

Metagenomic Insights Into Ecological and Phylogenetic Significances of Candidatus Natranaeroarchaeales, a Novel Abundant Archaeal Order in Soda Lake Sediment

Heng Zhou

Institute of Microbiology Chinese Academy of Sciences

Dahe Zhao (✉ zhaodh@im.ac.cn)

CAS Institute of Microbiology: Institute of Microbiology Chinese Academy of Sciences

<https://orcid.org/0000-0003-0312-6824>

Shengjie Zhang

Institute of Microbiology Chinese Academy of Sciences

Qiong Xue

Institute of Microbiology Chinese Academy of Sciences

Manqi Zhang

Institute of Microbiology Chinese Academy of Sciences

Haiying Yu

Institute of Microbiology Chinese Academy of Sciences

Jian Zhou

Institute of Microbiology Chinese Academy of Sciences

Ming Li

Institute of Microbiology Chinese Academy of Sciences

Hua Xiang

Institute of Microbiology Chinese Academy of Sciences

Research

Keywords: Archaeal microbiome, soda lake, culture-independent approach, environmental adaptation, biogeochemical cycle of carbon and sulfur, origination of methanogenesis

Posted Date: March 3rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-261970/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Archaea were originally discovered in extreme environments, and thrive in many extreme habitats including soda lakes with high pH and salinity. Characteristic and diverse archaeal community played a significant role in biogeochemical cycles; however, the archaeal community and their functions are still less-studied in the intricate sediment of soda lakes.

Results

In this article, the archaeal community of the deep sediment (40-50 cm depth) of five artificially-separated ponds with a salinity range from 7.0% to 33.0% in a soda saline lake was systematically surveyed using culture-independent metagenomics combined with the next-generation sequencing of the archaeal 16S rRNA amplicons. Nine archaeal phyla were detected, which accounted for 2.2% to 35.73% of microbial community in the five deep sediments. Besides the well-known class *Halobacteria*, one novel archaeal order (*Candidatus* Natranaeroarchaeales) of the class *Thermoplasmata* was even more abundant than *Halobacteria* in some deep sediment samples. Of 69 dereplicated archaeal metagenome-assembled genomes (MAGs), 30 MAGs belonged to *Ca.* Natranaeroarchaeales. Different genera of the *Ca.* Natranaeroarchaeales preferred to inhabit in the different salinities, and the divergent halophilic adaptation strategies (salt-out or salt-in) suggested the fast evolution adaptation within this lineage. Most high-quality MAGs had the genes of Wood-Ljungdahl pathway, organic acid fermentation and sulfur respiration, suggesting the putative functions in carbon fixation and sulfur reduction. Interestingly, heterodisulfide reductase and F420-non-reducing hydrogenase complex HdrABC-MvhADG were widely distributed in *Ca.* Natranaeroarchaeales and may play the core roles in energy metabolism from hydrogen. The regeneration of CoM-S-S-CoB was coupled to succinate or 2-oxoglutarate production in *Ca.* Natranaeroarchaeales instead of methanogenesis in the close related *Methanomassiliicoccales*. It suggested that methyl-coenzyme M reductase in *Methanomassiliicoccales* may be laterally transferred from other methanogens.

Conclusion

A novel archaeal order *Ca.* Natranaeroarchaeales of *Thermoplasmata* was found by culture-independent approaches. This order was the most abundant archaeal lineage in the deep sediment of soda lakes, with the characteristic environmental adaptation and biogeochemical potentials in carbon fixation and sulfur reduction. The difference in fermentation products coupled to energy metabolism between methanogens and *Ca.* Natranaeroarchaeales provided additional insights into the origination of methanogenesis in *Thermoplasmata* from the energy metabolism perspective.

Background

Archaea is proposed as one of three domains in 1990 based on molecular comparisons of their 16S rRNA gene sequences and cellular membrane lipids [1]. Along with more and more archaea identified, the currently recognized groups in the archaeal tree of life include four superphyla (Euryarchaeota, TACK, Asgard and DPANN) from previous two phyla (*Euryarchaeota* and *Crenarchaeota*) [2]. Very recently, the newly discovered Asgard archaea *Heimdallarchaeota* was identified as the closest relatives of the eukaryotic nuclear lineage, and this finding even rewrote the tree of life [3, 4]. Archaea were originally found and described in extreme environments, which defined the limits of life on Earth [5]. More and more evidences supported the viewpoint that archaea were widely distributed in global habitats, and universally participated the global biogeochemical cycles, such as methanogenesis [6].

Soda lake is a type of poly-extreme environment with high alkalinity and salinity due to high concentration of (bi)carbonate [7, 8]. It was classified further into normal soda lake and soda-saline lake based on the concentration of sulfate or chloride [9]. It has been reported that class *Halobacteria* and *Candidatus* Nanohaloarchaeota were the abundant archaeal members in the hypersaline brine [10–14]. Both lineages were also predominant in the neutral saline lakes and saltern [14–16]. *Halanaeroarchaeum sulfurireducens* and some natronoarchaea strains were able to perform sulfur respiration under the hypersaline anaerobic conditions [10, 17], while haloarchaeon strain HG 1 could oxidize thiosulfate to tetrathionate [18]. *Haloferax mediterranei* and some other *Haloferax* species had the capability of denitrification [19, 20]. These works demonstrated that *Halobacteria* participated in the biogeochemical cycle in the hypersaline soda lakes and saline lakes. Very recently, culture-dependent and -independent researches revealed that *Ca.* Nanohaloarchaeota had the capability of poly- or oligo-saccharide degradation in hypersaline brines of saltern and soda lakes [12, 21]. Besides, the methylotrophic methanogenesis was evidently a dominant process, and many haloalkaliphilic methylotrophic methanogens were isolated from soda lakes, including *Methanolobus*, *Methanosalsum*, *Methanocalculus* and *Methanonatronarchaeum* [22, 23].

Owing to the development of metagenomic technology, the metabolic and ecological implication of uncultured majority could be elucidated from the assembled genomes. The first metagenomic investigation on the brine of soda lakes recovered 24 genomes of novel members from *Halobacteria*, *Ca.* Nanohaloarchaeota and *Bacteroidetes*. The primary organic carbon degradation by *Halobacteria* and the fermentative lifestyle of *Ca.* Nanohaloarchaeota were deduced from the draft genomes [10]. The sediments of soda lakes exhibited much more complex microbial community and higher biodiversity [12]. By the similarly metagenomic methods, the first haloalkaliphilic members of the Candidate Phyla Radiation (CPR) was discovered in the surface sediment of soda lakes, meanwhile the key enzymes of the Wood-Ljungdahl pathway were found within more bacterial phyla [24]. In the following work, they concentrated on the microbial-mediated sulfur cycle which was classically driven by colorless sulfur-oxidizing bacteria (SOB), anoxygenic purple sulfur bacteria, heterotrophic SOB, and lithoautotrophic sulfate reducers [25]. As one of important findings, they discovered the potential for elemental sulfur/sulfite reduction in the “*Candidatus* Woesearchaeota”, one unculturable archaeal phylum.

Recently, we have deciphered the abundant taxa and their favorable pathways in the brine and surface sediment of soda-saline lakes in Inner Mongolia of China [12]. In this article, we focus on the characteristic archaeal microbiome in the deep sediments of the same region. We found diverse archaea in the deep sediment using high-throughput sequencing based 16S amplicon and metagenomics. Most importantly, we revealed the dominant presence of one novel soda lake-specific order (proposed as *Candidatus* Natranaeroarchaeales in this study) in class *Thermoplasmata*. Here we will systematically elucidate the adaptation mechanism, metabolic potentials, ecological implication and phylogenetic significance of this novel order.

Results

Diverse archaeal lineages were detected in the deep sediment of soda-saline lakes

From the five artificially-separated ponds in a soda-saline lake named Habor Lake (DK), four deep sediment (40–50 cm depth) samples of each pond were collected (20 samples in total). DK and Sed indicated Habor Lake and the sediment samples. The numbers after DK and Sed showed the salinity of brine and that of pore water in the sediment. The porosities were approximately 20–26%. The total salinities of pore waters in the deep sediment ranged from 7.0% up to 33.0%, while pH values were 9.72–9.96 (Table 1). To have a glance at the archaeal microbiome of deep sediment, the variable regions V3-V4 of 16S rRNA gene were amplified using archaea-specific primers and the high-throughput sequencing was performed. A total of 3,566,000 reads were obtained across all samples (Table S1). After clustering sequences at 99% similarity level and resampling to a minimum of 131,986 reads per sample, a total of 8598 OTUs were obtained, in which 1920 archaeal OTUs were assigned (Table S2). The rarefaction curve showed that all twenty curves tend to be flat, so it suggested that the sequencing depth was enough to cover most of the local species (Fig. S1a). The Good's coverage (99.17–100%) also supported this conclusion (Fig. S1b). Both Chao1 and Shannon indexes indicated that the alpha diversity index tended to decrease along with elevated salinities (Fig. S1c, d). The influence of deep sediment salinity on biodiversity tendency was similar as that of brine and surface sediment [12]. The principal coordinate analysis (PCoA) (Fig. S2a) and non-metric multidimensional scaling (NMDS) (Fig. S2b) showed that deep sediment salinity influenced the archaeal composition profiles.

Table 1 Geochemical characteristic of the deep sediments

Sediment name*	DK1Sed7	DK3Sed8	DK15Sed18	DK24Sed24	DK33Sed33
Porosity (%)	0.26 ± 0.02	0.22 ± 0.01	0.25 ± 0.01	0.23 ± 0.01	0.20 ± 0.01
Salinity# (%)	7	8.6	18	24	33
pH#	9.92 ± 0.07	9.92 ± 0.07	9.93 ± 0.03	9.96 ± 0.01	9.72 ± 0.12
Cl ⁻ (g/kg)	0.10 ± 0.02	0.13 ± 0.02	0.38 ± 0.02	0.85 ± 0.36	0.80 ± 0.27
CO ₃ ²⁻ (g/kg)	0.0035 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.06 ± 0.00	0.15 ± 0.01
HCO ₃ ⁻ (g/kg)	0.13 ± 0.03	0.25 ± 0.00	0.50 ± 0.15	0.30 ± 0.03	0.53 ± 0.01
SO ₄ ²⁻ (mmol/g)	3.29 ± 0.06	4.06 ± 0.04	1.26 ± 0.00	0.88 ± 0.07	2.30 ± 0.03
SO ₃ ²⁻ (mmol/g)	0.88 ± 0.04	0.64 ± 0.03	0.58 ± 0.03	0.85 ± 0.08	0.85 ± 0.03
NH ₄ ⁺ (μmol/g)	2.76 ± 0.47	3.29 ± 0.30	7.80 ± 0.37	3.72 ± 0.51	4.96 ± 0.95
NO ₃ ⁻ (μmol/g)	740.74 ± 824.18	179.46 ± 7.77	207.70 ± 6.81	176.17 ± 12.49	199.36 ± 14.27
NO ₂ ⁻ (μmol/g)	222.54 ± 143.89	62.10 ± 2.78	118.63 ± 9.27	39.34 ± 7.95	45.04 ± 21.26

Salinity, pH and concentration of inorganic ions of pore water were measured.

* DK and Sed indicated Habor Lake and the sediment samples. The numbers after DK and Sed showed the salinity of brine and that of pore water in the sediment.

