Supplementary Materials for Magnuson et al. "Active lithoautotrophic and methane-oxidizing microbial community in an anoxic, sub-zero, and hypersaline High Arctic spring".

Extended Materials and Methods

i. Site description and sample collection

Lost Hammer (LH) spring discharges through a precipitated mineral salt tufa as described in previous publications (1-5) (Figure S1). LH emits gases composed of methane (50%), nitrogen (35%), carbon dioxide (10%), and trace hydrogen, helium, and short-chain alkanes (1). The spring sediments and water contain high concentrations of sulfate (100 000 mg/kg) as well as sulfide (<50 mg/kg), ammonia (2.55 mg/kg), nitrate/nitrite (2.87 mg/kg), and iron (13 000 mg/kg) (1). Physical and geochemical parameters in the outlet (Table S1) have remained highly stable since 2005, allowing comparison among samples collected in different years.

For this study, sediment samples (top ~10 cm) were collected in July 2017 and July 2019 with an ethanol-sterilized scoop, and stored in sterile Falcon tubes filled to maximum to avoid aerobic headspace. Sediment from July 2017 was used for metagenomic sequencing, and sediment from July 2019 was used for RNA and SAG sequencing. In parallel, sediment for RNA extractions was mixed with Zymo Research DNA/RNA Shield (Irvine, CA, USA). Samples were kept at <5°C during transportation to Montreal, after which they were stored at -20°C for DNA and RNA extraction and at -5°C (unfrozen) for SAG sequencing. Physical and geochemical parameters of the overlying water were measured *in situ* with a YSI Professional Plus Multiparameter instrument (Yellow Springs, OH, USA) and a PyroScience Piccolo2 oxygen meter (Aachen, Germany). Ortho-phosphate and ammonia were measured *in situ* with CHEMetrics Inc. (Midland, VA, USA) test kits.

ii. DNA extraction, metagenome sequencing, and metagenome data analyses

DNA was extracted from two 5 g portions of sediment with a Qiagen DNeasy PowerMax Soil Kit (Hilden, Germany). The resulting DNA from each sediment sample was concentrated with a Thermo Fisher Scientific SpeedVac Vacuum Concentrator (Waltham, MA, USA) and sequenced on a HiSeq2500 (2x126 base reads) (Illumina, San Diego, CA, USA) at The Centre for Applied Genomics (Toronto, ON, Canada). Low-quality reads and bases were trimmed with Trimmomatic (v0.38, settings LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15) (6). Remaining reads were classified with Kaiju (v1.7.3, default settings, nr_euk database) (7) and phyloFlash (v3.4) (8). The reads from the two sediment samples were co-assembled with Megahit (v1.1.3, setting meta-sensitive) (9) as well as assembled separately with metaSPAdes (v3.13.0, default settings) (10). Reads were mapped to each assembly with BBMap (v38.26, minid=0.95) and contigs longer than 5000 bp were binned with MetaBAT (v2.12.1) (11). Bin completeness and contamination was estimated with CheckM (v1.0.12) (12). The Megahit co-assembly and resulting bins were selected for downstream analysis based on sequencing statistics as determined by metaQUAST (v5.0.1) (13) and number of high- and medium-quality bins. Sequencing and assembly statistics can be found in Tables S2 and S3. The metagenome was annotated with the Joint Genome Institute's IMG/M system using KEGG, COG, and pfam databases (14, 15). Bins were classified with the Genome Taxonomy Database Toolkit (v1.3.0, reference data R05-RS95) (16). Phylogenomic trees were constructed with Anvi'o (v6.2) using the Bacteria_71 collection of single copy genes (17). Amino acid sequences for all genes were concatenated, with a total alignment length of 24 451 bp, and approximately-maximum-likelihood trees were constructed in FastTree with default settings within Anvi'o, with midpoint rooting. Additional analyses were as follows: FeGenie was used to identify iron-related genes (18); *amoA* and *pmoA* were distinguished by HMMer using FunGene Hidden Markov Models (19); hydrogenases were classified with hydDB (20). Reductive and oxidative DsrAB were classified as follows: DsrAB amino acid sequences were aligned against reference sequences from Muller et al. (21) using MUSCLE with default settings (22). Maximum likelihood phylogenetic trees were constructed in CLC Genomics Workbench (v. 12.0.3) using the WAG protein substitution model and 1000 bootstraps (Figure S9). DsrAB were classified as reductive or oxidative based on phylogenetic clustering and the presence of accessory proteins as in Anantharaman et al. (23).

iii. Single cell sorting, genome amplification, sequencing, and data analyses

Sediment was shipped on ice to the Single Cell Genomics Center at the Bigelow Laboratory for Ocean Sciences (East Boothbay, ME, USA) for SYTO-9 fluorescence-activated single-cell sorting (FACS), DNA extraction, and genome amplification with WGA-X (as described in Stepanauskas et al. (24)). The 16S rRNA gene was PCR amplified from the genomes (primers 27F (25) and 1492R (26)) and sequenced at the Centre de Recherche at Université Laval (Quebec City, QC, Canada) on an Applied Biosystems 3730xl DNA Analyzer (Foster City, CA, USA). Obtained 16S rRNA sequences were annotated by BLAST with the SILVA rRNA database (release 138) (27). For whole genome sequencing, libraries were prepared with a Nextera XT DNA Library Prep Kit and sequenced on a MiSeq (Illumina) with MiSeq Reagent Kit v3 (600 cycles, 2x300 base reads). Low quality reads and bases were trimmed with BBDuk (minimum Phred quality score 15, minimum length 30 bp) and contaminant human reads were removed with DeconSeq (28). Genomes were assembled with SPAdes (v3.13.1, settings --sc --careful) (29) and screened for contamination using JGI's Kmer Frequency Analysis tool and by read classification with Kaiju (v1.7.3) (7). Average nucleotide identity of the SAGs against other SAGs and metagenome bins was calculated with FastANI (v1.3) (30). Genome annotation was done as described for the metagenome.

iv. mRNA extraction, sequencing, and analysis

RNA was extracted in triplicate with a Zymo Research ZymoBIOMICS DNA/RNA Miniprep Kit from approximately 3 g sediment per extraction. Extracted samples were then treated with the Invitrogen Turbo DNA-free kit (Carlsbad, CA, USA) to remove contaminating DNA. The treated samples were then pooled and concentrated with a New England Biolabs Monarch RNA Cleanup Kit (Ipswitch, MA, USA). Ribosomal RNA was depleted with a New England BioLabs NEBNext rRNA Depletion Kit (Bacteria) and a sequencing library was prepared with a New England BioLabs Ultra II RNA Library Prep Kit. The generated cDNA library was sequenced at The Center for Applied Genomics at the Hospital for Sick Children (Toronto, ON, Canada) on a NovaSeq 6000 (Illumina) with an S Prime 100-cycle flow cell (2x100 base reads). Reads were trimmed with Trimmomatic (settings LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN 50) and rRNA reads were removed with SortMeRNA (v4.2.0) (31). Human contaminant reads were removed with BBMap using the BBTools RemoveHuman masked reference genome. Reads were classified with Kaiju (v1.7.3), and reads classified within common sequencing contaminant groups as in Sheik et al. (2018) (32) were removed (~81,500 reads removed representing 77 genera; the genera with the most reads removed were

Pseudomonas (~12,000 reads), *Streptococcus* (~7000 reads), and *Ralstonia* (~6000 reads)). Remaining reads were aligned to metagenome contigs and SAG scaffolds with bowtie2 (v2.3.5.1, with setting -very-sensitive-local) (33). Reads aligned to CDS regions were counted with HTSeq (v0.12.4) (34) and transcripts per million reads (tpm) was calculated to normalize expression values for each gene. For comparison of stress response genes in SAGs to those in related genomes, protein-coding genes from reference genomes were queried against LH SAG protein-coding genes with BLASTp using a cut-off e-value of 1e-15.

v. Sequencing data availability

Sequencing reads, metagenome, MAGs, and SAGs were deposited in NCBI under BioProject PRJNA699472. JGI annotations of the metagenome and SAGs are available under GOLD Study ID Gs0135943. SAGs were also deposited individually in JGI (Table S4).

vi. Gibbs free energy calculations

Gibbs free energy values were calculated using the method reported in Jones et al. (2018) (35) with the following parameters based on measurements reported in Table S1: 278.1° K, pH 6.4, 7.39 M ionic strength, 1.02×10^{-7} M CH_4, 3.81×10^{-4} M NH_3, 8.61×10^{-6} M H_2S, 1.57×10^{-5} M O_2, 2.31×10^{-5} M NO_3^-, 1.04 M SO_4^{2-}, 4.11×10^{-5} M CO_2, 3.16×10^{-7} M H^+, 3.12×10^{-5} M NO_2^-, and 0.23 M Fe^{3+}.

