Full title: Geochemical and metagenomic characterization of Jinata Onsen, a Proterozoic-analog hot spring, reveals novel microbial diversity including iron-tolerant phototrophs and thermophilic lithotrophs

Running title: Microbes of Jinata Onsen

Lewis M. Ward[1,2,3]*, Airi Idei[4], Mayuko Nakagawa[2,5], Yuichiro Ueno[2,5,6], Woodward W. Fischer[3], Shawn E. McGlynn[2]*

1. Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138 USA
2. Earth-Life Science Institute, Tokyo Institute of Technology, Meguro, Tokyo, 152-8550, Japan
3. Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125 USA
4. Department of Biological Sciences, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan
5. Department of Earth and Planetary Sciences, Tokyo Institute of Technology, Meguro, Tokyo, 152-8551, Japan
6. Department of Subsurface Geobiological Analysis and Research, Japan Agency for Marine-Earth Science and Technology, Natsushima-cho, Yokosuka 237-0061, Japan

Correspondence: lewis_ward@fas.harvard.edu or mcglynn@elsi.jp

Keywords: geomicrobiology, Proterozoic, microaerobic, astrobiology, ferruginous, thermophiles

Abstract

Hydrothermal systems, including terrestrial hot springs, contain diverse geochemical conditions that vary over short spatial scales due to progressive interaction between the reducing hydrothermal fluids, the oxygenated atmosphere, and in some cases seawater. At Jinata Onsen, on Shikinejima Island, Japan, an intertidal, anoxic, iron-rich hot spring mixes with the oxygenated atmosphere and seawater over short spatial scales, creating a diversity of chemical potentials and redox pairs over a distance ~10 m. We characterized the geochemical conditions along the outflow of Jinata Onsen as well as the microbial communities present in biofilms, mats, and mineral crusts along its traverse via 16S rRNA amplicon and genome-resolved shotgun metagenomic sequencing. The microbial community changed significantly downstream as temperatures and dissolved iron concentrations decreased and dissolved oxygen increased. Near the spring source, visible biomass is limited relative to downstream, and primary productivity may be fueled by oxidation of ferrous iron and molecular hydrogen by members of the Zetaproteobacteria and Aquificae. Downstream, the microbial community is dominated by oxygenic Cyanobacteria. Cyanobacteria are abundant and active even at ferrous iron concentrations of ~150 μM, which challenges the idea that iron toxicity limited cyanobacterial expansion in Precambrian oceans. Several novel lineages of Bacteria are also present at Jinata Onsen, including previously uncharacterized members of the Chloroflexi and Caldithrichaeota phyla, positioning Jinata Onsen as a valuable site for future characterization of these clades.

Introduction

A major theme of environmental microbiology has been the enumeration of microbial groups that are capable of exploiting diverse chemical potentials (i.e. chemical disequilibria) that occur in nature (e.g. 8, 27, 115). Hot springs, with their varied chemical compositions, provide reservoirs of novel microbial diversity, where environmental and geochemical conditions select for lineages and metabolisms distinct from other Earth-surface environments (e.g. 4, 10, 28, 130, 131). In addition to their value as sources of microbial diversity, hot springs also provide valuable test beds for understanding microbial community processes driven by different suites of metabolisms (e.g. 52)—this in turn allows these systems to serve as process analogs and to provide a window into biosphere function during early times in Earth history, for example when the O_2 content of surface waters was low or non-existent. In contrast to most surface ecosystems which are fueled almost entirely by oxygenic photosynthesis by plants, algae, and Cyanobacteria, hot spring microbial communities are commonly supported by lithotrophic or anoxygenic phototrophic organisms that derive energy and electrons for carbon fixation by oxidizing geologically sourced electron donors such as Fe^{2+}, sulfide, arsenite, and molecular hydrogen (e.g. 36, 65, 71, 109, 130). These communities may therefore provide insight into the function of microbial communities on the early Earth or other planets, in which oxygenic photosynthesis may be absent or less significant and anoxygenic photosynthetic or lithotrophic metabolisms may play a larger role, resulting in overall lower rates of primary productivity (e.g. 14, 66, 101, 135, 136, 137).

Here, we present a geomicrobiological characterization of a novel Precambrian Earth process analog site: Jinata Onsen, on Shikinejima Island, Tokyo Prefecture, Japan. While a small number of metagenome-assembled genomes have previously been recovered from Jinata (129, 131), we describe here the first overall characterization of the geochemistry and microbial community of this site. This site supports sharp gradients in geochemistry that in some ways recapitulate spatially environmental transitions which occurred temporally during Proterozoic

time. The modern, sulfate-rich, well-oxygenated ocean that we see today is a relatively recent state, typical only of only the last few hundred million years (e.g. 79). For the first half of Earth history, until ~2.3 billion years ago (Ga), the atmosphere and oceans were anoxic (54), and the oceans were largely rich in dissolved iron but poor in sulfur (124). Following the Great Oxygenation Event ~2.3 Ga, the atmosphere and surface ocean accumulated some oxygen, and the ocean transitioned into a stratified state with oxygenated surface waters and anoxic deeper waters, rich in either dissolved iron or sulfide (92). At Jinata Onsen, this range of geochemical conditions is recapitulated over just a few meters, providing an ideal space-for-time analog to test hypotheses of how microbial diversity and productivity may have varied as environmental conditions changed through Earth history.

At Jinata hot spring, anoxic, iron-rich hydrothermal fluids feed a subaerial spring that flows into a small bay, and mixes with seawater over the course of a few meters. Over its course the waters transition from low-oxygen, iron-rich conditions analogous to some aspects of the early Proterozoic oceans, toward iron-poor and oxygen-rich conditions typical of modern coastal oceans. In upstream regions of the stream where oxygenic Cyanobacteria are at very low abundance, biomass is visibly sparse; however, downstream, biomass accumulates in the form of thick microbial mats containing abundant Cyanobacteria. Visible differences in accumulation and appearance of biomass across the temperature and redox gradient establish the hypothesis that microbial community composition, as well as the magnitude and metabolic drivers of primary productivity, varies along the spring flow. To begin testing this hypothesis and to provide a baseline description of the geochemistry and microbiology of this site in support of future investigation, we performed geochemical measurements, 16S rRNA amplicon sequencing, and genome-resolved metagenomic sequencing to recover draft genomes of diverse novel microbial lineages that inhabit Jinata Onsen.

Materials and Methods:

Geological context and sedimentology of Jinata:

Jinata Onsen is located at 34.318 N, 139.216 E on the island of Shikinejima, Tokyo Prefecture, Japan. Shikinejima is part of the Izu Islands, a chain of volcanic islands that formed in the last few million years along the northern edge of the Izu-Bonin-Mariana Arc (58). Shikinejima is formed of Late Paleopleistocene- to-Holocene non-alkaline felsic volcanics and Late-Miocene to Pleistocene non-alkaline pyroclastic volcanic flows, with Jinata Onsen located on a small bay on the southern side of the island (Figure 1).

Sample collections:

Five sites were sampled at Jinata Onsen: the Source Pool, Pool 1, Pool 2, Pool 3, and the Outflow (Figure 1, Figure 2). During the first sampling trip in January 2016, two whole community DNA samples were collected from each site for 16S rRNA amplicon sequencing. During the second sampling trip, additional DNA was collected from the Source Pool and Pool 2 for shotgun metagenomic sequencing along with gas samples for qualitative analysis. Samples for quantitative gas analysis were collected in October 2017 and April 2018.

Samples were collected as mineral scrapings of loosely attached, fluffy iron oxide coating from surfaces and clasts upstream (Source Pool and Pool 1) and as samples of microbial mat downstream (Pools 2 and 3, and Outflow) using sterile forceps and spatulas (~0.25 cm^3 of material). Immediately after sampling, cells were lysed and DNA preserved with a Zymo Terralyzer BashingBead Matrix and Xpedition Lysis Buffer. Lysis was achieved by attaching tubes to the blade of a cordless reciprocating saw (Black & Decker, Towson, MD) and operating for 1 minute. Aqueous geochemistry samples consisted of water collected with sterile syringes

and filtered through a 0.2 µm filter. Gas samples were collected near sites of ebullition emerging from the bottom of the Source Pool; collection was done into serum vials by water substitution, and then sealed underwater to prevent contamination by air.

Geochemical analysis:

Dissolved oxygen (DO), pH, and temperature measurements were performed *in situ* using an Extech DO700 8-in-1 Portable Dissolved Oxygen Meter (FLIR Commercial Systems, Inc., Nashua, NH). Iron concentrations were measured using the ferrozine assay (114) following acidification with 40 mM sulfamic acid to inhibit iron oxidation by O_2 or oxidized nitrogen species (69). Ammonia/ammonium concentrations were measured using a TetraTest NH_3/NH_4^+ Kit (TetraPond, Blacksburg, VA) following manufacturer's instructions but with colorimetry of samples and NH_4Cl standards quantified with a Thermo Scientific Nanodrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, MA) at 700 nm to improve sensitivity and accuracy. Anion concentrations were measured via ion chromatography on a Shimadzu Ion Chromatograph (Shimadzu Corp., Kyoto, JP) equipped with a Shodex SI-90 4E anion column (Showa Denko, Tokyo, JP).

Presence of H_2 and CH_4 in gas samples was initially qualitatively determined by comparison to standards with a Shimadzu GC-14A gas chromatograph within 12 hours of collection to minimize oxidation of reduced gases. Subsequent gas samples were analyzed following methods from (116). In brief, samples were analyzed using a gas chromatograph (GC-4000, GL Sciences) equipped with both a pulsed discharge detector (PDD) and a thermal conductivity detector (TCD). The GC was equipped with a ShinCarbon ST packed column (2 m × 2.2 mm ID, 50/80 mesh) connected to a HayeSepo Q packed column (2 m × 2.2 mm ID, 60/80 mesh) to separate O_2, N_2, CO_2, and light hydrocarbons. Temperature was held at 40°C for 6 minutes before ramping up to 200°C at 20°C/min. This temperature was held for 6 minutes before ramping up to 250˚C at 50˚C/min before a final hold for 15 minutes. The value of standard errors (SE) were determined by replicate measurement of samples. The detection limit was on the order of 1nmol/cc for H_2 and CH_4.

Water samples for dissolved inorganic carbon (DIC) and dissolved organic carbon (DOC) concentration measurements were collected with sterile syringes and transferred after filtering through a 0.2 µm filter to pre-vacuumed 30 mL serum vials which were sealed with butyl rubber septa and aluminum crimps.

DIC and DOC concentrations in water samples were analyzed by measuring CO_2 in the headspace of the sampled vials after the reaction of sample with either phosphoric acid for DIC or potassium persulfate for DOC with a Shimadzu GC-14A gas chromatograph. Sodium bicarbonate standards and glucose standards were used for making calibration curves to quantify DIC and DOC concentrations, respectively.

16S rRNA and metagenomic sequencing and analysis:

Sequencing and analysis of 16S rRNA amplicon data followed methods from (130). Following return to the lab, bulk environmental DNA was extracted and purified with a Zymo Soil/Fecal DNA extraction kit. The V4-V5 region of the 16S rRNA gene was PCR amplified using archaeal and bacterial primers 515F (GTGCCAGCMGCCGCGGTAA) and 926R (CCGYCAATTYMTTTRAGTTT) (15). DNA was quantified with a Qubit 3.0 fluorimeter (Life Technologies, Carlsbad, CA) following the manufacturer's instructions following DNA extraction and PCR steps. Successful amplification of all samples was verified by viewing on a gel after initial pre-barcoding PCR (30 cycles). Duplicate PCR reactions were pooled and

reconditioned for five cycles with barcoded primers. Samples for sequencing were submitted to Laragen (Culver City, CA) for analysis on an Illumnia MiSeq platform. Sequence data were processed using QIIME version 1.8.0 (15). Raw sequence pairs were joined and quality-trimmed using the default parameters in QIIME. Sequences were clustered into de novo operational taxonomic units (OTUs) with 99% similarity using UCLUST open reference clustering protocol (26). Then, the most abundant sequence was chosen as representative for each *de novo* OTU (125). Taxonomic identification for each representative sequence was assigned using the Silva-132 database (94) clustered separately at 99% and at 97% similarity. Singletons and contaminants (OTUs appearing in the negative control datasets) were removed. 16S rRNA sequences were aligned using MAFFT (62) and a phylogeny constructed using FastTree (93). Alpha diversity was estimated using the Shannon Index (100) and Inverse Simpson metric (1/D) (48, 104). Assessment of sampling depth was estimated using Good's Coverage (37). All statistics were calculated using scripts in QIIME and are reported at the 99% and 97% OTU similarity levels. Multidimensional scaling (MDS) analyses and plots to evaluate the similarity between different samples and environments were produced in R using the vegan and ggplot2 packages (86, 95, 139).

Following initial characterization via 16S rRNA sequencing, four samples were selected for shotgun metagenomic sequencing: JP1-A and JP3-A from the first sampling trip, and JP1L-1 and JP2-1 from the second sampling trip. Purified DNA was submitted to SeqMatic LLC (Fremont, CA) for library preparation and 2x100bp paired-end sequencing via Illumina HiSeq 4000 technology. Samples JP1-A and JP3-A shared a single lane with two samples from another project, while JP1L-1 and JP2-1 shared a lane with one sample from another project.

Raw sequence reads from all four samples were co-assembled with MegaHit v. 1.02 (77) and genome bins constructed based on nucleotide composition and differential coverage using MetaBAT (59), MaxBin (140), and CONCOCT (2) prior to dereplication and refinement with DAS Tool (103) to produce the final bin set. Genome bins were assessed for completeness, contamination, and strain-level heterogeneity using CheckM (89), tRNA sequences found with Aragorn (73), and presence of metabolic pathways of interest predicted with MetaPOAP (134). Coverage was extracted using bbmap (11) and samtools (76). Genes of interest (e.g. coding for ribosomal, photosynthesis, iron oxidation, and electron transport proteins) were identified from assembled metagenomic data locally with BLAST+ (12 and were screened against outlier (e.g. likely contaminant) contigs as determined by CheckM using tetranucleotide, GC, and coding density content. Translated protein sequences of genes of interest were aligned with MUSCLE (25), and alignments manually curated in Jalview (138). Phylogenetic trees were calculated using RAxML (110) on the Cipres science gateway (83). Node support for phylogenies was calculated with transfer bootstraps by BOOSTER (74). Trees were visualized with the Interactive Tree of Life viewer (Letunic and Bork 2016). Because sequencing depth of each sample in the full metagenome was uneven, relative abundance of genes of interest between metagenomic datasets was normalized to the coverage of *rpoB* genes in each raw dataset as mapped onto the coassembly. Like the 16S rRNA gene, *rpoB* is a highly conserved, vertically-inherited gene useful for taxonomic identification of organisms but has the added advantage that it is only known to occur as a single copy per genome (17) and is more readily assembled in metagenomic datasets (e.g. 131). Presence and classification of hydrogenase genes was determined with HydDB (107). Taxonomic assignment of MAGs was made based on placement in a reference phylogeny built with concatenated ribosomal protein sequences following Hug et al. (50) and

confirmed using GTDB-Tk (90). Optimal growth temperatures of MAGs was predicted based on proteome-wide 2-mer amino acid composition following methods from (78).

