ISME

www.nature.com/ismej

ARTICLE OPEN

Unique H_2-utilizing lithotrophy in serpentinite-hosted systems

Masaru Konishi Nobu[1,8✉], Ryosuke Nakai[1,2,8], Satoshi Tamazawa[1,3], Hiroshi Mori[4], Atsushi Toyoda[4], Akira Ijiri[5], Shino Suzuki[6,7], Ken Kurokawa[4], Yoichi Kamagata[1] and Hideyuki Tamaki[1✉]

Serpentinization of ultramafic rocks provides molecular hydrogen (H_2) that can support lithotrophic metabolism of microorganisms, but also poses extremely challenging conditions, including hyperalkalinity and limited electron acceptor availability. Investigation of two serpentinization-active systems reveals that conventional H_2-/CO_2-dependent homoacetogenesis is thermodynamically unfavorable in situ due to picomolar CO_2 levels. Through metagenomics and thermodynamics, we discover unique taxa capable of metabolism adapted to the habitat. This included a novel deep-branching phylum, "*Ca*. Lithacetigenota", that exclusively inhabits serpentinite-hosted systems and harbors genes encoding alternative modes of H_2-utilizing lithotrophy. Rather than CO_2, these putative metabolisms utilize reduced carbon compounds detected in situ presumably serpentinization-derived: formate and glycine. The former employs a partial homoacetogenesis pathway and the latter a distinct pathway mediated by a rare selenoprotein—the glycine reductase. A survey of microbiomes shows that glycine reductases are diverse and nearly ubiquitous in serpentinite-hosted environments. "*Ca*. Lithacetigenota" glycine reductases represent a basal lineage, suggesting that catabolic glycine reduction is an ancient bacterial innovation by Terrabacteria for gaining energy from geogenic H_2 even under hyperalkaline, CO_2-poor conditions. Unique non-CO_2-reducing metabolisms presented here shed light on potential strategies that extremophiles may employ for overcoming a crucial obstacle in serpentinization-associated environments, features potentially relevant to primordial lithotrophy in early Earth.

The ISME Journal (2023) 17:95–104; https://doi.org/10.1038/s41396-022-01197-9

INTRODUCTION

Serpentinite-hosted systems are rare and extreme habitats in which a hydrothermal process, serpentinization, alters ultramafic mantle rocks and yields hyperalkaline fluid rich in molecular hydrogen (H_2) and reduced one-carbon compounds [1–8]. These fluids are often electron acceptor depleted—oxygen, nitrate, sulfate, etc. are absent (i.e., anoxic) and even the least favorable exogenous acceptor, carbon dioxide (CO_2), is limiting due to the high alkalinity. Though previous studies explore the diversity of organisms in serpentinite-hosted systems, we have little insight into how indigenous H_2-utilizing microorganisms combat the unique metabolic challenges in situ. One recent study shows strategies that methane-generating archaea employ to oxidize H_2 in situ [9], but how other microorganisms (i.e., H_2-utilizing anaerobic bacteria) overcome the electron acceptor limitation is poorly understood. Further, given that life is theorized to have emerged as H_2-utilizing lithotrophs in early Earth serpentinite-hosted systems [10–24], modern lithotrophs inhabiting such ecosystems may represent valuable extant windows into the metabolism of primordial organisms. In this study, we pair metagenomics and thermodynamics to characterize uncultured putative anaerobic H_2 utilizers inhabiting alkaline H_2-rich serpentinite-hosted systems (Hakuba Happo hot springs in Hakuba, Japan, and The Cedars springs in California, USA; pH ~10.9 and ~11.9, respectively [25–27]) and elucidate novel, potentially ancient, lithotrophic strategies.

Thermodynamics and geochemistry

The two primary strategies for utilizing H_2 under anoxic conditions without favorable exogenous electron acceptors are methanogenesis and homoacetogenesis. As bacteria were detected in both Hakuba and The Cedars, yet archaea were absent in Hakuba, we focused our analyzes on metabolic strategies supporting bacterial H_2 utilization (i.e., homoacetogenesis). To evaluate whether homoacetogenesis is viable in situ, we examined the in situ geochemical environment and the thermodynamics of H_2/formate utilization and homoacetogenesis. The spring waters of both Hakuba and The Cedars contained H_2 (e.g., 201–664 μM in Hakuba [27]). Formate, another compound thought to be abiotically generated through serpentinization, was also detected in Hakuba (8 μM in drilling well #3 [28]) and The Cedars (6.9 μM in GPS1). Acetate has also been detected in situ (4 μM in Hakuba [28] and 69.3 μM in The Cedars GPS1), suggesting these ecosystems may host novel H_2- and/or formate-utilizing homoacetogens.

[1]Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Ibaraki, Tsukuba 305-8566, Japan. [2]Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 2-17-2-1, Tsukisamu-Higashi, Sapporo 062-8517, Japan. [3]Horonobe Research Institute for the Subsurface Environment (H-RISE), Northern Advancement Center for Science & Technology, 5-3 Sakaemachi, Horonobe, Teshio, Hokkaido 098-3221, Japan. [4]National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan. [5]Kochi Institute for Core Sample Research, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 200 Monobe Otsu, Nankoku, Kochi, Japan. [6]Institute for Extra-Cutting-Edge Science and Technology Avant-garde Research (X-star), JAMSTEC, Natsushima 2-15, Yokosuka, Kanagawa 237-0061, Japan. [7]Institute of Space and Astronautical Science (ISAS), Japan Aerospace Exploration Agency (JAXA), 3-1-1 Yoshinodai, Chuo-ku, Sagamihara, Kanagawa 252-5210, Japan. [8]These authors contributed equally: Masaru Konishi Nobu, Ryosuke Nakai. ✉email: m.nobu@aist.go.jp; tamaki-hideyuki@aist.go.jp

Received: 28 April 2021 Revised: 8 January 2022 Accepted: 17 January 2022
Published online: 7 October 2022

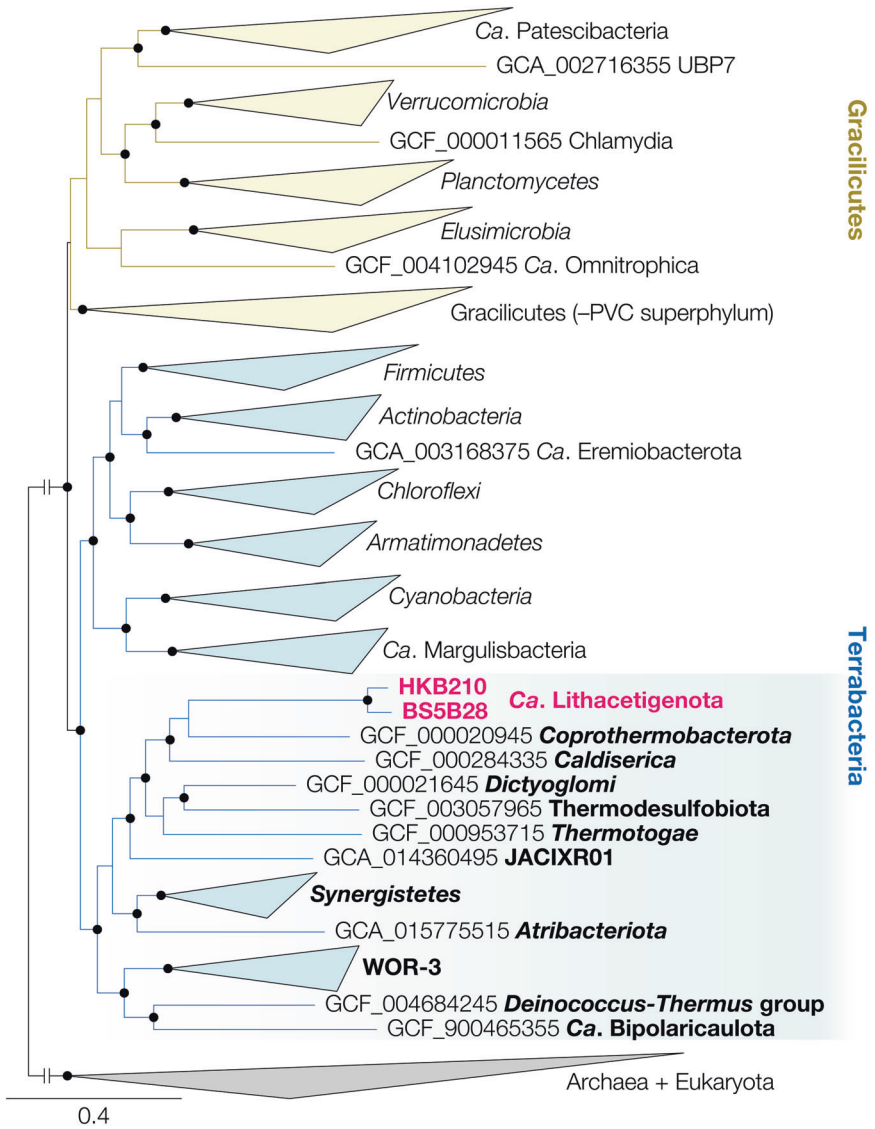
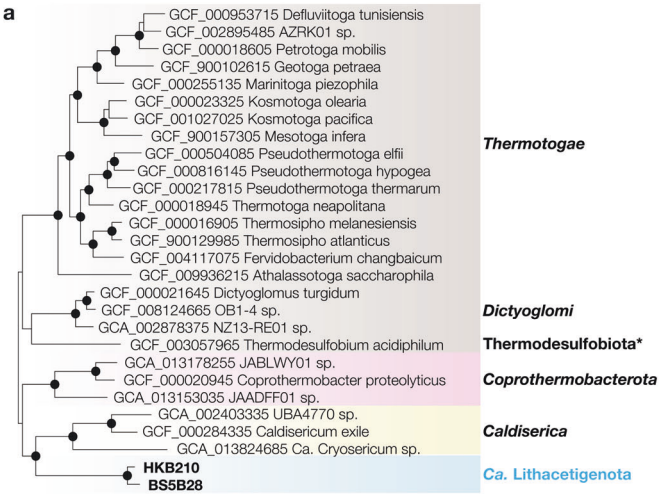


