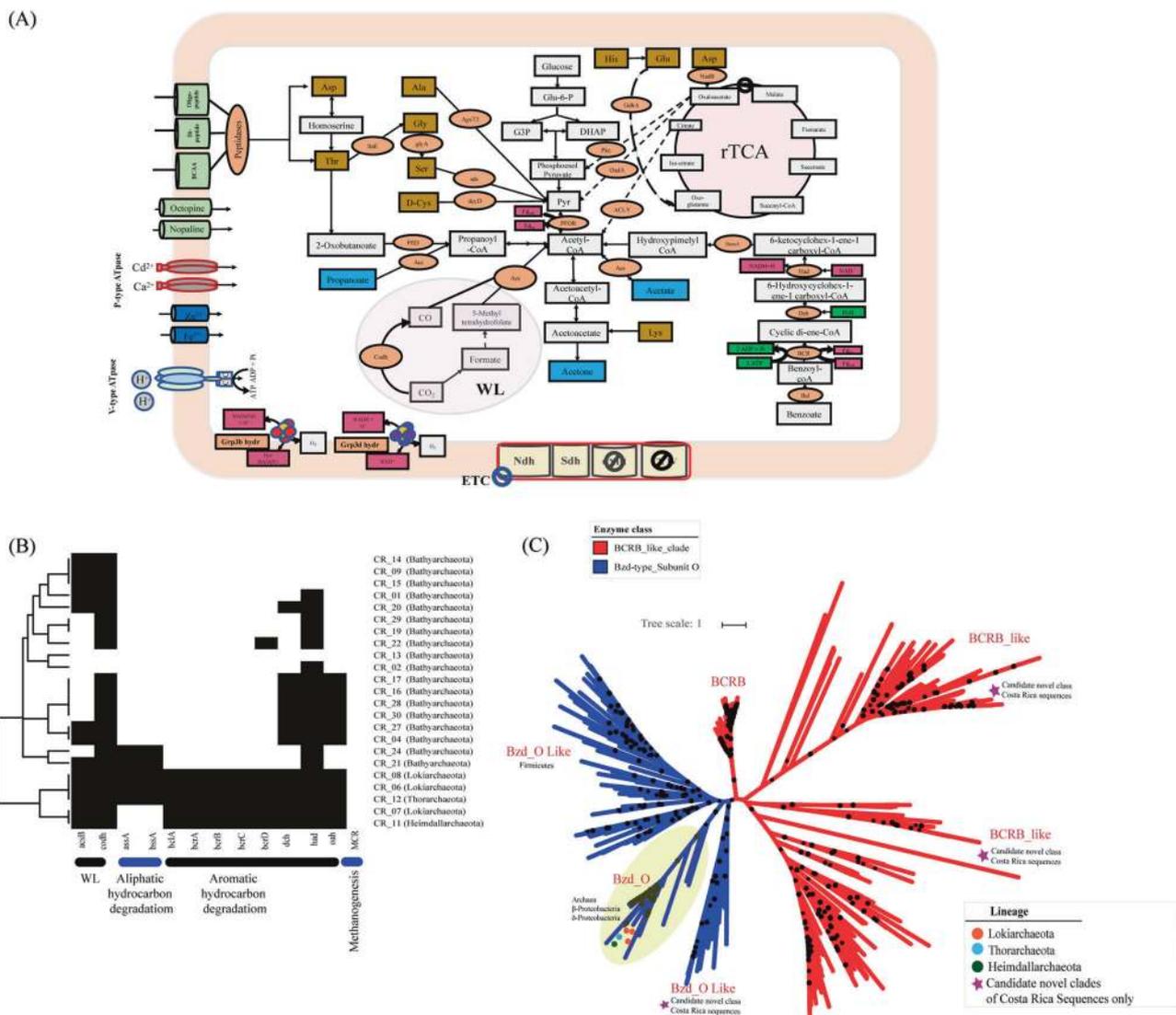
Fig. 3 Lokiarchaeotal metabolism and degradation of hydrocarbons. a Metabolic reconstruction of the CR_Lokiarchaeota bin (CR_06). Central metabolic pathways are shown in gray boxes, carbon fixation pathways (WL and rTCA cycles) are shown in pink, electron transport chain (ETC) proteins are shown in yellow, fermentation products are shown in blue boxes, amino acids are shown in brown boxes, enzymes and enzyme complexes are shown in orange circles, energy carriers are shown in red, energy molecules are shown in dark green, metabolite and amino acid transporters are shown in light green, and cation transporters are shown in dark blue. **b** Distribution patterns of the hydrocarbon degradation pathways among the constructed CR_Archaea genomes. Genomes were clustered based on presence (black)/absence (white) profiles using Euclidean distance and average linkage method. *X*-axis represents enzymes included in the analysis, bottom bars are the pathways that these enzymes represent, and *Y*-axis includes bin names/phylogenetic affiliations. **c** Maximum-likelihood tree of the benzoyl-CoA reductase subunit B. The tree was calculated using the best fit substitution model (VT + F + R7) that describes the evolutionary relationships between BCRB families. The tree was made using reference sequences under the KEGG entry (K04113) collected from AnnoTree [44] and branch location was tested using 1000 ultrafast bootstraps and approximate Bayesian computation, branches with bootstrap support >80% were marked by black circles. Blue and red clades highlight sequences belong to Bzd_O and BCR_B subfamilies, respectively. Scale bar indicates substitutions per site. Sequences from CR_Lokiarchaeota bins was marked with red circles, CR_Thorarchaeota bin was marked with light blue, and CR_Heimdallarchaeota was marked with green. Candidate novel clades present in the Costa Rica metagenomic datasets were marked with purple stars. The tree is available with full bootstrap values in Newick format in the Supplementary Data.