Supplementary Results and Discussion

Discrepancy between recovered MAGs and SAGs

Using the criteria of average nucleotide identity >95%, only one SAG was found to correspond to MAGs. Typically, more overlap between the two datasets would be expected, particularly for those taxa abundant in the microbial community (36); conversely, the SAGs appear to be enriched for taxa at low abundance in the metagenome. In the following sections we will discuss possible reasons as to why this occurred.

Firstly, multiple selection steps occurred during generation of the SAGs that may have excluded or did exclude genomes from taxa abundant in the metagenome and MAGs. While 365 events (cells) were sorted during single cell sorting, only the 95 wells with the greatest estimated genome amplification (based on fluorescence during the amplification reaction) were selected for further analysis. This may have resulted in a biased selection of cells: for example, selecting for those genomes most responsive to the MDA reaction or those most resistant to environmental fluxes that may have occurred during transportation or sorting. Additionally, after an initial round of 16S rRNA Sanger sequencing, some wells with highly similar (>98%) 16S rRNA sequences were excluded from subsequent genomic sequencing to maximize sequencing coverage on the remaining wells. This included several *Halomonas* and *Desulfobulbaceae* genomes represented by MAGs, therefore potentially reducing the number of SAGs that may have corresponded to MAGs. As a result of these filtering steps, SAGs corresponding to MAGs may have been excluded and low-abundance taxa including archaeal genomes may have been enriched.

Secondly, 20% of reads map to the high- and medium-quality MAGs, due in part to stringent quality control during binning (i.e. only contigs >5000 bp were binned). As a result, while the

MAGs broadly represent the taxonomic groups present in the metagenome, they represent only a portion of the sum diversity. Therefore, this restrictive binning process in combination with the relatively high number of SAGs corresponding to taxa at low abundance in the metagenome may also have contributed to the discrepancy.

Thirdly, the samples collected for metagenomic analysis and for SAG sequencing were collected in different years (2017 vs. 2019, respectively). This was necessitated by the small amount of sediment removed during sampling to avoid disturbing the site in combination with the low biomass of the sediment, requiring relatively high amounts of sample for processing. While the physical and geochemical parameters of the spring have remained stable for the ~15 years that we have studied the spring, including the measurements taken in 2017 and 2019, we can’t exclude the possibility of minor changes in the spring that could have affected the microbial community sampled. Additionally, due to the low biomass and high salinity of the sediment, it is difficult to extract DNA. Biases introduced by the differing processing steps during MAG and SAG sequencing (for example, DNA extraction vs. separation of cells from sediment) may therefore have had outsize effects.

To summarize, we suggest several factors potentially contributing to the discrepancy between MAGs and SAGs: 1) Selection during generation of SAGs, 2) Stringent binning criteria, and 3) Differences in input samples and biases introduced by differing experimental procedures. We conclude by noting that although there is a discrepancy between the MAGs and SAGs, there is overlap between the two samples additional to what was discussed in the manuscript. In addition to the corresponding SAG and MAG noted in the manuscript, two additional SAGs had 16S rRNA sequences >98% identical to unbinned 16S rRNA sequences, and nearly all taxonomic groups represented by the SAGs were also present in the metagenome reads and 16S sequences as classified by kaiju and PhyloFlash (with the sole exception of Iainarchaeota reads). Additionally, comparison of 16S rRNA sequences in the SAGs to previous 16S rRNA sequencing from over ten years ago (1) identified common sequences (>98%) between the two datasets, including sequences for *Halomonas*, ANME-1, and *Iainarchaeota*, suggesting that the discrepancy is more likely due to the potential technical factors discussed above rather than significant changes in the microbial community. The relative abundances of the microbial community represented in the metagenome are consistent with those observed in previous CARD-FISH and 16S rRNA and metagenomic sequencing (1, 2). We therefore suggest that the metagenome and MAGs more accurately represent the taxa abundant in the microbial community, whereas the SAGs are disproportionately enriched in low-abundance bacteria and archaea due to the potential factors discussed above.

Supplementary Figures and Tables

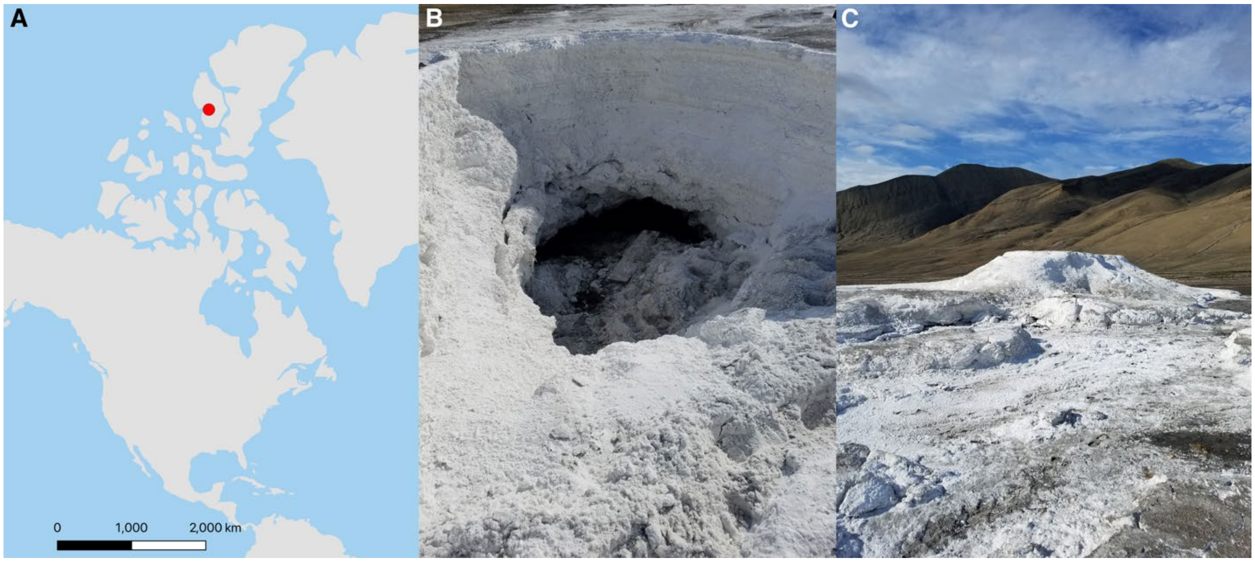

Figure S1. A. Location of the LH spring on Axel Heiberg Island in the Canadian High Arctic. Map generated in QGIS with the Natural Earth dataset. B. View looking into the LH salt tufa to the spring source (July 2019). C. View of the LH tufa and outflow channels (July 2019).