Results

Site description

The source water of Jinata Onsen emerges with low dissolved oxygen concentrations near our limit of detection, is iron-rich, and gently bubbles gas from the spring source (Table 1, Figure 1, Figure 2). Temperatures at the source are ~63°C. Water emerges into the Source Pool, which has no visible microbial mats or biofilms (Figure 2D). Surfaces are instead coated with a fluffy red precipitate, likely a poorly ordered or short range-ordered ferric iron oxide phase such as ferrihydrite. Flow from the source is—at least in part—tidally charged, with the highest water levels and flow rates occurring at high tide. At low tide, flow rates drop and the water level of the Source Pool can drop by decimeters and portions of the Source Pool can drain during spring low tides. Downstream, the spring water collects into a series of pools (Pool 1-3) (Figure 2C,E-F), which cool sequentially (Figure 3, Supplemental Table 1). Pool 1 contains iron oxides like the Source Pool, but also develops macroscopic microbial streamers that are coated in iron oxides and thin veil-like layers of microorganisms overlaying iron oxide sediments—structures similar to those typically made by marine iron-oxidizing Zetaproteobacteria (e.g. 35). Streamers are very fine (mm-scale) and delicate (break apart on contact with forceps) but can reach several centimeters in length. Cyanobacteria in Pool 2 and Pool 3 display high levels of photosynthetic activity as revealed by high dissolved oxygen concentration (~234 µM), low dissolved inorganic carbon concentrations, and the accumulation of visible O_2 bubbles on the surface and within the fabric of the mat. Downstream pools (Pools 2 and 3) mix with seawater during high tide due to wave action, but this seawater influence does not appear to influence the Source Pool or Pool 1. Samples were collected and temperatures were measured at high tide, reflecting the lowest temperatures experienced by microbes in the pools—at low tide, hot spring input is dominant and temperatures rise (observed range at each site in Supplemental Table 1). Subaqueous surfaces in Pools 2 and 3 are covered in thick microbial mats. In Pool 2, the mat is coated in a layer of fluffy iron oxide similar to that in the Source Pool, with dense microbial mat below (Figure 2E). Pool 3 contains only patchy iron oxides, with mostly exposed microbial mats displaying a finger-like morphology. These “fingers” were 0.5-1 cm in diameter and up to ~5 cm long and were closely packed and carpeting surfaces of Pool 3 below the high tide line, potentially related to turbulent mixing from wave action during high tide (Figure 2F). The Outflow is the outlet of a channel connecting Pool 2 to the bay. Its hydrology is dominantly marine with small admixtures of inflowing spring water (Figure 2G).

Jinata hot spring was visited twice for observation and community DNA sampling in 2016 (January and September), and again for observation and gas sampling in October 2017 and April 2018. These visits corresponded to a range of tidal conditions, including a spring low and high tide in September 2016. General features of the spring were consistent across this period (including abundance and distribution of iron minerals and microbial mats), differing primarily in an apparent tidal dependence in flow rate and water level of the spring and the amount of seawater influence on Pool 3. These differences in flow and mixing led to variation in water temperatures of 3-10 °C (Supplemental Table 1). At high tide, the flow rate of the spring increases, as does seawater influx to Pool 3. During the spring low tide, the spring flow stagnated and the water level of Source Pool and Pool 1 dropped by decimeters, with some portions draining entirely. During less extreme low tides observed on other dates, the spring flow was low

but nonzero and the water level of the Source Pool did not drop significantly. While there is substantial variability in the flow rate from the spring based on tides (and resulting shifts in water level and temperature), the overall geochemistry of the source water and the microbial community appeared largely similar between expeditions.

Geochemistry

Geochemical measurements along the flow path of Jinata Onsen revealed a major shift from hot, low-oxygen, high-iron source water to cooler, more oxygen-rich water with less dissolved iron downstream. Geochemistry measurements of Jinata source water are summarized in Table 1, while geochemical gradients along the stream outflow are summarized in Figure 3 and Supplemental Table 1. Source waters were slightly enriched in chloride relative to seawater (~23.2 g/L in Jinata source water versus ~19.4 g/L in typical seawater), depleted in sulfate (~1.6 g/L in Jinata versus ~2.7 g/L in seawater) but approached seawater concentrations downstream as mixing increased. Water emerging from the source was 63°C, very low in dissolved oxygen (~4.7 μM), at pH 5.4, and contained substantial concentrations of dissolved iron (~250 μM Fe^{2+}). Dissolved organic carbon (DOC) in the source water was high (~1.31 mM). It is unknown whether this is produced *in situ* or whether the source water emerges with high DOC. Both DOC and DIC decrease along the outflow of the spring (Supplemental Table 1). After emerging from the source, the spring water exchanges gases with the air due to mixing associated with water flow and gas ebullition, and DO rose to 39 μM at the surface of the Source Pool. As water flows downstream from the Source Pool, it cools slightly, exchanges gases with the atmosphere, and intermittently mixes with seawater below Pool 1.

Both H_2 and CH_4 were qualitatively detected in bubbles from the Source Pool following initial sampling in September 2016. However, during subsequent analyses to quantify the gas composition in October 2017 and April 2018 the gas was determined to contain CO_2, CH_4, N_2 (Supplemental Table 2). This subsequent non-detection of H_2 may be related to temporal variability in the gas composition at Jinata (e.g. following tidal influence; significant variability was observed in the CO_2:N_2 ratio between two sampling dates, Supplemental Table 2) or may reflect oxidation of H_2 between sampling and analysis. The detection limit of H_2 for these later measurements was ~1 nmol/cc (in the gas phase of our quantitative gas analyses, or ~1 nM in the aqueous phase, (3)), well above the energetic and ecological limits for hydrogenotrophic metabolisms (e.g. 53) leaving open the possibility of biologically significant H_2 fluxes at Jinata around the time of sampling. The oxidation of H_2 coupled to O_2 reduction is a thermodynamically favorable process even at very low substrate concentrations (e.g. $\Delta_rG' < -375$ kJ/mol with substrate concentrations of 0.1 nM $H_{2(aq)}$ and 0.1 μM $O_{2(aq)}$, well below our limit of detection) (32)). Consistent with this thermodynamic favorability, biology has been shown to make use of this metabolism in environments such as hot springs with H_2 concentrations near our detection limits (21) and in Antarctic soils where microbes rely on uptake of trace atmospheric H_2 at concentrations around 190 ppbv (53). Therefore the trace amounts of H_2 which may be present in the source water at Jinata may be sufficient to support lithoautotrophy near the hot spring source in organisms possessing the genetic capacity for hydrogen oxidation as discussed below. Improved quantification of H_2 concentrations and measurement of hydrogenase activity and the productivity of hydrogenotrophic microbes will be needed in future to determine the relative contribution of hydrogen oxidation to productivity at Jinata.

16S rRNA and genome-resolved metagenomic sequencing

16S rRNA and metagenomic sequencing of microbial communities at Jinata Onsen revealed a highly diverse community. In total, 16S rRNA amplicon sequencing recovered 456,737 sequences from the 10 samples at Jinata (Supplemental Tables 3-5). Reads per sample following filtering for quality and removal of chimeras ranged from 2,076 for Pool 3 Sample B to 96,268 for Pool 1 Sample A (median 32,222, mean 35,479, and standard deviation 26,014). On average 65% of the microbial community was recovered from Jinata samples at the 99% OTU level based on the Good's Coverage statistic of the 16S rRNA gene (ranging from 50% coverage in the Outflow Sample A to 80% in the Pool 1 Sample A) and 82% at the 97% OTU level (69% for Pool 2 Sample B to 93% for the Pool 1 Sample B). MDS analysis (Supplemental Figure 1) demonstrates that samples from the same site are highly similar, and adjacent sites (e.g. Source Pool and Pool 1, Outflow and Pool 3) also show a high degree of similarity. However, there is a substantial transition in microbial community diversity between the most distant samples (e.g. Source Pool and Outflow).

Shotgun metagenomic sequencing of four samples from Jinata Onsen recovered 121 GB of data, forming a 1.48 Gb coassembly consisting of 1,531,443 contigs with an N50 of 1,494 bp. Nucleotide composition and differential coverage-based binning of the coassembly via multiple methods followed by dereplication and refinement resulted in a final set of 161 medium- or high-quality metagenome-assembled genomes (MAGs) following current standards (i.e. completeness >50% and contamination <10%) (7). These MAGs are from diverse phyla of Bacteria and Archaea (Figure 4); metagenome and MAG statistics with tentative taxonomic assignments for recovered MAGs are available in Supplementary Table 6, while MAGs of particular interest due to their potential contribution to primary productivity at this site or due to substantial genetic or metabolic novelty are discussed in depth below and shown in phylogenetic trees alongside reference strains in Figures 5-7.

Discussion

As Jinata spring water flows from source to ocean, it transitions from hot, low-oxygen, high-iron water to cooler, iron-depleted, oxygen-rich water in downstream regions (Figure 3). Following this geochemical transition is a major shift in the composition of the microbial community, from a high-temperature, putatively lithotrophic community which produces little visible biomass upstream, to a lower temperature, community with well-developed, thick microbial mats downstream. This shift in community composition is summarized in Figure 3, with complete diversity data in the Supplemental Information (including OTU counts per samples in Supplemental Table 4 and relative abundance binned at the class level in Supplemental Table 5). Below, we discuss the overall physiological and taxonomic trends across the spring sites as inferred from diversity and genomic analysis.

Potential for iron and hydrogen oxidation

The hot spring water emerging at the Source Pool at Jinata contains abundant dissolved Fe^{2+} and trace H_2 (though measurements of gas content varied, as discussed above) (Table 1). Although rates of carbon fixation were not measured, the appearance of zetaproteobacterial veils and streamers and molecular evidence for lithoautotrophic microbes suggests that these electron donors may fuel productivity and determine the microbial community upstream at the Source Pool and Pool 1, where microbial mats are not well developed. The low accumulation of visible biomass in upstream regions of Jinata are similar to other microbial ecosystems fueled by iron oxidation (e.g. Oku-Okuhachikurou Onsen, 130, Fuschna Spring, 44, and Jackson Creek, 96), in which lithotrophic communities appear capable of accumulating less organic carbon than

communities fueled by oxygenic photosynthesis (including those in downstream regions at Jinata).

Results of 16S rRNA sequencing indicate that the most abundant organisms in the Source Pool are members of the Aquificae family Hydrogenothermaceae (32% of reads in the Source Pool and 11.5% of reads in Pool 1). Members of this family of marine thermophilic lithotrophs are capable of iron and hydrogen oxidation, as well as heterotrophy (118) and may be utilizing Fe^{2+}, H_2, or dissolved organic carbon at Jinata. The seventh most abundant OTU in the Source Pool samples is a novel sequence 89% similar to a strain of *Persephonella* found in an alkaline hot spring in Papua New Guinea. *Persephonella* is a genus of thermophilic, microaerophilic hydrogen oxidizing bacteria within the Hydrogenothermaceae (38). Despite their abundance as assessed by 16S rRNA sequencing (Figure 3), only four partial Aquificae MAGs were recovered from Jinata of which only one (J026) was reasonably complete (~94%). Two Aquificae MAGs recovered Group 1 NiFe hydrogenase genes, which may be used in hydrogenotrophy; the absence of hydrogenases from the other MAGs may be related to their low completeness, or could reflect a utilization of iron or other electron donors and not H_2 in these organisms.

The other most abundant organisms near the source are members of the Zetaproteobacteria—a group typified by the neutrophilic, aerobic iron-oxidizing genus *Mariprofundus* common in marine systems (29). Zetaproteobacteria accounted for 24% of 16S rRNA sequences in the Source Pool and 26.5% in Pool 1. All Zetaproteobacteria characterized to date are obligate iron- and/or hydrogen-oxidizing lithoautotrophs (85), suggesting that these organisms may play a substantial role in driving carbon fixation in the Source Pool and Pool 1.

Members of the Mariprofundaceae have been observed to have an upper temperature limit for growth of 30 °C (30), while Zetaproteobacteria are found at Jinata at temperatures up to 63 °C. This currently represents a unique high-temperature environment for these organisms. In particular, the third most abundant OTU in the Source Pool and Pool 1 sample A is an unknown sequence that is 92% identical to a sequence from an uncultured zetaproteobacterium from a shallow hydrothermal vent in Papua New Guinea (82). This sequence likely marks a novel lineage of high-temperature iron-oxidizing Zetaproteobacteria.

The relative abundance of Hydrogenothermaceae drops off to less than 1% of sequences where microbial mats become well developed downstream of Pool 1, but Zetaproteobacteria continue to make up ~1-4% percent of reads in Pool 2 and Pool 3 where dissolved iron concentrations are still significant (Figure 3). It may be that the relative abundance change is due more to the increase in abundance of other organisms, rather than a drop in the number of Zetaproteobacteria or their ability to make a living oxidizing iron. This hypothesis awaits confirmation by a technique such as qPCR. In contrast, the absence of Hydrogenothermaceae downstream may be a real signal driven by the rapid disappearance of trace H_2 as an electron donor. However, in both cases, a drop in relative abundance is likely related to the increasing total biomass (i.e. number of cells) downstream as Cyanobacteria become more productive, leading to sequences from Hydrogenothermaceae and Zetaproteobacteria being diluted out by increased numbers of Cyanobacteria, Chloroflexi, and other sequences.