under CR conditions with a positive $\Delta rG'$ value ($\Delta rG' = 196.3$ [kJ/mol]) (Fig. 3 and Table S1), suggesting that these genes are selectively maintained. Due to the genome incompleteness (86% complete), we cannot rule out the possibility for the presence of complete aromatic hydrocarbon mineralization pathways using one or more oxidized substrates as electron sinks (Table 1). More likely, however, Lokiarchaetoa are gaining energy through syntrophic interactions with partners capable of oxidizing the biodegradation intermediates.

Here, we inferred the identity of the potential syntrophic partners under marine subsurface conditions by comparing all the possible metabolic and thermodynamic scenarios, gauging each scenario based on the presence of the metabolic pathways in our metagenomic datasets and the thermodynamic feasibility under each condition (Fig. 3b, Tables S1 and S6). We calculated the Gibbs free energy of coupled reactions under a wide range of substrate concentration conditions: reaction 1–5 (Table S1, Fig. S4). The best conditions suggest that the degradation of benzoate (the central metabolic intermediate in aromatic hydrocarbon degradation pathways) potentially occurs under the following metabolic conditions (a, b, and c): (a) benzoate mineralization to CO_2 and H_2O coupled with nitrite reduction to ammonia ($\Delta rG' = -1206.3$ [kJ/mol]); (b) benzoate mineralization to CO_2 and H_2O coupled with sulfite reduction to hydrogen sulfide ($\Delta rG' = -373.6$ [kJ/mol]); and (c) benzoate mineralization to CO_2 and H_2O coupled with nitrate reduction to nitrite ($\Delta rG' = -119.9$ [kJ/mol]). Coupling with sulfate reduction is not thermodynamically favorable (Table S1, Figs. S4 and S5).

The type and complexity of the exchanged substrates are another key factor that may shape the syntrophic relationship and the identity of the syntrophic partners of Lokiarchaeota. The presence of genes encoding for membrane bound electron bifurcating classes of [NiFe] hydrogenases, groups 3b and 3d, which couple the oxidation of $NADH^+$ and $NADPH^+$ with H_2 evolution (Fig. S6) [56] also suggest there is syntrophic exchange of hydrogen between *Asgard* and their partners. In addition, we located genes encoding for β-oxidation enzymes (enoyl-CoA hydratase, acyl CoA dehydrogenase, and acetyl CoA acetyl transferase) and various fermentation pathways (acetate and formate) in Lokiarchaeota CR_06, which suggests that short chain fatty acid and different fermentation products could also be syntrophically exchanged (Table S4). The diverse nature of substrates could facilitate the interactions between a broader range of partners of diverse metabolic capabilities and support the conclusions driven from our thermodynamic calculations.

It is worth noting that efficient substrate and electron exchange between syntrophic partners require the presence of either biological conduits (e.g., type IV pili or flagella) or some sort of electron shuttle allowing extracellular electron transfers (e.g., multiheme cytochromes) [57, 58]. Hence, we screened CR_06 for these mechanisms and identified two candidate mechanisms for interspecies substrate and electron exchange. First, flagellar proteins are encoded by CR_06, suggesting flagella may be a potential structure mediating interspecies interactions. Second, CR_06 harbors a gene encoding for an oxidoreductase belonging to electron transfer flavoprotein-quinone oxidoreductase (CR_06_contig-100_4953_2), ETF-QO/FixC family, which potentially mediates the transfer of electrons across membranes.

Other metabolic features of Lokiarchaeota (CR_06)

The genomic analysis of Lokiarchaeota (CR_06) MAG also suggests versatile catabolic capacities potentially targeting detrital proteins and short chain fatty acids (e.g., propanoate, butyrate), which are abundant in benthic marine sediments [59, 60]. CR_06 MAG has a relatively large number of peptidases encoding genes (92 peptidases/1 Mbp) with diverse catalytic residues (e.g., aspartic, metallo, serine, etc.), which potentially degrade detrital proteins (Fig. S7). It also contains genes encoding for various classes of facilitated and active transporters, which are dedicated to shuttle oligo/dipeptides and single amino acids (e.g., polar and branched chain amino acids) across the cell membrane. Also, CR_06 encoded for enzymes enabling the utilization of wide range of amino acids (e.g., aspartate, threonine, alanine, glycine, serine, cysteine, histidine, glutamine, and lysine), which are then channeled to the central metabolic pathways and ultimately produce energy via fermentation (Fig. 3a).