The archaeal composition profiles at the phylum level were presented in Fig. S3. A large number of archaeal phyla were detected, including *Euryarchaeota*, *Nanoarchaeota*, *Thaumarchaeota*, *Diapherotrites*, *Crenarchaeota*, *Hadesarchaeota*, *Altiarchaeota*, *Hydrothermarchaeota*, *Asgardaeota* and unclassified lineages. *Euryarchaeota* was the most abundant archaeal phylum, and *Nanoarchaeota*, *Diapherotrites* and *Crenarchaeota* accounted for a considerable proportion in the archaeome of deep sediments. At class-level, a surprising result showed that *Thermoplasmata* and *Halobacteria* in *Euryarchaeota* were the most abundant archaeal classes, and the abundance of *Thermoplasmata* was even higher than that of *Halobacteria* in the deep sediments with total salinities greater than 18% (Fig. 1). Besides, we also found diverse methanogens, including *Methanomicrobia* (mainly consisted of *Methanomicrobiales*, *Methanosarcinales* and uncultured), *Methanobacteriales* (in *Methanobacteria*), *Methanofastidiosales* (in *Thermococci*) and *Methanomassiliicoccales* (in *Thermoplasmata*) (Table S2).

Genomes recovery of novel archaeal lineage from metagenomes

To unbiasedly evaluate the archaeal community in the polyextreme deep sediment, high-throughput metagenomic sequencing of the five deep sediments (mixing four samples from the same pond) was performed. A total of 88.35 Gb clean data was obtained in the five metagenomes (Table S3). Through extracting draft genomes (bins) from metagenomes, 69 archaeal genomes and 518 bacterial genomes (taxonomic classifications were assigned based on the Genome Database Taxonomy (GTDB)) was recovered, and the archaeal communities were further quantified in the deep sediments (Table S4). Archaea accounted for 2.2% to 35.73% of the microbial community in the deep sediments (Fig. 2a), and the most abundance was in the sample DK33Sed33 with the highest salinity. The relative abundance of *Thermoplasmatota* was almost equal to or much higher than *Halobacteriota* (Fig. 2a). In the archaeal microbiome composed of 69 MAGs, there were 31 *Thermoplasmatota* MAGs, 28 *Halobacteriota* MAGs, and 10 other archaeal MAGs (from phyla *Hadarchaeota*, *Iainarchaeota*, *Nanoarchaeota*, *Thermoproteota*, and PWEA01, detailed shown in Fig. 2b). Interestingly, *Thermoplasmatota* and *Halobacteriota* MAGs were in the majority in the relative abundance and the MAGs number (Fig.1 and Fig. 2).

Considering the high abundance of *Thermoplasmatota* in deep sediments and most of which were classified as uncultured, the detailed phylogenetic position of these 31 MAGs was further analyzed by constructing an evolutionary tree with 770 *Thermoplasmatota* genomes (classified based on GTDB) as reference. Until now, *Thermoplasmatota* was classified into 21 order level taxonomic units using 122 single-copy conserved proteins. Thirty of the 31 *Thermoplasmatota* MAGs were assigned to the order PWKY01 by GTDB, while the other one belonged to the order *Methanomassiliicoccales* (Fig. 3). The order PWKY01 accounted for the most diversity of *Thermoplasmatota* and was also the most abundant archaeal order in our deep sediment samples (Fig. S4a).

There were 12 released PWKY01 MAGs in GTDB, and all of them were assembled from the sediment metagenomes of soda lake (Table S5). It was obvious that order PWKY01 could be regarded as the haloalkaliphilic and anaerobic environment-specific lineage. All 30 PWKY01 MAGs were classified into four genera (SKVC01, PWKY01, B1SED10-34 and PWHR01), and the relative abundance of each MAGs was estimated in the archaeome. The 30 PWKY01 MAGs was renamed HAT1 to HAT30 for short (Table S4). Three SKVC01 MAGs (HAT24, HAT25 and HAT26) exhibited the high relative abundance in DK1Sed7 and DK3Sed8 (relative low salinity within our samples, 7-8%), while 4 PWKY01 MAGs (HAT27, HAT28, HAT29 and HAT30) in DK3Sed8 and DK15Sed18 (higher salinity, 8-18%). B1SED10-34 and PWHR01 tended to inhabit in DK15Sed18 and DK24Sed24 (18-24%), even some PWHR01 MAGs exhibited the highest relative abundance in DK33Sed33 (33%) with the highest salinity, such as HAT12 and HAT13 (Fig. 3). Therefore, these four genera (SKVC01, PWKY01, B1SED10-34 and PWHR01) tended to distribute in different salinities in the soda lake deep sediments (Fig. 3 and Fig. S4b).

To identify the relationship between the *Thermoplasmatota* OTUs (assigned based on SILVA) and the *Thermoplasmatota* MAGs (assigned based on GTDB), the 16S rRNA gene sequences was extracted from MAGs and alignment with the 16S amplicon sequence data was performed. The phylogenetic tree of class *Thermoplasmatota* OTUs based on the 16S rRNA amplicon sequences was constructed and the unclassified taxa was clustered into five clades (Fig. S5a). The relative abundances of unclassified

lineages were much higher, and Clade IV and V were the dominant lineages (Fig. S5b). The otu7254, otu8240 and otu2742 in clade IV showed 100% identity with order PWKY01 MAGs HAT6, HAT24 and HAT30, respectively (Fig. S5a & Table S6). Evidently, the order PWKY01 was identical to clade IV detected by 16S rRNA amplicon, and both high-throughput sequencing approaches revealed that this novel order was the abundant lineage in the deep sediment. Based on the results above, we proposed that this order was named *Candidatus* Natranaeroarchaeales (Natr.an.ae.ro.ar.chae.a'les. N.Gr. n. natron derived from Arabic natrun soda (sodium carbonate); Gr. pref. an not; Gr. n. aer air; N.L. neut. n. archaeum [from Gr. adj. archaios, -e, -on] ancient archaeon; L. fem. pl. suff. -ales, ending to denote an order; N.L. fem. pl. n. Natranaeroarchaeales, the order denoting the soda-requiring anaerobic archaea, affiliated to *Thermoplasmatota*).

Environmental adaptation mechanism of *Ca. Natranaeroarchaeales*

To further understand the environmental adaptation mechanism of *Ca. Natranaeroarchaeales*, the high- ($\geq 90\%$ complete, $< 5\%$ contamination) and medium-quality ($90\% > \text{complete} \geq 50\%$, $< 10\%$ contamination) MAGs of the four genera were selected to perform the isoelectric point calculation of predicted proteomes and functional annotation.

The isoelectric point profiles of predicted proteomes of 13 high- and medium-quality *Ca. Natranaeroarchaeales* MAGs obtained in this article were different from *Halobacteria* and nonhalophiles, but similar to that of halophilic bacteria (Fig. S6). Based on the isoelectric point profiles and average electric points, they were separated into four classes (Fig. 4). Like the reference species *Sipribacter salinus* and *Halomonas elongate* in class I (Fig. 4a), HAT24 may use dissolved solutes rather than inorganic salts (KCl) to resist the high osmotic pressure; however, no gene involved in the biosynthesis or transporting of frequently-used ectoine, trehalose and glycine betaine (Table S8). Maybe other dissolved solutes were employed by HAT24. The average *pI* (6.03) was most close to neutral pH (Table S7). In class II, III, and IV, the peaks of isoelectric point profiles move gradually to acidified pH, and the average *pI* keeps decreased (Fig. 4bcd & Table S7). Similar to the reference species *Natranaerobiusthermophilus* in class II and *Salinibacter ruber* in class IV, the species represented by MAGs in class II, III, and IV seemed to live in high saline environments by absorbing inorganic salts ("salt-in" strategy) rather than dissolved solutes.

To explore the molecular basis for halophilic and alkaliphilic adaptation, the related functional genes were analyzed. The marker genes for biosynthesis of ectoine, glycine betaine, and trehalose (most used solutes by halophilic bacteria) were not complete (Table S8), while genes for betaine and choline transport were less found in *Ca. Natranaeroarchaeales*. Therefore, the "salt-in" strategy may be mainly adopted by most *Ca. Natranaeroarchaeales* members to maintain osmotic pressure balance, and the dissolved organic solutes might be also used by some taxa (for example HAT24). Most high- and medium-quality MAGs contained *trkAHG* genes (detailed in Supplementary Results, Fig. S7), whose products functioned in potassium uptake. As for the adaptation to alkalinity, the most important mechanism is to keep intracellular neutral pH. Multicomponent $\text{Na}^+:\text{H}^+$ antiporter (MnhABCDEFGF) was

found in all of high- and medium-quality MAGs (Fig. S7), and they could maintain intracellular pH homeostasis.

A mixotrophic lifestyle of *Ca. Natranaeroarchaeales*

To decipher the metabolic potentials and putative ecological functions of *Ca. Natranaeroarchaeales*, the encoding genes of 11 high-quality MAGs (including 4 MAGs recovered in this article and 7 MAGs downloaded from GTDB) and 14 medium-quality MAGs (9 in this article and 5 from GTDB) were further investigated (Fig. 5 and S8). These 25 MAGs were classified into four genera named SKVC01, PUNK01, PWKY01 and PWHR01 by GTDB-tk pipeline according GTDB database and manual modification from phylogenomic tree (Fig. 3).

The most remarkable feature was that almost all of high-quality MAGs (nine out of eleven) from four genera contained anaerobic carbon-monoxide dehydrogenase catalytic subunit (AcsA) which was the key enzyme in the Wood-Ljungdahl pathway (WL pathway) (Fig. 5). AcsA usually constitutes Cdh/Acs complex and performs CO utilization and CO₂ fixation. Seven out of nine AcsA containing MAGs also had formate-tetrahydrofolate ligase (Fhs), methylenetetrahydrofolate dehydrogenase (NADP⁺) / methenyltetrahydrofolatecyclohydrolase (FolD) and methylenetetrahydrofolate reductase (MTHFR), while the other two MAGs had two of Fhs, FolD and MTHFR catalyzing the one carbon metabolism in WL pathway (Fig. 5). Of the nine high-quality MAGs, CSSed10_214 and T1Sed10_119m (belonging to PWKY01 and PWHR01, respectively) had formate dehydrogenase (NADP⁺) (FdhAB) which catalyzed the transformation of CO₂ into formate (Fig. 5). One medium-quality MAGs B1Sed10_107R1 belonging to genus PWKY01 also had all six above-mentioned genes in WL pathway (Table S8). Notably, like *Thermoplasmatales*, MG-II, MG-III, and *Thermoprofundales* (MBG-D), the genes encoding Cdh/Acs complex were defective even absent in the available genomes. It was widely accepted that some subunits of Cdh/Acs complexes from *Thermoplasmata* might share low similarity with those from other archaea [26]. Therefore, it was derived that some members (at least these three MAGs CSSed10_214, T1Sed10_119m and B1Sed10_107R1) had the potential of assimilating inorganic carbons via WL pathway.

Part of product acetyl-CoA could flow into gluconeogenesis, considering the complete pathway in almost all high-quality MAGs (Fig. 5 & S9). Carbonate could be utilized because of the presence of phosphate pentose pathway, although Embden-Meyerhof pathway was blocked (lacking phosphofructose kinase) while Entner-Doudoroff were incomplete (detailed in Supplementary Results, Fig. S9 & S10). In almost all high-quality MAGs, acetyl-CoA could enter into incomplete reduced citrate cycle and was converted to succinate and 2-oxoglutarate (Fig. 5), and could further turn into glutamate and glutamine (detailed in Supplementary Results, Fig. S11). Acetyl-CoA was also the important precursor for the biosynthesis of membrane lipid via mevalonate pathway (detailed in Supplementary Results, Fig. S12).

In summary, the *Ca. Natranaeroarchaeales* members had the potential of inorganic carbon fixation, and all taxa may utilize the extracellular organic acids. Thus, this lineage may have a mixotrophic lifestyle.