Four MAGs affiliated with the Zetaproteobacteria were recovered from Jinata with completeness estimates by CheckM ranging from ~80 to ~97% (J005, J009, J030, and J098). While these MAGs did not recover 16S rRNA genes, RpoB- and concatenated ribosomal protein-based phylogenies illustrated that members of this group at Jinata Onsen do not belong to the characterized genera *Mariprofundus* or *Ghiorsea*, but instead form separate basal lineages within the Zetaproteobacteria (Figure 5). Despite their phylogenetic distinctness, these MAGs

largely recovered genes associated with aerobic iron oxidation as expected given the physiology of other Zetaproteobacteria. These include a terminal O_2 reductase from the C-family of heme copper oxidoreductases for respiration at low O_2 concentrations and Cyc2 cytochrome genes implicated in ferrous iron oxidation in Zetaproteobacteria and other taxa (e.g. Chlorobi) (41, 42, 61). Hydrogenase catalytic subunit genes (neither [NiFe] nor [FeFe]) were not recovered in zetaproteobacterial MAGs even at high completeness, suggesting that these organisms are not hydrogenotrophic. Consistent with the obligately autotrophic lifestyle of previously characterized Zetaproteobacteria, J009 and J098 encode carbon fixation via the Calvin cycle. However, J005 and J030 which did not recover genes for carbon fixation via the Calvin cycle (such as the large and small subunits of rubisco, phosphoribulose kinase, or carboxysome proteins). The high completeness of these MAGs (~94-97%) makes it unlikely that these genes would all fail to be recovered (MetaPOAP False Negative estimates 10^{-5}-10^{-7}). The absence of carbon fixation pathways from these genomes together with the availability of abundant dissolved organic carbon in Pool 1 (~1.3 mM) suggest that these organisms may be heterotrophic, a lifestyle not previously observed for members of the Zetaproteobacteria.

Seven MAGs were recovered from the enigmatic bacterial phylum Calditrichaeota (J004, J008, J042, J070, J075, J140, and J141) (Figure 6). While few members of Calditrichaeota have been isolated or sequenced, the best known of these is *Caldithrix abyssi* (84); this taxon was characterized as an anaerobic thermophile capable of lithoheterotrophic H_2 oxidation coupled to denitrification and organoheterotrophic fermentation (1, 81). The Calditrichaeota MAGs reported here are up to 97% complete (J004) and contain members with variable putative metabolic capabilities, potentially including aerobic hydrogen- or iron-oxidizing lithoautotrophy. In the Calditrichaeota MAGs recovered from Jinata Onsen, aerobic respiration via A-family heme copper oxidoreductases could potentially be coupled to autotrophic hydrogen oxidation (via the Group 1d NiFe hydrogenase in J042) or iron oxidation (via the *pioA* gene in J075); however, *Caldithrix abyssi* appears incapable of aerobic respiration despite encoding an A-family heme copper oxidoreductase (70). A MAG from a member of Calditrichaeota has previously been recovered from Chocolate Pots hot spring in Yellowstone National Park (34); together with the data presented here this suggests that this phylum may be a common member of microbial communities in iron-rich hot springs. Unlike previously described Calditrichaeota which are all heterotrophic (81), most of the Calditrichaeota MAGs reported here possess a putative capacity for carbon fixation via the Calvin cycle. J004 is closely related to *Caldithrix abyssi*, while the other MAGs form two distinct but related clades (Figure 6).

Oxygenic photosynthesis

Cyanobacteria are nearly absent from near the Source Pool, but are observed in low numbers in Pool 1 and become abundant starting in Pool 2. The most abundant Cyanobacteria present are predominantly members of the Nostocales. This group includes *Leptolyngbya* and *Phormidium*, genera of filamentous non–heterocystous Cyanobacteria that are present in other hot springs of similar temperatures (e.g. 6, 98, 130). Diverse cyanobacterial MAGs were recovered, including members of the orders Pleurocapsales (J083), Chroococcales (J003 and J149), and Oscillatoriales (J007, J055, and J069). In the Outflow samples, chloroplast sequences are abundant, most closely related to the diatom *Melosira*.

Cyanobacteria are sometimes underrepresented in 16S rRNA amplicon sequencing datasets as a result of poor DNA yield or amplification biases (e.g. 88, 122), but the low

abundance of Cyanobacteria near the Source Pool was confirmed by fluorescent microscopy, in which cells displaying cyanobacterial autofluorescence were observed abundantly in samples from the downstream samples but not in the Source Pool (Supplemental Figure 2). Thick microbial mats first appear in Pool 2 when Cyanobacteria become abundant, suggesting that oxygenic photosynthesis fuels more net carbon fixation than lithotrophy in these environments. Previously, it has been suggested that high ferrous iron concentrations are toxic to Cyanobacteria, and that this would have greatly reduced their productivity under ferruginous ocean conditions such as those that may have persisted through much of the Archean era (117). The abundant Cyanobacteria observed to be active at Jinata under high iron concentrations suggest that Cyanobacteria can adapt to ferruginous conditions, and therefore iron toxicity might not inhibit Cyanobacteria over geological timescales. Indeed, the soluble iron concentrations observed at Jinata are higher (150-250 μM) than predicted for the Archean oceans (<120 μM, 49) or observed at other iron-rich hot springs (~100-200 μM, 91, 130), making Jinata an excellent test case for determining the ability of Cyanobacteria to adapt to high iron concentrations. Culture-based physiological experiments may be useful to determine whether Jinata Cyanobacteria utilize similar strategies to other iron-tolerant strains (e.g. by those in Chocolate Pots Hot Spring, 91, or the ferric iron-tolerant *Leptolyngbya*-relative *Marsacia ferruginose*, 9) or whether Jinata strains possess unique adaptations that allow them to grow at higher iron concentrations than known for other environmental Cyanobacteria strains. This will in turn provide insight into whether iron tolerance is due to evolutionarily conserved strategies or whether this is a trait that has evolved convergently multiple times.

Diverse novel Chloroflexi from Jinata Onsen

In addition to the primary phototrophic and lithotrophic carbon fixers at Jinata, 16S rRNA and metagenomic data sets revealed diverse novel lineages within the Chloroflexi phylum. A total of 23 Chloroflexi MAGs were recovered, introducing substantial genetic and metabolic diversity that expands our understanding of this group. While the best known members of this phylum are Type 2 Reaction Center-containing lineages such as *Chloroflexus* and *Roseiflexus* within the class Chloroflexia (e.g. 121), phototrophy is not a synapomorphy of the Chloroflexi phylum or even the Chloroflexia class (e.g. 126) and most of the diversity of the phylum belongs to several other classes made up primarily of nonphototrophic lineages (131). The bulk of Chloroflexi diversity recovered from Jinata belongs to "subphlyum I", a broad group of predominantly nonphototrophic lineages that was originally described based on the class- or order-level lineages Anaerolineae and Caldilineae (141), but also encompasses the related groups Ardenticatenia, Thermoflexia, and *Candidatus* Thermofonsia (22, 63, 131).

16S rRNA analysis indicates that members of Anaerolineae and *Candidatus* Thermofonsia (annotated by Silva and GTDB-Tk as the order SBR1031) are fairly abundant at Jinata, with Anaerolineae at ~3% relative abundance in the Source and Pool 1 and *Ca.* Thermofonsia at ~3.5% relative abundance in Pool 2 and Pool 3. Three MAGs recovered from Jinata (J082, J097, and J130) are associated with the Anaerolineae class as determined by RpoB and concatenated ribosomal protein phylogenies, along with seven associated with *Ca.* Thermofonsia (J027, J033, J036, J038, J039, J064, and J076). Particularly notable among these MAGs is J036, a close relative of the phototrophic *Ca.* Roseilinea gracile (68, 119, 120). J036 contains a 16S rRNA gene that is 96% similar to that of *Ca*. Roseilinea gracile, and two-way AAI estimates (97) showed 73.6% similarity between the two strains, indicating these strains are probably best classified as distinct species within the same genus. Unlike other phototrophs in the Chloroflexi phylum that are capable of photoautotrophy via the 3-hydroxypropionate bicycle

or the Calvin Cycle (67, 102), J036 and *Ca.* Roseilinea gracile do not encode carbon fixation and are likely photoheterotrophic. Previous analyses suggested that the Roseilinea lineage belongs to the Anaerolineae (68) or Thermofonsia (131); however, our updated phylogeny presented here places J036 and Roseilinea in a separate lineage along with J033 and J162, diverging just outside of the Anaerolineae+Thermofonsia clade, suggesting that these strains may instead be yet another class- or order-level lineage within the broader "subphylum I" of Chloroflexi (Figure 7), an interpretation supported by analysis via GTDB-Tk which places these genomes outside of characterized clades (Supplemental Table 6).

The Chloroflexi class Ardenticatenia was first described from an isolate from an iron-rich Japanese hydrothermal field (63) and has since been recovered from sulfidic hot springs as well (132). Members of Ardenticatenia were present at up to 1.2% relative abundance in Pool 3 in 16S amplicon data. A MAG closely related to *Ardenticatena maritima* was recovered from Jinata Onsen, J129. While *Ardenticatena maritima 110S* contains a complete denitrification pathway (47), MAG J129 did not recover any denitrification genes. This could be related to the relatively low completeness of this MAG (~70%), but False Negative estimates by MetaPOAP (134) indicates that the probability that all four steps in the canonical denitrication pathway would fail to be recovered in J129 given their presence in the source genome is less than 0.8%, suggesting that most if not all denitrification genes are absent and that the capacity for denitrification is not universal within members of *Ardenticatena*. This would be consistent with broad trends in the apparently frequent modular horizontal gene transfer of partial denitrification pathways between disparate microbial lineages to drive rapid adaption and metabolic flexibility of aerobic organisms in microoxic and anoxic environments, for reasons that are still not well established (19, 113).

Members of the Chloroflexi class Caldilineae were present at up to 0.5% abundance at Jinata in the 16S rRNA dataset. Members of the Caldilineae have previously been isolated from intertidal hot springs in Iceland (57) and Japanese hot springs (99). Characterized organisms in this class are filamentous, anaerobic, or facultatively aerobic heterotrophs (39, 57, 99); and therefore these taxa may play a role in degrading biomass within low-oxygen regions of microbial mats at Jinata. Three MAGs were recovered that form a deeply branching lineage with the Caldilineae class (J095, J111, and J123), sister to the previously characterized genera *Caldilinea* and *Litorilinea*. Like other members of the Caldilineae, these strains encode aerobic respiration via A-family heme copper oxidoreductases and both a *bc* complex III and an alternative complex III, and are therefore likely at least facultatively aerobic. J095 also encodes carbon fixation via the Calvin cycle as well as a Group 1f NiFe hydrogenase, suggesting a potential capability for lithoautotrophy by hydrogen oxidation, expanding the known metabolic diversity of this class and the Chloroflexi phylum as a whole.

MAG J114 branches at the base of subphylum I of the Chloroflexi, potentially the first member of a novel class-level lineage. The divergence between Anaerolineae and Caldilineae has been estimated to have occurred on the order of 1.7 billion years ago (102). The phylogenetic placement of J114 suggests that it diverged from other members of subphylum I even earlier, and it may be a good target for future investigation to assess aspects of the early evolution of the Chloroflexi phylum. J114 encodes aerobic respiration via an A-family heme copper oxidoreductase and an alternative complex III like many other nonphototrophic Chloroflexi lineages (e.g. 126, 131) as well as a Group 1f NiFe hydrogenase and carbon fixation via the

Calvin Cycle, suggesting the capacity for aerobic hydrogen-oxidizing autotrophy—a lifestyle not previously described for members of the Chloroflexi.

Conclusions

To our knowledge, this is the first overall geomicrobiological characterization of Jinata Onsen, providing baseline descriptions of geochemistry and microbial diversity in order to establish a series of testable hypotheses which can be addressed by future studies. We have also provided genome-resolved metagenomics sequencing of this site focusing on members of the microbial community predicted to be responsible for the bulk of primary productivity in this system along with other organisms belonging to novel or under-characterized lineages. However, this is just a subset of the diverse microbial populations at Jinata Onsen; many more MAGs from across the tree of life were recovered than are discussed in detail here but which may be of use to others (Figure 4, Supplemental Table 6).

The diversity of iron oxidizing bacteria at Jinata is different than in other Fe^{2+} -rich springs and environments. For example, in freshwater systems such as Oku-Okuhachikurou Onsen in Akita Prefecture, Japan (130), and Budo Pond in Hiroshima, Japan (60), iron oxidation is driven primarily by the activity of chemoautotrophs such as members of the Gallionellaceae. In contrast, at Chocolate Pots hot spring in Yellowstone National Park, USA, iron oxidation is primarily abiotic, driven by O_2 produced by Cyanobacteria, with only a small contribution from iron oxidizing bacteria (34, 123). The distinct iron-oxidizing community at Jinata Onsen may be related to the salinity of the spring water, or biogeographically by access to the ocean, as Zetaproteobacteria are typically found in marine settings, particularly in deep ocean basins associated with hydrothermal iron sources (30). Despite the taxonomically distinct iron oxidizer communities between Jinata and Oku-Okuhachikurou Onsen, both communities support only limited visible biomass in regions dominated by iron oxidizers (130), perhaps reflecting the shared biochemical and bioenergetic challenges of iron oxidation incurred by diverse iron oxidizing bacteria including Gallionellaceae and Zetaproteobacteria (5, 30, 130). Future work focused on isolation and physiological characterization of microbes, quantification of rates and determination of microbial drivers of carbon fixation and aerobic and anaerobic heterotrophy, and carbon isotope profiling of organic and inorganic species along the flow path of the hot spring will be necessary to fully characterize the activity of microbes at Jinata and to fully compare this system to other areas with high dissolved ferrous iron concentrations (e.g. Oku-Okuhachikurou Onsen, 130, Fuschna Spring, 44, Jackson Creek, 96, and Chocolate Pots Hot Spring, 34, 123).

The relatively high concentrations of dissolved organic carbon (DOC) measured in Pool 1 (~1.3 mM) may stimulate heterotrophic activity by the microbial community at Jinata, coupled to aerobic or anaerobic respiration (such as dissimilatory iron reduction, as observed in other iron-rich hot springs, e.g. 33), resulting in the drawdown of DOC downstream. The source of this DOC is unclear; future work will be necessary to determine whether DOC is present in the source water or if it is produced *in situ* by the microbial community in the Source Pool and Pool 1. Future work is also needed to evaluate the potential for dissimilatory iron reduction and other anaerobic metabolisms at this site.