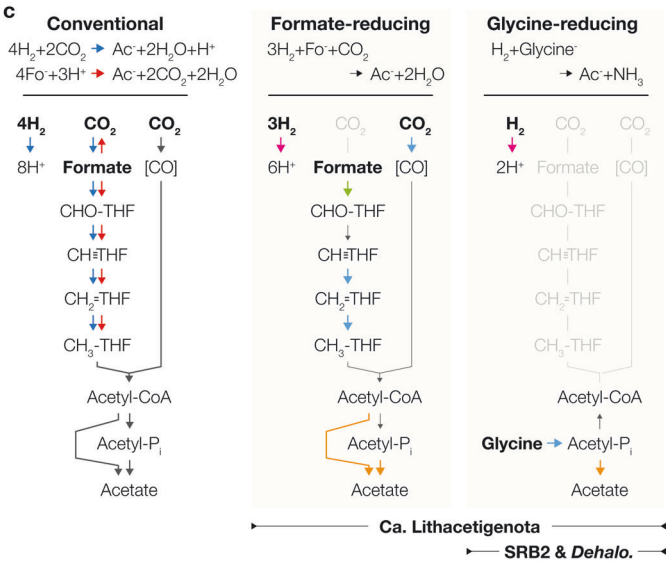
Fig. 1 Ribosomal protein tree including high-quality MAGs from 74 GTDB-defined phylum-level lineages. Representative genomes (highest quality based on a score defined as completeness – 5*contamination, both estimated by CheckM) were chosen for bacterial classes that contain at least one genome the meet the following criteria: (i) cultured organisms with ≥90% completeness, ≤5% contamination (as estimated by CheckM), and ≤ 20 contigs; (ii) uncultured organisms with ≥85% completeness, ≤3% contamination, and ≤20 contigs; and (iii) *Ca.* Patescibacteria with ≥60% completeness and ≤1 contig. Universally conserved ribosomal proteins were collected from each genome, aligned with MAFFT v7.394, trimmed with BMGE v1.12 (-m BLOSUM30 -g 0.67 -b 3), and concatenated. A maximum likelihood tree was calculated using IQ-TREE v2.1.3 with the UDM0064LCLR model (-m Poisson+UDM0064LCLR), ultrafast bootstrap approximation, and SH-like approximate likelihood ratio test (-B 1000 -alrt 1000; bootstrap values are recalculated with BOOSTER using the -tbe option). Branches with ≥90% ultrafast bootstrap support and ≥80% SH-alrt support are indicated with black circles. Archaeal and eukaryotic genomes were used as an outgroup. The inter-domain branch was shortened with a break to 1/10 of the calculated length for illustrative purposes. Phylogenetic groups corresponding to "Gracilicutes" and "Terrabacteria" are indicated yellow and blue respectively. *Ca.* Lithacetigenota are highlighted (magenta). See Supplementary Fig. S4 for full tree.

Thermodynamic calculations using newly measured and published geochemical data (Tables S1 and S2) confirmed that H_2 and formate are reductants in situ (i.e., $H_2 = 2H^+ + 2e^-$/$Formate^- = H^+ + CO_2 + 2e^-$): the Gibbs free energy yields (ΔG) for oxidation (coupled with physiological electron carriers $NADP^+$, NAD^+, and ferredoxin) are less than −4.78 kJ per mol H_2 and −24.92 kJ per mol formate in Hakuba, and −10.73 and −22.03 in The Cedars respectively (H_2 concentration was not available for The Cedars so the highest concentration observed in Hakuba [664 µM] was used; see Supplementary Results). However, serpentinite-hosted systems impose a unique challenge to homoacetogenesis—a key substrate, CO_2, is at extremely low concentrations due to the high alkalinity. We estimate that the aqueous CO_2 concentration is below 0.0006 nM in Hakuba (pH 10.7 and <0.1 µM TIC) and 0.003 nM in The Cedars (pH 11.9 and 35 µM TIC) (Table S2) [25, 27]. In Hakuba, H_2/CO_2-driven acetogenesis (ΔG of −3.71 kJ per mol acetate) cannot support microbial energy generation ($\Delta G \leq -20$ kJ per mol is necessary [29]; Fig. S1). Moreover, in both Hakuba and The Cedars, one of the first steps in CO_2-reducing homoacetogenesis, reduction of CO_2 to formate, is unfavorable based on the thermodynamics presented above ($\Delta G > +24.92$ or +22.02 kJ per mol formate). Thus, catabolic reduction of CO_2 to acetate is thermodynamically challenging in situ and may only run if investing ATP (e.g., Calvin–Benson–Bassham cycle [−6 ATP; ΔG of −361.68 kJ per mol acetate in Hakuba] or reductive tricarboxylic acid [−1 ATP; −61.68 kJ per mol]). Under CO_2 limitation,



b

	(1)	(2)	(3)	(4)	(5)	(6)
Thermotogae (1)	53.7	44.1	44.7	45.0	46.2	44.3
Caldiserica (2)	44.3	53.0	44.8	44.8	45.4	43.8
Ca. Lithacetigenota (3)	44.4	44.5	84.5	45.2	46.6	44.6
Coprothermobacterota (4)	45.1	45.1	45.5	66.0	46.3	44.4
Dicytoglomi (5)	46.3	45.7	46.8	46.3	74.1	46.7
Thermodesulfobiota* (6)	44.6	43.8	44.9	44.5	46.6	86.8