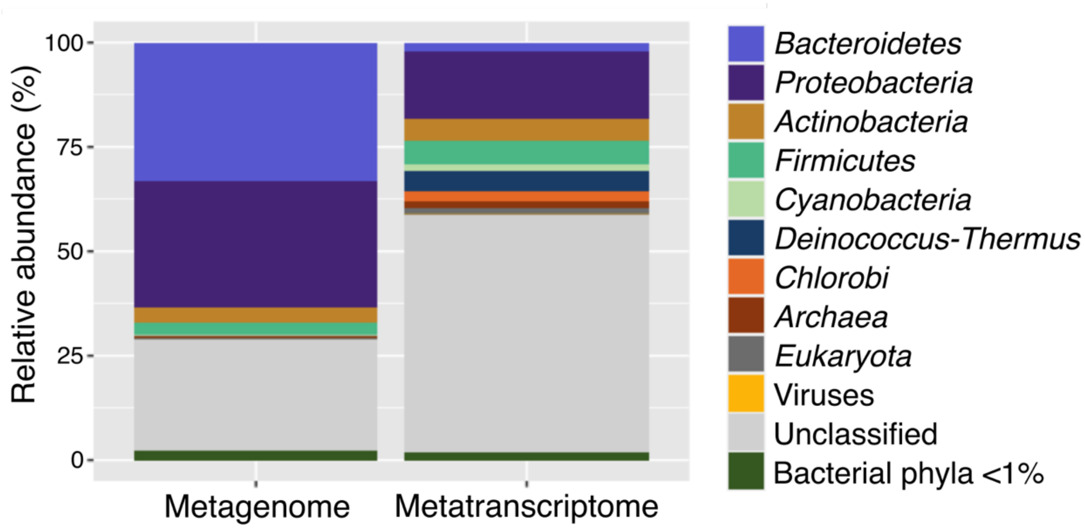

Figure S2. Taxonomic classification of metagenome and metatranscriptome reads. Reads were classified by Kaiju using the NCBI non-redundant database. Metagenome reads are an average of the relative abundance in duplicate samples.

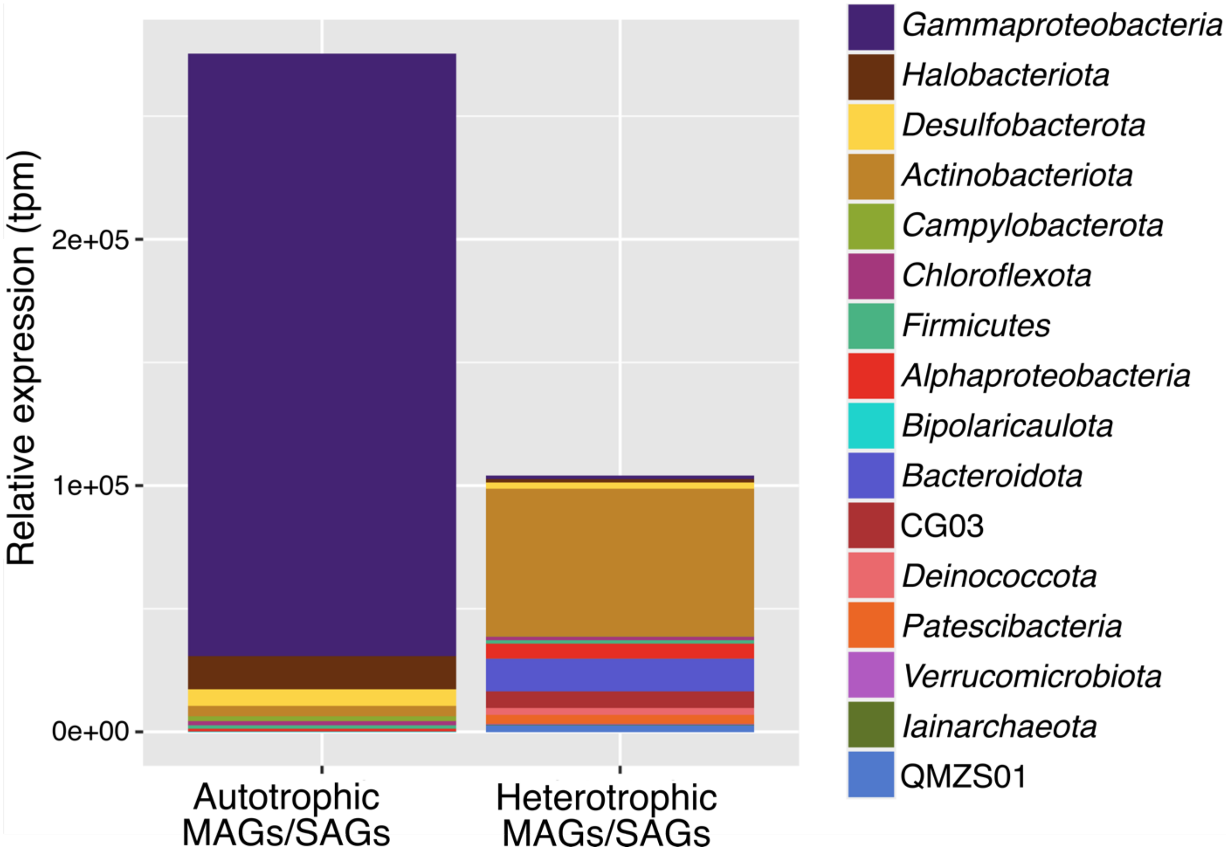

Figure S3. Relative expression (tpm) attributed to MAGs and SAGs containing CO$_2$-fixation genes (autotrophic MAGs/SAGs) and those without (heterotrophic MAGs/SAGs). Unbinned genes with mapped transcripts were omitted from this analysis.