MAG CR_06 suggested the capacity to breakdown short chain fatty acids e.g., propanoate, oxobutanoate as other potential substrates. Both propanoate and oxobutanoate are converted to propanoyl CoA via formate acetyl transferase and acetyl synthase, respectively. Then, the resulting propanoyl CoA is converted to acetyl-CoA via the malonyl-CoA pathway. CR_06 as well as other Asgard phyla showed different autotrophic capacities enabling carbon fixation using both the Wood–Ljungdahl (WL) pathway and the reverse tricaboxylic acid (rTCA) cycle (Fig. 3a).

Metabolic features of CR_Bathyarchaeota

The other main lineage of Archaea in these sediments is Bathyarchaeota. Bathyarchaeota CR_14 is described in detail since it has the highest quality at >95% complete and 4% contaminated. Phylogenomic analysis showed that CR_14 is clustered together with other Bathyarcaheota MAGs from CR in two distinct clades within the Bathyarchaeota phylum (Fig. 1). Also, relatively low similarity scores were observed between CR_14 and other reference Bathyarchaeota MAGs, which suggest that CR_14 belongs to a novel Bathyarchaeota class (Table S2).

Metabolic analysis revealed that CR_14 harbors genes encoding for incomplete methylotrophic methanogenesis pathway, yet is missing the key genes encoding for MCR complex (*mcrABCDG*). The absence of this complex suggests that CR_14 is incapable of methanogenesis and
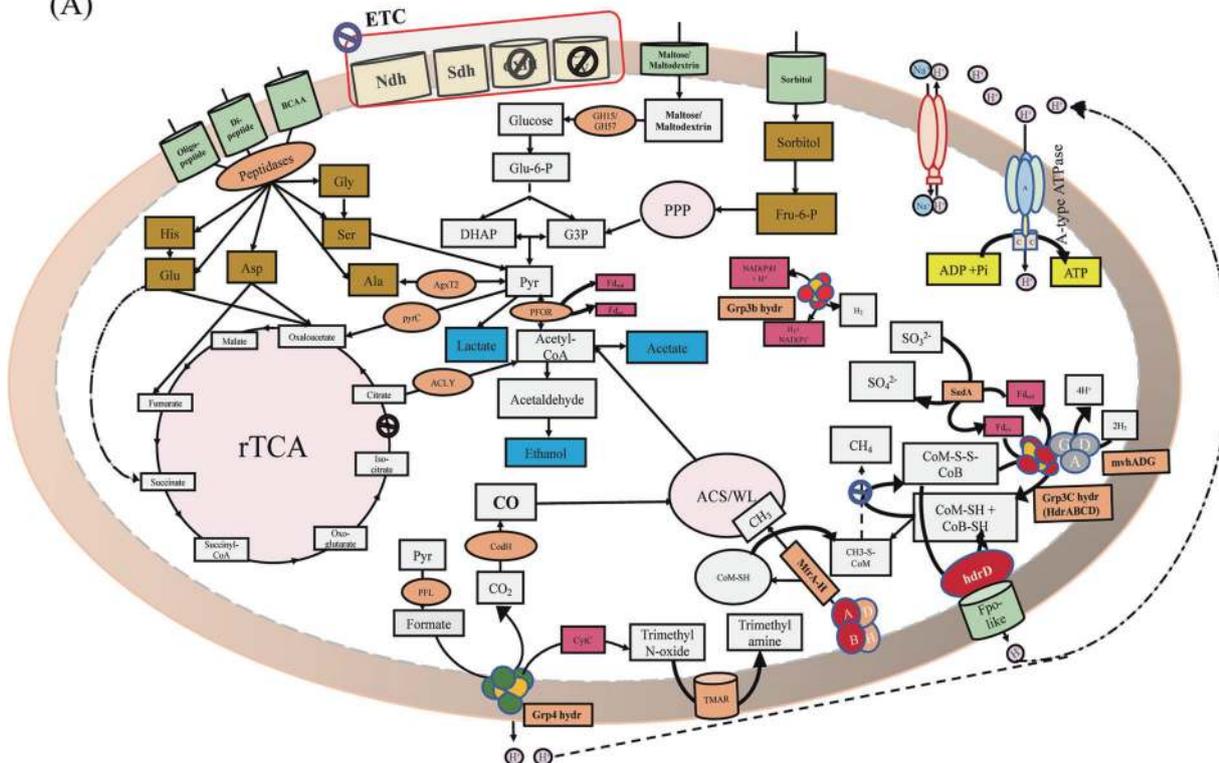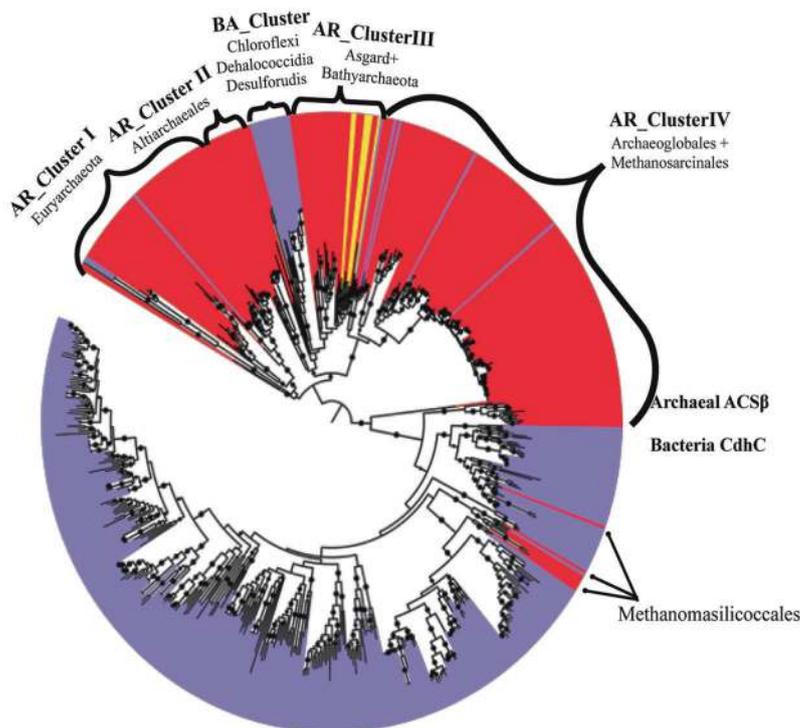
methane metabolic genes may be present to perform different functions, where they recycle methyl groups from different methylated compounds and replace the functions of methyl branch in Wood–Ljungdahl pathway (Fig. 4a). CR_14 showed the potential capability to use formate as electron and hydrogen donor through using Group 4