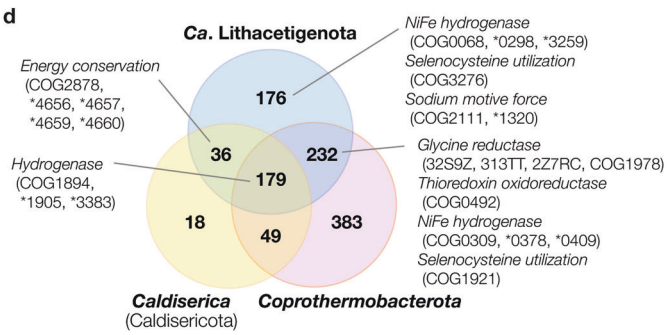


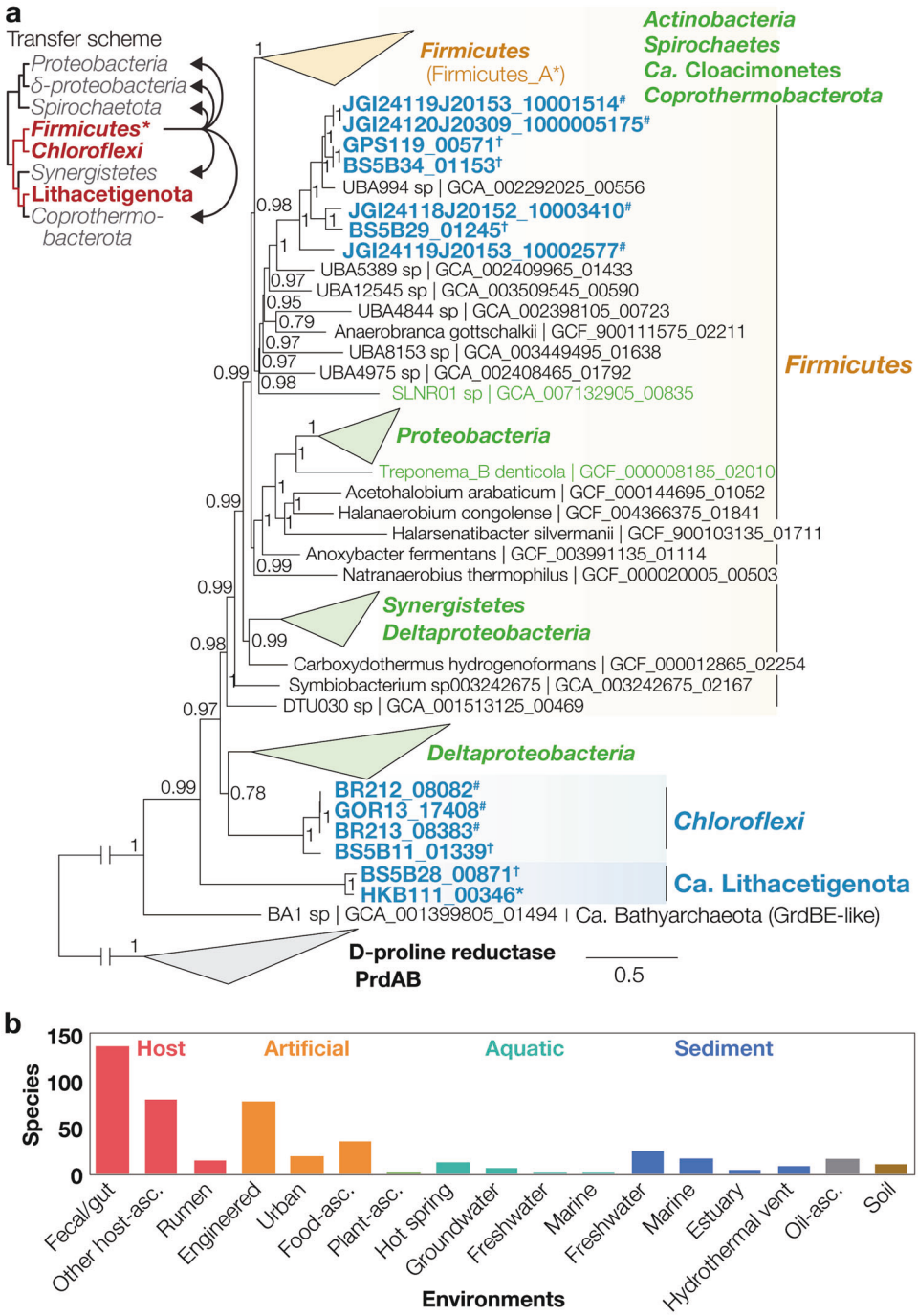
Fig. 2 "*Ca*. Lithacetigenota" phylogeny, lithotrophic acetate generation pathways, and comparative genomics with neighboring phyla. a A maximum likelihood tree was calculated for a concatenated alignment of universally conserved ribosomal protein sequences from representative genomes of individual phyla (aligned with MAFFT v7.394 [default parameters] and trimmed with BMGE v1.12 (−m BLOSUM30 −g 0.67 −b 3) using IQ-TREE v2.1.3 with the UDM0064LCLR model (-m Poisson+UDM0064LCLR), ultrafast bootstrap approximation, and SH-like approximate likelihood ratio test (-B 1000 -alrt 1000; bootstrap values are recalculated with BOOSTER using the -tbe option). Branches with ≥90% ultrafast bootstrap support and ≥80% SH-alrt support are indicated with black circles. Phylum names are shown for NCBI taxonomy (italicized) or GTDB classification (*). **b** The average inter-phylum AAI (as calculated by CompareM) was calculated using GTDB species representatives. **c** Putative metabolic pathways potentially adapted to the CO_2-limited hyperalkaline conditions encoded by "*Ca*. Lithacetigenota" members and others: formate- and glycine-reducing acetate generation. Arrow colors indicate oxidative (pink), reductive (blue), ATP-yielding (orange), and ATP-consuming (green) steps. **d** Venn diagram of COGs/NOGs (as predicted by eggnog-mapper) fully conserved across all members of each phylum (genomes included in GTDB release 95 with completeness ≥85% and contamination ≤5%). COGs/NOGs related to lithotrophy and alkaliphily are highlighted. * "COG" abbreviated.

autotrophs are known to accelerate CO_2 uptake through HCO_3^- dehydration to CO_2 (carbonic anhydrase) or carbonate mineral dissolution, but both only modify kinetics and are not effective in changing the maximum CO_2 concentration (determined by equilibrium with carbonate species). In addition, in Hakuba, the CO_3^{2-} concentration is too low (84.7 nM CO_3^{2-}) to cause carbonate mineral precipitation (e.g., [CO_3^{2-}] must exceed 38.5 μM given K_s of 5×10^{-9} for $CaCO_3$ and [Ca^{2+}] of 0.13 mM).

Based on thermodynamic calculations, the energy obtainable from H_2/CO_2-driven homoacetogenesis is too small to support life in many serpentinite-hosted systems, yet acetate is detected in some of these ecosystems (Fig. S2 and Table S2; note that we cannot exclude the possibility that acetate may be produced abiotically by water–rock reactions [30]). Thus, CO_2-independent electron-disposing metabolism may have been necessary for extremophilic organisms to gain energy from H_2 in the hyperalkaline fluids of hydrothermal systems. Here, we explore the metabolic capacities of organisms living in serpentinite-hosted systems to gain insight into potential metabolic strategies for utilizing H_2 under the extreme conditions in situ.

Diverse putative H_2- and formate-utilizing organisms

Through metagenomic exploration of the two serpentinite-hosted systems (Table S3), we discover a plethora of phylogenetically novel organisms encoding genes for H_2 and formate metabolism (19 bins with 73.2–94.8% completeness and 0.0–8.1% contamination [86.1% and 3.8% on average respectively]; available under NCBI BioProject PRJNA453100) despite challenges in acquisition of genomic DNA (15.7 and 18.9 ng of DNA from 233 and 720 L of filtered Hakuba Happo spring water, respectively; RNA was below the detection limit). We find metagenome-assembled genomes (MAGs) affiliated with lineages of *Firmicutes* (e.g., *Syntrophomonadaceae* and uncultured family SRB2), *Actinobacteria*, and candidate division NPL-UPA2 [31] (Fig. S3). We also recovered MAGs for a novel lineage, herein referred to as "*Ca*. Lithacetigenota", that inhabits both Hakuba and The Cedars and, to our knowledge, no other ecosystems (Figs. 1, 2a, S3, and S4). The average amino acid identity (AAI) between *Ca*. Lithacetigenota and neighboring phyla (*Coprothermobacterota*, *Dictyoglomi*, Thermodesulfobiota [GTDB-defined phylum], *Thermotogae*, and *Caldiserica*) was comparable to the average interphylum AAI among the neighboring phyla (45.33 ± 0.86% vs 45.17 ± 0.99%), suggesting that *Ca*. Lithacetigenota represents a novel phylum-level lineage (Fig. 2a, b). These genomes encode enzymes for oxidizing H_2 and formate (i.e., hydrogenases and formate dehydrogenases [32–39]; see Supplementary Results), suggesting that organisms in situ can employ H_2 and formate as electron donors.