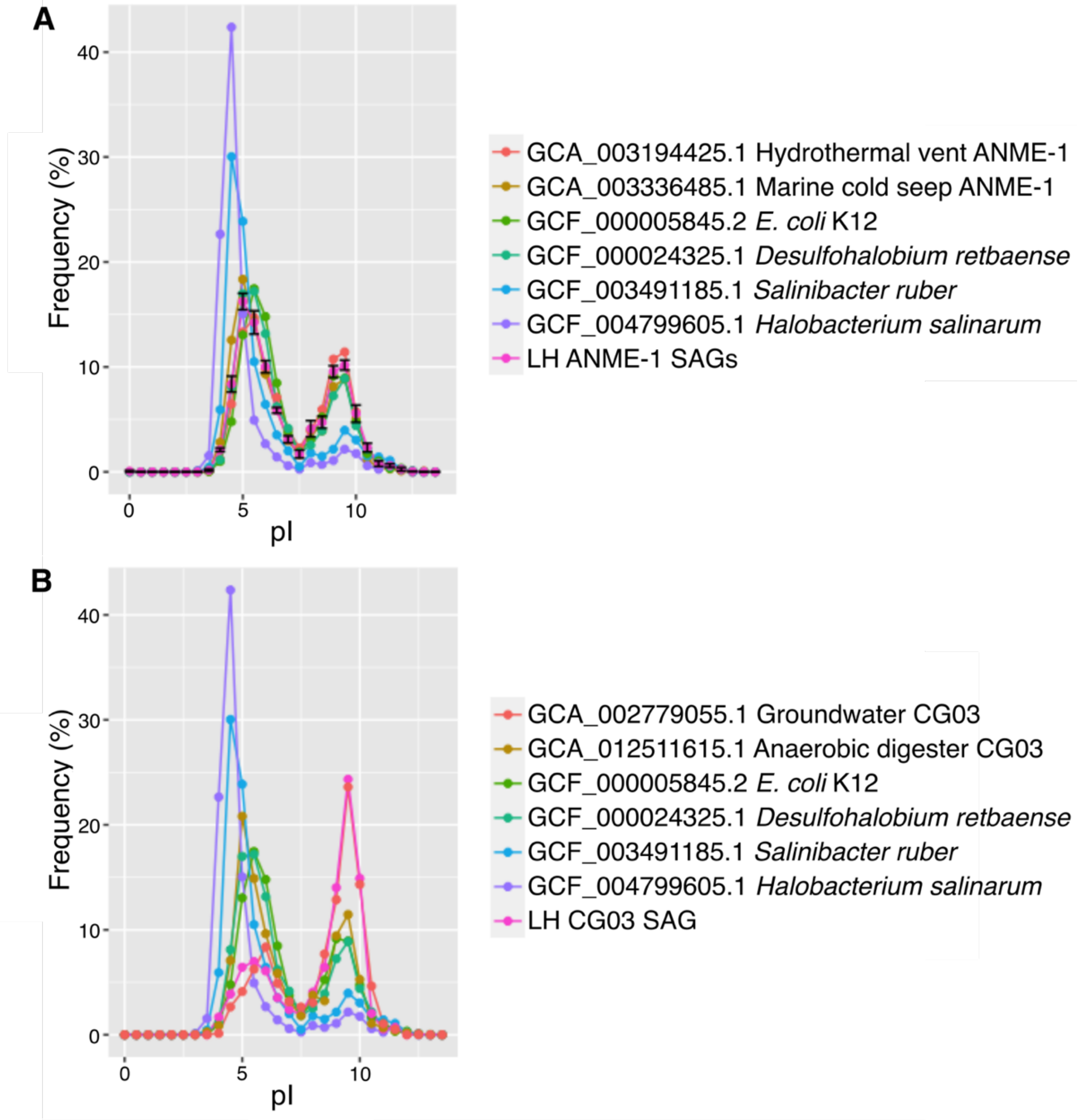

Figure S4. Comparison of predicted protein isoelectric points for A. ANME-1 SAGs S10-S14 and B. CG03 SAG S2. Protein isoelectric points were calculated with the ExPASy Compute pI/MW online tool. Genomes for comparison (based on similar analysis in Nigro et al. (37)) include microorganisms known to accumulate salts through the "salting-in" osmoregulation strategy (*Salinibacter ruber* and *Halobacterium salinarum*) and those that do not accumulate salts (*E. coli* K12 and *Desulfohalobium retbaense*).

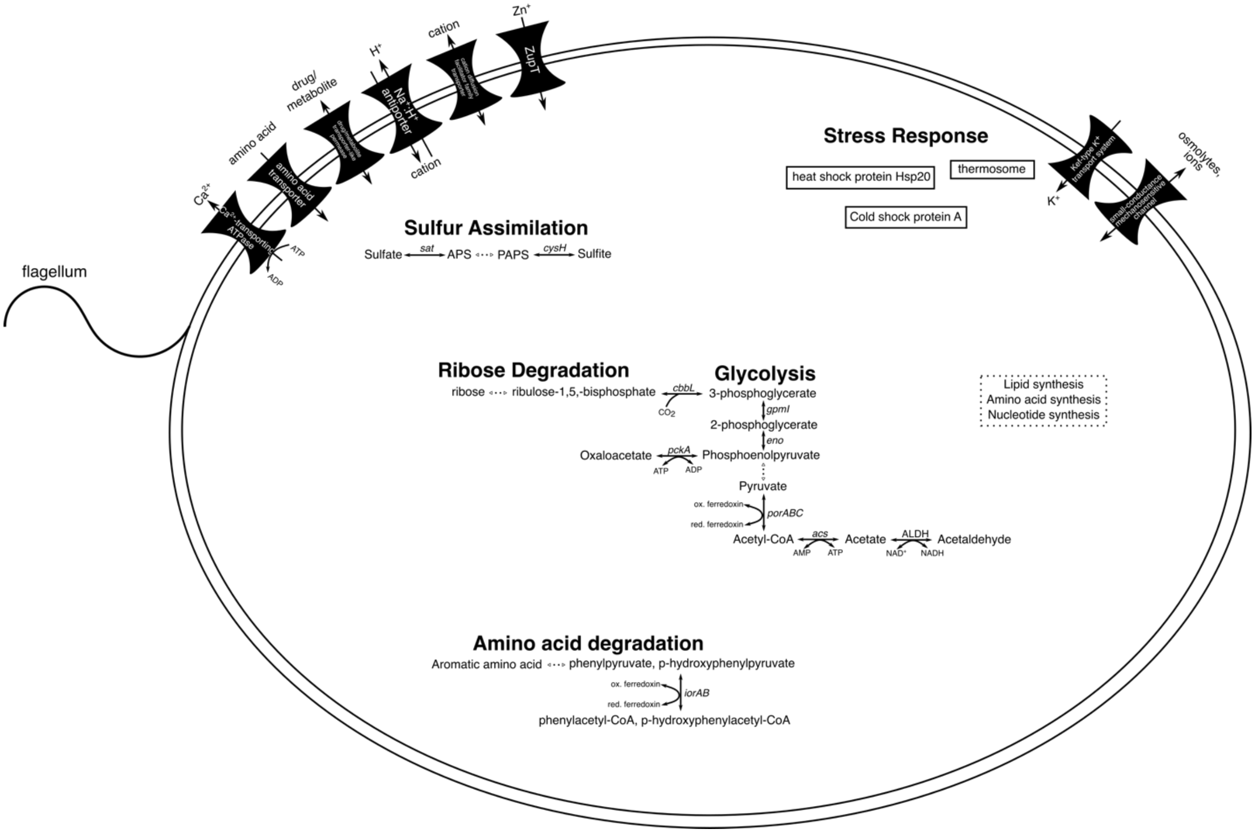

Figure S5. Metabolic reconstruction of *Iainarchaeia* SAG S9 (670822 bp, 52.9% completeness, 0% contamination). Solid lines represent genes present within the genome, and dashed lines indicate steps or pathways absent from the genome. Table S15 contains a complete list of annotated genes represented in this figure.