Energy metabolism coupling with sulfur respiration

The genes involved in sulfur metabolism were rich in *Ca. Natranaeroarchaeales*. Sulfate could be reduced to sulfite via the assimilatory sulfate reduction pathway by 3'-phosphoadenosine 5'-phosphosulfate synthase (PAPSS), sulfate adenylyltransferase (Sat), adenylylsulfate kinase (CysC) and phosphoadenosine phosphosulfate reductase (CysH) in most high-quality MAGs of the four genera, although the assimilatory (CysJI and Sir) or dissimilatory (DsrAB and AsrABC) sulfite reduction was less found. Sir was found in only one high-quality MAG HAT24, and anaerobic sulfite reductase subunit B (AsrB) was found in three high-quality MAGs HAT25, CSSed162cmB_145, and CSSed10_245. No CysJI or DsrAB were found in both high- and medium- quality MAGs. Some *Ca. Natranaeroarchaeales* had the potential of reduction of S^0 and thiosulfate and disproportionation of thiosulfate (Fig. 5). All of the high-quality MAGs had at least one subunit of sulfhydrogenase HydGBAD (functioning in the reduction of S^0 and elemental sulfur), thiosulfate reductase (PhsA) or thiosulfate/3-mercaptopyruvate sulfurtransferase (TST). The hydGBAD genes clustered in one operon in high-quality MAGs HAT25, CSSed162cmB_145, B1Sed10_191R1, and medium-quality MAGs CSSed165cm_322 and T1Sed10_113R1 (containing two clusters). Almost all (ten of eleven) high-quality MAGs contained Phs and TST, and more than half (nine of fourteen) medium-quality MAGs also had Phs. Phs transforms thiosulfate into sulfite and sulfide with quinol (CoQ) as the cofactor, while TST transfers the sulfur atom onto one thiol group to form a disulfide group and also produces sulfite and sulfide (Table S8). The CoQ could be reduced by NADH dehydrogenase, and this process coupled with the formation of transmembrane proton gradient for ATP regeneration (Fig. 5). All high-quality *Ca. Natranaeroarchaeales* MAGs contained some subunits of NADH dehydrogenase (NuoEFIKLMN) and almost complete V/A type ATPases (AtpABCDEFIK) which took part in ATP regeneration via electromotive force (Fig. 5 & Table S8) coupled with thiosulfate reduction.

As we know, sulfide could be assimilated into amino acids (cysteine and methionine) biosynthesis, or released to the environment. No more than one-third of all 25 high- and medium- quality MAGs contained cysteine synthase (CysK) (Fig. S11), so we proposed the above-mentioned MAGs or closely related species may generate energy by sulfur respiration and may involve in the sulfur cycle at the same time, e.g. sulfidogenesis in the anaerobic conditions of soda-saline lake sediment. Only one doubt was the reduction of sulfite. Maybe sulfite would be exported similar as sulfide.

Hydrogenotrophic acidogenesis coupled with energy metabolism by CoM and CoB

There were diverse hydrogenase and related proteins annotated in *Ca. Natranaeroarchaeales*, including energy-converting hydrogenase (EhbAEHI), coenzyme F420 hydrogenase (FrhABG), NADP-reducing hydrogenase (HndAC), Bidirectional [NiFe] hydrogenase (HoxEHY), Membrane-bound hydrogenase (MbhJKL) and F420-non-reducing hydrogenase (MvhADG) (Table S8). Besides, almost all high-quality MAGs (except HAT15) had complete or nearly complete (lacking one subunit) hydrogenase maturation proteins (HypABCDEFGF). It was suggested that hydrogen may be the main source of electron donor. Notably, while some subunits were missing in several hydrogenases in these MAGs, the F420-non-

reducing hydrogenase (MvhADG) containing three subunits was present in almost all high-quality MAGs (except HAT15) and most medium-quality MAGs of genera PWKY01 and PWHR01 (Fig. 5 & Table S8).

Interestingly, genes *mvhADG* formed a gene operon with heterodisulfide reductase genes (*hdrABC*) in at least nine MAGs (seven shown in Fig. 6 and HAT29 and HAT17 in Fig. S13, contigs listed in Table S9), and these six proteins constituted a complex catalyzing the reduction of CoM-S-S-CoB and ferredoxin by molecular hydrogen. Complex HdrABC-MvhADG coupled with methane biosynthesis in some methanogens, for example *Methanomassiliicoccales* ([27-29], discussed below). However, none of *Ca. Natranaeroarchaeales* contained a methyl-coenzyme M reductase (Table S8). Interestingly, the oxidation of HS-CoM and HS-CoB was replaced by a fumarate reductase (TfrAB) (Fig. 5). It was supported that *trfAB* genes are also located in one operon with *hdrABC* and *mvhADG* and/or at the downstream of *hdrB* in thirteen MAGs (seven shown in Fig. 6, while HAT25, CSSen162cmB_145, B1Sed10_107R1, T1Sed10_92, HAT28 and HAT29 in Fig. S13). Same as the incomplete reductive citrate cycle, *tfrA* and *tfrB* (or *tfrB*-homolog) genes distributed in almost all high-quality MAGs and most medium-quality MAGs (Fig. 6 & Fig. S14). More details on *trfAB* were in the Supplementary Results.

The clustered genes in one operon usually were involved in the same life process. Following the incomplete reductive citrate cycle, succinate was transferred to succinyl-CoA, and then was converted into 2-oxoglutarate with reductive ferredoxin as cofactor. The reductive ferredoxin may be produced from an electron bifurcation catalyzed by HdrABC-MvhADG complex (Fig. 5). Thus, the production of 2-oxoglutarate was coupled with hydrogen utilization by HdrABC-MvhADG complex in gene distribution and metabolic functions. Alternatively, succinate was produced meanwhile reductive ferredoxin was oxidized by Na⁺-translocating ferredoxin:NAD⁺ oxidoreductase which was also annotated in almost high-quality MAGs and some medium-quality MAGs (Fig. 5 & Table S8). In this process, transmembrane Na⁺ gradient could be formed for ATP regeneration and NAD⁺ could be reduced for NADH regeneration in sulfur respiration (Fig. 5). So *Ca. Natranaeroarchaeales* could perform hydrogenotrophic acidogenesis and sulfur respiration (described above). In addition, most of *hdrABC-mvhADG-tfrAB* operon containing MAGs had four subunits of 2-oxoglutarate/2-oxoacid ferredoxin oxidoreductase (KorDABC) (Fig. 5), and this enzyme could catalyze the reversible transformation between succinyl-CoA and 2-oxoglutarate [30]. Genes *korDABC* located at the downstream of *hdrABC-mvhADG-tfrAB* operon and shared the same promoter (Fig. 6, S13 & S14). This result suggested that 2-oxoglutarate may be the one of the final products. Besides, one or both subunits of *fumAB* involving in the formation of fumarate from malate were located at the upstream of *hdrABC-mvhADG-tfrAB* operon in HAT24, HAT26, CSSed10_245 and CSSed10_214 (Fig. 6). Genes *fumAB* was separated by seven genes from *hdrABC-mvhADG-tfrAB* operon in MAG B1Sed10_191R1 (Fig. S13). It also supported the relationship between hydrogen metabolism and incomplete reductive TCA. The evidence of reversed TCA flow also included that none MAGs had the potential of phosphorylation (lacking cytochrome c reductase and cytochrome c oxidase) and photophosphorylation (lacking all of photosystem I, II, bacteriorhodopsin (K04641) and halorhodopsin (K04642)). This is reasonable considering the sediment is an anaerobic and dark environment.

Discussion

There were extensive researches on microbial community composition, structure and function in soda lakes, but a large proportion of microbes, including archaea and bacteria, were less characterized, and even still be undiscovered. Due to the high-throughput 16S rRNA amplicon and metagenomic sequencing technologies, we found a novel uncultured order (recommended name *Ca. Natranaeroarchaeals*) in the class *Thermoplasmata* which was reclassified into the phylum *Thermoplasmatota* in GTDB (Fig. 2&3). Class *Halobacteria* and phylum *Ca. Nanohaloarchaea* were found to be the most dominant archaea in the hypersaline brines and surface sediment of soda lakes [10, 12, 24]. Our result demonstrated that soda lake-specific order *Ca. Natranaeroarchaeals* was the most abundant archaeal lineage in the deep sediments (Fig. S4 & S5).

Up to now, representative members of three orders including *Thermoplasmatales* [31], *Aciduliprofundales* [32] and *Methanomassiliicoccales* [33] in phylum *Thermoplasmatota* (GTDB database) were purely cultured. The type strains were certainly important for the identification of the taxonomy and metabolism of novel taxa. Although microbial culture was indispensable to identify the novel taxon and metabolic functions, the culture-independent methods were useful and effective strategies to detect the uncultured majority in the nature. As we know, more and more orders in phylum *Thermoplasmatota* were discovered based on high-throughput sequencing method, including *Candidatus* Poseidoniales [34], *Candidatus* Thermoprofundales [26], *Candidatus* Gimiplasmatales [35] and *Candidatus* Lunaplasmalacustris [36]. *Candidatus* Natranaeroarchaeales in this work was also found by this strategy (Fig. 3). Interestingly, *Thermoplasma acidophilum* was firstly isolated from a coal refuse pile, and was a thermophilic and acidophilic archaeon [31]. Another archaeon *Aciduliprofundum boonei* was isolated from deep-sea hydrothermal vents, and was a ubiquitous thermoacidophilic lineage in that inhabits [32]. Order *Candidatus* Poseidoniales (MGII) preferred for the photic zone [34] and *Candidatus* Thermoprofundales (MBG-D) were abundant archaea of surficial and deep sediment in the marine [37, 38]. So many lineages of *Thermoplasmatota* were present or predominant in the harsh environments, and it suggested that *Thermoplasmatota* adopted the tactics of evolutionary adaptation into the different environments.

Ca. Natranaeroarchaeales exhibited an excellent ability to adapt the hypersaline and alkaline sediments possible as a result of the energetically favorable salt-in strategy (Fig. 4) and characteristic metabolic potentials (Fig. 5). Microorganisms would consume less energy to accumulate inorganic KCl for osmotic balance than compatible solutes, meanwhile their intracellular proteins contained more acidic amino acids [39]. Salt-in strategy is widely adopted by halophiles. It was experimentally confirmed that class *Halobacteria* [40, 41], *Spiribacter ruber* [42, 43] and *Natronaerobius thermophilus* [44, 45] used salt-in strategy, while haloalkaliphilic CPR (*Candidatus* Neelsonbacteria and *Candidatus* Vogelbacteria) [24], *Ca. Nanohaloarchaeota* [10] and *Methanonatronarchaeia* [23] were evidently inferred to use this strategy from their acidic proteomes. Obviously, most of salt-in halophiles form the separate lineages from closely related salt-out halophiles or non-halophiles [46]. However, different members in *Ca. Natranaeroarchaeales* have the favorite ranges of salinity, and their proteomes enriched more acidic amino acids along with increased salinity ranges (Figs. 3 & 4). It indicated that salt-in strategy played a

crucial role in (hyper)saline adaptation. The gradually divergent isoelectric point profiles of predicted proteomes between different genera or among the same genus within *Ca. Natranaeroarchaeales* suggested that this clade employed the strategy of fast evolutionary to adapt himself into the increased salinity. It was reported that the type species *Thermoplasma acidophilum* and *Cuniculiplasma divulgatum* of class *Thermoplasmata* lacked cell wall[31, 47]. Although no member of *Ca. Natranaeroarchaeales* was isolated, we were of the opinion that the characteristics of cell wall deficiency may exist, and heterologous DNA containing new genes could easily pass through cell membrane and integrate into the genome. This possibility was also supported by the interdomain lateral gene transfer of bacterial-type H₄F Wood-Ljungdahl pathway of *Candidatus Gimiplasmatales* and *Methanomassiliicoccales* from the *Firmicutes* [35].