"*Ca*. Lithacetigenota" has unique site-adapted metabolism

Inspection of the serpentinite-hosted environment-exclusive phylum "*Ca*. Lithacetigenota" reveals specialization to H_2-driven lithotrophy potentially suitable for the low-CO_2 in situ conditions (Fig. 2c). We discover that The Cedars-inhabiting population (e.g., MAG BS5B28, 94.8% completeness and 2.9% contamination) harbors genes for H_2 oxidation ([NiFe] hydrogenase Hox), a nearly complete Wood-Ljungdahl pathway, and an oxidoreductase often associated with acetogenesis—NADH:ferredoxin oxidoreductase Rnf [40, 41] (Tables S4 and S5). One critical enzyme, the formate dehydrogenase, is missing from all three "*Ca*. Lithacetigenota" MAGs from The Cedars (and unbinned contigs), indicating that these bacteria can neither perform H_2/CO_2-driven nor formate-oxidizing acetogenesis (Fig. 2c). However, even without the formate dehydrogenase, the genes present can form a coherent pathway that uses formate rather than CO_2 as a starting point for the "methyl branch" of the Wood–Ljungdahl pathway (i.e., formate serves as an electron acceptor; Fig. 2c). This is a simple yet potentially effective strategy for performing homoacetogenesis while circumventing the unfavorable reduction of CO_2 to formate. Coupling H_2 oxidation with this formate-reducing pathway is thermodynamically viable as it halves the usage of CO_2 ($3H_2 + Formate^- + CO_2 = Acetate^- + 2H_2O$; ΔG of -29.62 kJ per mol acetate) and, as a pathway, is simply an intersection between the conventional H_2/CO_2-driven and formate-disproportionating acetogenesis (Fig. 2c and S5). Although use of formate as an electron acceptor for formate-oxidizing acetogenesis is quite common, no previous homoacetogens have been observed to couple H_2

Fig. 3 Evolution and distribution of glycine reductases. a Phylogeny of serpentinite-hosted microbiome glycine reductase subunit GrdBE homologs (Hakuba Happo hot spring*, The Cedars springs†, and other serpentinite-hosted system metagenomes#) and a brief scheme for evolutionary history of Grd. Grd-related COG1978 homologs were collected from the representative species genomes in GTDB, filtered using a GrdB motif conserved across members of phyla known to perform glycine-reducing Stickland reaction (see Methods and Supplementary Fig. S6) and clustered with 75% amino acid sequence similarity using CD-HIT (-c 0.75). GrdB-related sarcosine reductase subunits were excluded by identification of a GrdF motif conserved across sequences that form a distinct cluster around the biochemically characterized *Peptoclostridium acidaminophilum* GrdF. GrdE neighboring GrdB were collected. D-proline reductase subunits PrdBA (homologous to GrdB and GrdE respectively) was used as an outgroup. GrdB+PrdB and GrdE+PrdA were aligned (MAFFT v7.394) and trimmed (BMGE v1.12 -m BLOSUM30 -g 0.05) separately, then concatenated. A maximum likelihood tree was calculated using IQ-TREE v2.1.3 (-m LG+C20+G+F) and 1000 ultrafast bootstrap replicates (bootstrap values are recalculated with BOOSTER). Branches with ≥95% ultrafast bootstrap support are indicated with pink circles. Serpentinite-hosted system-derived sequences are shown in blue and taxa that may have gained GrdB through horizontal transfer are shown in green. Though the GrdB motif did not match, the closest (and only) detectable archaeal homolog (COG1978) identified in *Ca*. Bathyarchaeota is included. An axis break is used for the branch connecting GrdBE (and the *Ca*. Bathyarchaeota homolog) and outlier PrdBA for readability (10% of actual length). See Supplementary Fig. S6 for complete tree and full branch length between GrdBE and PrdBA. In the brief scheme of Grd evolution (top left), the cladogram topology is based on Fig. S4. Vertical transfer (red lines in cladogram) and horizontal transfer (black arrows) inferred from tree structures are shown. Phyla that may have acquired Grd vertically (red) and horizontally (gray) are indicated. GTDB phyla belonging to *Firmicutes* were grouped together. * GTDB-defined phylum-level lineage nomenclature. **b** Number of glycine reductase-encoding GTDB-defined species representatives (GTDB r95) associated with different environments. Only genomes with both GrdB and GrdE were included.

oxidation with acetogenesis from formate, likely because CO_2 has a much higher availability than formate in most ecosystems.

The Hakuba-inhabiting "*Ca*. Lithacetigenota" (HKB210 and HKB111) also encodes Hox for H_2 oxidation but lacks genes for homoacetogenesis (no homologs closely related to The Cedars population genes were detected even in unbinned metagenomic contigs). We suspect that this population forgoes the above H_2/formate-driven homoacetogenesis because the estimated energy yield of the net reaction in situ (ΔG of −19.94 kJ per mol acetate) is extremely close to the thermodynamic threshold of microbial catabolism (slightly above −20 kJ per mol) and, depending on the actual threshold for "*Ca*. Lithacetigenota" and/or even slight changes in the surrounding conditions (e.g., ΔG increases by 1 kJ per mol if H_2 decreases by 20 µM decreases in Hakuba), the metabolism may be unable to recover energy. Through searching the physicochemical environment for alternative exogenous electron acceptors and MAGs for electron-disposing pathways, we detected a low concentration of glycine in situ (5.4 ± 1.6 nM; Table S6) and found genes specific to catabolic glycine reduction (see next paragraph). We suspect that some portion of this glycine is likely geochemically generated in situ, given that (a) glycine is often detected as the most abundant amino acid produced by both natural and laboratory-based serpentinization (e.g., H_2 + Formate = Formaldehyde ⇒ Formaldehyde + NH_3 = Glycine) [10, 16, 42–47] and (b) no other amino acid was consistently detectable (if glycine was cell-derived, other amino acids ought to also be consistently detected).

For utilization of the putatively abiotic glycine, the Hakuba "*Ca*. Lithacetigenota" encodes glycine reductases (Grd; Fig. 3 and S6; Tables S4 and S5)—a unidirectional selenoprotein for catabolic glycine reduction [48, 49]. Based on the genes available, this population likely specializes in coupling H_2 oxidation and glycine reduction (H_2 + Glycine$^-$ → Acetate$^-$ + NH_3; Fig. 2c). Firstly, the genomes encode NADP-linked thioredoxin reductases (NADPH + $Thioredoxin_{ox}$ → $NADP^+$ + $Thioredoxin_{red}$) that can bridge electron transfer from H_2 oxidation (H_2 + NADP → NADPH + H^+) to glycine reduction (Glycine$^-$ + $Thioredoxin_{red}$ → Acetyl-$_{Pi}$ + NH_3 + $Thioredoxin_{ox}$). Secondly, though glycine reduction is typically coupled with amino acid oxidation (*i.e*., Stickland reaction in *Firmicutes* and *Synergistetes* [48, 50]), similar metabolic couplings have been reported for some organisms (i.e., formate-oxidizing glycine reduction [via Grd] [51] and H_2-oxidizing trimethylglycine reduction [via Grd-related betaine reductase] [52]). Thirdly, Grd is a rare catabolic enzyme, so far found in organisms that specialize in amino acid (or peptide) catabolism, many of which are reported to use glycine for the Stickland reaction (e.g., *Peptoclostridium* of *Firmicutes* and *Aminobacterium* of *Synergistetes* [53]). Lastly, the population lacks any discernable fermentative (propionate [methylmalonyl-CoA pathway], butyrate [reverse beta oxidation], lactate [lactate dehydrogenase], and alanine [alanine dehydrogenase]) and respiratory (aerobic [terminal oxidases], nitrate [nitrate reductase, nitrite reductase, nitric oxide reductase, nitrous oxide reductase], sulfate [dissimilatory sulfate reductase and sulfite reductase], other sulfurous compounds [molybdopterin-binding protein family sulfurous compound reductases], and metals [outer membrane cytochrome OmcB]) electron disposal pathways and oxidative organotrophy (Tables S4 and S5). Although the BS5B28 genome encodes a bifunctional alcohol/aldehyde dehydrogenase and aldehyde:ferredoxin oxidoreductase, no complete sugar or amino acid degradation pathways could be identified, suggesting that these genes have a physiological role unrelated to ethanol fermentation. Further, though formate and glycine transporters were absent in the genomes, a survey of transporters (annotated in UniProtKB 2021_03 [54]) revealed that no alkaliphiles (organisms with optimum pH ≥ 9.5 in the DSMZ BacDive database [55]) encoded known formate transporters (focA; TIGR04060) or amino acid permeases (PF00324) (ABC transporters were not considered as substrate specificity for these complexes cannot be annotated reliably), indicating that alkaliphiles likely employ unknown transport proteins. Reflecting the lack of other catabolic pathways, the Hakuba "*Ca*. Lithacetigenota" MAGs display extensive genome streamlining, comparable to that of *Aurantimicrobium* [56, 57], "*Ca*. Pelagibacter" [58], and *Rhodoluna* [59] in aquatic systems, as also reported for other organisms inhabiting serpentinite-hosted systems [60, 61] (Fig. S7). Thermodynamic calculations show that H_2-oxidizing glycine reduction is favorable in situ ($\Delta G^{\circ\prime}$ of −70.37 kJ per mol glycine [ΔG of −85.84 in Hakuba]; Fig. S1). Further, based on the pathway identified, this putative metabolism is >10 times more efficient in recovering energy from H_2 (1 mol ATP per mol H_2) than acetogenesis utilizing H_2/CO_2 (0.075 mol ATP per mol H_2 based on the pathway *Acetobacterium woodii* utilizes) or H_2/formate (0.075 mol ATP per mol H_2, assuming no energy recovery associated with the formate dehydrogenase). We also detect glycine reductases in The Cedars "*Ca*. Lithacetigenota", indicating that it may also perform this metabolism (ΔG of −76.87 in The Cedars, assuming 201 µM H_2).