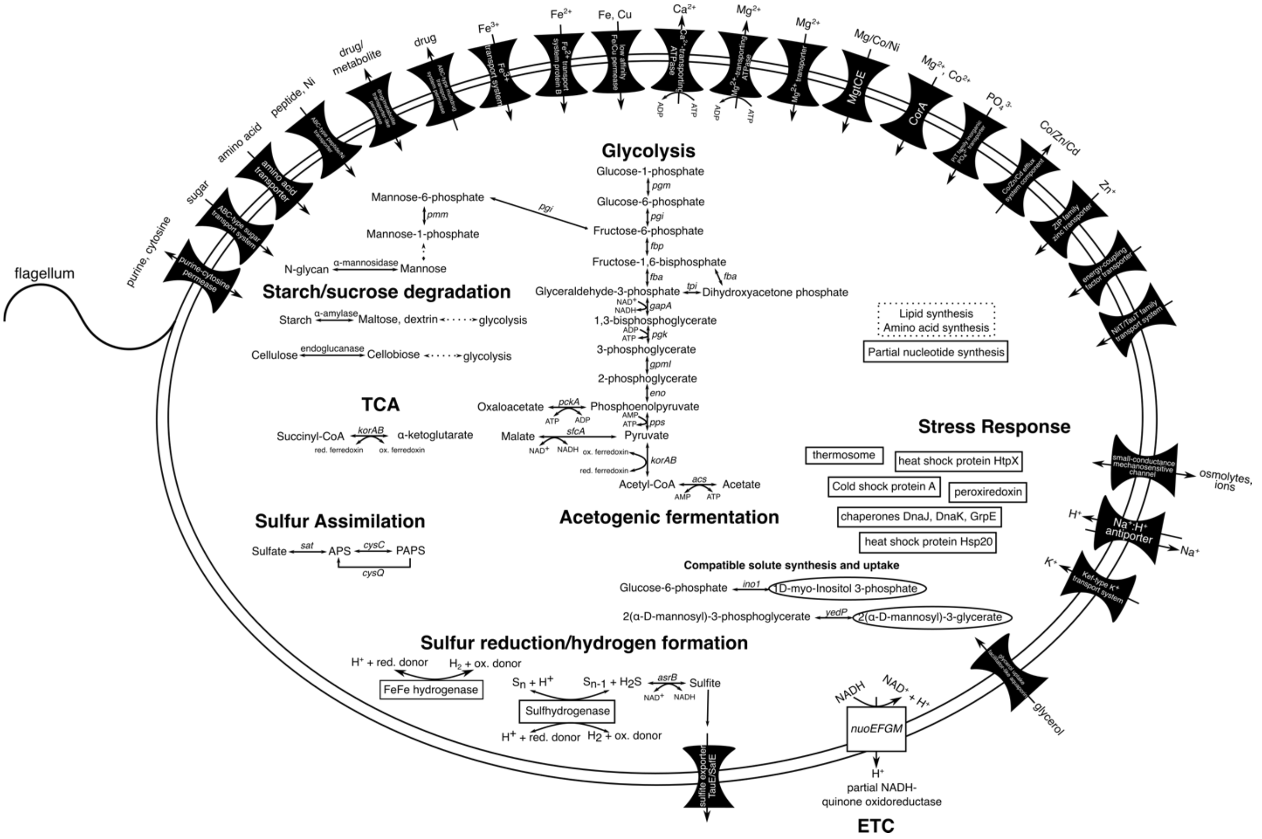

Figure S6. Metabolic reconstruction of QMZS01 SAGs S4-S8 (between 563912-779153 bp, 57.7-76.2% completeness, 0.8-2.2% contamination). Solid lines represent genes present within the genome, and dashed lines indicate steps or pathways absent from the genome. Table S16 contains a complete list of annotated genes represented in this figure.

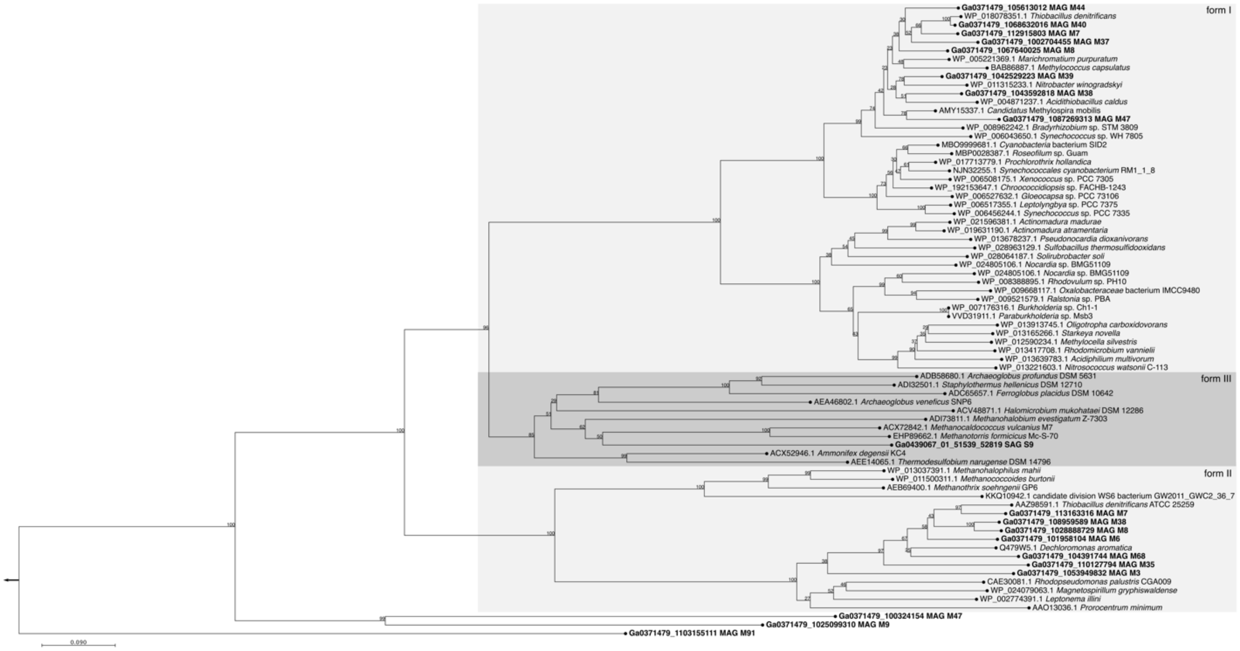
form I
Ga0371479_105613012 MAG M44
WP_018078351.1 Thiobacillus denitrificans
Ga0371479_1068632016 MAG M40
Ga0371479_112915803 MAG M7
Ga0371479_1002704455 MAG M37
Ga0371479_1067640025 MAG M8
WP_005221369.1 Marichromatium purpuratum
BAB86887.1 Methylococcus capsulatus
Ga0371479_1042529223 MAG M39
WP_011315233.1 Nitrobacter winogradskyi
Ga0371479_1043592818 MAG M38
WP_004871237.1 Acidithiobacillus caldus
AMY15337.1 Candidatus Methylospira mobilis
Ga0371479_1087269313 MAG M47
WP_008962242.1 Bradyrhizobium sp. STM 3809
WP_006043650.1 Synechococcus sp. WH 7805
MBO9999681.1 Cyanobacteria bacterium SID2
MBP0028387.1 Roseofilum sp. Guam
WP_017713779.1 Prochlorothrix hollandica
NJN32255.1 Synechococcales cyanobacterium RM1_1_8
WP_006508175.1 Xenococcus sp. PCC 7305
WP_192153647.1 Chroococcidiopsis sp. FACHB-1243
WP_006527632.1 Gloeocapsa sp. PCC 73106
WP_006517355.1 Leptolyngbya sp. PCC 7375
WP_006456244.1 Synechococcus sp. PCC 7335
WP_021596381.1 Actinomadura madurae
WP_019631190.1 Actinomadura atramentaria
WP_013678237.1 Pseudonocardia dioxanivorans
WP_028963129.1 Sulfobacillus thermosulfidooxidans
WP_028064187.1 Solirubrobacter soli
WP_024805106.1 Nocardia sp. BMG51109
WP_024805106.1 Nocardia sp. BMG51109
WP_008388895.1 Rhodovulum sp. PH10
WP_009668117.1 Oxalobacteraceae bacterium IMCC9480
WP_009521579.1 Ralstonia sp. PBA
WP_007176316.1 Burkholderia sp. Ch1-1
VVD31911.1 Paraburkholderia sp. Msb3
WP_013913745.1 Oligotropha carboxidovorans
WP_013165266.1 Starkeya novella
WP_012590234.1 Methylocella silvestris
WP_013417708.1 Rhodomicrobium vannielii
WP_013639783.1 Acidiphilium multivorum
WP_013221603.1 Nitrosococcus watsonii C-113
form III
ADB58680.1 Archaeoglobus profundus DSM 5631
ADI32501.1 Staphylothermus hellenicus DSM 12710
ADC65657.1 Ferroglobus placidus DSM 10642
AEA46802.1 Archaeoglobus veneficus SNP6
ACV48871.1 Halomicrobium mukohataei DSM 12286
ADI73811.1 Methanohalobium evestigatum Z-7303
ACX72842.1 Methanocaldococcus vulcanius M7
EHP89662.1 Methanotorris formicicus Mc-S-70
Ga0439067_01_51539_52819 SAG S9
ACX52946.1 Ammonifex degensii KC4
AEE14065.1 Thermodesulfobium narugense DSM 14796
form II
WP_013037391.1 Methanohalophilus mahii
WP_011500311.1 Methanococcoides burtonii
AEB69400.1 Methanothrix soehngenii GP6
KKQ10942.1 candidate division WS6 bacterium GW2011_GWC2_36_7
AAZ98591.1 Thiobacillus denitrificans ATCC 25259
Ga0371479_113163316 MAG M7
Ga0371479_108959589 MAG M38
Ga0371479_1028888729 MAG M8
Ga0371479_101958104 MAG M6
Q479W5.1 Dechloromonas aromatica
Ga0371479_104391744 MAG M68
Ga0371479_110127794 MAG M35
Ga0371479_1053949832 MAG M3
CAE30081.1 Rhodopseudomonas palustris CGA009
WP_024079063.1 Magnetospirillum gryphiswaldense
WP_002774391.1 Leptonema illini
AAO13036.1 Prorocentrum minimum
Ga0371479_100324154 MAG M47
Ga0371479_1025099310 MAG M9
Ga0371479_1103155111 MAG M91
0.090