Another fact is that the *Ca. Natranaeroarchaeales* had the metabolic potentials of sulfur respiration, hydrogen utilization and organic acids production (Fig. 5). There was high concentration of sulfate and sulfite in the deep sediment (Table 1), and sulfide accumulated at a high concentration in the deep sediment of soda lakes in both Cock Soda Lake and Mono Lake [25, 48]. Therefore, some *Ca. Natranaeroarchaeales* members may drive the biogeochemical cycle of sulfur in the deep sediment of soda lakes. In addition, *Ca. Natranaeroarchaeales* had a mixotrophic lifestyle via hydrogen utilization, WL pathway, gluconeogenesis and incomplete rTCA cycle (Fig. 4). The molecular hydrogen was widespread in nature, and has been discovered associated with sedimentary rocks, salt deposits and ground water [49]. The genesis of hydrogen was complex, such as degassing from the Earth's core and mantle, the reaction of water with ultrabasic rocks, decomposition of organic matter and biological activity. The hydrogen could be consumed as electron carrier by diverse microorganisms in many habitats like deep-sea [50, 51]. The hydrogen-utilizing potential was found in the microbes of brine and surface sediment of soda lakes [12], so the hydrogen cycle may be flourish and *Ca. Natranaeroarchaeales* may participate in this process. In addition, the acetate or 2-oxoglutarate produced by *Ca. Natranaeroarchaeales* could provide the substrates for other microbes in the deep sediments. However, it is still essential to confirm these metabolic capabilities and ecological significance of the *Ca. Natranaeroarchaeales* with culture-dependent methods in the future.

The most related cultured lineage to *Ca. Natranaeroarchaeales* was *Methanomassiliicoccales* (Fig. 3), which was the seventh order of methanogenic archaea [33, 52, 53]. HdrABC-MvhADG complex catalyzed a ferredoxin- and hydrogen- dependent reductive reaction of CoM-S-S-CoB in hydrogenotrophic methanogenic archaea [27], and this complex also played a core role in hydrogen utilization and energy metabolism of *Methanomassiliicoccales* [28, 29], same as *Ca. Natranaeroarchaeales*. Obviously, the main difference was that: (1) regeneration of CoM-S-S-CoB coupled with methane biosynthesis by methyl-coenzyme M reductase in *Methanomassiliicoccales* [54], while coupled with succinate and oxoglutarate production in *Ca. Natranaeroarchaeales* (Fig. 5) by thiol-fumarate reductase [55]. There was excess (bi)carbonate in the extremely alkaline environments and the produced organic acids could be exactly buffered [56]. (2) reductive ferredoxin provided reducing power for HyrD-Fpo complex coupled H⁺ translocation in *Methanomassiliicoccales* [29] while reductive ferredoxin and NADH provided reducing

power for Rnf and NADH dehydrogenase coupled Na^+ and H^+ translocation respectively in *Ca. Natranaeroarchaeales* (Fig. 5, [57]). The concentration of proton (H^+) in the soda lakes was lower than neutral pH, while sodium (Na^+) was much higher than most habitats [9]. Transmembrane electrochemical gradient of Na^+ (sodium-motive force) was employed as energy intermediate to adapt into the high external pH [58, 59]. We inferred HdrABC-MvhADG complex may exist in the last common ancestor of *Methanomassiliicoccales* and *Ca. Natranaeroarchaeales*, and its presence in non-methanogens offered a clue that methyl-coenzyme M reductase in *Methanomassiliicoccales* was transferred laterally from other methanogens.

In summary, *Ca. Natranaeroarchaeales* was a typical polyextremophile which could inhabit in a large range of salinities, high pH and anaerobic condition. This lineage would play significant roles in the biogeochemical cycle in the harsh environments of the earth. Besides, there were Cl-rich brines with pH 4.96–9.13 in the subsurface of Mars [5]. Although no life was found in Mars, *Ca. Natranaeroarchaeales* may be a better candidate to inhabit the subsurface of Mars and change the geochemical compositions considering the characteristic lifestyle of autotroph and organic acid fermentation.

Conclusion

With the development of culture-independent high-throughput sequencing technologies, more and more previously unknown microbial lineages come to light, especially archaea. Here we systematically investigated the archaeal microbiome of deep sediments of soda lakes by 16S rRNA gene amplicon and metagenomics. The high diversity of archaeal taxa was detected in the deep sediment of soda lake, and they distributed across *Euryarchaeota*, DPANN superphylum, TACK superphylum, *Asgardaeota* and unclassified lineages. One soda lake sediment-specific uncultured archaeal order *Ca.*

Natranaeroarchaeales belonging to the class *Thermoplasmata* was found by both 16S rRNA gene amplicon and metagenomic technologies. This order was the most abundant archaeal lineage in the deep sediment of soda lakes, and even showed higher quantity than class *Halobacteria*. Members of *Ca. Natranaeroarchaeales* had the potential of carbon fixation, sulfur respiration and acid production, indicating their putative capability of driving carbon and sulfur cycle in the soda lake. Considering the excellent adaptation into extreme environments, this order may be a better candidate to drive the geochemical cycle of the extraterrestrial planets. In addition, this research shed lights on the origination of methanogenesis in *Thermoplasmatota* from the perspective of energy metabolism.

Methods

Site description, sample collection and physicochemical characterization

Five crystallizer ponds with different salinities were selected in soda-saline Habor lake in Inner Mongolia, China [9, 12]. Firstly, the salinity of brine was measured, and then the sediments at 40-50 cm depth were sampled in October, 2019. The deep sediment samples were named DK1Sed7, DK3Sed8, DK15Sed18,

DK24Sed24 and DK33Sed33 according to their salinities (Table 1). Four duplicates were designed for each pond at different positions as far as possible.

The samples were immediately brought to the laboratory in an icebox. Each sample was divided into two approximately equal parts. One part was used for the measurement of physicochemical properties, and the other part was stored at -80°C for high-throughput next-generation sequencing. For metagenome sequence, four replicated samples were mixed into one sample. The porosity of the sediment was determined by comparing the wet weight and dry weight of sediment [60]. The pore water was obtained by centrifuging. The salinity of pore water was measured by a handheld refractometer (Beijing Wanchengbeizeng Precision Instrument Co., Ltd., Beijing, China), and the pH value was measured by a PB-10 digital pH-meter (Sartorius, Germany). The total inorganic ions were extracted by 5 times of double distilled water (by weight). The concentration of CO_3^{2-} and HCO_3^- were measured by MD600 Photometer (Lovibond R Water Testing, Dortmund, Germany), and SO_4^{2-} , SO_3^{2-} and Cl^- were measured by Aquakem TM250 discrete photometric Autoanalyzer (Thermo Fisher Scientific, MA, United States).

DNA extraction and 16S rRNA gene amplicon and metagenomic sequencing

About 1 g sediment was used to extract DNA by ALFA-SEQ Advance Soil DNA kit according to the manufacturer's instruction (Guangdong Magigene Biotechnology Co., Ltd. Guangzhou, China). The concentration and purity were measured using NanoDrop One (Thermo Fisher Scientific, MA, USA). The V3-V4 region of 16S rRNA gene was amplified with the archaea specific primers 349F (5'-GYGCASCAGKCGMGA AW -3') and 806R (5'-GGACTACVSGGGTATCTAAT -3') [61] with a 12bp barcode. At last, the high-throughput sequencing was performed on an Illumina Hiseq2500 platform and 250 bp paired-end reads were generated (Guangdong Magigene Biotechnology Co., Ltd. Guangzhou, China).

Metagenomic sequencing were generated using NEB Next® Ultra™ DNA Library Prep Kit for Illumina® (New England Biolabs, USA) following the manufacturer's recommendations and index codes were added. The library quality was assessed on the Qubit 3.0 Fluorometer (BIO-RAD) and Agilent 4200 system. At last, the library was sequenced on an Illumina Hiseq X-ten platform and 150 bp paired-end reads were generated (Guangdong Magigene Biotechnology Co., Ltd. Guangzhou, China).

Bioinformatic analysis

The 16S rRNA amplicon sequences analysis was performed by the 2019.10 distribution of the QIIME2 software suite [62]. After removing the barcode and primer sequences, the reads were denoised by DADA2 [63], and the qiime feature-table summarize command was used to produce a visualization file that showed the spread of sequence depths across the samples, which was used to identify a lower boundary on the sequence depth and then (if desired) filter out low sequence depth samples with the qiime feature-table filter-samples command with the -p-min-frequency parameter. A classifier artifact on the SILVA database (SILVAngs 132, <http://www.arb-silva.de>) [64] was trained, and then was used to perform taxonomic assignment. The reads were clustered at 99% sequence identity.

To standardize sequencing depth, each sample was rarefied to 131,986 reads (the lowest sequence number across all samples). The rarefied sequences were calculated for Good's coverage (an estimator of sampling completeness: percentage of the total species is represented in a sample), alpha diversity indexes by the vegan package in R 3.6.3 [65]. The relative abundance of the archaeal microbial profile was calculated by the read count of OTU in each sample. The sequences of 16S rRNA gene were performed multiple sequence alignment using MAFFT program (v7.407) [66], and the maximum-likelihood phylogenetic tree was constructed using standard model selection followed by tree inference (-m TEST) in IQ-TREE (version 1.6.12) [67] with a bootstrap of 1000.

The quality control of metagenomic raw reads was conducted using Read_qc module in metaWRAP (V2.12.1) [68] to trim the barcode and remove low-quality reads. Clean reads from one sample were assembled using metaWRAP::Assembly module with MEGAHIT (v1.1.2) [69]. Metagenome-assembled genomes were obtained using metaWRAP::Binning module with software MaxBin2 [70], metaBAT2 [71] and CONCOCT [72] with default settings, and then were improved by Bin_refinement module. The software checkM (v1.0.12) was used to evaluate the genome completeness and contamination [73] (Parks et al., 2015). MAGs from five sediment samples were dereplicated by dRep (v2.2.1) [74]. The abundance of MAGs across five sediment samples was estimated by Quant_bins module in metaWRAP pipeline [68]. Protein-coding sequences were predicted by Prodigal (v2.6.3) [75] (Hyatt et al., 2010). The functions of these proteins were annotated against to database eggNOG_5.0 with diamond (v0.9.26.127) method by emapper-2.0.0 [76, 77]. The metabolic pathways were reconstructed using online tool KEGG Mapper with KEGG Orthology annotation (<https://www.genome.jp/kegg/>). Ribosomal RNA was predicted using SSU-ALIGN 0.1.1 [78]. The sequences of 16S rRNA amplicon belonged to *Thermoplasmata* (according to SILVA) were aligned to that from metagenome sequence using the programs NCBI Nucleotide-Nucleotide BLAST 2.6.0+ [79], and the identify value equal to 100% was used. The taxonomic assignment for each MAG was performed by the Genome Taxonomy Database Toolkit (GTDB-tk) [80-82]. The reference genomes of all 770 from phylum *Thermoplasmata* and 445 from other archaeal lineages in Genome Taxonomy Database (GTDB) were download [82]. The tree was based on a concatenated alignment of 122 ubiquitous single-copy proteins [80] and visualized using iTOL (v5.6.3) [83].

The isoelectric point of each protein was calculated individually using the Python-based command line "Protein isoelectric point calculator" (IPC Python) in a Linux platform [84]. *Salinibacter ruber* [43], *Natranaerobius thermophilus* [85], *Spiriobacter salinus* [86], *Halomonas elongate* [87], *Bacillus subtilis* [88], *Escherichia coli* [89], *Saccharomyces cerevisiae* [90] and *Sulfolobus islandicus* [91] were used as references. The isoelectric point profiles (bin width 0.1) of different predicted proteomes were clustered based on the correlation distances with a bootstrap of 1000 replications using R package Pvcust [92]. The relatedness of TfrB and TfrB-homologs was conducted by TBtools with blastP under default parameters [93].