◀ **Fig. 4 Bathyarchaeotal metabolism and acetyl CoA synthase phylogeny. a** Metabolic reconstruction of the Bathyarchaeota bin CR_14. Central metabolic pathways found in the genome (glycolysis, carbonyl branch of Wood–Ljungdahl, and methanogenesis-related genes) are shown in gray boxes, carbon fixation pathways (ACS/WL, PPP, and rTCA cycles) are shown in pink, electron transport chain (ETC) proteins are shown in yellow, fermentation products are shown in blue boxes, amino acids are shown in brown boxes, enzymes and enzyme complexes are shown in orange circles, energy carriers are shown in red, and metabolite and amino acid transporters are shown in light green. **b** Maximum-likelihood tree of the acetyl CoA synthase β subunit (ACSβ/CdhC). The tree was calculated using the best fit substitution model (LG + R9) that describes the evolutionary relationships between ACS families. The tree was made using reference sequences under the KEGG entry (K00193) collected from AnnoTree [44] and branch location was tested using 1000 ultrafast bootstraps and approximate Bayesian computation, branches with bootstrap support >80% were marked by black circles. Blue and red clades highlight sequences belong to bacterial (CdhC) and archaeal (ACSβ) versions, respectively. Scale bar indicates substitutions per site. Sequences from CR_Lokiarchaeota and CR_Bathyarchaeota bins were shaded with yellow. The tree is available with full bootstrap values in Newick format in the Supplementary Data.

hydrogenases (formate hydrogenlyases) (Fig. S6) [56], and these electrons reduce trimethylamine N-oxide (TMAO) to trimethylamine via trimethylamine-oxide reductase or anaerobic dimethyl sulfoxide reductase (TMAO/DMSO reductase) [61]. Together the presence of genes encoding for trimethylamine-specific corrinoid protein as well as diverse classes of methyltransferases, CoB—CoM heterodisulfide reductase/F420 nonreducing hydrogenase (hdrABCD and mvhADG) suggest the capability of CR_14 to recycle coenzyme M (CoM) and coenzyme B (CoB) and transfer the methyl group from trimethylamine to CoM-SH [19, 62]. We located genes encoding the tetrahydromethanopterin S-methyltransferase (mtrA-H), suggesting that mtrA protein transfers the methyl group from CoM to 5,6,7,8-tetrahydromethanopterin (H4MPT) and assimilates the methyl group into acetyl-CoA via the beta subunit of CODH/ACS complex, replacing the function of the methyl branch of Wood–Ljungdahl pathway. This agrees with our finding that only the genes encoding for the carbonyl branch of Wood–Ljungdahl coupled with acetate fermentation genes (acetogenesis) were present in CR_14 and the genes encoding for methyl branch were completely missing for the same pathway. Comparative genomic analyses suggested the presence of other Bathyarchaeota lineages that are capable of coupling methylotrophy with acetogenesis (e.g., B26_1 from marine subsurface sediments and RBG 13_38_9 from Rifle sediments) (Fig. S8). Yet, CR_14 was the only Bathyarchaeota MAG having the capability to transfer methyl groups from mono-, di-, trimethylated compounds. These collective metabolic features in Bathyarchaeota CR_14 suggest this genome may represent a lineage that bridges the gap

between methanogenic and acetogenic Bathyarchaeota through adopting a noncanonical acetogenic lifestyle (Table S5) [19, 20].