Throughout Earth history, the metabolic opportunities available to life, and the resulting organisms and metabolisms responsible for driving primary productivity, have been shaped by the geochemical conditions of the atmosphere and oceans. The modern, sulfate-rich, well-oxygenated oceans we see today reflect a relatively recent state—one typical of only the last few

hundred million years (e.g. 79). For the first half of Earth history, until ~2.3 billion years ago (Ga), the atmosphere and oceans were anoxic (54), and the oceans were largely rich in dissolved iron but poor in sulfur (124). At this time, productivity was low and fueled by metabolisms such as methanogenesis and anoxygenic photosynthesis (14, 66, 135). Following the expansion of oxygenic photosynthesis by Cyanobacteria and higher primary productivity around the Great Oxygenation Event ~2.3 Ga (20, 31, 128, 137), the atmosphere and surface ocean accumulated some oxygen, and the ocean transitioned into a state with oxygenated surface waters but often anoxic deeper waters, rich in either dissolved iron or sulfide (13, 55, 56, 92). At Jinata Onsen, this range of geochemical conditions is recapitulated over just a few meters, providing a useful test case for probing the shifts of microbial productivity over the course of Earth history. In particular, the concomitant increase in visible biomass at Jinata as the community shifts from lithotrophy toward water-oxidizing phototrophy (i.e. oxygenic photosynthesis) is consistent with estimates for greatly increased primary production following the evolution and expansion of Cyanobacteria around the GOE (20, 101, 106, 128, 135, 137).

The dynamic abundances of redox-active compounds including oxygen, iron, and hydrogen at Jinata may not only be analogous to conditions on the early Earth, but may have relevance for potentially habitable environments on Mars as well. Early Mars is thought to have supported environments with metabolic opportunities provided by the redox gradient between the oxidizing atmosphere and abundant electron donors such as ferrous iron and molecular hydrogen sourced from water/rock interactions (e.g. 51), and production of these substrates may continue today (24, 111), potentially supporting past or present life in the Martian subsurface (112). Understanding the potential productivity of microbial communities fueled by lithotrophic metabolisms is critical for setting expectations of the presence and size of potential biospheres on other worlds and early in Earth history (e.g. 135, 136, 137). Uncovering the range of microbial metabolisms present under the environmental conditions at Jinata, and their relative contributions to primary productivity, may therefore find application to predicting environments on Mars most able to support productive microbial communities.

Data availability:

Raw 16S rRNA, raw metagenomic sequence data, and MAGs have been uploaded and made publicly available on NCBI under Project Number PRJNA392119 (genome accession numbers can be found in Supplemental Table 6).

Acknoweldgements:

LMW acknowledges support from NASA NESSF (#NNX16AP39H), NSF (#OISE 1639454), NSF GROW (#DGE 1144469), Lewis and Clark Fund for Exploration and Field Research in Astrobiology, the ELSI Origins Network, and an Agouron Institute postdoctoral fellowship. WWF acknowledges the support of NASA Exobiology award #NNX16AJ57G, the David and Lucile Packard Foundation, and the Simons Collaboration on the Origins of Life. MN and YU acknowledge the support of the Japan Society for the Promotion of Science (JP17K14412, JP17H06105)

References

1. Alauzet, C. and Jumas-Bilak, E., 2014. The phylum Deferribacteres and the genus Caldithrix. In *The prokaryotes* (pp. 595-611). Springer, Berlin, Heidelberg.
2. Alneberg, J., Bjarnason, B.S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., Loman, N.J., Andersson, A.F. and Quince, C., 2013. CONCOCT: clustering contigs on coverage and composition. *arXiv preprint arXiv:1312.4038*.

3. Amend, J.P. and Shock, E.L., 2001. Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and Bacteria. *FEMS Microbiol. Rev.*, *25*(2), pp.175-243.
4. Beam, J.P., Jay, Z.J., Schmid, M.C., Rusch, D.B., Romine, M.F., de M Jennings, R., Kozubal, M.A., Tringe, S.G., Wagner, M. and Inskeep, W.P., 2016. Ecophysiology of an uncultivated lineage of Aigarchaeota from an oxic, hot spring filamentous 'streamer'community. *ISME J.*, *10*(1), p.210.
5. Bird, L.J., Bonnefoy, V. and Newman, D.K., 2011. Bioenergetic challenges of microbial iron metabolisms. *Trends Microbiol.*, *19*(7), pp.330-340.
6. Bosak, T., Liang, B., Wu, T.D., Templer, S.P., Evans, A., Vali, H., Guerquin‐Kern, J.L., Klepac‐Ceraj, V., Sim, M.S. and Mui, J., 2012. Cyanobacterial diversity and activity in modern conical microbialites. *Geobiology*, *10*(5), pp.384-401.
7. Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., et al., 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.*, *35*(8), p.725.
8. Broda, E., 1977. Two kinds of lithotrophs missing in nature. *Z. Allg. Mikrobiol.*, *17*(6), pp.491-493.
9. Brown, I.I., Bryant, D.A., Casamatta, D., et al., 2010. Polyphasic characterization of a thermotolerant siderophilic filamentous cyanobacterium that produces intracellular iron deposits. *Appl. Environ. Microbiol.*, *76*(19), pp.6664-6672.
10. Bryant, D.A., Costas, A.M.G., Maresca, J.A., et al., 2007. Candidatus Chloracidobacterium thermophilum: an aerobic phototrophic acidobacterium. *Science*, *317*(5837), pp.523-526.
11. Bushnell, B. "BBMap short read aligner." *University of California, Berkeley, California. URL http://sourceforge. net/projects/bbmap* (2016).
12. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. and Madden, T.L., 2009. BLAST+: architecture and applications. *BMC Bioinf.*, *10*(1), p.421.
13. Canfield, D.E., 1998. A new model for Proterozoic ocean chemistry. *Nature*, *396*(6710), p.450.
14. Canfield, D.E., Rosing, M.T. and Bjerrum, C., 2006. Early anaerobic metabolisms. *Philos. Trans. R. Soc., B*, *361*(1474), pp.1819-1836.
15. Caporaso, J.G., Kuczynski, J., Stombaugh, J., et al., 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods*, *7*(5), p.335.
16. Caporaso, J.G., Lauber, C.L., Walters, W.A., et al., 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J*, *6*(8), p.1621.
17. Case, R.J., Boucher, Y., Dahllöf, I., Holmström, C., Doolittle, W.F. and Kjelleberg, S., 2007. Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Appl. Environ. Microbiol.*, *73*(1), pp.278-288.
18. Chang, Y.J., Land, M., Hauser, L., et al., 2011. Non-contiguous finished genome sequence and contextual data of the filamentous soil bacterium Ktedonobacter racemifer type strain (SOSP1-21 T). *Stand. Genomic Sci.*, *5*(1), p.97.
19. Chen, J. and Strous, M., 2013. Denitrification and aerobic respiration, hybrid electron transport chains and co-evolution. *Biochim. Biophys. Acta, Bioenerg.*, *1827*(2), pp.136-144
20. Crockford, P.W., Hayles, J.A., Bao, H., Planavsky, N.J., Bekker, A., Fralick, P.W., Halverson, G.P., Bui, T.H., Peng, Y. and Wing, B.A., 2018. Triple oxygen isotope evidence for limited mid-Proterozoic primary productivity. *Nature*, *559*(7715), p.613.

683 21. D'Imperio, S., Lehr, C.R., Oduro, H., Druschel, G., Kühl, M. and McDermott, T.R., 2008.
684 Relative importance of H2 and H2S as energy sources for primary production in geothermal
685 springs. *Appl. Environ. Microbiol.*, *74*(18), pp.5802-5808.
686 22. Dodsworth, J.A., Gevorkian, J., Despujos, F., Cole, J.K., Murugapiran, S.K., Ming, H., Li,
687 W.J., Zhang, G., Dohnalkova, A. and Hedlund, B.P., 2014. Thermoflexus hugenholtzii gen.
688 nov., sp. nov., a thermophilic, microaerophilic, filamentous bacterium representing a novel
689 class in the Chloroflexi, Thermoflexia classis nov., and description of Thermoflexaceae fam.
690 nov. and Thermoflexales ord. nov. *Int. J. Syst. Evol. Microbiol.*, *64*(6), pp.2119-2127.
691 23. Dunfield, P.F., Tamas, I., Lee, K.C., Morgan, X.C., McDonald, I.R. and Stott, M.B., 2012.
692 Electing a candidate: a speculative history of the bacterial phylum OP10. *Environ.*
693 *Microbiol.*, *14*(12), pp.3069-3080.
694 24. Dzaugis, M., Spivack, A.J. and D'Hondt, S., 2018. Radiolytic H2 Production in Martian
695 Environments. *Astrobiology*, *18*(9), pp.1137-1146.
696 25. Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high
697 throughput. *Nucleic Acids Res.*, *32*(5), pp.1792-1797.
698 26. Edgar, R.C., 2010. Search and clustering orders of magnitude faster than
699 BLAST. *Bioinformatics*, *26*(19), pp.2460-2461.
700 27. Ehrenreich, A. and Widdel, F., 1994. Anaerobic oxidation of ferrous iron by purple bacteria,
701 a new type of phototrophic metabolism. *Appl. Environ. Microbiol.*, *60*(12), pp.4517-4526.
702 28. Eloe-Fadrosh, E.A., Paez-Espino, D., Jarett, J., et al., 2016. Global metagenomic survey
703 reveals a new bacterial candidate phylum in geothermal springs. *Nat. Commun.*, *7*, p.10476.
704 29. Emerson, D., Rentz, J.A., Lilburn, T.G., Davis, R.E., Aldrich, H., Chan, C. and Moyer, C.L.,
705 2007. A novel lineage of proteobacteria involved in formation of marine Fe-oxidizing
706 microbial mat communities. *PloS One*, *2*(8), p.e667.
707 30. Emerson, D., Fleming, E.J. and McBeth, J.M., 2010. Iron-oxidizing bacteria: an
708 environmental and genomic perspective. *Annu. Rev. Microbiol.*, *64*, pp.561-583.
709 31. Fischer, W.W., Hemp, J. and Johnson, J.E., 2016. Evolution of oxygenic
710 photosynthesis. *Annu. Rev. Earth Planet. Sci.*, *44*, pp.647-683.
711 32. Flamholz, A., Noor, E., Bar-Even, A. and Milo, R., 2011. eQuilibrator—the biochemical
712 thermodynamics calculator. *Nucleic Acids Res.*, *40*(D1), pp.D770-D775.
713 33. Fortney, N.W., He, S., Converse, B.J., Beard, B.L., Johnson, C.M., Boyd, E.S. and Roden,
714 E.E., 2016. Microbial Fe (III) oxide reduction potential in Chocolate Pots hot spring,
715 Yellowstone National P ark. *Geobiology*, *14*(3), pp.255-275.
716 34. Fortney, N.W., He, S., Converse, B.J., Boyd, E.S. and Roden, E.E., 2018. Investigating the
717 Composition and Metabolic Potential of Microbial Communities in Chocolate Pots Hot
718 Springs. *Front. Microbiol.*, *9*.
719 35. Fullerton, H., Hager, K.W., McAllister, S.M. and Moyer, C.L., 2017. Hidden diversity
720 revealed by genome-resolved metagenomics of iron-oxidizing microbial mats from Lō’ ihi
721 Seamount, Hawai’i. *ISME J*, *11*(8), p.1900.
722 36. Giovannoni, S.J., Revsbech, N.P., Ward, D.M. and Castenholz, R.W., 1987. Obligately
723 phototrophic Chloroflexus: primary production in anaerobic hot spring microbial mats. *Arch.*
724 *Microbiol.*, *147*(1), pp.80-87.
725 37. Good, I.J., 1953. The population frequencies of species and the estimation of population
726 parameters. *Biometrika*, *40*(3-4), pp.237-264.
727 38. Götz, D., Banta, A., Beveridge, T.J., Rushdi, A.I., Simoneit, B.R.T. and Reysenbach, A.L.,
728 2002. Persephonella marina gen. nov., sp. nov. and Persephonella guaymasensis sp. nov., two

novel, thermophilic, hydrogen-oxidizing microaerophiles from deep-sea hydrothermal vents. *Int. J. Syst. Evol. Microbiol.*, *52*(4), pp.1349-1359.

39. Grégoire, P., Bohli, M., Cayol, J.L., et al., 2011. Caldilinea tarbellica sp. nov., a filamentous, thermophilic, anaerobic bacterium isolated from a deep hot aquifer in the Aquitaine Basin. *Int. J. Syst. Evol. Microbiol.*, *61*(6), pp.1436-1441.
40. Grouzdev, D.S., Rysina, M.S., Bryantseva, I.A., Gorlenko, V.M. and Gaisin, V.A., 2018. Draft genome sequences of 'Candidatus Chloroploca asiatica'and 'Candidatus Viridilinea mediisalina', candidate representatives of the Chloroflexales order: phylogenetic and taxonomic implications. *Stand. Genomic Sci.*, *13*(1), p.24.
41. Han, H., Hemp, J., Pace, L.A., Ouyang, H., Ganesan, K., Roh, J.H., Daldal, F., Blanke, S.R. and Gennis, R.B., 2011. Adaptation of aerobic respiration to low O2 environments. *Proc. Natl. Acad. Sci.*, *108*(34), pp.14109-14114.
42. He, S., Barco, R.A., Emerson, D. and Roden, E.E., 2017. Comparative genomic analysis of neutrophilic iron (II) oxidizer genomes for candidate genes in extracellular electron transfer. *Front. Microbiol.*, *8*, p.1584.
43. Hedlund, B.P., Murugapiran, S.K., Huntemann, M., et al., 2015. High-quality draft genome sequence of Kallotenue papyrolyticum JKG1T reveals broad heterotrophic capacity focused on carbohydrate and amino acid metabolism. *Genome Announcements*, *3*(6), pp.e01410-15.
44. Hegler, F., Lösekann-Behrens, T., Hanselmann, K., Behrens, S. and Kappler, A., 2012. Influence of seasonal and geochemical changes on iron geomicrobiology of an iron-carbonate mineral water spring. *Appl. Environ. Microbiol.*, pp.AEM-01440.
45. Hemp, J., Ward, L.M., Pace, L.A. and Fischer, W.W., 2015. Draft genome sequence of Levilinea saccharolytica KIBI-1, a member of the Chloroflexi class Anaerolineae. *Genome Announcements*, *3*(6), pp.e01357-15.
46. Hemp, J., Ward, L.M., Pace, L.A. and Fischer, W.W., 2015. Draft genome sequence of Ornatilinea apprima P3M-1, an anaerobic member of the Chloroflexi class Anaerolineae. *Genome Announcements*, *3*(6), pp.e01353-15.
47. Hemp, J., Ward, L.M., Pace, L.A. and Fischer, W.W., 2015. Draft genome sequence of Ardenticatena maritima 110S, a thermophilic nitrate-and iron-reducing member of the Chloroflexi class Ardenticatenia. *Genome Announcements*, *3*(6), pp.e01347-15.
48. Hill, M.O., 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology*, *54*(2), pp.427-432.
49. Holland, H.D., 1984. *The chemical evolution of the atmosphere and oceans*. Princeton University Press.
50. Hug, L.A., Baker, B.J., Anantharaman, K., et al., 2016. A new view of the tree of life. *Nat. Microbiol.*, *1*(5), p.16048.
51. Hurowitz, J.A., Fischer, W.W., Tosca, N.J. and Milliken, R.E., 2010. Origin of acidic surface waters and the evolution of atmospheric chemistry on early Mars. *Nat. Geosci.*, *3*(5), p.323.
52. Inskeep, W.P., Ackerman, G.G., Taylor, W.P., Kozubal, M., Korf, S. and Macur, R.E., 2005. On the energetics of chemolithotrophy in nonequilibrium systems: case studies of geothermal springs in Yellowstone National Park. *Geobiology*, *3*(4), pp.297-317.
53. Ji, M., Greening, C., Vanwonterghem, I., et al., 2017. Atmospheric trace gases support primary production in Antarctic desert surface soil. *Nature*, *552*(7685), p.400.
54. Johnson, J.E., Gerpheide, A., Lamb, M.P. and Fischer, W.W., 2014. O2 constraints from Paleoproterozoic detrital pyrite and uraninite. *Geol. Soc. Am. Bull.*, *126*(5-6), pp.813-830.