Given the phylogenetic and metabolic uniqueness of these populations, we report provisional taxonomic assignment to "*Ca*. Lithacetigenota" phyl. nov., "*Ca*. Lithacetigena glycinireducens" gen. nov., sp. nov. (HKB111 and HKB210), and "*Ca*. Psychracetigena formicireducens" gen. nov., sp. nov. (BS525, BS5B28, and GPS1B18) (see Supplementary Results). Based on a concatenated ribosomal protein tree, this serpentinite-hosted ecosystem-associated

candidate phylum is closely related to the deepest-branching group of bacterial phyla in "Terrabacteria", one of the two major of lineages *Bacteria* (Fig. 1). Comparative genomics shows that "*Ca*. Lithacetigenota" shares 623 core functions (based on Bacteria-level COGs/NOGs predicted by eggnog-mapper shared by the two highest quality Hakuba and The Cedars MAGs HKB210 and BS5B28; Fig. 2d). When compared with the core functions of two closest related phyla (*Caldiserica* and *Coprothermobacterota*), 176 functions were unique to "*Ca*. Lithacetigenota", including those for NiFe hydrogenases (and their maturation proteins), selenocysteine utilization (essential for Grd), and sodium:proton antiporter for alkaliphily. With *Coprothermobacterota*, 232 functions were shared, including Grd, thioredoxin oxidoreductase (essential for electron transfer to Grd), and additional proteins for NiFe hydrogenases and selenocysteine utilization, pointing toward importance of H_2 metabolism and glycine reduction for these closely related phyla. More importantly, among bacterial phyla in the deep-branching group, "*Ca*. Lithacetigenota" represents the first lineage inhabiting hyperalkaliphilic serpentinite-hosted ecosystems, suggesting that these organisms may be valuable extant windows into potential physiologies of primordial organisms who are thought to have lived under hyperalkaline conditions (albeit with 4 billion years of evolution in between; see discussion regarding Grd below).

Widespread glycine reduction in serpentinite-hosted systems

Uncultured members of *Chloroflexi* (Chloroflexota) class *Dehalococcoidia* inhabiting The Cedars and *Firmicutes* (Firmicutes_D) class SRB2 in Hakuba and The Cedars also possess glycine reductases (Table S5). In addition, these populations encode hydrogenases and formate dehydrogenases, suggesting that they may also link H_2 and formate metabolism to glycine reduction. Closely related glycine reductases were also detected in other studied serpentinite-hosted systems (47–94% amino acid similarity in Tablelands, Voltri Massif, and Coast Range Ophiolite) [1, 2, 7]. Phylogenetic analysis of the glycine-binding "protein B" subunits GrdB and GrdE reveals close evolutionary relationships between glycine reductases from distant/remote sites (Fig. 3a and S6). Note that Tablelands spring glycine reductase sequences were not included in the analysis as they were only detected in the unassembled metagenomic reads (4460690.3; 69.7–82.2% similarity to Hakuba SRB2). Overall, "*Ca*. Lithacetigenota", *Dehalococcoidia*, and SRB2 glycine reductases are all detected in at least two out of the seven metagenomically investigated systems despite the diverse environmental conditions (e.g., temperature). Thus, we propose glycine as an overlooked thermodynamically and energetically favorable electron acceptor for H_2 oxidation in serpentinite-hosted systems. We suspect that glycine reduction may be a valuable catabolic strategy as the pathway requires few genes/proteins (a hydrogenase, Grd, acetate kinase, and thioredoxin oxidoreductase) and conveniently provides acetate, ammonia, and ATP as basic forms of carbon, nitrogen, and energy.

Phylogenetic analysis of glycine reductases (Fig. 3a and S6) shows that the novel homologs recovered from serpentinite-hosted systems represent deep-branching lineages distantly related from those detectable in published genomes (GTDB r95 species representatives). Further comparison of the topology with a ribosomal protein-based genome tree (Fig. 2a) indicates that the two deep-branching serpentinite-hosted system-affiliated lineages (*Ca*. Lithacetigenota and novel *Chloroflexi* family) and *Firmicutes* vertically inherited glycine reductases. Thus, catabolic glycine reduction can be traced back to the concestor of these three lineages, suggesting the metabolism at least dates back to the ancestor of "Terrabacteria". We further identified an archaeal GRD homolog (in Miscellaneous Crenarchaeota Group [MCG] or *Ca*. Bathyarchaeota member BA-1; Fig. 3a and S6), but whether this gene functions as a glycine reductase (GrdB motif not fully conserved) and, further, truly belong to this clade (source is metagenome-assembled genome) remains to be verified. Reconstruction of the ancestral Grd sequence and estimation of its pH preference (via AcalPred) showed that the ancestral enzyme likely had good efficiency under alkaline conditions (pH > 9; probability of 0.9973 and 0.9858 for GrdB and GrdE, respectively). Thus, the currently available data suggest that Grd (and catabolic glycine reduction) is an ancient bacterial catabolic innovation in an alkaline habitat, dating back to one of the deepest nodes in the bacterial tree.

While we detect glycine reductases in many serpentinite-hosted systems, examination of genomes derived from other natural ecosystems shows that only 107 species (species representatives in GTDB r95; 0.35% of all GTDB species) inhabiting such habitats encode GrdBE (Fig. 3b and S6). This is a level comparable to rare artificial contaminant-degrading enzymes (*e.g.*, tetrachloroethane dehalogenase pceA—65 species [encoding KEGG KO K21647 based on AnnoTree with GTDB r95 and default settings [62]]; dibenzofuran dioxygenase—258 species [K14599 and K14600]). Most glycine reductase homologs are found in species affiliated with host-associated (mostly human body and rumen) or artificial habitats (360 species), the majority of which belong to the phylum *Firmicutes*. We suspect that glycine reduction has low utility in most natural ecosystems (e.g., no excess glycine via abiotic generation and no severe nutrient/electron acceptor limitation) and has been repurposed by some anaerobes for the fermentative Stickland reaction in organic-rich ecosystems (e.g., host-associated ecosystems) where excess amino acids are available but access to favorable electron acceptors is limited (Fig. 3b) (notably, glycine is the dominant amino acid in collagen [>30%], the most abundant protein in vertebrate bodies).

Other characteristics of putative indigenous homoacetogens

In contrast with members of "*Ca*. Lithacetigenota", several other putative homoacetogenic populations encode the complete Wood–Ljundgahl pathway (Tables S4 and S5), indicating that other forms of acetogenesis may also be viable in situ. One putative homoacetogen in The Cedars, NPL-UPA2, lacks hydrogenases but encodes formate dehydrogenases. Although the NPL-UPA2 population cannot perform H_2/formate-driven acetogenesis, it may couple formate oxidation with formate-reducing acetogenesis—another thermodynamically viable metabolism (ΔG of −50.90 kJ per mol acetate in The Cedars; Fig. S5). The pathway uses CO_2 as a substrate but has lower CO_2 consumption compared to H_2/CO_2 homoacetogenesis and can produce intracellular CO_2 from formate. A recent study also points out that methanogens inhabiting serpentinite-hosted environments oxidize formate presumably to generate intracellular CO_2 [9]. In Hakuba, an *Actinobacteria* population affiliated with the uncultured class UBA1414 (MAG HKB206) encodes hydrogenases and a complete Wood–Ljungdahl pathway (Table S5) and, thus, may be capable of H_2/formate or the above formate-disproportionating acetogenesis (Fig. 2c and S5). Indeed, the UBA1414 population was enriched in Hakuba-derived cultures aiming to enrich acetogens using the H_2 generated by the metallic iron–water reaction [63] (Fig. S8). Many populations encoding a complete Wood–Ljungdahl pathway possess monomeric CO dehydrogenases (CooS unassociated with CODH/ACS subunits; NPL-UPA2, *Actinobacteria*, *Syntrophomonadaceae* [Hakuba and The Cedars], and *Dehalococcoidia* [The Cedars]; Table S4). Although CO is below the detection limit in Hakuba (personal communication with permission from Dr. Konomi Suda), another study shows that CO metabolism takes place in an actively serpentinizing system with no detectable CO [64]. Given that CO is a known product of serpentinization [7, 64], it may be an important substrate for thermodynamically favorable acetogenesis in situ. However, further investigation is necessary to verify this (e.g., need to measure CO at multiple time points).