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Date availability

The 16S rRNA amplicon sequences of 20 deep sediment samples was upload to the BioProject PRJNA695142 in the National Center for Biotechnology Information (NCBI, <https://www.ncbi.nlm.nih.gov/>). The raw data of metagenomic sequence for five deep sediment have been uploading to the NCBI under accession number SRR13089369, SRR13089372, SRR13089371, SRR13089373 and SRR13089370 (BioProject PRJNA679647). The genomes of 31 *Ca. Thermoplasmatota* MAGs have been upload to the NCBI under the Project ID PRJNA682239 (accession number: JAFDTB000000000, JAFDTC000000000, JAFDTD000000000, JAFDTE000000000, JAFDTF000000000, JAFDTG000000000, JAFDTH000000000, JAFDTI000000000, JAFDTJ000000000, JAFDTK000000000, JAFDTL000000000, JAFDTM000000000, JAFDTN000000000, JAFDTO000000000, JAFDTP000000000, JAFDTQ000000000, JAFDTR000000000, JAFDTS000000000, JAFDTT000000000, JAFDTU000000000, JAFDTV000000000, JAFDTW000000000, JAFDTX000000000, JAFDTY000000000, JAFDTZ000000000, JAFDUA000000000, JAFDUB000000000, JAFDUC000000000, JAFDUD000000000, JAFDUE000000000, JAFDUF000000000) and the National Omics Data Encyclopedia (NODE, <https://www.biosino.org/node>) under the Project ID OEP001447. The nucleotide and amino acid sequences of predicted genes of 31 MAGs are available in FigShare (<https://doi.org/10.6084/m9.figshare.13378673.v1> and <https://doi.org/10.6084/m9.figshare.13372472.v1>). The newick format files of maximum likelihood trees based on 16S rRNA amplicon sequence and concatenated 122 ubiquitous single-copy archaeal proteins in MAGs for Fig. 2, 3 and S6 are available in FigShare (<https://doi.org/10.6084/m9.figshare.13474689.v1>).

Competing interests

The authors declare that they have no competing interests.

Funding

This study was funded by the National Natural Science Foundation of China (grant number 91751201, 32000046) and partially supported by National Science and Technology Foundation Project of China (grant number 2017FY100300).

Authors' contributions

HX designed and supervised the study. HZ, DZ, SZ, MZ and JZ collected the samples. HZ and MZ measured the physicochemical characteristics. HZ, DZ, SZ and HY performed bioinformatic and

statistical analyses. HZ and DZ prepared the figures and wrote the manuscript under the guidance of HX. QX and ML participated in discussions and revisions. All authors read and approved the final manuscript.

Acknowledgments

Not applicable

Authors' information

¹State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China; ²College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China

References

1. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences*. 1990;87(12):4576-9.
2. Spang A, Caceres EF, Ettema TJ. Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. *Science*. 2017;357(6351).
3. Williams TA, Cox CJ, Foster PG, Szöllősi GJ, Embley TM. Phylogenomics provides robust support for a two-domains tree of life. *Nature ecology & evolution*. 2020;4(1):138-47.
4. Doolittle WF. Evolution: Two Domains of Life or Three? *Current Biology*. 2020;30(4):R177-R9.
5. Merino N, Aronson HS, Bojanova DP, Feyhl-Buska J, Wong ML, Zhang S, Giovannelli D. Living at the Extremes: Extremophiles and the Limits of Life in a Planetary Context. *Frontiers in Microbiology*. 2019;10:780.
6. Evans PN, Boyd JA, Leu AO, Woodcroft BJ, Parks DH, Hugenholtz P, Tyson GW. An evolving view of methane metabolism in the Archaea. *Nat Rev Microbiol*. 2019;17(4):219-32.
7. Schagerl M, Burian A. The ecology of African soda lakes: driven by variable and extreme conditions. *Soda Lakes of East Africa*: Springer; 2016. p. 295-320.
8. Grant WD, Sorokin DY. Distribution and Diversity of Soda Lake Alkaliphiles. 2011.
9. Boros E, Kolpakova M. A review of the defining chemical properties of soda lakes and pans: An assessment on a large geographic scale of Eurasian inland saline surface waters. *PLoS One*. 2018;13(8):e0202205.
10. Vavourakis CD, Ghai R, Rodriguez-Valera F, Sorokin DY, Tringe SG, Hugenholtz P, Muyzer G. Metagenomic insights into the uncultured diversity and physiology of microbes in four hypersaline soda lake brines. *Frontiers in microbiology*. 2016;7:211.
11. Mesbah NM, Abou-El-Ela SH, Wiegel J. Novel and unexpected prokaryotic diversity in water and sediments of the alkaline, hypersaline lakes of the Wadi An Natrun, Egypt. *Microbial Ecology*. 2007;54(4):598-617.

12. Zhao D, Zhang S, Xue Q, Chen J, Zhou J, Cheng F, Li M, Zhu Y, Yu H, Hu S, Zheng Y, Liu S, Xiang H. Abundant Taxa and Favorable Pathways in the Microbiome of Soda-Saline Lakes in Inner Mongolia. *Frontiers in Microbiology*. 2020;11:1740.
13. Grant S, Grant WD, Jones BE, Kato C, Li L. Novel archaeal phylotypes from an East African alkaline saltern. *Extremophiles*. 1999;3(2):139-45.
14. Narasingarao P, Podell S, Ugalde JA, Brochier-Armanet C, Emerson JB, Brocks JJ, Heidelberg KB, Banfield JF, Allen EE. De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. *The ISME journal*. 2012;6(1):81-93.
15. Oren A. Halophilic microbial communities and their environments. *Current Opinion in Biotechnology*. 2015;33:119-24.
16. Ventosa A, Fernández AB, León MJ, Sánchez-Porro C, Rodríguez-Valera F. The Santa Pola saltern as a model for studying the microbiota of hypersaline environments. *Extremophiles*. 2014;18(5):811-24.
17. Sorokin DY, Messina E, La Cono V, Ferrer M, Ciordia S, Mena MC, Toshchakov SV, Golyshin PN, Yakimov MM. Sulfur respiration in a group of facultatively anaerobic natronoarchaea ubiquitous in hypersaline soda lakes. *Frontiers in Microbiology*. 2018;9:2359.
18. Sorokin DY, Tourova TP, Muyzer G. Oxidation of thiosulfate to tetrathionate by an haloarchaeon isolated from hypersaline habitat. *Extremophiles*. 2005;9(6):501-4.
19. Torregrosa-Crespo J, Pire C, Martínez-Espinosa RM, Bergaust L. Denitrifying haloarchaea within the genus *Haloferax* display divergent respiratory phenotypes, with implications for their release of nitrogenous gases. *Environmental Microbiology*. 2019;21(1):427-36.
20. Torregrosa-Crespo J, Pire C, Richardson DJ, Martínez-Espinosa RM. Exploring the Molecular Machinery of Denitrification in *Haloferax mediterranei* Through Proteomics. *Frontiers in microbiology*. 2020;11:605859.
21. La Cono V, Messina E, Rohde M, Arcadi E, Ciordia S, Crisafi F, Denaro R, Ferrer M, Giuliano L, Golyshin PN. Symbiosis between nanohaloarchaeon and haloarchaeon is based on utilization of different polysaccharides. *Proceedings of the National Academy of Sciences*. 2020;117(33):20223-34.
22. Sorokin DY, Berben T, Melton ED, Overmars L, Vavourakis CD, Muyzer G. Microbial diversity and biogeochemical cycling in soda lakes. *Extremophiles*. 2014;18(5):791-809.
23. Sorokin DY, Makarova KS, Abbas B, Ferrer M, Golyshin PN, Galinski EA, Ciordia S, Mena MC, Merkel AY, Wolf YI. Discovery of extremely halophilic, methyl-reducing euryarchaea provides insights into the evolutionary origin of methanogenesis. *Nat Microbiol*. 2017;2(8):17081.
24. Vavourakis CD, Andrei A-S, Mehrshad M, Ghai R, Sorokin DY, Muyzer G. A metagenomics roadmap to the uncultured genome diversity in hypersaline soda lake sediments. *Microbiome*. 2018;6.
25. Vavourakis CD, Mehrshad M, Balkema C, van Hall R, Andrei A-S, Ghai R, Sorokin DY, Muyzer G. Metagenomes and metatranscriptomes shed new light on the microbial-mediated sulfur cycle in a Siberian soda lake. *BMC Biol*. 2019;17(1).
26. Zhou Z, Liu Y, Lloyd KG, Pan J, Yang Y, Gu J-D, Li M. Genomic and transcriptomic insights into the ecology and metabolism of benthic archaeal cosmopolitan, Thermoprofundales (MBG-D archaea).

The ISME journal. 2019;13(4):885-901.

27. Kaster A-K, Moll J, Parey K, Thauer RK. Coupling of ferredoxin and heterodisulfide reduction via electron bifurcation in hydrogenotrophic methanogenic archaea. *Proceedings of the National Academy of Sciences*. 2011;108(7):2981-6.
28. Lang K, Schuldes J, Klingl A, Poehlein A, Daniel R, Brune A. New mode of energy metabolism in the seventh order of methanogens as revealed by comparative genome analysis of "Candidatus Methanoplasma termitum". *Applied and environmental microbiology*. 2015;81(4):1338-52.
29. Kröninger L, Berger S, Welte C, Deppenmeier U. Evidence for the involvement of two heterodisulfide reductases in the energy-conserving system of *Methanomassiliicoccus luminyensis*. *The FEBS journal*. 2016;283(3):472-83.
30. Tersteegen A, Linder D, Thauer RK, Hedderich R. Structures and functions of four anabolic 2-oxoacid oxidoreductases in *Methanobacterium thermoautotrophicum*. *European Journal of Biochemistry*. 1997;244(3):862-8.
31. Darland G, Brock TD, Samsonoff W, Conti SF. A thermophilic, acidophilic mycoplasma isolated from a coal refuse pile. *Science*. 1970;170(3965):1416-8.
32. Reysenbach A, Liu Y, Banta A, Beveridge T, Kirshtein J, Schouten S, Tivey M, Von Damm K, Voytek M. A ubiquitous thermoacidophilic archaeon from deep-sea hydrothermal vents. *Nature*. 2006;442(7101):444-7.
33. Dridi B, Fardeau M-L, Ollivier B, Raoult D, Drancourt M. *Methanomassiliicoccus luminyensis* gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *International journal of systematic and evolutionary microbiology*. 2012;62(8):1902-7.
34. Rinke C, Rubino F, Messer LF, Youssef N, Parks DH, Chuvochina M, Brown M, Jeffries T, Tyson GW, Seymour JR. A phylogenomic and ecological analysis of the globally abundant Marine Group II archaea (Ca. Poseidoniales ord. nov.). *Isme Journal*. 2018;13.
35. Hu W, Pan J, Wang B, Guo J, Li M, Xu M. Metagenomic insights into the metabolism and evolution of a new Thermoplasmata order (Candidatus Gimiplasmatales). *Environ Microbiol*. 2020.
36. Zinke LA, Evans PN, Schroeder A, Parks DH, Varner RK, Rich V, Tyson G, Emerson JB. Evidence for non-methanogenic metabolisms in globally distributed archaeal clades basal to the Methanomassiliicoccales. *bioRxiv*. 2020.
37. Lloyd KG, Schreiber L, Petersen DG, Kjeldsen KU, Lever MA, Steen AD, Stepanauskas R, Richter M, Kleindienst S, Lenk S. Predominant archaea in marine sediments degrade detrital proteins. *Nature*. 2013;496(7444):215.
38. Teske A, Sørensen KB. Uncultured archaea in deep marine subsurface sediments: have we caught them all? *Isme Journal*.
39. Oren A. Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes. *Frontiers in microbiology*. 2013;4:315.
40. CHRISTIAN J, WALTHO J. Solute concentrations within cells of halophilic and non. 1962.