55. Johnston, D.T., Wolfe-Simon, F., Pearson, A. and Knoll, A.H., 2009. Anoxygenic photosynthesis modulated Proterozoic oxygen and sustained Earth's middle age. *Proc. Natl. Acad. Sci.*, *106*(40), pp.16925-16929.
56. Johnston, D.T., Poulton, S.W., Dehler, C., Porter, S., Husson, J., Canfield, D.E. and Knoll, A.H., 2010. An emerging picture of Neoproterozoic ocean chemistry: Insights from the Chuar Group, Grand Canyon, USA. *Earth Planet. Sci. Lett.*, *290*(1-2), pp.64-73.
57. Kale, V., Björnsdóttir, S.H., Friðjónsson, Ó.H., Pétursdóttir, S.K., Ómarsdóttir, S. and Hreggviðsson, G.Ó., 2013. Litorilinea aerophila gen. nov., sp. nov., an aerobic member of the class Caldilineae, phylum Chloroflexi, isolated from an intertidal hot spring. *Int. J. Syst. Evol. Microbiol.*, *63*(3), pp.1149-1154.
58. Kaneoka, I., Isshiki, N. and Zashu, S., 1970. K-Ar ages of the Izu-Bonin islands. *Geochem. J.*, *4*(2), pp.53-60.
59. Kang, D.D., Froula, J., Egan, R. and Wang, Z., 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ*, *3*, p.e165.
60. Kato, S., Kikuchi, S., Kashiwabara, T., Takahashi, Y., Suzuki, K., Itoh, T., Ohkuma, M. and Yamagishi, A., 2012.Prokaryotic abundance and community composition in a freshwater iron1-rich microbial mat at circumneutral pH. *Geomicrobiol. J.*, *29*(10), pp.896-905.
61. Kato, S., Ohkuma, M., Powell, D.H., Krepski, S.T., Oshima, K., Hattori, M., Shapiro, N., Woyke, T. and Chan, C.S., 2015. Comparative genomic insights into ecophysiology of neutrophilic, microaerophilic iron oxidizing bacteria. *Front. Microbiol.*, *6*, p.1265.
62. Katoh, K., Misawa, K., Kuma, K.I. and Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.*, *30*(14), pp.3059-3066.
63. Kawaichi, S., Ito, N., Kamikawa, R., Sugawara, T., Yoshida, T. and Sako, Y., 2013. Ardenticatena maritima gen. nov., sp. nov., a ferric iron-and nitrate-reducing bacterium of the phylum ‘Chloroflexi’isolated from an iron-rich coastal hydrothermal field, and description of Ardenticatenia classis nov. *Int. J. Syst. Evol. Microbiol.*, *63*(8), pp.2992-3002.
64. Kawaichi, S., Yoshida, T., Sako, Y. and Nakamura, R., 2015. Draft genome sequence of a heterotrophic facultative anaerobic thermophilic bacterium, Ardenticatena maritima strain 110ST. *Genome Announcements*, *3*(5), pp.e01145-15.
65. Kawasumi, T., Igarashi, Y., Kodama, T. and Minoda, Y., 1980. Isolation of strictly thermophilic and obligately autotrophic hydrogen bacteria. *Agric. Biol. Chem.*, *44*(8), pp.1985-1986.
66. Kharecha, P., Kasting, J. and Siefert, J., 2005. A coupled atmosphere–ecosystem model of the early Archean Earth. *Geobiology*, *3*(2), pp.53-76.
67. Klatt, C.G., Bryant, D.A. and Ward, D.M., 2007. Comparative genomics provides evidence for the 3‐hydroxypropionate autotrophic pathway in filamentous anoxygenic phototrophic bacteria and in hot spring microbial mats. *Environ. Microbiol.*, *9*(8), pp.2067-2078.
68. Klatt, C.G., Wood, J.M., Rusch, D.B., et al., 2011. Community ecology of hot spring cyanobacterial mats: predominant populations and their functional potential. *ISME J.*, *5*(8), p.1262.
69. Klueglein, N. and Kappler, A., 2013. Abiotic oxidation of Fe (II) by reactive nitrogen species in cultures of the nitrate‐reducing Fe (II) oxidizer A cidovorax sp. BoFeN1–questioning the existence of enzymatic Fe (II) oxidation. *Geobiology*, *11*(2), pp.180-190.

70. Kublanov, I.V., Sigalova, O.M., Gavrilov, S.N., et al. 2017. Genomic analysis of Caldithrix abyssi, the thermophilic anaerobic bacterium of the novel bacterial phylum Calditrichaeota. *Front. Microbiol.*, *8*, p.195.
71. Kulp, T.R., Hoeft, S.E., Asao, M., Madigan, M.T., Hollibaugh, J.T., Fisher, J.C., Stolz, J.F., Culbertson, C.W., Miller, L.G. and Oremland, R.S., 2008. Arsenic (III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. *Science*, *321*(5891), pp.967-970.
72. Kuznetsov, B.B., Ivanovsky, R.N., Keppen, O.I., et al., 2011. Draft genome sequence of the anoxygenic filamentous phototrophic bacterium Oscillochloris trichoides subsp. DG-6. *J. Bacteriol.*, *193*(1), pp.321-322.
73. Laslett, D. and Canback, B., 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.*, *32*(1), pp.11-16.
74. Lemoine, F., Entfellner, J.B.D., Wilkinson, E., Correia, D., Felipe, M.D., Oliveira, T. and Gascuel, O., 2018. Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature*, *556*(7702), p.452.
75. Letunic, I. and Bork, P., 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.*, *44*(W1), pp.W242-W245.
76. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R., 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), pp.2078-2079.
77. Li, D., Liu, C.M., Luo, R., Sadakane, K. and Lam, T.W., 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, *31*(10), pp.1674-1676.
78. Li, G., Rabe, K.S., Nielsen, J. and Engqvist, M.K., 2019. Machine learning applied to predicting microorganism growth temperatures and enzyme catalytic optima. *BioRxiv*, p.522342.
79. Lyons, T.W., Reinhard, C.T. and Planavsky, N.J., 2014. The rise of oxygen in Earth's early ocean and atmosphere. *Nature*, *506*(7488), p.307.
80. Makita, H., Tanaka, E., Mitsunobu, S., Miyazaki, M., Nunoura, T., Uematsu, K., Takaki, Y., Nishi, S., Shimamura, S. and Takai, K., 2017. Mariprofundus micogutta sp. nov., a novel iron-oxidizing zetaproteobacterium isolated from a deep-sea hydrothermal field at the Bayonnaise knoll of the Izu-Ogasawara arc, and a description of Mariprofundales ord. nov. and Zetaproteobacteria classis nov. *Arch. Microbiol.*, *199*(2), pp.335-346.
81. Marshall, I.P., Starnawski, P., Cupit, C., Fernández Cáceres, E., Ettema, T.J., Schramm, A. and Kjeldsen, K.U., 2017. The novel bacterial phylum Calditrichaeota is diverse, widespread and abundant in marine sediments and has the capacity to degrade detrital proteins. *Environ. Microbiol. Rep.*9(4), pp.397-403.
82. Meyer-Dombard, D., Amend, J.P. and Osburn, M.R., 2013. Microbial diversity and potential for arsenic and iron biogeochemical cycling at an arsenic rich, shallow-sea hydrothermal vent (Tutum Bay, Papua New Guinea). *Chem. Geol*, *348*, pp.37-47.
83. Miller, M.A., Pfeiffer, W. and Schwartz, T., 2010, November. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Gateway Computing Environments Workshop (GCE), 2010* (pp. 1-8).
84. Miroshnichenko, M.L., Kostrikina, N.A., Chernyh, N.A., Pimenov, N.V., Tourova, T.P., Antipov, A.N., Spring, S., Stackebrandt, E. and Bonch-Osmolovskaya, E.A., 2003. Caldithrix

abyssi gen. nov., sp. nov., a nitrate-reducing, thermophilic, anaerobic bacterium isolated from a Mid-Atlantic Ridge hydrothermal vent, represents a novel bacterial lineage. *Int. J. Syst. Evol. Microbiol.*, *53*(1), pp.323-329.

85. Mori, J.F., Scott, J.J., Hager, K.W., Moyer, C.L., Küsel, K. and Emerson, D., 2017. Physiological and ecological implications of an iron-or hydrogen-oxidizing member of the Zetaproteobacteria, Ghiorsea bivora, gen. nov., sp. nov. *ISME J*, *11*(11), p.2624.
86. Oksanen, J., Blanchet, F.G., Friendly, M., et al., 2016. Vegan: community ecology package. R package version 2.3-5. R Foundation, Vienna, Austria.
87. Pace, L.A., Hemp, J., Ward, L.M. and Fischer, W.W., 2015. Draft genome of Thermanaerothrix daxensis GNS-1, a thermophilic facultative anaerobe from the Chloroflexi class Anaerolineae. *Genome Announcements*, *3*(6), pp.e01354-15.
88. Parada, A.E., Needham, D.M. and Fuhrman, J.A., 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.*, *18*(5), pp.1403-1414.
89. Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P. and Tyson, G.W., 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.*, pp.gr-186072.
90. Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.A. and Hugenholtz, P., 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.*
91. Pierson, B.K., Parenteau, M.N. and Griffin, B.M., 1999. Phototrophs in high-iron-concentration microbial mats: physiological ecology of phototrophs in an iron-depositing hot spring. *Appl. Environ. Microbiol.*, *65*(12), pp.5474-5483.
92. Poulton, S.W., Fralick, P.W. and Canfield, D.E., 2010. Spatial variability in oceanic redox structure 1.8 billion years ago. *Nat. Geosci.*, *3*(7), p.486.
93. Price, M.N., Dehal, P.S. and Arkin, A.P., 2010. FastTree 2–approximately maximum-likelihood trees for large alignments. *PloS One*, *5*(3), p.e9490.
94. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.*, *41*(D1), pp.D590-D596.
95. R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
96. Roden, E., McBeth, J.M., Blothe, M., Percak-Dennett, E.M., Fleming, E.J., Holyoke, R.R., Luther III, G.W. and Emerson, D., 2012. The microbial ferrous wheel in a neutral pH groundwater seep. *Front. Microbiol.*, *3*, p.172.
97. Rodriguez-R, L.M. and Konstantinidis, K.T., 2014. Bypassing cultivation to identify bacterial species. *Microbe*, *9*(3), pp.111-118.
98. Roeselers, G., Norris, T.B., Castenholz, R.W., Rysgaard, S., Glud, R.N., Kühl, M. and Muyzer, G., 2007. Diversity of phototrophic bacteria in microbial mats from Arctic hot springs (Greenland). *Environ. Microbiol.*, *9*(1), pp.26-38.
99. Sekiguchi, Y., Yamada, T., Hanada, S., Ohashi, A., Harada, H. and Kamagata, Y., 2003. Anaerolinea thermophila gen. nov., sp. nov. and Caldilinea aerophila gen. nov., sp. nov., novel filamentous thermophiles that represent a previously uncultured lineage of the domain Bacteria at the subphylum level. *Int. J. Syst. Evol. Microbiol.*, *53*(6), pp.1843-1851.
100. Shannon, C., 1948. A Mathematical Theory of Communication-The Bell System Technical Journal, vol. 27, pag. 379-423, 623-656.