Another interesting adaptation observed for all putative homoacetogens detected in Hakuba and The Cedars was possession of an unusual CODH/ACS complex. Although *Bacteria*

and *Archaea* are known to encode structurally distinct forms of CODH/ACS (designated as Acs and Cdh respectively for this study), all studied Hakuba/The Cedars putative homoacetogens encode genes for a hybrid CODH/ACS that integrate archaeal subunits for the CO dehydrogenase (AcsA replaced with CdhAB) and acetyl-CoA synthase (AcsB replaced with CdhC) and bacterial subunits for the corrinoid protein and methyltransferase components (AcsCDE) (Table S4). The *Firmicutes* lineages also additionally encode the conventional bacterial AcsABCDE. Given that all of the identified putative homoacetogens encode this peculiar hybrid complex, we suspect that such CODH/ACS's may have features adapted to the high-pH low-CO_2 conditions (e.g., high affinity for CO_2 and/or CO). In agreement, a similar hybrid CODH/ACS has also been found in the recently isolated "*Ca*. Desulforudis audaxviator" inhabiting an alkaline (pH 9.3) deep subsurface environment with a low CO_2 concentration (below detection limit [65, 66]) [67].

Implications for primordial biology

The last universal common ancestor (LUCA) is hypothesized to have evolved within alkaline hydrothermal mineral deposits at the interface of serpentinization-derived fluid and ambient water (e.g., Hadean weakly acidic seawater) [22–24]. Although such interfaces no longer exist (i.e., ancient Earth lacked O_2 but most water bodies contain O_2 on modern Earth), modern anoxic terrestrial and oceanic ecosystems harboring active serpentinization [1–8] may hold hints for how primordial organisms utilized H_2 under hyperalkaline CO_2-depleted conditions (e.g., post-LUCA H_2-utilizing organisms that ventured away from the interface towards the alkaline fluids). Our findings suggest that unconventional modes of lithotrophy that take advantage of geogenic reduced carbon compounds (e.g., formate and glycine) as exogenous electron acceptors may have been viable approaches to circumventing thermodynamic issues and obtaining energy from H_2 oxidation in situ. The strategies we discover are largely exclusive to the bacterial domain (archaeal CO_2 reduction does not involve formate as an intermediate and, to our knowledge, glycine reduction is limited to *Bacteria*) and originated deep in the bacterial tree, suggesting they may have been relevant in the divergence towards the bacterial and archaeal domains. Notably, the estimated alkaliphily of the ancestral Grd also points towards the relevance of this metabolism in ancient alkaline habitats.

CONCLUSION

Through the investigation of serpentinite-hosted systems, we discover a novel habitat-exclusive alkaliphilic phylum that belongs to a deep-branching group of bacterial phyla and likely relies on putative site-adapted H_2-oxidizing lithotrophy (e.g., coupled with formate or glycine reduction) thermodynamically favorable in the electron acceptor- and CO_2-depleted conditions in situ. The consistent presence of catabolic glycine reductases across these habitats also indicates that glycine may be an important electron acceptor, potentially abiotically generated via serpentinization. Moreover, the identified glycine reductases represent hitherto overlooked deep-branching lineages that point toward antiquity of catabolic and potentially alkaline glycine reduction, suggesting thermodynamic and phylogenetic relevance to ancient metabolism in serpentinite-hosted systems. Further investigation of microbiology in these habitats may reveal novel organisms, metabolic strategies, and rare enzymes adapted to polyextreme conditions in Earth's modern and ancient subsurface environments.

METHODS

Sampling site and sample collection

The Hakuba Happo samples for geochemical and microbiological analysis were artificially pumped from a drilling well (700 m in depth), which was previously described and named Happo #3 (36°42′N 137°48′E [27]). For microbiological analysis, two spring water samples were taken at different time points, 233 L taken in July 2016 (labeled HKB701) and 720 L taken in October 2016 (labeled HKB702), respectively. To collect microbial cells, samples were filtered through a 0.1-µm Omnipore membrane filter (Merck Millipore) using a 90 mm diameter stainless-steel filter holder (Merck Millipore) attached to FDA Viton tubing (Masterflex) at a sampling site. After filtration, filters were immediately transferred to sterile tubes and frozen in a dry ice-ethanol bath, transported in dry ice, and stored at −80 °C until DNA extraction. Only in October 2017, water samples for NH_3 and amino acid analysis were collected from the same well Happo #3, transferred to dry-heat-sterilized nitrogen-purged 100 ml glass vials, and stored at 4 °C.

Geochemical analysis

The water temperature of hot spring water was measured using a thermometer (CT-430WP, Custom Ltd.) at a site. The pH, oxidation reduction potential (ORP), electrical conductivity (EC), and dissolved oxygen (DO) level were determined with portable devices, including a pH meter (D-23, Horiba), an ORP meter (RM-30P, TOA-DKK), an EC meter (CM-31P, TOA-DKK), and a DO meter (DO-31P, TOA-DKK), correspondingly (note that the water sample was pumped up from underground and immediately used for these measurements before cooling). The ion concentrations of Na^+, K^+, Ca^{2+}, and NO_3^- were determined using portable sensors (LAQUAtwin series, Horiba). The in situ NH_3 concentration was determined by measuring aqueous NH_4^+ and gaseous NH_3 (purged with N_2 gas, gas dissolved into deionized water, and measured dissolved NH_4^+) of a sample stored as described above using high-performance liquid chromatography (HPLC; Prominence; Shimadzu), then adding the two together. For amino acid quantification, the sample was concentrated under a stream of nitrogen gas and then analyzed following Shimadzu protocol no. L323 (https://www.an.shimadzu.co.jp/hplc/prominence/l323.pdf) using HPLC with minor modifications (fluorescence detector RF-20Axs; sodium hypochlorite solution was not added for detection of proline). The Cedars spring concentrations of formate and acetate were determined by Isotope-Ratio-Monitoring Liquid Chromatography Mass Spectrometry; Thermo- Finnigan Delta Plus XP isotope-ratio mass spectrometer connected to LC IsoLink, as described by Heuer et al. [68] and Ijiri et al. [69].

Thermodynamic calculations

The Hakuba and Cedars calculations Gibbs free energy yield (ΔG) are based on $\Delta G°_f$ and $\Delta H°_f$ values at 298 K, respective pH (10.7 and 11.9), and adjustment to the in situ temperatures (48 and 17 °C) through the Gibbs–Helmholtz equation [70]. The effect of pressure was approximated as described by Wang et al. [71]. For both Hakuba and The Cedars calculations, the glycine concentration (5.4 nM) was based on measurements from Hakuba. Formate, acetate, and NH_3 concentrations were based on respective measurements from Hakuba (8 µM formate, 4 µM acetate, and 2.9 µM NH_3) and The Cedars (6.9 µM formate [average of 6.777 and 7.079 µM measured on September 2017], 69.3 µM acetate [average of 69.601 and 68.967 µM measured on September 2017], and 1 µM NH_3 [below detection limit]). For Hakuba, the H_2 concentration measured in Hakuba drilling well #3 (DNA source) was used (201 µM H_2). For The Cedars, the highest detected H_2 concentration in Hakuba was used (664 µM H_2 in drilling well #1). See also Tables S1 and S2 and Supplementary Equations.