41. Reistad R. On the composition and nature of the bulk protein of extremely halophilic bacteria. *Archiv für Mikrobiologie*. 1970;71(4):353-60.
42. Oren A, Heldal M, Norland S, Galinski EA. Intracellular ion and organic solute concentrations of the extremely halophilic bacterium *Salinibacter ruber*. *Extremophiles*. 2002;6(6):491-8.
43. Mongodin EF, Nelson K, Daugherty S, Deboy R, Wister J, Khouri H, Weidman J, Walsh D, Papke R, Perez GS. The genome of *Salinibacter ruber*: convergence and gene exchange among hyperhalophilic bacteria and archaea. *Proceedings of the National Academy of Sciences*. 2005;102(50):18147-52.
44. Mesbah NM, Cook GM, Wiegel J. The halophilic alkalithermophile *Natranaerobius thermophilus* adapts to multiple environmental extremes using a large repertoire of Na⁺ (K⁺)/H⁺ antiporters. *Molecular microbiology*. 2009;74(2):270-81.
45. Bardavid RE, Oren A. Acid-shifted isoelectric point profiles of the proteins in a hypersaline microbial mat: an adaptation to life at high salt concentrations? *Extremophiles*. 2012;16(5):787-92.
46. Gunde-Cimerman N, Plemenitaš A, Oren A. Strategies of adaptation of microorganisms of the three domains of life to high salt concentrations. *FEMS microbiology reviews*. 2018;42(3):353-75.
47. Olga, V., Golyshina, Peter, N., Golyshin, Nadine, I., Goldenstein, Heinrich. The novel extremely acidophilic, cell-wall-deficient archaeon *Cuniculiplasma divulgatum* gen. nov., sp. nov. represents a new family, Cuniculiplasmataceae fam. nov., of the order Thermoplasmatales. *International Journal of Systematic & Evolutionary Microbiology*. 2016.
48. Edwardson CF, Hollibaugh JT. Metatranscriptomic analysis of prokaryotic communities active in sulfur and arsenic cycling in Mono Lake, California, USA. *The ISME journal*. 2017;11(10):2195-208.
49. Zgonnik V. The occurrence and geoscience of natural hydrogen: A comprehensive review. *Earth-Science Reviews*. 2020:103140.
50. Gregory SP, Barnett MJ, Field LP, Milodowski AE. Subsurface microbial hydrogen cycling: Natural occurrence and implications for industry. *Microorganisms*. 2019;7(2):53.
51. Adam N, Perner M. Microbially mediated hydrogen cycling in deep-sea hydrothermal vents. *Frontiers in microbiology*. 2018;9:2873.
52. Iino T, Tamaki H, Tamazawa S, Ueno Y, Ohkuma M, Suzuki K-i, Igarashi Y, Haruta S. Candidatus *Methanogranum caenicola*: a novel methanogen from the anaerobic digested sludge, and proposal of Methanomassiliicoccaceae fam. nov. and Methanomassiliicoccales ord. nov., for a methanogenic lineage of the class Thermoplasmata. *Microbes and environments*. 2013;ME12189.
53. Borrel G, Parisot N, Harris HM, Peyretailade E, Gaci N, Tottey W, Bardot O, Raymann K, Gribaldo S, Peyret P. Comparative genomics highlights the unique biology of Methanomassiliicoccales, a Thermoplasmatales-related seventh order of methanogenic archaea that encodes pyrrolysine. *BMC genomics*. 2014;15(1):679.
54. Kröninger L, Steiniger F, Berger S, Kraus S, Welte CU, Deppenmeier U. Energy conservation in the gut microbe *Methanomassiliicoccus luminyensis* is based on membrane-bound ferredoxin oxidation coupled to heterodisulfide reduction. *The FEBS journal*. 2019;286(19):3831-43.

55. Heim S, Kunkel A, Thauer RK, Hedderich R. Thiol : Fumarate reductase (Tfr) from *Methanobacterium thermoautotrophicum* - Identification of the catalytic sites for fumarate reduction and thiol oxidation. *European Journal of Biochemistry*. 1998;253(1):292-9.
56. Sorokin DY, Banciu HL, Muyzer G. Functional microbiology of soda lakes. *Current opinion in microbiology*. 2015;25:88-96.
57. Welte C, Deppenmeier U. Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*. 2014;1837(7):1130-47.
58. Mulkidjanian AY, Dibrov P, Galperin MY. The past and present of sodium energetics: may the sodium-motive force be with you. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*. 2008;1777(7-8):985-92.
59. Avetisyan AV, Dibrov PA, Semeykina AL, Skulachev VP, Sokolov MV. Adaptation of *Bacillus FTU* and *Escherichia coli* to alkaline conditions: the Na⁺-motive respiration. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*. 1991;1098(1):95-104.
60. Lee S, Ferse SC, Ford A, Wild C, Mangubhai S. Effect of sea cucumber density on the health of reef-flat sediments. Wildlife Conservation Society; 2017.
61. Hristov AN, Callaway TR, Lee C, Dowd SE. Rumen bacterial, archaeal, and fungal diversity of dairy cows in response to ingestion of lauric or myristic acid. *Journal of Animal Science*. 2012;90(12):4449-57.
62. Hall M, Beiko RG. 16S rRNA Gene Analysis with QIIME2. *Methods Mol Biol*. 2018;1849:113-29.
63. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. *Nature methods*. 2016;13(7):581-3.
64. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*. 2012;41(D1):D590-D6.
65. Kindt R. Vegan: community ecology package. Ordination methods, diversity analysis and other functions for community and vegetation ecologists. *World*. 2020;11.
66. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*. 2013;30(4):772-80.
67. Nguyen L-T, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution*. 2015;32(1):268-74.
68. Kang DD, Froula J, Egan R, Wang Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *Peerj*. 2015;3(8):e1165.
69. Dinghua L, Chi-Man L, Ruibang L, Kunihiro S, Tak-Wah L. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*. 2015;31(10):1674-6.

70. Wu Y-W, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*. 2016;32(4):605-7.
71. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *Peerj*. 2019;7:e7359.
72. Alneberg J, Bjarnason BS, De Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. Binning metagenomic contigs by coverage and composition. *Nature methods*. 2014;11(11):1144-6.
73. 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 Research*. 2015;25(7):1043-55.
74. 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 Journal*. 2017.
75. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *Bmc Bioinformatics*. 2010;11(1):119.
76. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, Von Mering C, Bork P. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Molecular biology and evolution*. 2017;34(8):2115-22.
77. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*. 2019;47(D1):D309-D14.
78. Nawrocki EP, Farm HJ. SSU-ALIGN User's Guide. 2010.
79. Coordinators NR. Database resources of the national center for biotechnology information. *Nucleic Acids Research*. 2016;44(D1):D7-D19.
80. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Oxford University Press; 2020.
81. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarszewski A, Chaumeil P-A, Hugenholtz P. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature biotechnology*. 2018;36(10):996-1004.
82. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. A complete domain-to-species taxonomy for Bacteria and Archaea. *Nature biotechnology*. 2020:1-8.
83. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res*. 2019;47(W1):W256-W9.
84. Kozłowski LP. IPC–isoelectric point calculator. *Biology direct*. 2016;11(1):55.
85. Zhao B, Mesbah NM, Dalin E, Goodwin L, Nolan M, Pitluck S, Chertkov O, Brettin TS, Han J, Larimer FW, Land ML, Hauser L, Kyrpides N, Wiegell J. Complete genome sequence of the anaerobic,

- halophilic alkalithermophile *Natranaerobius thermophilus* JW/NM-WN-LF. *J Bacteriol.* 2011;193(15):4023-4.
86. Leon MJ, Ghai R, Fernandez AB, Sanchez-Porro C, Rodriguez-Valera F, Ventosa A. Draft Genome of *Spiribacter salinus* M19-40, an Abundant Gammaproteobacterium in Aquatic Hypersaline Environments. *Genome Announc.* 2013;1(1).
 87. Vreeland RH, Litchfield C, Martin E, Elliot E. *Halomonas elongata*, a new genus and species of extremely salt-tolerant bacteria. *International journal of systematic and evolutionary microbiology.* 1980;30(2):485-95.
 88. Nakamura LK, Roberts MS, Cohan FM. Relationship of *Bacillus subtilis* clades associated with strains 168 and W23: a proposal for *Bacillus subtilis* subsp. *subtilis* subsp. nov. and *Bacillus subtilis* subsp. *spizizenii* subsp. nov. *Int J Syst Bacteriol.* 1999;49 Pt 3:1211-5.
 89. Trevena WB, Willshaw GA, Cheasty T, Domingue G, Wray C. Transmission of Vero cytotoxin producing *Escherichia coli* O157 infection from farm animals to humans in Cornwall and west Devon. *Commun Dis Public Health.* 1999;2(4):263-8.
 90. Shao Y, Lu N, Wu Z, Cai C, Wang S, Zhang LL, Zhou F, Xiao S, Liu L, Zeng X, Zheng H, Yang C, Zhao Z, Zhao G, Zhou JQ, Xue X, Qin Z. Creating a functional single-chromosome yeast. *Nature.* 2018;560(7718):331-5.
 91. Reno ML, Held NL, Fields CJ, Burke PV, Whitaker RJ. Biogeography of the *Sulfolobus islandicus* pan-genome. *Proc Natl Acad Sci U S A.* 2009;106(21):8605-10.
 92. Suzuki R, Shimodaira H. Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics.* 2006;22(12):1540-2.
 93. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular Plant.* 2020;13(8):1194-202.