Figure S7. Maximum likelihood phylogenetic tree of RbcL/CbbL sequences. The tree was constructed in CLC Genomics Workbench with 1000 bootstraps and WAG substitution model using reference sequences from NCBI. A RubisCO-like protein sequence from *Rhodopseudomonas palustris* TIE-1 (ACF00976.1) was used as an outgroup (direction indicated by arrow).

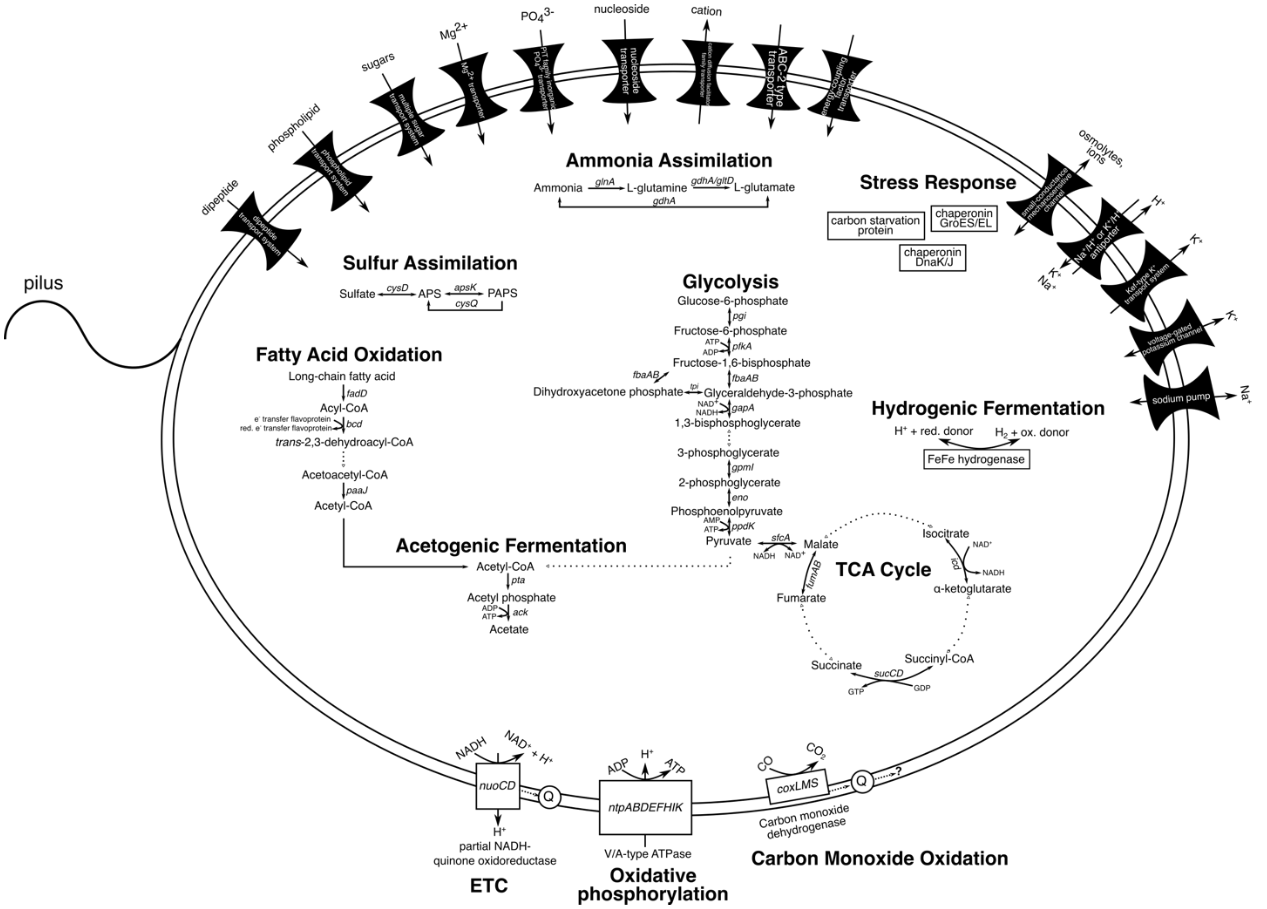

Figure S8. Metabolic reconstruction of candidate phylum CG03 SAG S2 (1039757 bp, 62.8% completeness, 0% contamination). Solid lines represent genes present within the genome, and dashed lines indicate steps or pathways absent from the genome. Table S17 contains a complete list of annotated genes represented in this figure.