911 101. Shih, P.M., Hemp, J., Ward, L.M., Matzke, N.J. and Fischer, W.W., 2017. Crown group
912 Oxyphotobacteria postdate the rise of oxygen. *Geobiology*, *15*(1), pp.19-29.
913 102. Shih, P.M., Ward, L.M. and Fischer, W.W., 2017. Evolution of the 3-hydroxypropionate
914 bicycle and recent transfer of anoxygenic photosynthesis into the Chloroflexi. *Proc. Natl.*
915 *Acad. Sci.*, *114*(40), pp.10749-10754.
916 103. Sieber, C.M., Probst, A.J., Sharrar, A., Thomas, B.C., Hess, M., Tringe, S.G. and
917 Banfield, J.F., 2018. Recovery of genomes from metagenomes via a dereplication,
918 aggregation and scoring strategy. *Nat. Microbiol.*, p.1.
919 104. Simpson, E.H. 1949. Measurement of diversity: *Nature*, v. 163, p. 688.
920 105. Singer, E., Emerson, D., Webb, E.A., et al., 2011. Mariprofundus ferrooxydans PV-1 the
921 first genome of a marine Fe (II) oxidizing Zetaproteobacterium. *PloS One*, *6*(9), p.e25386.
922 106. Sleep, N.H. and Bird, D.K., 2007. Niches of the pre‐photosynthetic biosphere and
923 geologic preservation of Earth's earliest ecology. *Geobiology*, *5*(2), pp.101-117.
924 107. Søndergaard, D., Pedersen, C.N. and Greening, C., 2016. HydDB: a web tool for
925 hydrogenase classification and analysis. *Sci. Rep.*, *6*, p.34212.
926 108. Sorokin, D.Y., Lücker, S., Vejmelkova, D., et al., 2012. Nitrification expanded:
927 discovery, physiology and genomics of a nitrite-oxidizing bacterium from the phylum
928 Chloroflexi. *ISME J*, *6*(12), p.2245.
929 109. Spear, J.R., Walker, J.J., McCollom, T.M. and Pace, N.R., 2005. Hydrogen and
930 bioenergetics in the Yellowstone geothermal ecosystem. *Proc. Natl. Acad. Sci.*, *102*(7),
931 pp.2555-2560.
932 110. Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-
933 analysis of large phylogenies. *Bioinformatics*, *30*(9), pp.1312-1313.
934 111. Stamenković, V., Ward, L.M., Mischna, M. and Fischer, W.W., 2018. O 2 solubility in
935 Martian near-surface environments and implications for aerobic life. *Nat. Geosci.*, p.1.
936 112. Stamenković, V., Beegle, L.W., Zacny, K., et al., 2019. The next frontier for planetary
937 and human exploration. *Nature Astronomy*, p.1.
938 113. Stein, L.Y. and Klotz, M.G., 2016. The nitrogen cycle. *Curr. Biol.*, *26*(3), pp.R94-R98.
939 114. Stookey, L.L., 1970. Ferrozine---a new spectrophotometric reagent for iron. *Anal.*
940 *Chem.*, *42*(7), pp.779-781.
941 115. Strous, M., Fuerst, J.A., Kramer, E.H., Logemann, S., Muyzer, G., van de Pas-Schoonen,
942 K.T., Webb, R., Kuenen, J.G. and Jetten, M.S., 1999. Missing lithotroph identified as new
943 planctomycete. *Nature*, *400*(6743), p.446.
944 116. Suda, K., Gilbert, A., Yamada, K., Yoshida, N. and Ueno, Y., 2017. Compound–and
945 position–specific carbon isotopic signatures of abiogenic hydrocarbons from on–land
946 serpentinite–hosted Hakuba Happo hot spring in Japan. *Geochim. Cosmochim. Acta*, *206*,
947 pp.201-215.
948 117. Swanner, E.D., Mloszewska, A.M., Cirpka, O.A., Schoenberg, R., Konhauser, K.O. and
949 Kappler, A., 2015. Modulation of oxygen production in Archaean oceans by episodes of Fe
950 (II) toxicity. *Nat. Geosci.*, *8*(2), pp.126-130.
951 118. Takai, K. and Nakagawa, S., 2014. The Family Hydrogenothermaceae. In *The*
952 *Prokaryotes* (pp. 689-699). Springer, Berlin, Heidelberg.
953 119. Tank, M., Thiel, V., Ward, D.M. and Bryant, D.A., 2017. A panoply of phototrophs: an
954 overview of the thermophilic chlorophototrophs of the microbial mats of alkaline siliceous
955 hot springs in Yellowstone National Park, WY, USA. In *Modern topics in the phototrophic*
956 *prokaryotes* (pp. 87-137). Springer, Cham.

957 120. Thiel, V., Hügler, M., Ward, D.M. and Bryant, D.A., 2017. The dark side of the
958 mushroom spring microbial mat: life in the shadow of chlorophototrophs. II. Metabolic
959 functions of abundant community members predicted from metagenomic analyses. *Front.*
960 *Microbiol.*, *8*, p.943.
961 121. Thiel, V., Tank, M. and Bryant, D.A., 2018. Diversity of chlorophototrophic bacteria
962 revealed in the omics era. *Annu. Rev. Plant Biol.*, *69*, pp.21-49.
963 122. Trembath-Reichert, E., Ward, L.M., Slotznick, S.P., Bachtel, S.L., Kerans, C.,
964 Grotzinger, J.P. and Fischer, W.W., 2016. Gene Sequencing-Based Analysis of Microbial-
965 Mat Morphotypes, Caicos Platform, British West Indies. *J. Sediment. Res.*, *86*(6),
966 pp.629-636.
967 123. Trouwborst, R.E., Johnston, A., Koch, G., Luther III, G.W. and Pierson, B.K., 2007.
968 Biogeochemistry of Fe (II) oxidation in a photosynthetic microbial mat: implications for
969 Precambrian Fe (II) oxidation. *Geochim. Cosmochim. Acta*, *71*(19), pp.4629-4643.
970 124. Walker, J.C. and Brimblecombe, P., 1985. Iron and sulfur in the pre-biologic
971 ocean. *Precambrian Res.*, *28*(3-4), pp.205-222.
972 125. Wang, Q., Garrity, G.M., Tiedje, J.M. and Cole, J.R., 2007. Naive Bayesian classifier for
973 rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ.*
974 *Microbiol.*, *73*(16), pp.5261-5267.
975 126. Ward, L.M., Hemp, J., Pace, L.A. and Fischer, W.W., 2015. Draft genome sequence of
976 Herpetosiphon geysericola GC-42, a nonphototrophic member of the Chloroflexi class
977 Chloroflexia. *Genome Announcements*, *3*(6), pp.e01352-15.
978 127. Ward, L.M., Hemp, J., Pace, L.A. and Fischer, W.W., 2015. Draft genome sequence of
979 Leptolinea tardivitalis YMTK-2, a mesophilic anaerobe from the Chloroflexi class
980 Anaerolineae. *Genome Announcements*, *3*(6), pp.e01356-15.
981 128. Ward, L.M., Kirschvink, J.L. and Fischer, W.W., 2016. Timescales of oxygenation
982 following the evolution of oxygenic photosynthesis. *Origins Life Evol. Biospheres*, *46*(1),
983 pp.51-65.
984 129. Ward, L.M., McGlynn, S.E. and Fischer, W.W., 2017. Draft Genome Sequences of a
985 Novel Lineage of Armatimonadetes Recovered from Japanese Hot Springs. *Genome*
986 *Announcements*, *5*(40), pp.e00820-17.
987 130. Ward, L.M., Idei, A., Terajima, S., Kakegawa, T., Fischer, W.W. and McGlynn, S.E.,
988 2017. Microbial diversity and iron oxidation at Okuoku‐hachikurou Onsen, a Japanese hot
989 spring analog of Precambrian iron formations. *Geobiology*, *15*(6), pp.817-835.
990 131. Ward, L.M., Hemp, J., Shih, P.M., McGlynn, S.E. and Fischer, W.W., 2018. Evolution of
991 phototrophy in the Chloroflexi phylum driven by horizontal gene transfer. *Front.*
992 *Microbiol.*, *9*, p.260.
993 132. Ward, L.M., McGlynn, S.E. and Fischer, W.W., 2018. Draft Genome Sequence of a
994 Divergent Anaerobic Member of the Chloroflexi Class Ardenticatenia from a Sulfidic Hot
995 Spring. *Genome Announcements*, *6*(25), pp.e00571-18.
996 133. Ward, L.M., McGlynn, S.E. and Fischer, W.W., 2018. Draft Genome Sequences of Two
997 Basal Members of the Anaerolineae Class of Chloroflexi from a Sulfidic Hot
998 Spring. *Genome Announcements*, *6*(25), pp.e00570-18.
999 134. Ward, L.M., Shih, P.M.,.and Fischer, W.W, 2018. MetaPOAP: Presence or Absence of
1000 Metabolic Pathways in Metagenome-Assembled Genomes. *Bioinformatics*, bty510.

135. Ward, L.M., Rasmussen, B. and Fischer, W.W., 2019. Primary productivity was limited by electron donors prior to the advent of oxygenic photosynthesis. *J. Geophys. Res.: Biogeosci.*, *124*(2), pp.211-226.

136. Ward, LM, V Stamenković, K Hand, and WW Fischer. 2019. Follow the Oxygen: Comparative Histories of Planetary Oxygenation and Opportunities for Aerobic Life. *Astrobiology*, in press.

137. Ward, L.M. and Shih, P.M., 2019. The evolution and productivity of carbon fixation pathways in response to changes in oxygen concentration over geological time. *Free Radical Biol. Med.*

138. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. and Barton, G.J., 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), pp.1189-1191.

139. Wickham, H., 2010. ggplot2: elegant graphics for data analysis. *J. Stat. Softw.*, *35*(1), pp.65-88.

140. Wu, Y.W., Tang, Y.H., Tringe, S.G., Simmons, B.A. and Singer, S.W., 2014. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome*, *2*(1), p.26.

141. Yamada, T. and Sekiguchi, Y., 2009. Cultivation of uncultured chloroflexi subphyla: significance and ecophysiology of formerly uncultured Chloroflexi 'subphylum i' with natural and biotechnological relevance. *Microbes Environ.*, *24*(3), pp.205-216.

Table 1: Geochemistry of source water at Jinata Onsen.

T	63°C
pH	5.4
DO	4.7 μM
Fe^{2+}	261 μM
NH_3/NH_4^+	87 μM
Cl^-	654 mM
SO_4^{--}	17 mM
NO_3^-	<1.6 μM
NO_2^-	<2.2 μM
HPO_4^-	<1 μM

Figure 1:
Location of Jinata Onsen on Shikinejima Island, Japan, and inset overview sketch of field site with sampling localities marked.

Figure 2:
Representative photos of Jinata. A) Panorama of field site, with Source Pool on the left (Pool 1 below), Pool 2 and 3 in the center, and Out Flow to the bay on the right. B) Undistorted view north up the canyon. C) Undistorted view south toward the bay, overlooking Pool 2. D) Source Pool, coated in floc-y iron oxides and bubbling with gas mixture containing CO_2, CH_4 and trace, potentially variable, H_2. E) Pool 2, with mixture of red iron oxides and green from Cyanobacteria-rich microbial mats. F) Close up of textured microbial mats in Pool 3. G) Close

up of Out Flow, where hot spring water mixes with ocean water. Reprinted with permission from (131).

Figure 3:
Summary of geochemical and microbiological trends along the flow path of Jinata Onsen. Top: panoramic view of Jinata Onsen, with Source Pool at left and flow of spring water toward the bay at right, with sampling locations indicated. Middle: geochemical transect across the spring, showing temperature (°C, left axis) and dissolved Fe(II) and O_2 (μM, right axis). Bottom: stacked bar chart of relative community abundance of dominant microbial phyla as determined by 16S rRNA amplicon sequencing. Sequence data were binned at the phylum level and duplicate samples at each site were averaged. Reads that could not be assigned to a phylum were discarded; all phyla that do not make up more than 2% of the community at any one site have been collapsed to "Other". Near the source, the community is predominantly made up of iron- and/or hydrogen-oxidizing organisms in the Proteobacteria and Aquificae phyla. As the hot spring water flows downstream, it equilibrates with the atmosphere and eventually mixes with seawater, resulting in downstream cooling, accumulation of oxygen, and loss of dissolved iron due to biological and abiotic processes. Oxygenic Cyanobacteria become progressively more abundant downstream Hydrogen- and iron-oxidizing lithotrophs dominate near the source, but phototrophic Cyanobacteria come to dominate downstream. Additional community diversity is found in Supplemental Table 4.

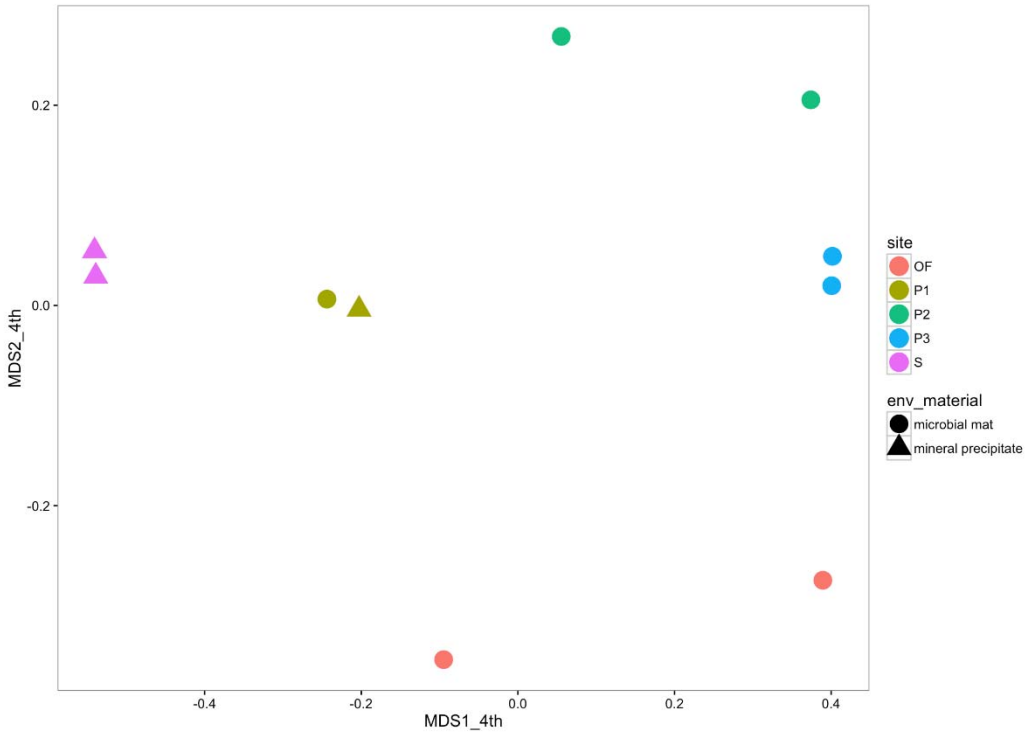
Figure 4:
Phylogeny of Bacteria and Archaea based on concatenated ribosomal proteins. Numbers in parentheses next to phylum labels refer to number of MAGs recovered from Jinata Onsen. Labels for phyla with two or fewer MAGs recovered from Jinata omitted for clarity. The reference alignment was modified from Hug et al. (50). Full list of MAGs recovered available in Supplemental Table 6.

Figure 5: Phylogeny of the Zetaproteobacteria, rooted with Alphaproteobacteria, built with concatenated ribosomal protein sequences. Data from (80), (85), (105), and other draft genomes available on Genbank. All nodes recovered TBE support values greater than 0.7. In cases where reference genomes have a unique strain name or identifier, this is included; otherwise Genbank WGS genome prefixes are used.

Figure 6: Phylogeny of the Calditrichaeota, rooted with Bacteroidetes, built with concatenated ribosomal protein sequences. Data from (70) and other draft genomes available on genomes have a unique strain name or identifier, this is included; otherwise Genbank WGS genome prefixes are used.

Figure 7: Detailed phylogeny of the Chloroflexi phylum, with class-level clades highlighted in gray, built with concatenated ribosomal protein sequences. The large basal class Dehalococcoidia, which was not observed in 16S rRNA or metagenome data from Jinata, is omitted for clarity. The phylogeny contains MAGs reported here, members of the Chloroflexi phylum previously described (18, 22, 40, 43, 45, 46, 47, 64, 72, 87, 108, 126, 127, 131, 132, 133), and members of the closely related phylum Armatimonadetes as an outgroup (23, 129).