Although the proposed scenario where CR_14 performs acetogenesis instead of methanogenesis is metabolically feasible, it is not clear why CR_14 invests a large amount of energy to maintain genes for methane metabolism. A plausible explanation for the presence of methanogenesis-related genes is the lack of dedicated methyltransferases that could transport methyl groups directly between methylated compounds (e.g., trimethylamine) to tetrahydromethanopterin (H4MPT). Therefore, encoding genes mediating the synthesis and cycling of CoM is necessary to use it as intermediate carrier to transport methyl groups from/to H4MPT. Also, the phylogenetic tree of the acetyl-CoA synthase beta subunit (AcsB) showed that CR_14 AcsB genes clustered together with genes recovered from other Bathyarchaeota lineages, Asgard phyla, Chloroflexi, and Altiarchaeales (Fig. 4b). The uniqueness of that clade stemmed from the previous hypothesis that Altiarchaeales members possess an ancient version of AcsB [63]. Likewise, members of Chloroflexi encoded for this archaeal version of ACSβ subunit similar to the ones present in Bathyarchaeota, Asgard phyla, and Altiarchaeales [64, 65]. Finally, this ACS type supports the hydrogen dependent, autotrophic lifestyle of the Asgard archaea [15]. Accordingly, we hypothesize that Bathyarchaeota possess the type of ACS capable of incorporating methyl groups into acetyl-CoA utilizing different carrier proteins (CoM, H4MPT, and may be other unknown carriers), allowing Bathyarchaeota to assimilate methyl groups originated from various methylated compounds.

Discussion

In this study, we employed a metagenomics-enabled genomics approach to recover MAGs from abundant archaeal phyla inhabiting deep sediments of CR margin and elucidate their potential ecological roles. A total of 31 MAGs belonging to archaeal phyla (Lokiarchaeota, Thorarchaeota, Heimdallarchaeota, Bathyarcheota, Thermoplasmatales, and Hadesarchaea) were successfully recovered from the metagenomic datasets of five different samples. Only 11 MAGs met our completion and contamination thresholds (>60% complete and <10% contamination) and were included for the detailed genomic analyses and metabolic reconstruction. More than 90% of the high-quality genomes were affiliated to Bathyarchaeota and Asgard phyla and the phylogenetic affiliations of the predicted proteins in each MAG confirmed the proper quality of the MAGs considered in this study. Remarkably, all the CR archaeal MAGs were enriched with ESP

encoding genes. The wide distribution of these eukaryotic homologs indicates that ESPs are more ubiquitous in anaerobic archaea than previously recognized.

Notably, the sediments used in this study were collected from much deeper depths compared with the sediments where Bathyarchaeota and Lokiarchaeota genomes were previously reported. As such, our analyses were focused on the MAGs belonging to Bathyarchaeota and Lokiarchaeota to understand their ecological potentials under these deep marine sediment conditions.

Community interaction and Lokiarchaeota metabolic interdependencies

Deep marine sediments are rich in aliphatic and aromatic hydrocarbons that provide the associated microbes with a significant portion of their energy and carbon needs [59, 66]. Previous studies showed that hydrocarbon degradation is restricted to limited bacterial and archaeal phyla (e.g., Aminicenantes, TA06, Aerophobetes, Atribacteria, Helarchaeota, and Bathyarchaeota) [17, 19, 67]. However, the full spectrum of archaea involved in the hydrocarbon degradation processes and the nature of their interactions are not fully elucidated. In the light of the above evidence, this study provides new views regarding the ecological roles and potential metabolic capacities of Lokiarchaeota. Lokiarchaeota CR_6 expanded the range of metabolic features encoded by this phylum and predicts metabolic functions enabling the utilization of aliphatic and aromatic hydrocarbons as carbon and energy sources. Considering the high energy demands for aromatic hydrocarbon breakdown under the energy limited conditions of subseafloor sediments [54], it was unexpected to find the ATP-dependent BCR complex (class Bzd) in CR_Asgard/CR_Lokiarchaeota genomes while they have fermentation and/or acetogenic lifestyles. This strongly suggests that acetogenesis or acetate fermentation cannot be the ultimate fate of the aromatic hydrocarbon degradation process due to the low energy yields of these pathways. Under these circumstances, we propose that the CR_Lokiarchaeota members have the capacity of completely mineralizing aromatic hydrocarbons in syntrophy with microbes capable of using nitrite, sulfite, and nitrate as electron sink under the subsurface settings to increase their energy budgets and sustain their energy requirement. After testing all possible partnership scenarios based on the presence/absence profiles of the candidate pathways and the thermodynamic feasibility (Fig. 5a, b), we found the potential partners for Lokiarchaeota belonging to diverse metabolic groups (e.g., nitrate reducers, nitrite reducers, sulfate reducers, sulfite reducers, and thiosulfate reducers), however only sulfite, nitrate, and nitrite reducers are thermodynamically favored under the subsurface conditions. In the same context, a diverse group of syntrophs might require sharing substrates of different qualities (e.g., acetate, propanoate, and other short chain fatty acids). Accordingly, this suggests the presence of Lokiarchaeota syntrophic partners capable of acetate oxidation and short chain fatty acid oxidation in addition to the hydrogenotrophic ones as previously proposed [15]. Therefore, we hypothesize that thermodynamic favorability and potential diversity of the shared metabolites will push the Lokiarchaeota to syntrophy, beyond individual metabolite exchange.