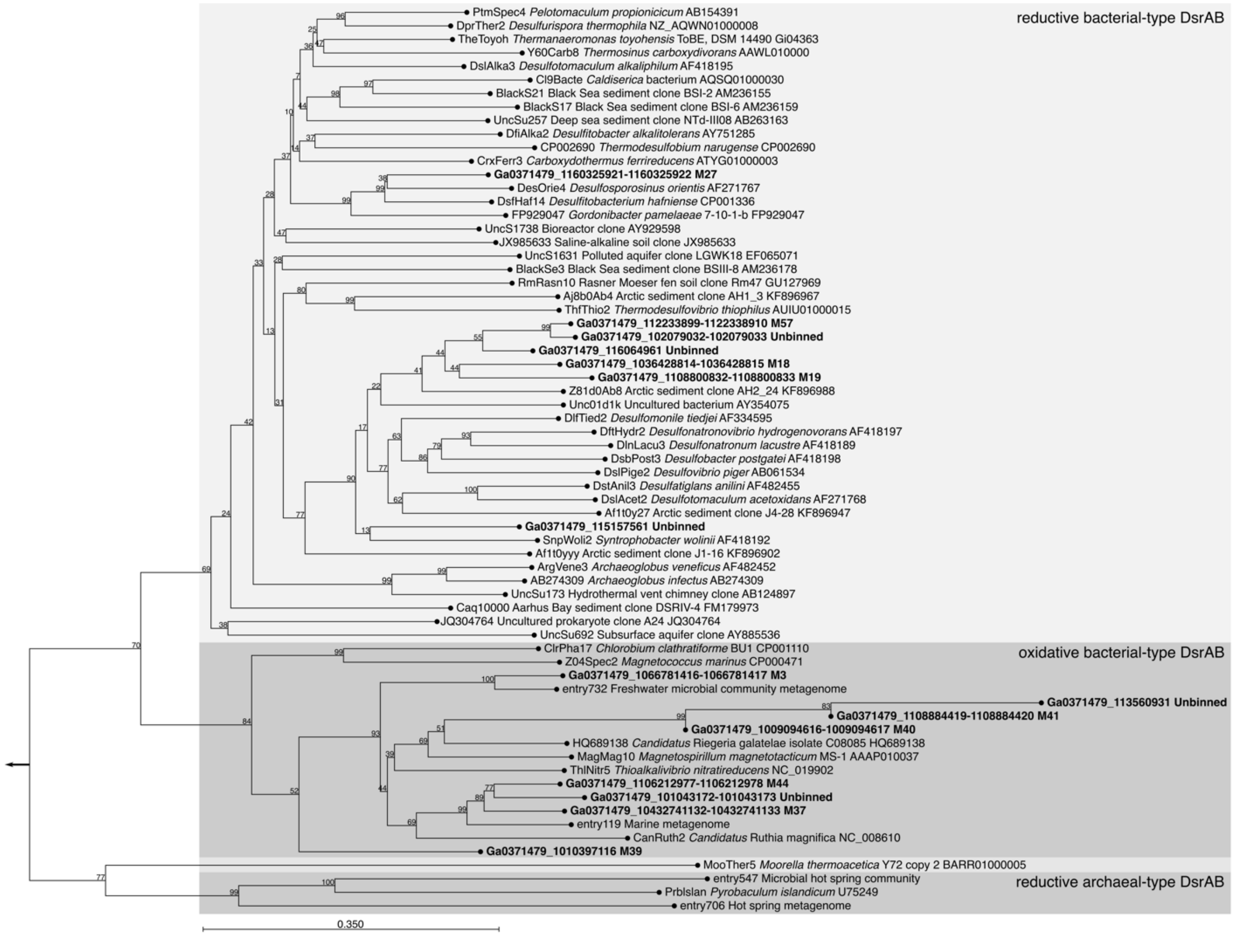
reductive bacterial-type DsrAB
PtmSpec4 Pelotomaculum propionicicum AB154391
DprTher2 Desulfurispora thermophila NZ_AQWN01000008
TheToyoh Thermanaeromonas toyohensis ToBE, DSM 14490 Gi04363
Y60Carb8 Thermosinus carboxydivorans AAWL010000
DslAlka3 Desulfotomaculum alkaliphilum AF418195
Cl9Bacte Caldiserica bacterium AQSQ01000030
BlackS21 Black Sea sediment clone BSI-2 AM236155
BlackS17 Black Sea sediment clone BSI-6 AM236159
UncSu257 Deep sea sediment clone NTd-III08 AB263163
DfiAlka2 Desulfitobacter alkalitolerans AY751285
CP002690 Thermodesulfobium narugense CP002690
CrxFerr3 Carboxydothermus ferrireducens ATYG01000003
Ga0371479_1160325921-1160325922 M27
DesOrie4 Desulfosporosinus orientis AF271767
DsfHaf14 Desulfitobacterium hafniense CP001336
FP929047 Gordonibacter pamelaeae 7-10-1-b FP929047
UncS1738 Bioreactor clone AY929598
JX985633 Saline-alkaline soil clone JX985633
UncS1631 Polluted aquifer clone LGWK18 EF065071
BlackSe3 Black Sea sediment clone BSIII-8 AM236178
RmRasn10 Rasner Moeser fen soil clone Rm47 GU127969
Aj8b0Ab4 Arctic sediment clone AH1_3 KF896967
ThfThio2 Thermodesulfovibrio thiophilus AUIU01000015
Ga0371479_112233899-1122338910 M57
Ga0371479_102079032-102079033 Unbinned
Ga0371479_116064961 Unbinned
Ga0371479_1036428814-1036428815 M18
Ga0371479_1108800832-1108800833 M19
Z81d0Ab8 Arctic sediment clone AH2_24 KF896988
Unc01d1k Uncultured bacterium AY354075
DlfTied2 Desulfomonile tiedjei AF334595
DftHydr2 Desulfonatronovibrio hydrogenovorans AF418197
DlnLacu3 Desulfonatronum lacustre AF418189
DsbPost3 Desulfobacter postgatei AF418198
DslPige2 Desulfovibrio piger AB061534
DstAnil3 Desulfatiglans anilini AF482455
DslAcet2 Desulfotomaculum acetoxidans AF271768
Af1t0y27 Arctic sediment clone J4-28 KF896947
Ga0371479_115157561 Unbinned
SnpWoli2 Syntrophobacter wolinii AF418192
Af1t0yyy Arctic sediment clone J1-16 KF896902
ArgVene3 Archaeoglobus veneficus AF482452
AB274309 Archaeoglobus infectus AB274309
UncSu173 Hydrothermal vent chimney clone AB124897
Caq10000 Aarhus Bay sediment clone DSRIV-4 FM179973
JQ304764 Uncultured prokaryote clone A24 JQ304764
UncSu692 Subsurface aquifer clone AY885536
oxidative bacterial-type DsrAB
ClrPha17 Chlorobium clathratiforme BU1 CP001110
Z04Spec2 Magnetococcus marinus CP000471
Ga0371479_1066781416-1066781417 M3
entry732 Freshwater microbial community metagenome
Ga0371479_113560931 Unbinned
Ga0371479_1108884419-1108884420 M41
Ga0371479_1009094616-1009094617 M40
HQ689138 Candidatus Riegeria galatelae isolate C08085 HQ689138
MagMag10 Magnetospirillum magnetotacticum MS-1 AAAP010037
ThlNitr5 Thioalkalivibrio nitratireducens NC_019902
Ga0371479_1106212977-1106212978 M44
Ga0371479_101043172-101043173 Unbinned
Ga0371479_10432741132-10432741133 M37
entry119 Marine metagenome
CanRuth2 Candidatus Ruthia magnifica NC_008610
Ga0371479_1010397116 M39
MooTher5 Moorella thermoacetica Y72 copy 2 BARR01000005
reductive archaeal-type DsrAB
entry547 Microbial hot spring community
PrbIslan Pyrobaculum islandicum U75249
entry706 Hot spring metagenome
0.350

Figure S9. Maximum likelihood phylogenetic tree of DsrAB. The tree was constructed in CLC Genomics Workbench with 1000 bootstraps and WAG substitution model using reference sequences from Muller et al. (2015). Sequence names include the gene IDs and MAG number where applicable. A sequence from eggNOG group COG2221 (*Campylobacter ureolyticus* JFJK01000015_gene476) was used as an outgroup (direction indicated by arrow).

Table S1 (xlsx). Physical and geochemical parameters of Lost Hammer water and sediment.

Located in Supplementary Tables xlsx file.

Table S2. Metagenome and metatranscriptome sequencing statistics.

DNA/RNA	**Sample (NCBI ID)**	**Sequencing platform**	**Read length**	**Raw read pairs**	**Read pairs after quality control**
DNA	SRR13628066	Illumina HiSeq2500	2x126	116844950	116746762
DNA	SRR13628065	Illumina HiSeq2500	2x126	107914225	107647735
RNA	SRR13633097	Illumina NovaSeq6000	2x100	9009952	4233149

Table S3. Metagenomic assembly statistics.

Assembler	Megahit v1.1.3
# contigs	1620755
# contigs >= 1000 bp	276578
# contigs >= 5000 bp	36050
# contigs >= 10000 bp	14841
# contigs >= 25000 bp	3946
# contigs >= 50000 bp	1251
Total length (bp)	1599970111
GC %	51.63
N50	2968
N75	1032
L50	67307
L75	264669
% reads mapping to contigs (average of two samples)	89.7
% reads mapping to contigs > 5000 bp	67.9
Average coverage of metatranscriptome against the metagenome (reads/base)	0.093 ± 59
Average coverage of metatranscriptome against CDS with mapped reads (reads/base)	13 ± 604

Table S4 (xlsx). SAG supplemental information. Located in Supplementary Tables xlsx file.

Table S5 (xlsx). MAG supplemental information. Located in Supplementary Tables xlsx file.

Table S6 (xlsx). Relative expression for KEGG, COG, and pfam IDs. Located in Supplementary Tables xlsx file.