MAGs described here are highlighted in green, MAGs previously reported from Jinata Onsen highlighted in pink. All nodes recovered TBE support values greater than 0.7. In cases where reference genomes have a unique strain name or identifier, this is included; otherwise Genbank WGS genome prefixes are used.



Supplemental Figure 1:
Multidimensional scaling plot of Jinata samples. Each point represents the recovered microbial community from a given sample, with sites identified by color and sample type by shape. Samples plotting close to each other are relatively more similar in community composition. Abundance data are transformed by the 4th root to down-weight the effect of abundant taxa. Stress value is 0.0658.

Supplemental Figure 2:
Microscopy images of sediment (Source Pool and Pool 1) or mat (Pool 2, Pool 3, and Out Flow). Left are light microscopy images. Center and right are fluorescence images. At center, blue signal is DAPI-stained (Excitation: 365nm, Emission: BP445~50nm). At right, red is autofluorescence signal of Cyanobacteria (BP395~440nm, LP470nm). Scale bars 50 μm.

Supplemental Table 1: Geochemistry and brief description at sampling sites along the flow path of Jinata Onsen as discussed in the text.

	pH	T (°C)	Fe(II) (μM)	DO (μM)	DIC (mM)	DOC (mM)	Descriptions

Source Pool	**5.4**	**60-63**	**260**	**4.7 (source) 39 (surface)**	**Not measured**	**Not measured**	**Fluffy red iron oxide precipitate**
Pool 1	**5.8**	**59-59.5**	**265**	**58**	**5.51 ± 0.28**	**1.31 ± 0.18**	**Reddish precipitate and streamers in shallower regions, more yellowish deeper**
Pool 2	**6.5**	**44.5-54**	**151**	**134**	**2.09 ± 0.11**	**0.76 ± 0.10**	**Iron oxide-coated microbial mats. Orange to orange-green.**
Pool 3	**6.7**	**37.3-46**	**100**	**175**	**1.79 ± 0.09**	**0.70 ± 0.10**	**Green or mottled orange-green microbial mats, commonly with 1-5cm finger-like morphology.**
Outflow	**6.5**	**27-32**	**45**	**234**	**Not measured**	**Not measured**	**Ocean water within mixing zone at high tide, with constant flow of spring water from Pool 2. Thin green microbial mats.**

Supplemental Table 2: Gas composition of bubbles collected from the Source Pool at Jinata Onsen.

	Average of gas compositions (percent composition)								
Sampling dates (mm/dd/yyyy)	Measurement number	N_2	SE	O_2	SE	CH_4	SE	CO_2	SE
10/03/2017	2	30.5	4.6	0.10	0.01	0.04	0.01	69.3	4.6
04/13/2018	4	55.5	5.5	0.07	0.04	0.05	0.01	44.4	5.0

Supplemental Table 3:
Diversity metrics of Jinata sequencing. Diversity metrics calculated for both 99% and 97% sequence identity cutoffs for assigning OTUs.

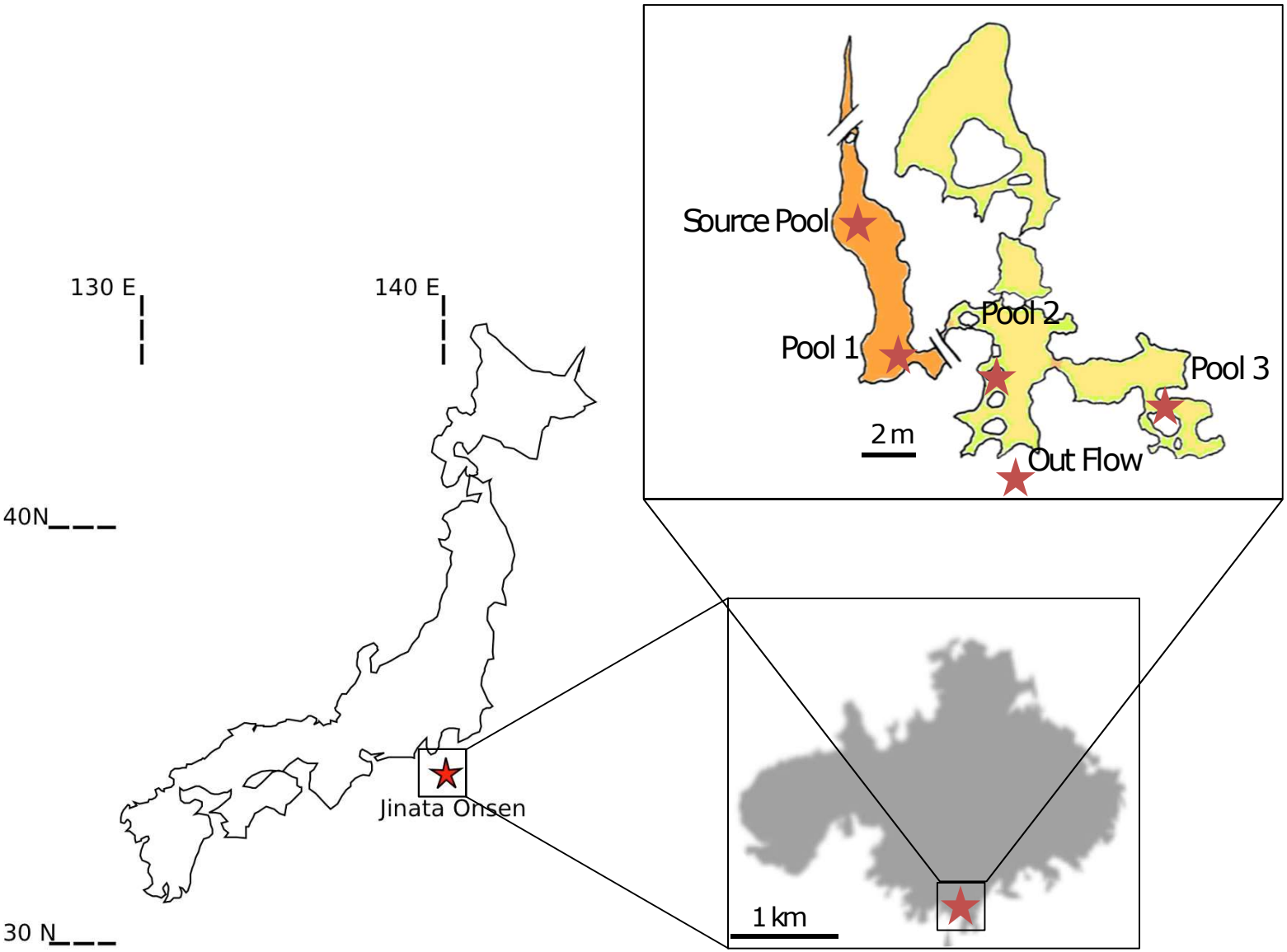
Sample:	Reads:	OTUs (99%):	Good Coverage (99%):	Shannon Index (99%):	Inverse Simpson (99%):	OTUs (97%):	Goods Coverage (97%):	Shannon Index (97%):	Inverse Simpson (97%):
Source Pool A	48680	18832	0.665	10.443	35.146	6951	0.907	7.996	17.361
Source Pool B	18235	7772	0.646	10.388	68.018	4139	0.844	8.651	31.822
Pool 1 A	96268	25305	0.788	10.172	53.546	11734	0.920	8.403	28.691
Pool 1 B	56672	13835	0.797	8.813	26.975	5598	0.933	6.818	13.827
Pool 2 A	35690	12489	0.713	9.625	22.248	7352	0.855	8.271	16.600
Pool 2 B	4454	2274	0.560	9.066	49.599	1729	0.689	8.104	27.390
Pool 3 A	28273	11334	0.665	10.046	35.705	6766	0.824	8.403	20.034
Pool 3 B	2076	1166	0.522	8.832	75.900	786	0.699	7.489	35.312
Outflow A	31980	18486	0.497	11.989	64.318	11994	0.712	10.538	34.881
Outflow B	32465	10896	0.713	9.281	25.691	5918	0.857	7.133	11.585

Supplemental Table 4:
16S rRNA data as OTU table with sequences.

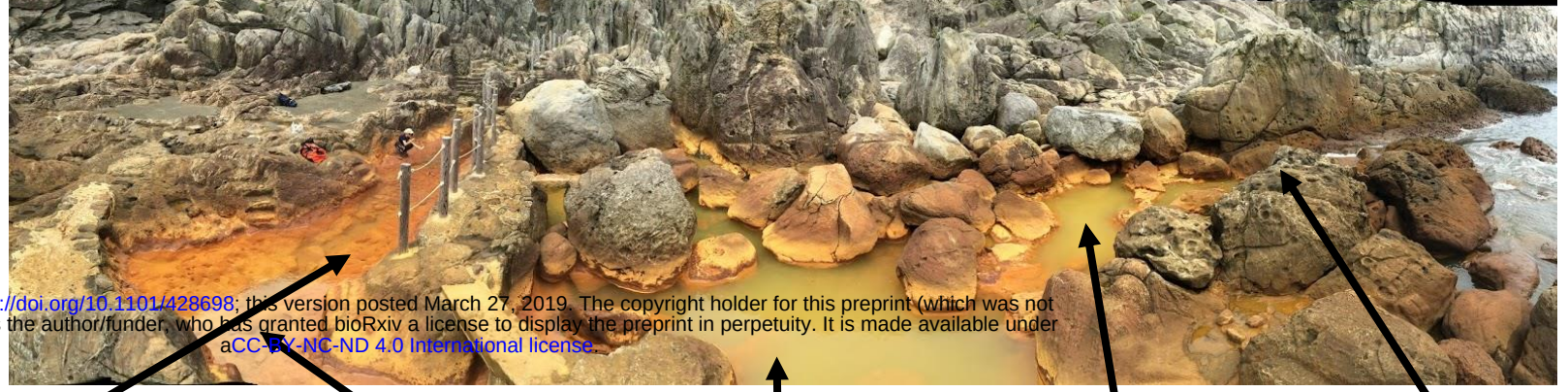
Supplemental Table 5:
16S rRNA data as relative abundance binned at the class level.

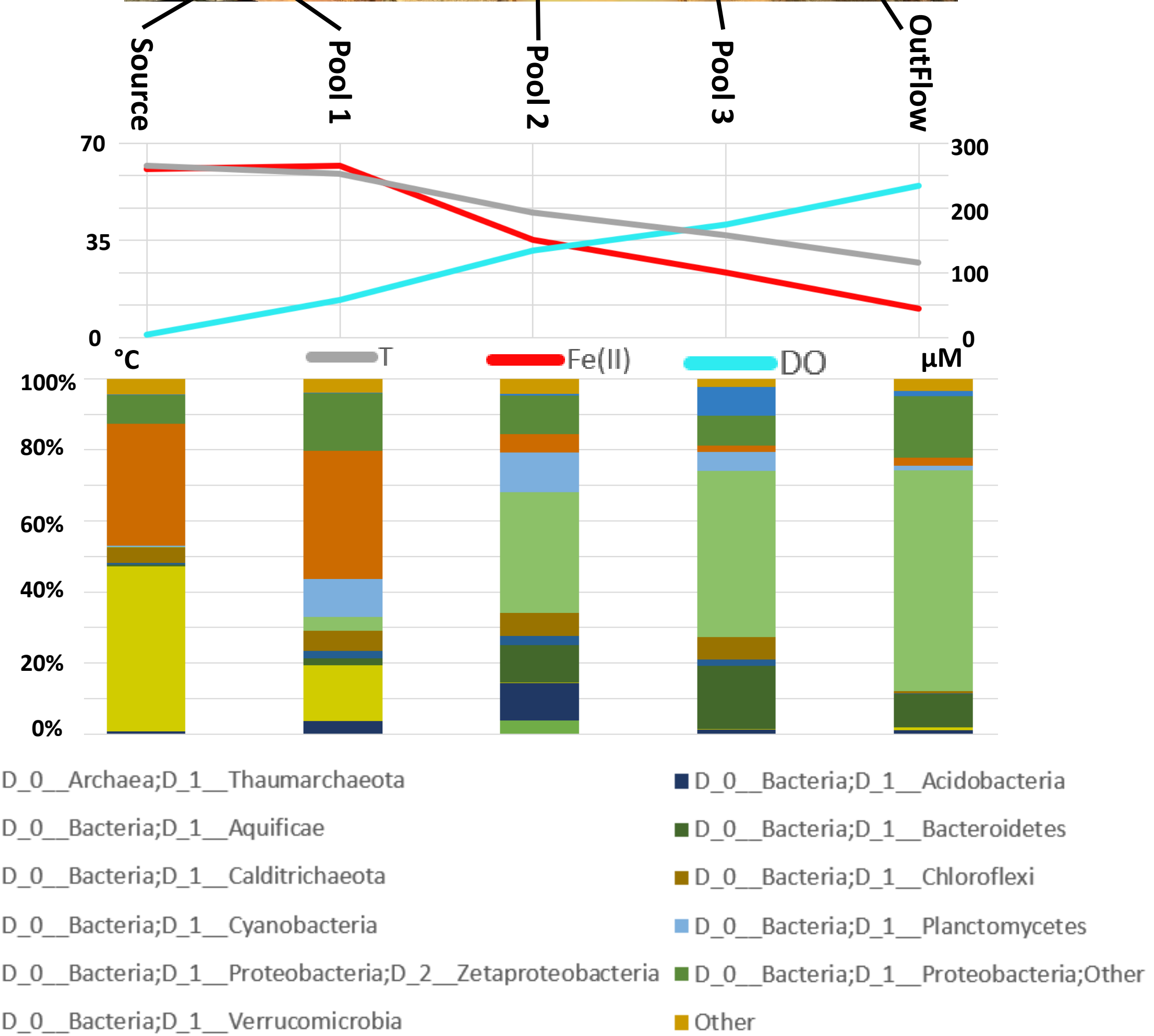
Supplemental Table 6:
High- and medium-quality metagenome-assembled genomes (MAGs) (>50% completeness and <10% contamination) recovered from Jinata Onsen. Predicted taxonomy based on placement in reference phylogeny as presented in Figure 4 and by GTDB-Tk (90). Optimal growth temperatures predicted following methods from (78).

Supplemental Table 7:
Presence of genes involved in aerobic respiration, hydrogen- and iron-oxidation, and carbon fixation in MAGs discussed in the text.