Figures

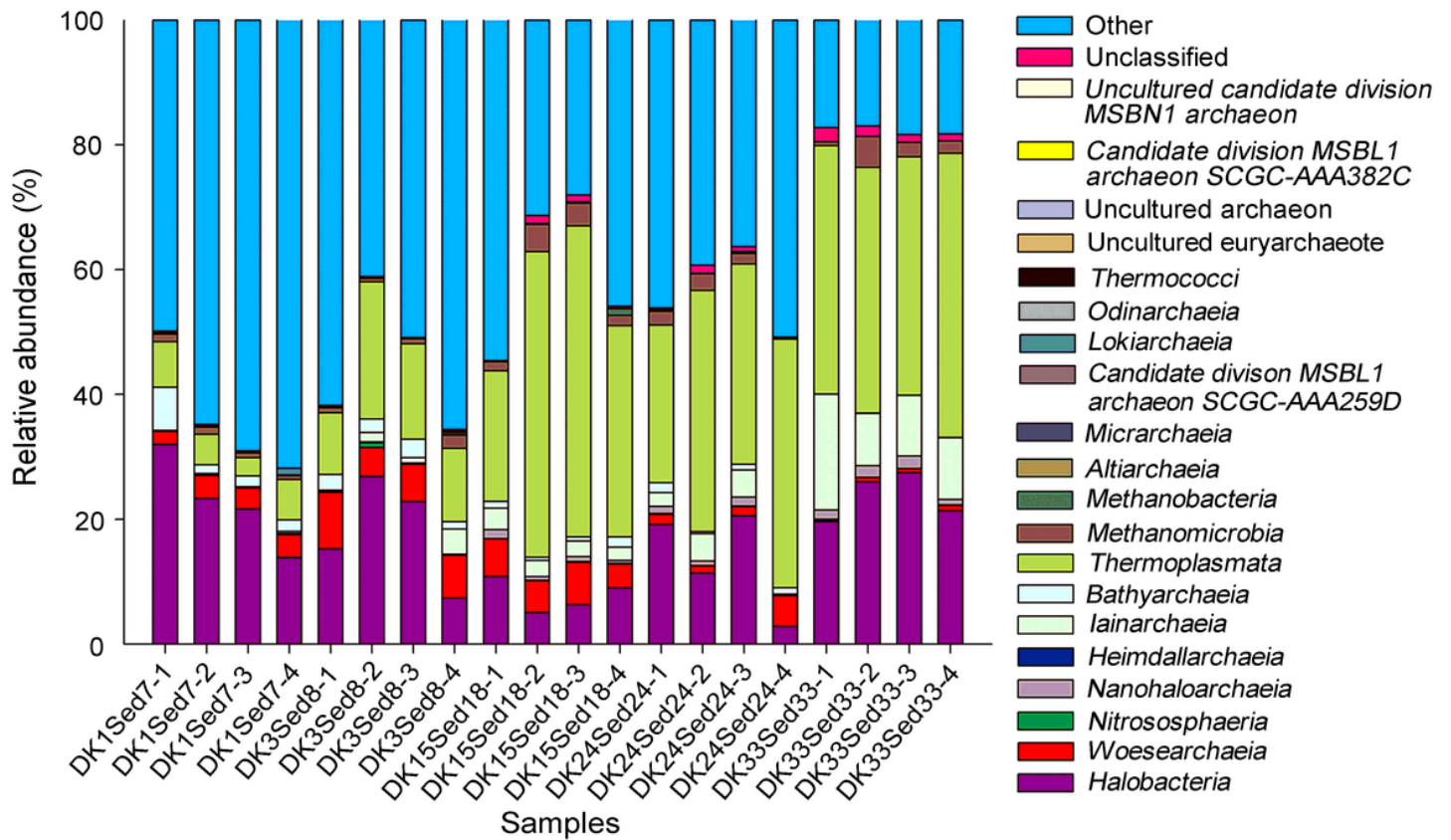


Figure 1

Archaea community composition profiles of deep sediment in soda lake with different salinity. The top lineages at class levels were exhibited based on the SILVA database release 136. There were four duplicates for each sediment.

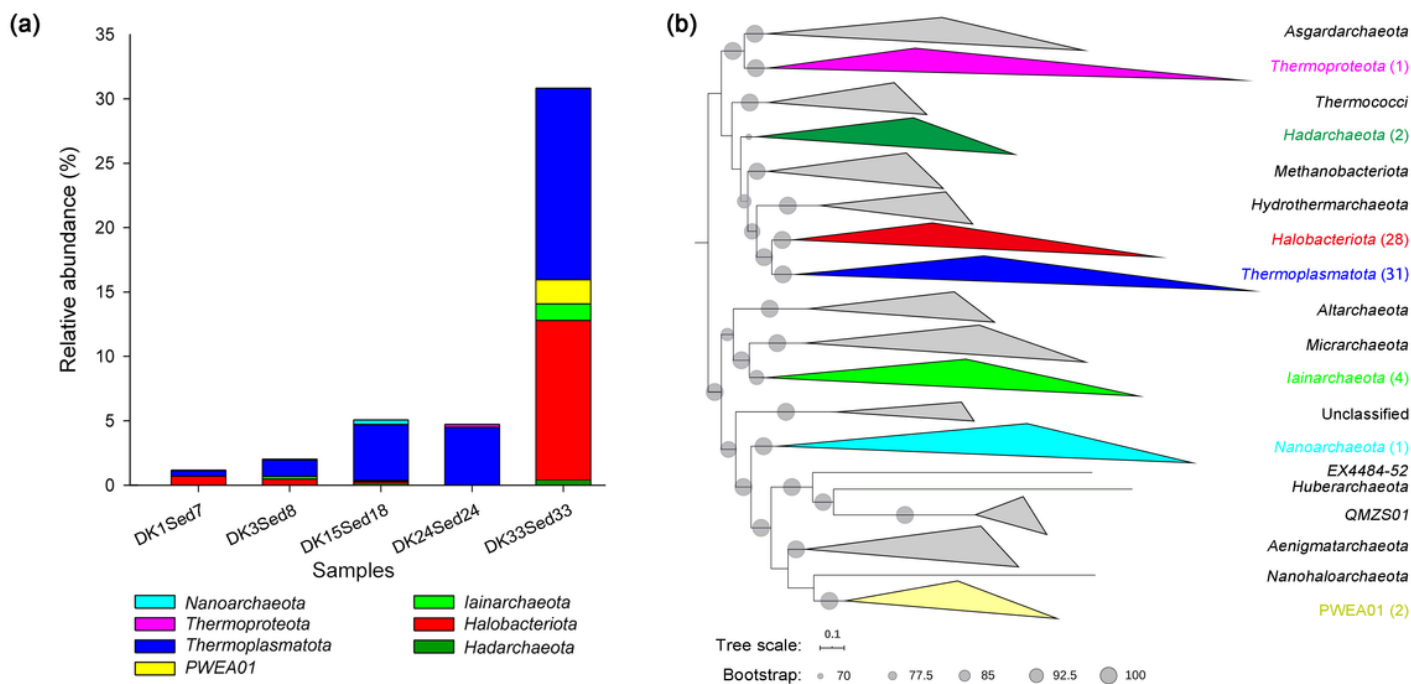


Figure 2

Archaeal community compositions of soda lake sediment based on metagenomics. (a) Relative abundance of archaea based on archaeal MAGs at phylum level. (b) The phylogeny of archaeal MAGs based on the concatenated 122 ubiquitous single-copy archaeal proteins according to GTDB. Bacteria were set as outer branch. The phyla found in this work were colored. The number of MAGs was added in the bracket behind the phylum name. Bootstrap resampling was 1000 times.

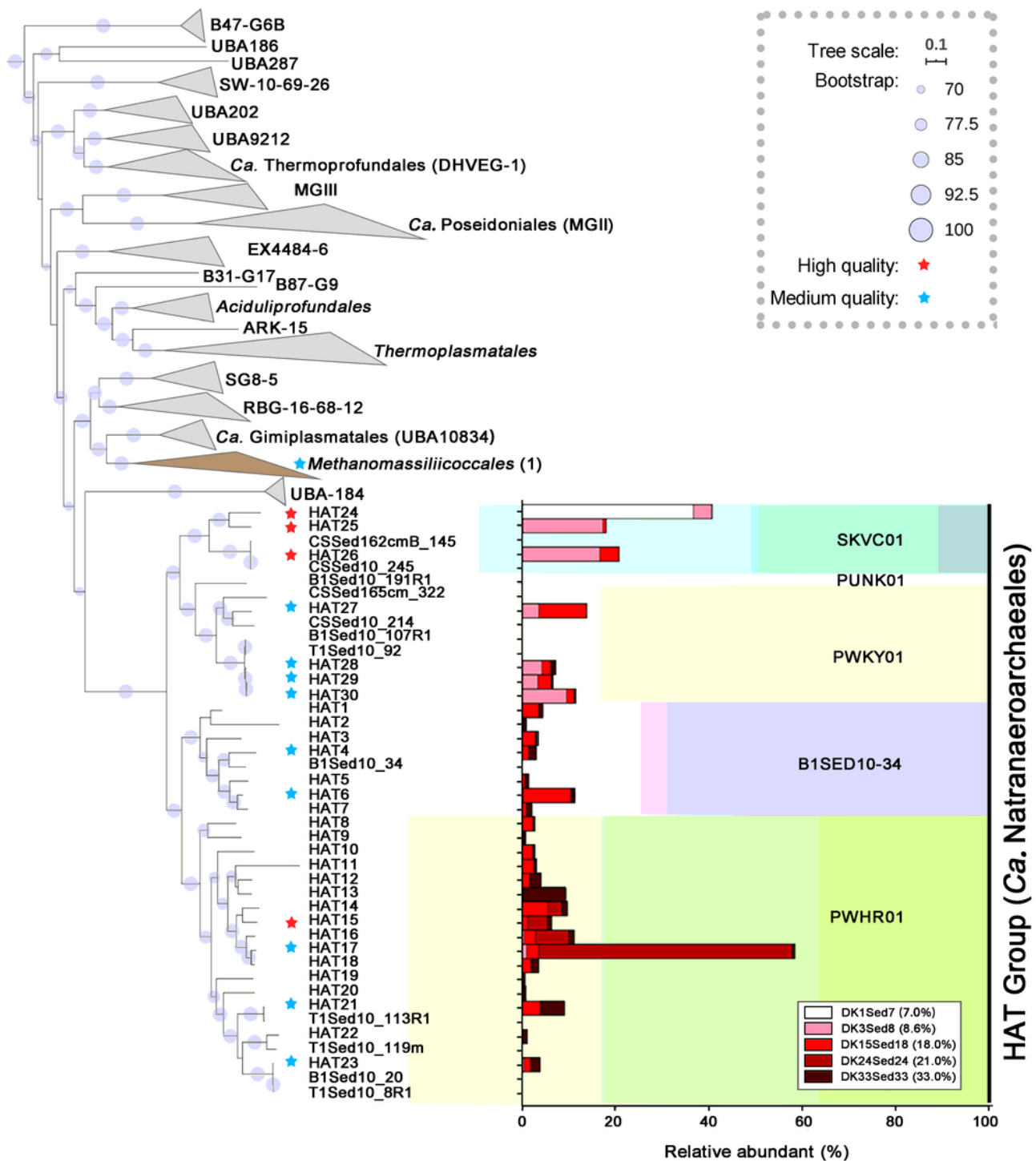


Figure 3

Phylogenetic tree placing *Ca. Natranaeroarchaeales* into phylum *Thermoplasmata* and the relative abundance of each MAG. This phylogenetic tree was based on the concatenated 122 ubiquitous single-copy archaeal proteins according to GTDB using maximum likelihood method. Red and blue star indicated the high- (completeness >90, contamination < 5%) and medium-quality (completeness >70, contamination < 10%) MAGs, respectively. Bootstrap resampling was 1000 times, and bootstrap value of

more than 70% were showed. The integrated bar graph showed the relative abundance of MAGs in each sediment. There were five genera marked by different background in this order according to GTDB and manually modified.