Metagenome sequencing, assembly, and binning

The filter was aseptically cut into 16 equal pieces using sterilized tweezers, and each piece was placed in the bead-beating tube (Lysing Matrix E tube; MP Biomedicals). After DNA extraction following the bead-beating method described previously [72], the 16 DNA samples were mixed and then stored at −80 °C until used. Sequence libraries were prepared with Nextera XT DNA Library Preparation kit (Illumina) with a genomic DNA fragment size ranging from 200 to 2000 bp. These libraries sequenced on HiSeq2500 sequencing platform (Illumina) with HiSeq Rapid SBS kit v2 (Illumina), generating paired-end reads up to 250 bp. The generated sequences were trimmed using Trimmomatic v0.33 [73] with a quality cutoff of 30, sliding window of 6 bp, and minimum length cutoff of 78 bp. The trimmed sequences were assembled using SPAdes v3.10.1 [74] with the "-meta" option and k-mer values of 21, 31, 41, 53, 65, and 77. The assembled contigs were binned using MaxBin2.2.1 [75, 76]. The completeness and contamination of each bin were checked using CheckM [77]. These bins were manually curated as described in our metagenomics study [40]. Genes were then annotated using Prokka v1.12 [78] and eggnog-

mapper [79]. For interpretation and comparison of microbial metabolism, bin genomes were also constructed from public metagenomic data generated from The Cedars [3] (trimmed with sliding window of 6, quality cutoff of 20, and a minimum length of 68 bp through Trimmomatic v0.33, normalized using BBMap 36.99 (https://jgi.doe.gov/data-and-tools/bbtools/) with target and minimum coverages of 40 and 2, assembled using SPAdes v3.10.1 with the "-meta" option and k-mer values of 21, 33, 45, 55, 67, and binned through MaxBin2.2.1) and were then analyzed collectively.

Phylogenomic and phylogenetic analysis

For tree construction, sequences were aligned with MAFFT [80] v7.453 (default parameters) and trimmed using trimAl [81] v1.2rev59 (-gt 0.9) or BMGE v1.12 (-m BLOSUM30) [82]. For ribosomal protein trees, a concatenated alignment of universally conserved ribosomal proteins [83] was used. Protein sequences were retrieved by downloading the GTDB [84] database and predicting protein sequences using Prokka [78] 1.14 (-kingdom Bacteria/Archaea -rnammer -addgenes -mincontiglen 200). Maximum likelihood trees were calculated using IQ-TREE v2.1.3 using the LG [85] model and C20 mixture model with 1000 ultrafast bootstrap replicates (-m LG+C20+G+F -B 1000 -tbe). For ribosomal protein tree calculation with IQ-TREE, a universal distribution mixture (UDM) model with 64 components and LCLR transformation constructed based on the HOGENOM and HSSP databases (-m Poisson+UDM0064LCLR) [86], ultrafast bootstrap approximation (-B 1000), and SH-like approximate likelihood ratio test (-alrt 1000) were used instead. The number of components for the mixture model was selected as the maximum number of components that did not result in estimation of mixture weights close to zero, as indicated by IQ-TREE. The UDM model was chosen over the C60 mixture model often used in ribosomal protein tree calculation as it has been shown to have improved model fit and performance [86, 87]. The model that integrates both the HOGENOM and HSSP databases was used. Bootstrap values were recalculated using BOOSTER [88] (-tbe for IQ-TREE).

For glycine reductase phylogeny, sequence clustering was performing through CD-HIT [89] v4.8.1. For glycine reductase GrdB and sarcosine reductase GrdF, conserved motifs were predicted by first identifying fully conserved residues in the sequence cluster including the biochemically characterized *Peptoclostridium acidaminophilum* GrdF ((YxNx(6)GGE x (34,38) CGD x(27,35) GPxF[NF]AGRYG x(150,181) IHGGYDRx(6)[IP]x(4)PxD x(19,20) TTGTGTx(7)F x(12) [HILV])), then identifying fully conserved residues in the phylogenetic clusters that include GrdB from phyla known to perform the Stickland reaction (*Firmicutes*, *Spirochaetes*, and *Synergistetes*) subtracting any sequences clusters that contain the GrdF motif above (YxNx(6)GGE x(34,38) CGD x(27,35) GPxF[NF]AGRYG x(157,178) AHGGxD[QTAP] x(8) RV[IL]PxD x(19,20) TxGNxTxV)). Ancestral sequence reconstruction for GrdBE was performed using IQ-TREE (-asr), an alignment of glycine reductase sequences included in the concatenated tree (including one archaeal sequence that had mismatches and excluding proline reductases) (MAFFT with default parameters), and a tree rooted with the archaeal GrdBE-like sequence and D-proline reductase subunits PrdBA sequences as the outgroups. The pH preference of the reconstructed sequence was estimated using AcalPred [90].

DATA AVAILABILITY

The datasets generated during and/or analyzed during the current study are available in the National Center for Biotechnology Information (NCBI) under BioProject PRJNA453100 and BioSamples SAMN08978938-SAMN08978962.

REFERENCES

1. Brazelton WJ, Thornton CN, Hyer A, Twing KI, Longino AA, Lang SQ, et al. Metagenomic identification of active methanogens and methanotrophs in serpentinite springs of the Voltri Massif, Italy. PeerJ. 2017;5:e2945.
2. Brazelton WJ, Nelson B, Schrenk MO. Metagenomic evidence for H_2 oxidation and H_2 production by serpentinite-hosted subsurface microbial communities. Front Microbiol. 2011;2:268.
3. Suzuki S, Ishii S, Hoshino T, Rietze A, Tenney A, Morrill PL, et al. Unusual metabolic diversity of hyperalkaliphilic microbial communities associated with subterranean serpentinization at The Cedars. ISME J. 2017;11:2584–98.
4. Kelley DS, Karson JA, Früh-Green GL, Yoerger DR, Shank TM, Butterfield DA, et al. A serpentinite-hosted ecosystem: the lost city hydrothermal field. Science. 2005;307:1428–34.
5. Tiago I, Verissimo A. Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinization. Environ Microbiol. 2013;15:1687–706.
6. Crespo-Medina M, Twing KI, Sánchez-Murillo R, Brazelton WJ, McCollom TM, Schrenk MO. Methane dynamics in a tropical serpentinizing environment: The Santa Elena Ophiolite, Costa Rica. Front Microbiol. 2017;8:916.
7. Twing KI, Brazelton WJ, Kubo MDY, Hyer AJ, Cardace D, Hoehler TM, et al. Serpentinization-influenced groundwater harbors extremely low diversity microbial communities adapted to high pH. Front Microbiol. 2017;8:308.
8. Neubeck A, Sun L, Müller B, Ivarsson M, Hosgörmez H, Özcan D, et al. Microbial community structure in a serpentine-hosted abiotic gas seepage at the Chimaera Ophiolite, Turkey. Appl Environ Microbiol. 2017;83:e03430–16.
9. Fones EM, Colman DR, Kraus EA, Stepanauskas R, Templeton AS, Spear JR, et al. Diversification of methanogens into hyperalkaline serpentinizing environments through adaptations to minimize oxidant limitation. ISME J. 2021;15:1121–35.
10. Schrenk MO, Brazelton WJ, Lang SQ. Serpentinization, carbon, and deep life. Rev Miner Geochem. 2013;75:575–606.
11. Komiya T, Maruyama S, Hirata T, Yurimoto H, Nohda S. Geochemistry of the oldest MORB and OIB in the Isua Supracrustal Belt, southern West Greenland: Implications for the composition and temperature of early Archean upper mantle. Isl Arc. 2004;13:47–72.
12. Sleep NH, Bird DK, Pope EC. Serpentinite and the dawn of life. Philos Trans R Soc Lond B Biol Sci. 2011;366:2857.
13. Martin WF, Sousa FL. Early microbial evolution: the age of anaerobes. Cold Spring Harb Perspect Biol. 2016;8:a018127.
14. Lane N, Martin, William F. The origin of membrane bioenergetics. Cell 2012;151:1406–16.
15. Battistuzzi FU, Feijao A, Hedges SB. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. BMC Evol Biol. 2004;4:44.
16. Shibuya T, Yoshizaki M, Sato M, Shimizu K, Nakamura K, Omori S, et al. Hydrogen-rich hydrothermal environments in the Hadean ocean inferred from serpentinization of komatiites at 300 °C and 500 bar. Prog Earth Planet Sci. 2015;2:46.
17. Camprubi E, Jordan SF, Vasiliadou R, Lane N. Iron catalysis at the origin of life. IUBMB life. 2017;69:373–81.
18. Decker K, Jungermann K, Thauer RK. Energy production in anaerobic organisms. Angew Chem Int Ed Engl. 1970;9:138–58.
19. Martin WF, Weiss MC, Neukirchen S, Nelson-Sathi S, Sousa FL. Physiology, phylogeny, and LUCA. MicrobialCell 2016;3:582–7.
20. Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M, Nelson-Sathi S, et al. The physiology and habitat of the last universal common ancestor. Nat Microbiol. 2016;1:16116.
21. Adam PS, Borrel G, Gribaldo S. Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. Proc Natl Acad Sci. 2018;115:E1166.
22. Lane N. Proton gradients at the origin of life. Bioessays. 2017;39:1600217.
23. Shibuya T, Russell MJ, Takai K. Free energy distribution and hydrothermal mineral precipitation in Hadean submarine alkaline vent systems: Importance of iron redox reactions under anoxic conditions. Geochim Cosmochim Acta. 2016;175:1–19.
24. Sojo V, Herschy B, Whicher A, Camprubí E, Lane N. The origin of life in alkaline hydrothermal vents. Astrobiology. 2016;16:181–97.
25. Morrill PL, Kuenen JG, Johnson OJ, Suzuki S, Rietze A, Sessions AL, et al. Geochemistry and geobiology of a present-day serpentinization site in California: The Cedars. Geochim Cosmochim Acta. 2013;109:222–40.
26. Suda K, Gilbert A, Yamada K, Yoshida N, Ueno Y. Compound– and position-specific carbon isotopic signatures of abiogenic hydrocarbons from on-land serpentinite-hosted Hakuba Happo hot spring in Japan. Geochim Cosmochim Acta. 2017;206:201–15.
27. Suda K, Ueno Y, Yoshizaki M, Nakamura H, Kurokawa K, Nishiyama E, et al. Origin of methane in serpentinite-hosted hydrothermal systems: the CH_4–H_2–H_2O hydrogen isotope systematics of the Hakuba Happo hot spring. Earth Planet Sci Lett. 2014;386:112–25.
28. Suda K. Origins of hydrocarbons in on-land serpentinization fields and insights into hadean hydrothermal systems: systematic study using stable isotopes. Doctor thesis, Tokyo Institute of Technology. 2016.
29. Schink B. Energetics of syntrophic cooperation in methanogenic degradation. Microbiol Mol Biol Rev. 1997;61:262–80.
30. Preiner M, Igarashi K, Muchowska KB, Yu M, Varma SJ, Kleinermanns K, et al. A hydrogen-dependent geochemical analogue of primordial carbon and energy metabolism. Nat Ecol Evol. 2020;4:534–42.
31. Suzuki S, Nealson KH, Ishii S. Genomic and in-situ transcriptomic characterization of the candidate phylum NPL-UPL2 from highly alkaline highly reducing serpentinized groundwater. Front Microbiol. 2018;9:3141.
32. Morandi P, Valzasina B, Colombo C, Curti B, Vanoni MA. Glutamate synthase: identification of the NADPH-binding site by site-directed mutagenesis. Biochemistry. 2000;39:727–35.
33. Schneider K, Schlegel HG. Purification and properties of soluble hydrogenase from Alcaligenes eutrophus H 16. Biochim Biophys Acta. 1976;452:66–80.

