**Supplemental Information**

**Supplemental Methods:**

 **Geochemical analysis**

Gas 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 O2, N2, CO2, 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 H2 and CH4.

 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 CO2 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 rDNA sequencing and analysis:**

Following return to the lab, microbial 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 from each extract using archaeal and bacterial primers 515F (GTGCCAGCMGCCGCGGTAA) and 926R (CCGYCAATTYMTTTRAGTTT) (106). DNA was quantified with a Qubit 3.0 fluorimeter (Life Technologies, Carlsbad, CA) according to manufacturer’s instructions following DNA extraction and PCR steps. All samples yielded PCR amplicons when viewed 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 (107). 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 (108). Then, the most abundant sequence was chosen as representative for each *de novo* OTU (109). Taxonomic identification for each representative sequence was assigned using the Silva-115 database (110) clustered separately at 99% and at 97% similarity. Singletons and contaminants (OTUs appearing in the negative control datasets) were removed. 16S rDNA sequences were aligned using MAFFT (111) and a phylogeny constructed using FastTree (112). Alpha diversity was estimated using the Shannon Index (113) and Inverse Simpson metric (1/D) (114, 115). Assessment of sampling depth was estimated using Good’s Coverage (116). 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 OHK environments were produced in R using the vegan and ggplot2 packages (117-119).

**Extended Discussion:**

 **Site description:**

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, 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 expeditions.

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. 120), 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.

**Hydrogen and iron oxidation**

Both H2 and CH4 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 CO2, CH4, N2 (Supplemental Table 2). This subsequent non-detection of H2 may be related to temporal variability in the gas composition at Jinata (e.g. following tidal influence; significant variability was observed in the CO2:N2 ratio between two sampling dates, Supplemental Table 2) or may reflect oxidation of H2 between sampling and analysis. The detection limit of H2 for these later measurements was ~1 nmol/cc, well above the energetic and ecological limits for hydrogenotrophic metabolisms (e.g. 121) leaving open the possibility of biologically significant H2 fluxes at Jinata around the time of sampling.

The abundance at Jinata of microbes with the genetic capacity for hydrogenotrophy suggests that H2 may be contributing to lithoautotrophy near the hot spring source, despite H2 concentrations being low (below the detection of ~1 nmol/cc in the gas phase of our quantitative gas analyses, or ~1 nM in the aqueous phase, 122). However, this is unsurprising, as the oxidation of H2 coupled to O2 reduction is a thermodynamically favorable process even at low substrate concentrations (e.g. **ΔrG' < -340** kJ/mol under conditions at Jinata with substrate concentrations at our limit of detection, 123). Consistent with this thermodynamic favorability, biology has been shown to make use of this metabolism in environments such as hot springs with H2 concentrations near our detection limits (124) and in Antarctic soils where microbes rely on uptake of trace atmospheric H2 at concentrations around 190 ppbv (121). Improved quantification of H2 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.

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 ~1-3% 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. In contrast, the absence of Hydrogenothermaceae downstream may be a real signal driven by the rapid disappearance of H2 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 swamped out by increased numbers of Cyanobacteria, Chloroflexi, and other sequences.

Genes recovered in Zetaproteobacteria MAGs include a terminal O2 reductase from the C-family of heme copper oxidoreductases for respiration at low O2 concentrations and Cyc2 cytochrome genes implicated in ferrous iron oxidation in Zetaproteobacteria and other taxa (e.g. Chlorobi) (125-127). 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. J098 did not recover a Cyc2 cytochrome gene; based on phylogenetic position this MAG captures a member of the most basal Zetaproteobacteria lineage recovered to date, which if correct may have diverged prior to the evolution of iron oxidation in this group. However, this MAG is also only 80% complete and so there is a significant probability of failure to recover this gene even if it were present in the source genome (MetaPOAP False Negative estimate 0.205).