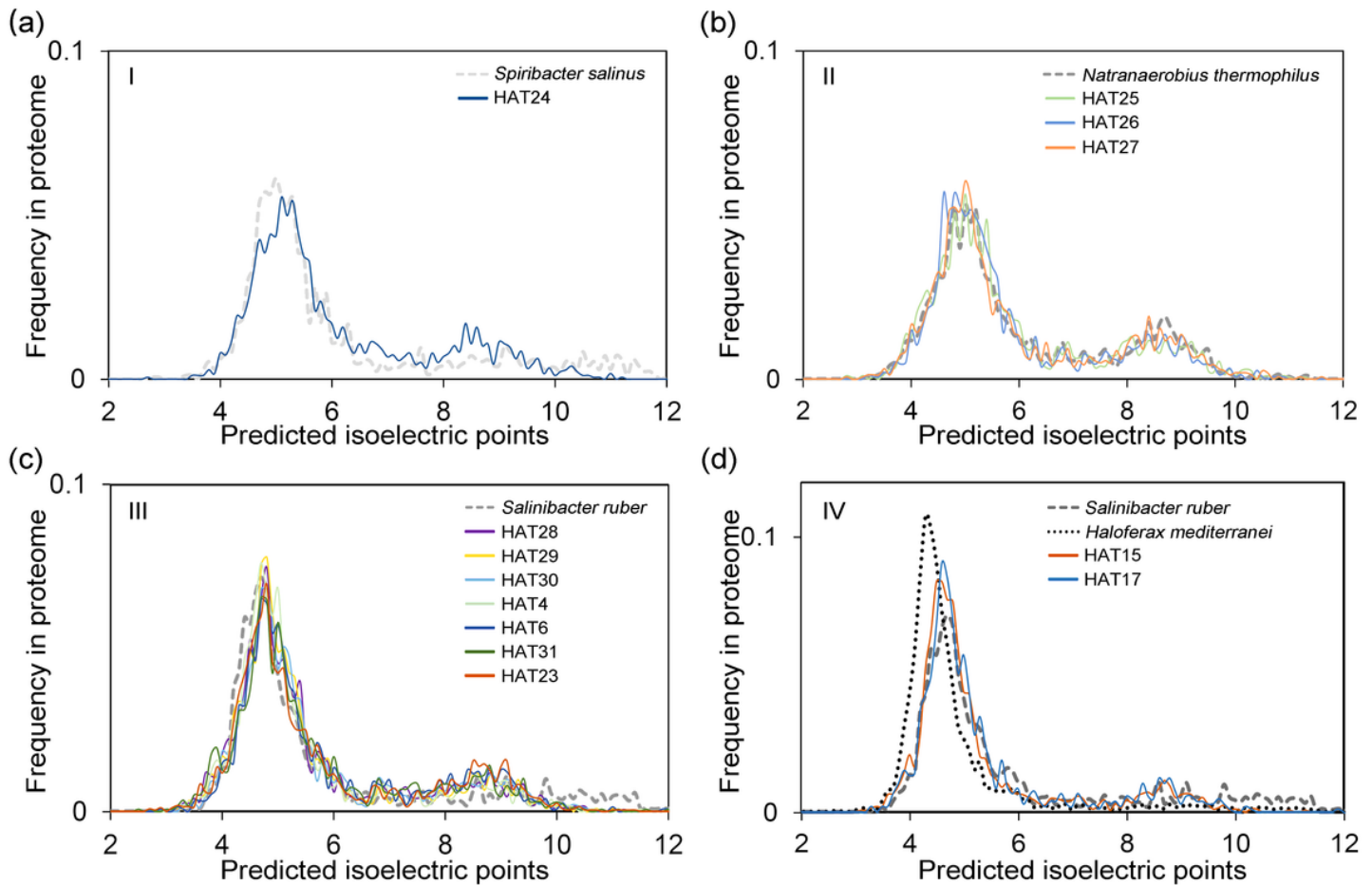


Figure 4

Isoelectric point profiles of predicted proteomes of Ca. Natranaeroarchaeales and reference species. According to Fig. S6, the isoelectric point (pI) profiles of four classes including Ca. Natranaeroarchaeales MAGs were exhibited in a (I), b (II), c (III) and d (IV). Avg pI means average isoelectric point. The genomes of the *Spiribacter salinus*, *Natranaerobius thermophilus*, *Salinibacter ruber*, *Haloferax mediterranei*, *Bacillus subtilis*, *Escherichia coli*, *Saccharomyces cerevisiae* and *Sulfolobus islandicus* were chosen as reference.

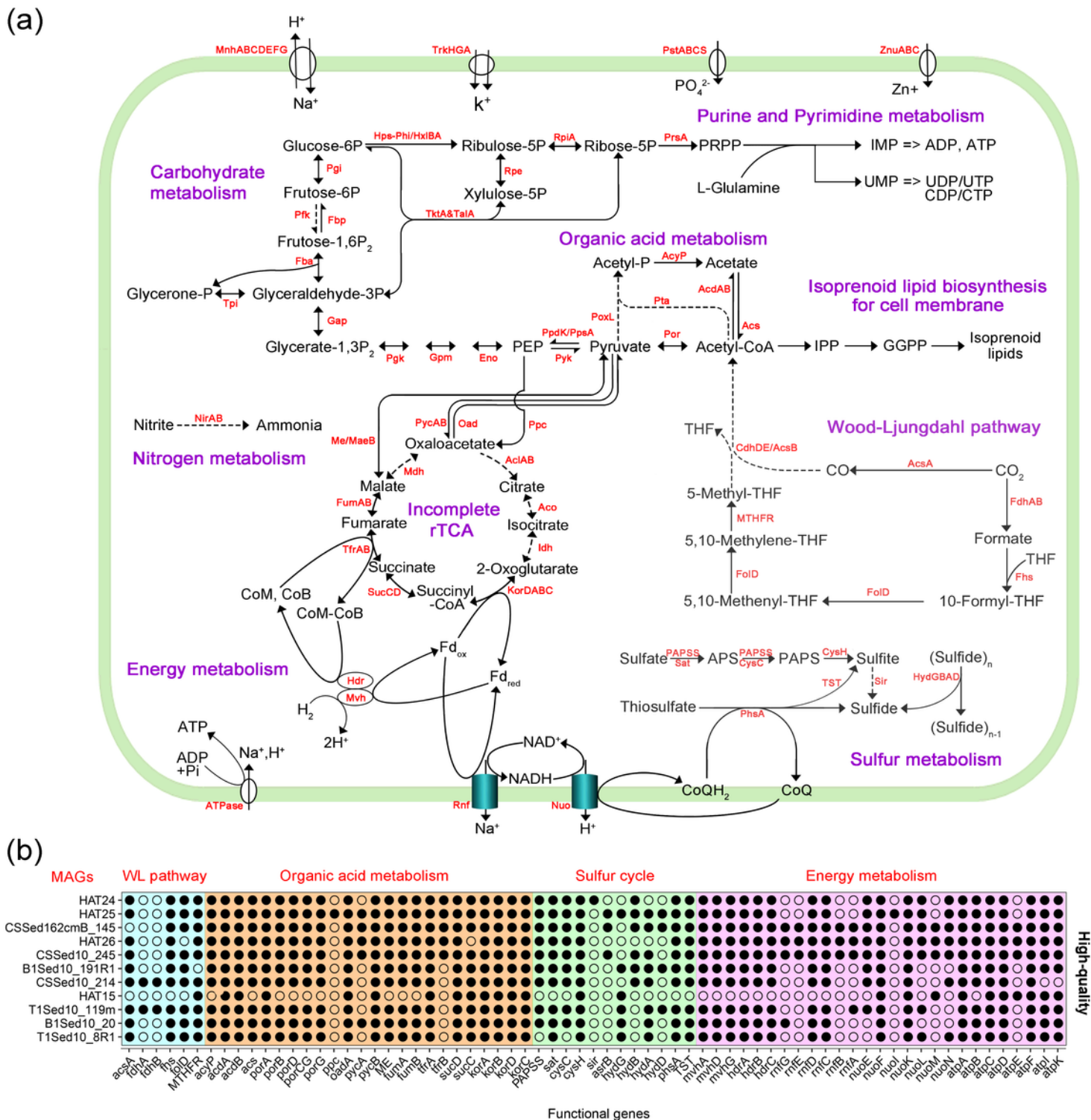


Figure 5

Metabolic potentials of *Ca. Natranaeroarchaeales* based on the genomic prediction. (a) Pathway reconstruction of the key processes on carbon, sulfur and energy metabolism in *Ca. Natranaeroarchaeales*. Dash line indicated the process was less or not presence in the MAGs, but some may be catalyzed by unidentified protein. (b) Dot plot showing the detailed gene information involving in Wood-Ljungdahl pathway, organic acid metabolism, sulfur cycle and energy metabolism in high-quality

MAGs (including 4 MAGs obtained in this article and 7 downloaded from GTDB). Solid and hollow dots indicated the presence and absence in the MAGs, respectively. More information (including medium- and low-quality MAGs) involving in metabolism was shown in Fig. S7-S12 and Table S8. Enzymes and compounds abbreviations were listed in supplementary materials.

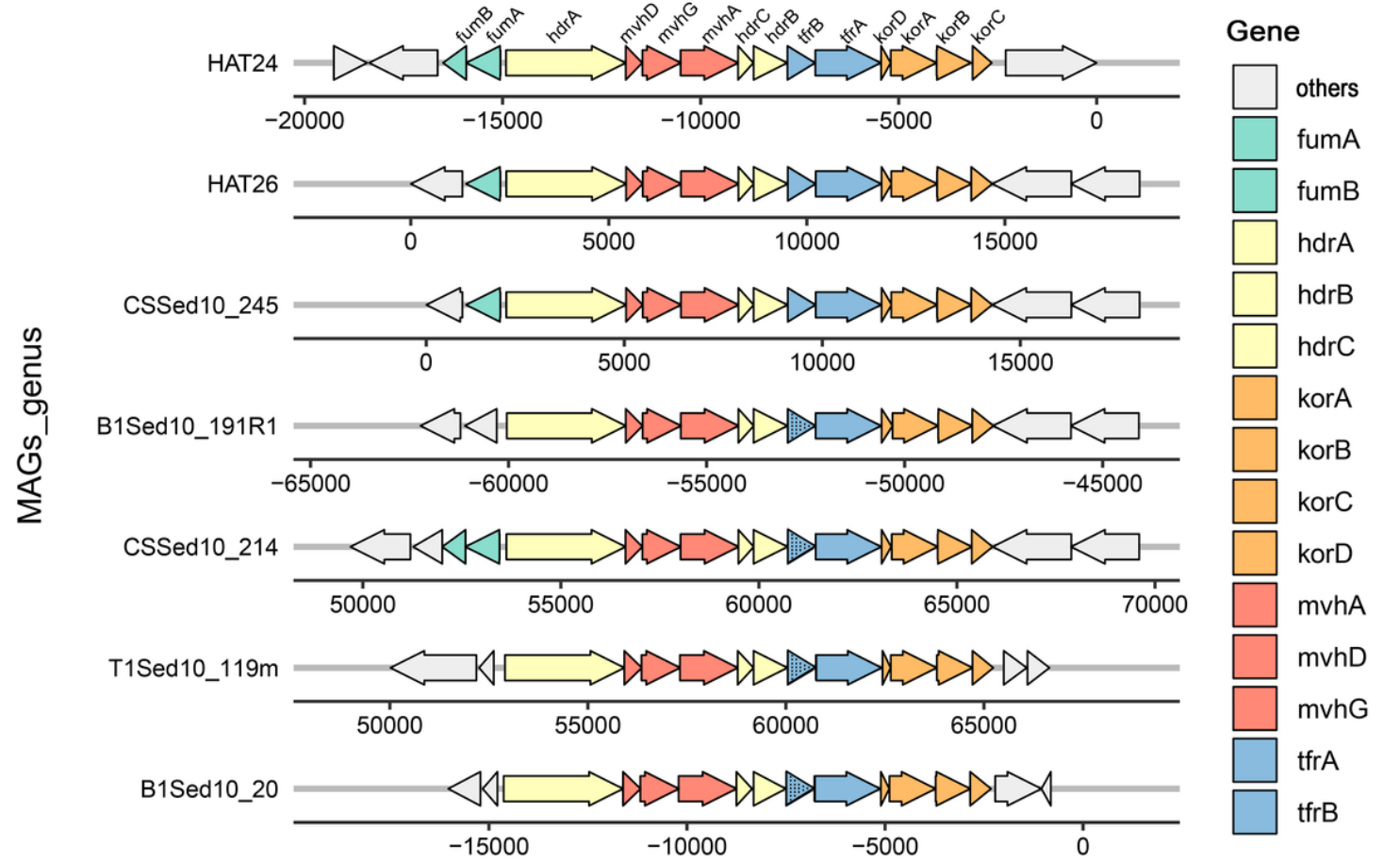


Figure 6

Gene arrangement of *hdrABC*, *mvhADG* and *tfrAB* in selected MAGs of *Ca. Natranaeroarchaeales*. Longitudinal coordinate exhibited the MAG names. The genes were aligned by *hdrB* gene. The number in the horizontal ordinate indicated the position in the contig, and minus meant contig sequence was reverse complement. The length and direction of the arrow showed the gene length and direction. The enzymes abbreviations were listed in supplementary materials.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementmaterials.pdf](#)
- [Supplementarytables.xlsx](#)