Bathyarchaeota CR_14 bridges the gap between acetogenic and methanogenic lineages

Bathyarchaeota have been shown to have both methanogenic and acetogenic capabilities [19, 20], yet the evolutionary trajectory and the ecological context driving the switch between lifestyles is not fully understood. Generally, Bathyarchaeota lineages relative to those of conventional methanogens are widely distributed in different marine benthic habitats [18], which could be facilitated by their acetogenesis capacity that degrades a wide range of organic compounds under thermodynamic favorable conditions [47].

Interestingly, phylogenomic analysis revealed that Bathyarchaeota MAGs from CR formed a novel class that are significantly divergent from previously reported acetogenic and methanogenic lineages. In these CR sediments, distinct environmental conditions may have captured the intermediate transition between acetogenic and methanogenic lifestyles. We propose that these environmental conditions include the presence of high levels of methylated compounds (e.g., methylated amines) in deep-sea environments [68] and the absence of dedicated pathways/carriers necessary to recycle these methylated compounds and directly shuttle the extracted methyl groups to the Wood–Ljungdahl pathway. The redox potential and availability of oxidized substrates in CR sediments may not favor the occurrence of full methanogenesis and the associated methanogenic archaea, as indicated by the absence of *mcrA* genes at the sampled depths (Fig. 5a).

In summary, this study presents archaeal MAGs recovered from deep subsurface sediments, some of which are of high quality and completeness. Numerous CR archaeal MAGs particularly those belonging to Asgard host eukaryotic signatures, yet, only the Asgard MAGs have homologs for cell division and cytoskeleton. We also report a metabolically novel Lokiarchaeota lineage capable of aliphatic and aromatic hydrocarbon degradation with a putative partnership with metabolically diverse syntrophic organisms. Also, we revealed the presence of an intermediate stage between acetogenic and methanogenic Bathyarchaeota that could convert methylated amines to acetate through linking methylotrophy to acetogenesis.

Fig. 5 Overall metabolism within the environment. a Heat map depicts the distribution patterns of the genes encoding for the key metabolic pathways across all sampled metagenomic datasets. Metagenomes were clustered based on presence (black)/absence (white) profiles using Euclidean distance and average linkage method. *Y*-axis represents pathways included in the analysis and *X*-axis includes metagenomic datasets analyzed. **b** Hypothetical models for the microbial interactions in the Costa Rica sediments. The potential substrates degraded by Lokiarchaeota, Bathyarchaeota, and other microbes in the environment were colored in gray, potential metabolic products (in blue), potential interactions between Lokiarchaeota and their syntrophic partners and possible shared metabolites.

Data availability

Author contributions JB, IF, and RLZ conceived the study. RLZ recovered the genomes from the metagenomic datasets. IF, RLZ analyzed the genomic data. AM and CH collected the samples. RZ

revised the thermodynamic calculations and IF, JB, and RLZ wrote the paper with input from all authors. All documents were edited and approved by all authors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Biddle JF, Lipp JS, Lever MA, Lloyd KG, Sørensen KB, Anderson R, et al. Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. Proc Natl Acad Sci USA. 2006;103:3846–51.

2. Schippers A, Neretin LN, Kallmeyer J, Ferdelman TG, Cragg BA, Parkes RJ, et al. Prokaryotic cells of the deep sub-seafloor biosphere identified as living bacteria. Nature. 2005;433:861–4.

3. Lipp JS, Morono Y, Inagaki F, Hinrichs K-U. Significant contribution of Archaea to extant biomass in marine subsurface sediments. Nature. 2008;454:991–4.

4. Buongiorno J, Turner S, Webster G, Asai M, Shumaker AK, Roy T, et al. Interlaboratory quantification of Bacteria and Archaea in deeply buried sediments of the Baltic Sea (IODP Expedition 347). FEMS Microbiol Ecol. 2017;93:fix007.

5. Biddle JF, Fitz-Gibbon S, Schuster SC, Brenchley JE, House CH. Metagenomic signatures of the Peru Margin subseafloor biosphere show a genetically distinct environment. Proc Natl Acad Sci USA. 2008;105:10583–8.

6. Youssef NH, Rinke C, Stepanauskas R, Farag I, Woyke T, Elshahed MS. Insights into the metabolism, lifestyle and putative evolutionary history of the novel archaeal phylum "Diapherotrites. ISME J. 2015;9:447–60.