Table S7 (xlsx). Taxonomic classification of genes of interest with mapped transcripts. Located in Supplementary Tables xlsx file.

Table S8 (xlsx). Gene presence and expression in medium-quality MAGs. Located in Supplementary Tables xlsx file.

Table S9 (xlsx). Complete list of genes with mapped transcripts. Located in Supplementary Tables xlsx file.

Table S10. Gibbs free energy of redox pairs present in LH. Values were calculated using the methodology in Jones et al. (2018) (35) (additional details in Supplementary Methods). These values should be considered preliminary estimates as some parameter concentrations are based on single-year measurements.

Redox pair	**△G reaction (kJ/mol electron^{-1})**
H_2/O_2	-111.9
H_2/NO_3^- (to NH_3)	-61.3
N_2/NO_3^- (to NO_2^-)	-68.1
H_2/SO_4^{2-}	-3.0
H_2/CO_2	-9.1
H_2S/O_2	-109.0
H_2S/NO_3^-	-93.8
Fe^{2+}/O_2	-16.0
NH_3/O_2	-52.9
NH_3/NO_3^-	-90.8
NH_3/SO_4^{2-}	-56.5
CH_4/O_2	-205.7
CH_4/NO_3^-	-32.4
CH_4/SO_4^{2-}	6.1

Table S11 (xlsx). Stress response gene comparison to related genomes. Located in Supplementary Tables xlsx file.

Table S12 (xlsx). Gene content of MAGs. Located in Supplementary Tables xlsx file.

Table S13 (xlsx). Gene content of SAGs. Located in Supplementary Tables xlsx file.

Table S14 (xlsx). ANME-1 composite genome gene content. Located in Supplementary Tables xlsx file.

Table S15 (xlsx). *Iainarchaeia* sp. S9 gene content. Located in Supplementary Tables xlsx file.

Table S16 (xlsx). QMZS01 composite genome gene content. Located in Supplementary Tables xlsx file.

Table S17 (xlsx). CG03 sp. S2 gene content. Located in Supplementary Tables xlsx file.

Table S18 (xlsx). Metagenome copy number and total relative expression of genes of interest. Located in Supplementary Tables xlsx file.

References

1. Niederberger TD, Perreault NN, Tille S, Lollar BS, Lacrampe-Couloume G, Andersen D, et al. Microbial characterization of a subzero, hypersaline methane seep in the Canadian High Arctic. ISME J. 2010;4(10):1326-39.

2. Lay CY, Mykytczuk NC, Yergeau E, Lamarche-Gagnon G, Greer CW, Whyte LG. Defining the functional potential and active community members of a sediment microbial community in a high-Arctic hypersaline subzero spring. Appl Environ Microbiol. 2013;79(12):3637-48.

3. Lamarche-Gagnon G, Comery R, Greer CW, Whyte LG. Evidence of *in situ* microbial activity and sulphidogenesis in perennially sub-0 degrees C and hypersaline sediments of a high Arctic permafrost spring. Extremophiles. 2015;19(1):1-15.

4. Battler MM, Osinski GR, Banerjee NR. Mineralogy of saline perennial cold springs on Axel Heiberg Island, Nunavut, Canada and implications for spring deposits on Mars. Icarus. 2013;224(2):364-81.

5. Ward MK, Pollard WH. A hydrohalite spring deposit in the Canadian high Arctic: a potential Mars analogue. Earth Planet Sci Lett. 2018;504:126-38.

6. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.

7. Menzel P, Ng KL, Krogh A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun. 2016;7(1):1-9.

8. Gruber-Vodicka HR, Seah BKB, Pruesse E. phyloFlash: rapid small-subunit rRNA profiling and targeted assembly from metagenomes. mSystems. 2020;5(5):1-16.

9. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015;31(10):1674-6.

10. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 2017;27(5):824-34.

11. Kang DD, Froula J, Egan R, Wang Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ. 2015;3:1-15.

12. 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(7):1043-55.

13. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics. 2018;34(13):i142-i150.

14. Chen IA, Chu K, Palaniappan K, Ratner A, Huang J, Huntemann M, et al. The IMG/M data management and analysis system v.6.0: new tools and advanced capabilities. Nucleic Acids Res. 2020;49(D1):D751–D763.

15. Mukherjee S, Stamatis D, Bertsch J, Ovchinnikova G, Sundaramurthi JC, Lee J, et al. Genomes OnLine Database (GOLD) v.8: overview and updates. Nucleic Acids Res. 2020;49(D1):D723-D733.

16. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics. 2019;36(6):1925–7.

17. Eren AM, Esen OC, Quince C, Vineis JH, Morrison HG, Sogin ML, et al. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ. 2015;3:1-29.

18. Garber AI, Nealson KH, Okamoto A, McAllister SM, Chan CS, Barco RA, et al. FeGenie: a comprehensive tool for the identification of iron genes and iron gene neighborhoods in genome and metagenome assemblies. Front Microbiol. 2020;11:1-23.

19. Fish JA, Chai B, Wang Q, Sun Y, Brown CT, Tiedje JM, et al. FunGene: the functional gene pipeline and repository. Front Microbiol. 2013;4:1-14.

20. Sondergaard D, Pedersen CN, Greening C. HydDB: a web tool for hydrogenase classification and analysis. Sci Rep. 2016;6:1-8.

21. Muller AL, Kjeldsen KU, Rattei T, Pester M, Loy A. Phylogenetic and environmental diversity of DsrAB-type dissimilatory (bi)sulfite reductases. ISME J. 2015;9(5):1152-65.

22. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792-7.

23. Anantharaman K, Hausmann B, Jungbluth SP, Kantor RS, Lavy A, Warren LA, et al. Expanded diversity of microbial groups that shape the dissimilatory sulfur cycle. ISME J. 2018;12(7):1715-28.

24. Stepanauskas R, Fergusson EA, Brown J, Poulton NJ, Tupper B, Labonte JM, et al. Improved genome recovery and integrated cell-size analyses of individual uncultured microbial cells and viral particles. Nat Commun. 2017;8(1):1-10.

25. Kuske CR, Barns SM, Grow CC, Merrill L, Dunbar J. Environmental survey for four pathogenic bacteria and closely related species using phylogenetic and functional genes. J Forensic Sci. 2006;51(3):548-58.

26. Miller CS, Handley KM, Wrighton KC, Frischkorn KR, Thomas BC, Banfield JF. Short-read assembly of full-length 16S amplicons reveals bacterial diversity in subsurface sediments. PLoS One. 2013;8(2):1-11.

27. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41:D590-596.

28. Schmieder R, Edwards R. Fast identification and removal of sequence contamination from genomic and metagenomic datasets. PLoS One. 2011;6(3):1-11.

29. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455-77.

30. Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun. 2018;9(1):1-8.

31. Kopylova E, Noe L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics. 2012;28(24):3211-7.

32. Sheik CS, Reese BK, Twing KI, Sylvan JB, Grim SL, Schrenk MO, et al. Identification and removal of contaminant sequences from ribosomal gene databases: lessons from the Census of Deep Life. Front Microbiol. 2018;9:1-14.

33. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357-9.

34. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166-9.

35. Jones RM, Goordial JM, Orcutt BN. Low energy subsurface environments as extraterrestrial analogs. Front Microbiol. 2018;9:1-18.

36. Alneberg J, Karlsson CMG, Divne AM, Bergin C, Homa F, Lindh MV, et al. Genomes from uncultivated prokaryotes: a comparison of metagenome-assembled and single-amplified genomes. Microbiome. 2018;6(1):1-14.

37. Nigro LM, Hyde AS, MacGregor BJ, Teske A. Phylogeography, salinity adaptations and metabolic potential of the candidate division KB1 bacteria based on a partial single cell genome. Front Microbiol. 2016;7:1-9.