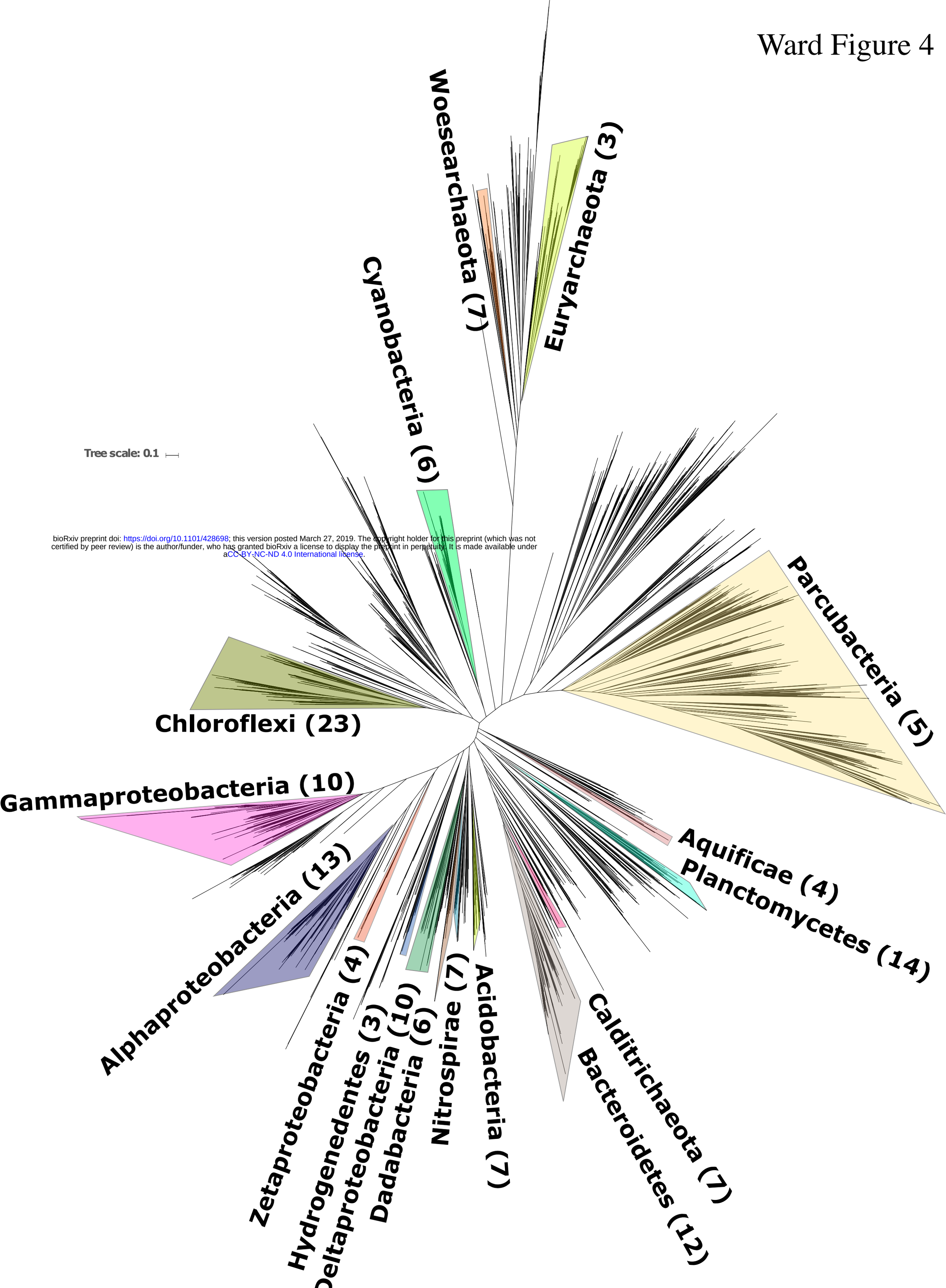

Source Pool
Pool 1
Pool 2
Pool 3
2 m
Out Flow
130 E
140 E
40N
30 N
Jinata Onsen
1 km

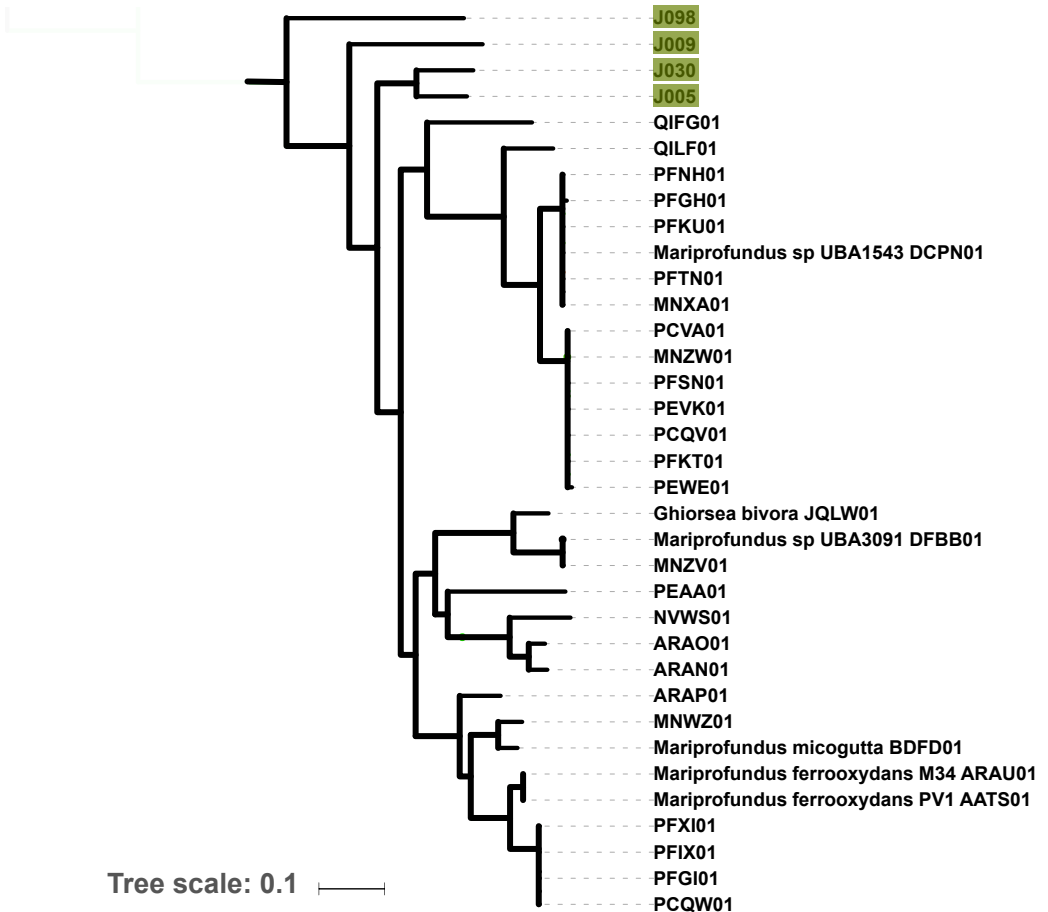
Ward Figure 2

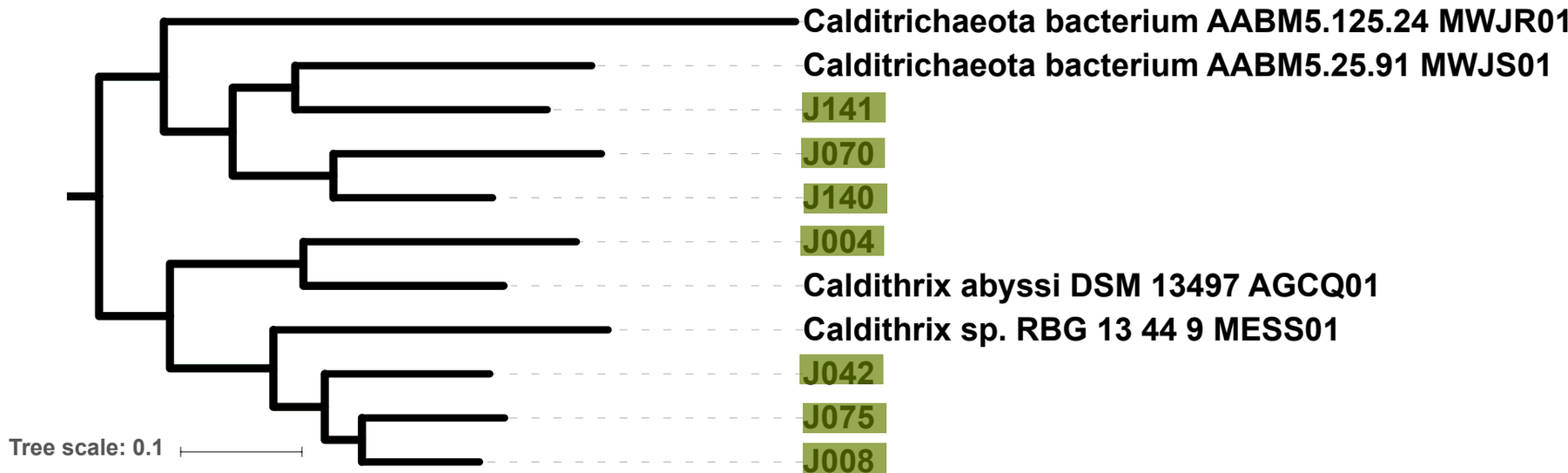


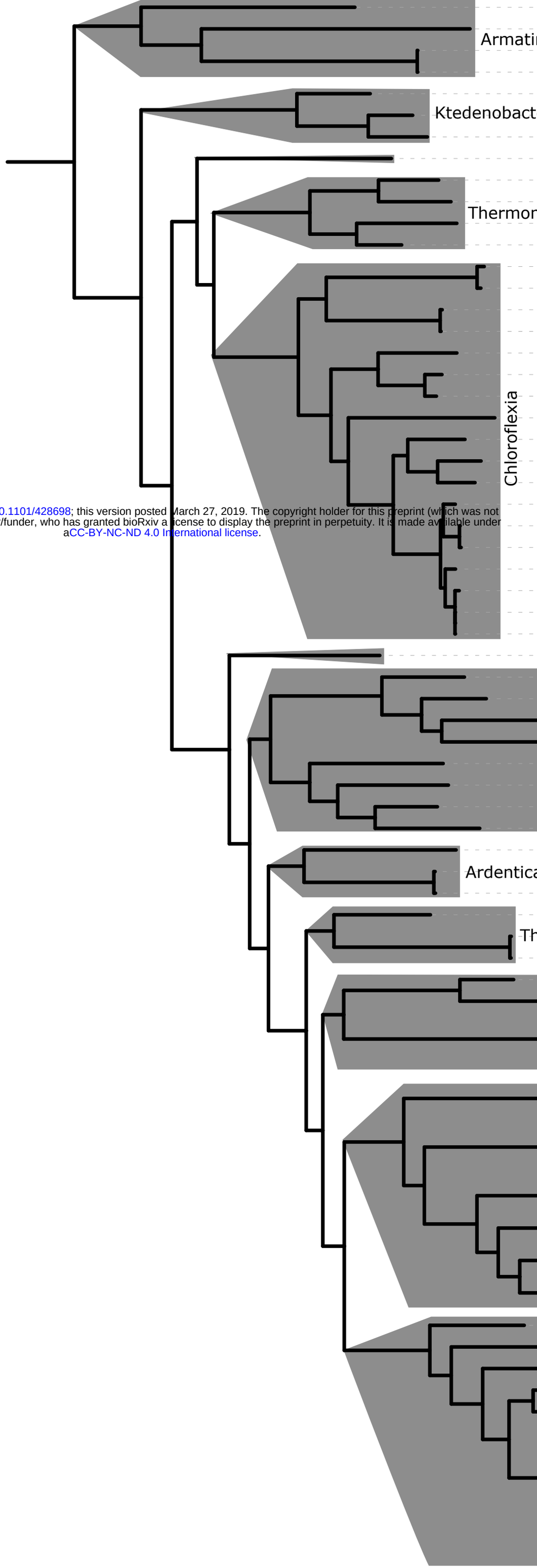


Ward Figure 4




J098
J009
J030
J005
QIFG01
QILF01
PFNH01
PFGH01
PFKU01
Mariprofundus sp UBA1543 DCPN01
PFTN01
MNXA01
PCVA01
MNZW01
PFSN01
PEVK01
PCQV01
PFKT01
PEWE01
Ghiorsea bivora JQLW01
Mariprofundus sp UBA3091 DFBB01
MNZV01
PEAA01
NVWS01
ARAO01
ARAN01
ARAP01
MNWZ01
Mariprofundus micogutta BDFD01
Mariprofundus ferrooxydans M34 ARAU01
Mariprofundus ferrooxydans PV1 AATS01
PFXI01
PFIX01
PFGI01
PCQW01
Tree scale: 0.1


Calditrichaeota bacterium AABM5.125.24 MWJR01
Calditrichaeota bacterium AABM5.25.91 MWJS01
J141
J070
J140
J004
Caldithrix abyssi DSM 13497 AGCQ01
Caldithrix sp. RBG 13 44 9 MESS01
J042
J075
J008
Tree scale: 0.1



Tree scale: 0.1

Chthonomonas calidirosea
Fimbriimonas ginsengisoli
NKPU01
NKPV01
Thermogemmatispora sp. PM5
Thermosporothrix hazakensis
Ktedonobacter racemifer
Thermobaculum terrenum
Thermomicrobiales sp. KI4
Thermomicrobium roseum
Nitrolancetus hollandicus
Sphaerobacter thermophilus
Herpetosiphon geysericola
Herpetosiphon aurantiacus
Kallotenue papyrolyticum
Chloroflexi bacterium JKG1
Kouleothrix aurantiaca
Roseiflexus castenholzii
Roseiflexus sp. RS1
Chlorothrix halophila
Oscillochloris trichoides
Viridilinea mediosalina
Chloroploca asiatica
Chloroflexus islandicus
Chloroflexus aggregans
J150
Chloroflexus sp. Y396
Chloroflexus aurantiacus
J043
Chloroflexus sp. Y400
J114
J117
Litorilinea aerophila
Caldilinea aerophila
Caldilinea sp. AAV1
QEXY01
J123
J111
J095
J118
Ardenticatena maritima
J129
J086
Thermoflexaceae sp. JAD2
Thermoflexus hugenholtzii
J033
J162
Roseilinea gracile
J036
PGTN01
J027
PGTG01
PGTL01
PGTM01
J076
PGTI01
J064
JP1 8
J039
J038
PGTF01
J130
J082
QEXX01
J097
QEXW01
Bellilinea caldifistulae
Thermanaerothrix daxensis
Ornatilinea apprima
Levilinea saccharolytica
Leptolinea tardivitalis
Anaerolinea thermophila
Anaerolinea thermolimosa