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 (90). A MAG from a member of Calditrichaeota has previously been recovered from Chocolate Pots hot spring in Yellowstone National Park (49); together with the data presented here this suggests that this phylum may be a common member of microbial communities in iron-rich hot springs.

Consistent with expectations of the nitrogen demand of highly productive oxygenic phototrophic ecosystems relative to poorly productive lithotrophic systems, the abundance of genes for biological nitrogen fixation via nitrogenase was 2.5 times higher in Pool 2 and Pool 3 than near the source (*nifD*/*rpoB* of 0.075 versus 0.03).

**Chloroflexi**

Previous analyses suggested that the Roseilinea lineage belongs to the Anaerolineae (42) or Thermofonsia (7) classes; 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-level lineage within the broader “subphylum I” of Chloroflexi (Figure 7).

Members of the Caldilineae have previously been isolated from intertidal hot springs in Iceland (128) and Japanese hot springs (129). Characterized organisms in this class are filamentous, anaerobic, or facultatively aerobic heterotrophs (128- 130); and therefore these taxa may play a role in degrading biomass within low-oxygen regions of microbial mats at Jinata.

The Chloroflexi class Ardenticatenia was first described from an isolate from an iron-rich Japanese hydrothermal field (40) and has since been recovered from sulfidic hot springs as well (102). A MAG closely related to *Ardenticatena maritima* was recovered from Jinata Onsen,J129. While *Ardenticatena maritima 110S* contains a complete denitrification pathway (98), 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 (75) 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 (131, 132).

J114 encodes aerobic respiration via an A-family heme copper oxidoreductase and an alternative complex III like many other nonphototrophic Chloroflexi lineages (e.g. 7, 38) 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. The alternative complex III encoded by J114 branches basally to a clade of alternative complex III sequences from other subphylum I Chloroflexi, potentially reflecting vertical inheritance of these genes from the last common ancestor of this clade; however, the A-family heme copper oxidoreductase encoded by J114 is in a more derived position closely related to sequences from members of the Caldilineae, and may have been acquired via horizontal gene transfer from a member of this group.