7. Farag IF, Youssef NH, Elshahed MS. Global distribution patterns and pangenomic diversity of the candidate phylum "Latescibacteria" (WS3). Appl Environ Microbiol. 2017;83:e00521–17.

8. Kubo K, Lloyd KG, Biddle F, Amann J, Teske R, Knittel A, et al. Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine sediments. ISME J. 2012;6:1949–65.

9. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J-F, et al. Insights into the phylogeny and coding potential of microbial dark matter. Nature. 2013;499:431–7.

10. Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, et al. Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature. 2015;521:173–9.

11. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, et al. A new view of the tree of life. Nat Microbiol. 2016;1:16048.

12. Castelle CJ, Banfield JF. Major new microbial groups expand diversity and alter our understanding of the tree of life. Cell. 2018;172:1181–97.

13. Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, et al. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nat Commun. 2016;7:13219.

14. Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Bäckström D, Juzokaite L, Vancaester E, et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature. 2017;541:353–8.

15. Spang A, Stairs CW, Dombrowski N, Eme L, Lombard J, Caceres EF, 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. 2019;4:1138–48.

16. MacLeod F, Kindler GS, Wong HL, Chen R, Burns BP. Asgard archaea: diversity, function, and evolutionary implications in a range of microbiomes. AIMS Microbiol. 2019;5:48–61.

17. Seitz KW, Dombrowski N, Eme L, Spang A, Lombard J, Sieber JR, et al. Asgard archaea capable of anaerobic hydrocarbon cycling. Nat Commun. 2019;10:1822.

18. Zhou Z, Pan J, Wang F, Gu J-D, Li M. Bathyarchaeota: globally distributed metabolic generalists in anoxic environments. FEMS Microbiol Rev. 2018;42:639–55.

19. Evans PN, Parks DH, Chadwick GL, Robbins SJ, Orphan VJ, Golding SD, et al. Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. Science. 2015;350:434–8.

20. He Y, Li M, Perumal V, Feng X, Fang J, Xie J, et al. Genomic and enzymatic evidence for acetogenesis among multiple lineages of the archaeal phylum Bathyarchaeota widespread in marine sediments. Nat Microbiol. 2016;1:16035.

21. Martino A, Rhodes ME, León-Zayas R, Valente IE, Biddle JF, House CH. Microbial diversity in sub-seafloor sediments from the Costa Rica margin. Geosciences. 2019;9:218.

22. Vannucchi P, Ujiie K, Stroncik N, IODP Exp. 334 Scientific Party, Yatheesh V. IODP expedition 334: An investigation of the sedimentary record, fluid flow and state of stress on top of the seismogenic zone of an erosive subduction margin. Sci Drill. 2013;15:23–30.

23. Peng Y, Leung HCM, Yiu SM, Chin FYL. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics. 2012;28:1420–8.

24. Wu Y-W, Tang Y-H, Tringe SG, Simmons BA, Singer SW. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. Microbiome. 2014;2:26.

25. Laczny CC, Sternal T, Plugaru V, Gawron P, Atashpendar A, Margossian HH, et al. VizBin—an application for reference-independent visualization and human-augmented binning of metagenomic data. Microbiome. 2015;3:1.

26. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25:1043–55.

27. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J. 2017;11:2864–8.

28. Darling AE, Jospin G, Lowe E, Matsen FA, Bik HM, Eisen JA. PhyloSift: phylogenetic analysis of genomes and metagenomes. PeerJ. 2014;2:e243.

29. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.

30. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform. 2010; 11:119.

31. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–74.

32. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: 2010 Gateway Computing Environments Workshop (GCE). New Orleans, LA: Institute of Electrical and Electronics Engineers; 2010. p. 1–8. https://ieeexplore.ieee.org/abstract/document/5676129.

33. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018;35:518–22.

34. Johnson LS, Eddy SR, Portugaly E. Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinform. 2010;11:431.

35. Kanehisa M, Sato Y, Morishima K. BlastKOALA and Ghost-KOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol. 2016;428:726–31.

36. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 2012;40:W445–51.

37. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol. 2019;37:420–3.

38. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26:2460–1.

39. Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res. 2018;46:D624–32.

40. Saier MH, Reddy VS, Tsu BV, Ahmed MS, Li C, Moreno-Hagelsieb G. The transporter classification database (TCDB): recent advances. Nucleic Acids Res. 2016;44:D372–9.

41. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30:1236–40.

42. Podell S, Gaasterland T. DarkHorse: a method for genome-wide prediction of horizontal gene transfer. Genome Biol. 2007;8:R16.

43. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7.

44. Mendler K, Chen H, Parks DH, Lobb B, Hug LA, Doxey AC. AnnoTree: visualization and exploration of a functionally annotated microbial tree of life. Nucleic Acids Res. 2019;47:4442–8.

45. Stagars MH, Ruff SE, Amann R, Knittel K. High diversity of anaerobic alkane-degrading microbial communities in marine seep sediments based on (1-methylalkyl)succinate synthase genes. Front Microbiol. 2015;6:1511.