**Supplemental references:**

1. Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M. and Gormley, N., 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME journal*, *6*(8), p.1621.
2. Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I. and Huttley, G.A., 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, *7*(5), p.335.
3. Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), pp.2460-2461.
4. Wang, Q., Garrity, G.M., Tiedje, J.M. and Cole, J.R., 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology*, *73*(16), pp.5261-5267.
5. 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 research*, *41*(D1), pp.D590-D596.
6. 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 research*, *30*(14), pp.3059-3066.
7. 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.
8. SHANNON, C., 1948. A Mathematical Theory of Communication-The Bell System Technical Journal, vol. 27, pag. 379-423, 623-656.
9. SIMPSON, E.H. 1949. Measurement of diversity: Nature, v. 163, p. 688.
10. Hill, M.O., 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology*, *54*(2), pp.427-432.
11. Good, I.J., 1953. The population frequencies of species and the estimation of population parameters. *Biometrika*, *40*(3-4), pp.237-264.
12. R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
13. Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O’Hara, R.B., Simpson, G.L., Solymos, P. and Stevens, M.H.H., 2016. Vegan: community ecology package. R package version 2.3-5. R Foundation, Vienna, Austria.
14. Wickham, H., 2010. ggplot2: elegant graphics for data analysis. *J Stat Softw*, *35*(1), pp.65-88.
15. Fortney, N.W., He, S., Converse, B.J., Beard, B.L., Johnson, C.M., Boyd, E.S. and Roden, E.E., 2016. Microbial F e (III) oxide reduction potential in C hocolate P ots hot spring, Y ellowstone N ational P ark. *Geobiology*, *14*(3), pp.255-275.
16. Ji, M., Greening, C., Vanwonterghem, I., Carere, C.R., Bay, S.K., Steen, J.A., Montgomery, K., Lines, T., Beardall, J., van Dorst, J. and Snape, I., 2017. Atmospheric trace gases support primary production in Antarctic desert surface soil. *Nature*, *552*(7685), p.400.
17. Amend, J.P. and Shock, E.L., 2001. Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and Bacteria. *FEMS microbiology reviews*, *25*(2), pp.175-243.
18. Flamholz, A., Noor, E., Bar-Even, A. and Milo, R., 2011. eQuilibrator—the biochemical thermodynamics calculator. *Nucleic acids research*, *40*(D1), pp.D770-D775.
19. D'Imperio, S., Lehr, C.R., Oduro, H., Druschel, G., Kühl, M. and McDermott, T.R., 2008. Relative importance of H2 and H2S as energy sources for primary production in geothermal springs. *Applied and environmental microbiology*, *74*(18), pp.5802-5808.
20. 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. *Proceedings of the National Academy of Sciences*, *108*(34), pp.14109-14114.
21. 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. *Frontiers in microbiology*, *6*, p.1265.
22. 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. *Frontiers in microbiology*, *8*, p.1584.
23. 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. *International journal of systematic and evolutionary microbiology*, *63*(3), pp.1149-1154.
24. 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. *International journal of systematic and evolutionary microbiology*, *53*(6), pp.1843-1851.
25. Grégoire, P., Bohli, M., Cayol, J.L., Joseph, M., Guasco, S., Dubourg, K., Cambar, J., Michotey, V., Bonin, P., Fardeau, M.L. and Ollivier, B., 2011. Caldilinea tarbellica sp. nov., a filamentous, thermophilic, anaerobic bacterium isolated from a deep hot aquifer in the Aquitaine Basin. *International journal of systematic and evolutionary microbiology*, *61*(6), pp.1436-1441.
26. Chen, J. and Strous, M., 2013. Denitrification and aerobic respiration, hybrid electron transport chains and co-evolution. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1827*(2), pp.136-144
27. Stein, L.Y. and Klotz, M.G., 2016. The nitrogen cycle. *Current Biology*, *26*(3), pp.R94-R98.



**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 | N2 | SE | O2 | SE | CH4 | SE | CO2 | 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** | 26057 | 9558 | 0.724 | 10.594 | 83.020 | 4632 | 0.884 | 8.196 | 23.035 |
| **Source Pool B** | 49340 | 14392 | 0.790 | 10.275 | 44.714 | 5530 | 0.932 | 7.229 | 12.835 |
| **Pool 1 A** | 97445 | 21166 | 0.848 | 10.128 | 56.287 | 10160 | 0.935 | 8.080 | 24.682 |
| **Pool 1 B** | 57250 | 10559 | 0.872 | 8.794 | 33.323 | 4766 | 0.945 | 6.414 | 12.005 |
| **Pool 2 A** | 41515 | 13114 | 0.759 | 9.754 | 24.340 | 7710 | 0.873 | 8.118 | 14.702 |
| **Pool 2 B** | 45171 | 17211 | 0.697 | 10.708 | 50.836 | 10525 | 0.832 | 8.980 | 25.783 |
| **Pool 3 A** | 45148 | 15988 | 0.722 | 10.287 | 33.295 | 9302 | 0.853 | 8.351 | 16.880 |
| **Pool 3 B** | 29778 | 12023 | 0.682 | 10.894 | 84.725 | 6625 | 0.837 | 8.553 | 31.520 |
| **Outflow A** | 32382 | 17741 | 0.542 | 11.931 | 57.572 | 11290 | 0.738 | 10.262 | 28.674 |
| **Outflow B** | 32651 | 8881 | 0.797 | 9.237 | 28.728 | 4210 | 0.909 | 6.373 | 9.850 |

**Supplemental Table 4:**

16S rDNA data as OTU table with sequences. Available as a single file at <https://figshare.com/s/9cfb9865b31431c59195>

**Supplemental Table 5:**

16S rDNA 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.