46. Flamholz A, Noor E, Bar-Even A, Milo R. eQuilibrator-the biochemical thermodynamics calculator. Nucleic Acids Res. 2012;40: D770–5.

47. Lever MA. Acetogenesis in the energy-starved deep biosphere—a paradox? Front Microbiol. 2011;2:284.

48. LaRowe DE, Amend JP. Catabolic rates, population sizes and doubling/replacement times of microorganisms in natural settings. Am J Sci. 2015;315:167–203.

49. Dick JM. Calculation of the relative metastabilities of proteins using the CHNOSZ software package. Geochem Trans. 2008;9:10.

50. Helgeson HC. Thermodynamics of hydrothermal systems at elevated temperatures and pressures. Am J Sci. 1969;267:729–804.

51. Levin LA, Orphan VJ, Rouse GW, Rathburn AE, Ussler W, Cook GS, et al. A hydrothermal seep on the Costa Rica margin: middle ground in a continuum of reducing ecosystems. Proc Biol Sci. 2012;279:2580–8.

52. Meckenstock RU, Boll M, Mouttaki H, Koelschbach JS, Cunha Tarouco P, Weyrauch P, et al. Anaerobic degradation of benzene and polycyclic aromatic hydrocarbons. J Mol Microbiol Biotechnol. 2016;26:92–118.

53. Song B, Ward BB. Genetic diversity of benzoyl coenzyme A reductase genes detected in denitrifying isolates and estuarine sediment communities. Appl Environ Microbiol. 2005;71:2036–45.

54. López Barragán MJ, Carmona M, Zamarro MT, Thiele B, Boll M, Fuchs G, et al. The bzd gene cluster, coding for anaerobic benzoate catabolism, in Azoarcus sp. strain CIB. J Bacteriol. 2004;186:5762–74.

55. Huang J-M, Baker BJ, Li J-T, Wang Y. New microbial lineages capable of carbon fixation and nutrient cycling in deep-sea sediments of the Northern South China Sea. Appl Environ Microbiol. 2019;85:e00523–19.

56. Greening C, Biswas A, Carere CR, Jackson CJ, Taylor MC, Stott MB, et al. Genomic and metagenomic surveys of hydrogenase distribution indicate H2 is a widely utilised energy source for microbial growth and survival. ISME J. 2016;10:761–77.

57. Kouzuma A, Kato S, Watanabe K. Microbial interspecies interactions: recent findings in syntrophic consortia. Front Microbiol. 2015;6:477.

58. Pirbadian S, Barchinger SE, Leung KM, Byun HS, Jangir Y, Bouhenni RA, et al. Shewanella oneidensis MR-1 nanowires are outer membrane and periplasmic extensions of the extracellular electron transport components. Proc Natl Acad Sci USA. 2014;111:12883–8.

59. Dong X, Greening C, Rattray JE, Chakraborty A, Chuvochina M, Mayumi D, et al. Metabolic potential of uncultured bacteria and archaea associated with petroleum seepage in deep-sea sediments. Nat Commun. 2019;10:1816.

60. Glombitza C, Jaussi M, Røy H, Seidenkrantz M-S, Lomstein BA, Jørgensen BB. Formate, acetate, and propionate as substrates for sulfate reduction in sub-arctic sediments of Southwest Greenland. Front Microbiol. 2015;6:846.

61. McCrindle SL, Kappler U, McEwan AG. Microbial dimethyl-sulfoxide and trimethylamine-N-oxide respiration. Adv Microb Physiol. 2005;50:147–98.

62. Evans PN, Boyd JA, Leu AO, Woodcroft BJ, Parks DH, Hugenholtz P, et al. An evolving view of methane metabolism in the Archaea. Nat Rev Microbiol. 2019;17:219–32.

63. Adam PS, Borrel G, Gribaldo S. Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. Proc Natl Acad Sci USA. 2018;115: E1166–73.

64. Hug LA, Castelle CJ, Wrighton KC, Thomas BC, Sharon I, Frischkorn KR, et al. Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. Microbiome. 2013;1:22.

65. Mehrshad M, Salcher MM, Okazaki Y, Nakano S-I, Šimek K, Andrei A-S, et al. Hidden in plain sight-highly abundant and diverse planktonic freshwater Chloroflexi. Microbiome. 2018;6:176.

66. Jørgensen BB, Boetius A. Feast and famine-microbial life in the deep-sea bed. Nat Rev Microbiol. 2007;5:770–81.

67. Liu Y-F, Qi Z-Z, Shou L-B, Liu J-F, Yang S-Z, Gu J-D, et al. Anaerobic hydrocarbon degradation in candidate phylum "Atribacteria" (JS1) inferred from genomics. ISME J. 2019;13:2377–90.

68. Mausz MA, Chen Y. Microbiology and ecology of methylated amine metabolism in marine ecosystems. Curr Issues Mol Biol. 2019;33:133–48.