34. Burgdorf T, van der Linden E, Bernhard M, Yuan Yin Q, Back JW, Hartog AF, et al. The soluble NAD^+-reducing [NiFe]-hydrogenase from *Ralstonia eutropha* H16 consists of six subunits and can be specifically activated by NADPH. J Bacteriol. 2005;187:3122–32.
35. de Luca G, de Philip P, Rousset M, Belaich JP, Dermoun Z. The NADP-reducing hydrogenase of *Desulfovibrio fructosovorans*: evidence for a native complex with hydrogen-dependent methyl-viologen-reducing activity. Biochem Biophys Res Commun. 1998;248:591–6.
36. Schut GJ, Adams MW. The iron-hydrogenase of *Thermotoga maritima* utilizes ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen production. J Bacteriol. 2009;191:4451–7.
37. de Bok FA, Hagedoorn PL, Silva PJ, Hagen WR, Schiltz E, Fritsche K, et al. Two W-containing formate dehydrogenases (CO_2-reductases) involved in syntrophic propionate oxidation by *Syntrophobacter fumaroxidans*. Eur J Biochem. 2003;270:2476–85.
38. Yamamoto I, Saiki T, Liu SM, Ljungdahl LG. Purification and properties of NADP-dependent formate dehydrogenase from *Clostridium thermoaceticum*, a tungsten-selenium-iron protein. J Biol Chem. 1983;258:1826–32.
39. Hidalgo-Ahumada CAP, Nobu MK, Narihiro T, Tamaki H, Liu WT, Kamagata Y, et al. Novel energy conservation strategies and behaviour of *Pelotomaculum schinkii* driving syntrophic propionate catabolism. Environ Microbiol. 2018;20:4503–11.
40. Nobu MK, Narihiro T, Rinke C, Kamagata Y, Tringe SG, Woyke T, et al. Microbial dark matter ecogenomics reveals complex synergistic networks in a methanogenic bioreactor. ISME J. 2015;9:1710–22.
41. Poehlein A, Schmidt S, Kaster AK, Goenrich M, Vollmers J, Thurmer A, et al. An ancient pathway combining carbon dioxide fixation with the generation and utilization of a sodium ion gradient for ATP synthesis. PLoS One. 2012;7:e33439.
42. Fuchida S, Mizuno Y, Masuda H, Toki T, Makita H. Concentrations and distributions of amino acids in black and white smoker fluids at temperatures over 200 °C. Org Geochem. 2014;66:98–106.
43. Haberstroh PR, Karl DM. Dissolved free amino acids in hydrothermal vent habitats of the Guaymas Basin. Geochim Cosmochim Acta. 1989;53:2937–45.
44. Svensson E, Skoog A, Amend JP. Concentration and distribution of dissolved amino acids in a shallow hydrothermal system, Vulcano Island (Italy). Org Geochem. 2004;35:1001–14.
45. Fox SW, Windsor CR. Synthesis of amino acids by the heating of formaldehyde and ammonia. Science. 1970;170:984–6.
46. Islam MN, Kaneko T, Kobayashi K. Determination of amino acids formed in a supercritical water flow reactor simulating submarine hydrothermal systems. Anal Sci. 2001;17:1631–34.
47. Inaba S. Primary formation path of formaldehyde in hydrothermal vents. Orig Life Evol Biosph. 2018;48:1–22.
48. Andreesen JR. Glycine reductase mechanism. Curr Opin Chem Biol. 2004;8:454–61.
49. Andreesen JR. Acetate via glycine: a different form of acetogenesis. In: Drake HL, (ed). Acetogenesis. Springer US, Boston, MA. 1994; pp.568–629.
50. Nisman B. The Stickland reaction. Bacteriol Rev. 1954;18:16–42.
51. Hormann K, Andreesen JR. Reductive cleavage of sarcosine and betaine by *Eubacterium acidaminophilum* via enzyme systems different from glycine reductase. Arch Microbiol. 1989;153:50–9.
52. Moune S, Manac'h N, Hirschler A, Caumette P, Willison JC, Matheron R. *Haloanaerobacter salinarius* sp. nov., a novel halophilic fermentative bacterium that reduces glycine-betaine to trimethylamine with hydrogen or serine as electron donors; emendation of the genus *Haloanaerobacter*. Int J Syst Bacteriol. 1999;49:103–12.
53. Hamdi O, Ben Hania W, Postec A, Bouallagui H, Hamdi M, Bonin P, et al. *Aminobacterium thunnarium* sp. nov., a mesophilic, amino acid-degrading bacterium isolated from an anaerobic sludge digester, pertaining to the phylum Synergistetes. Int J Syst Evol Microbiol. 2015;65:609–14.
54. The UniProt C. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 2021;49:D480–D9.
55. Reimer LC, Vetcininova A, Carbasse JS, Söhngen C, Gleim D, Ebeling C, et al. BacDive in 2019: bacterial phenotypic data for High-throughput biodiversity analysis. Nucleic Acids Res. 2019;47:D631–D6.
56. Nakai R, Baba T, Niki H, Nishijima M, Naganuma T. *Aurantimicrobium minutum* gen. nov., sp. nov., a novel ultramicrobacterium of the family *Microbacteriaceae*, isolated from river water. Int J Syst Evol Microbiol. 2015;65:4072–9.
57. Nakai R, Fujisawa T, Nakamura Y, Nishide H, Uchiyama I, Baba T, et al. Complete genome sequence of *Aurantimicrobium minutum* type strain KNC^T, a planktonic ultramicrobacterium isolated from river water. Genome Announc. 2016;4:e00616–16.
58. Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, et al. Genome streamlining in a cosmopolitan oceanic bacterium. Science. 2005;309:1242–5.
59. Hahn MW, Schmidt J, Taipale SJ, Doolittle WF, Koll U. *Rhodoluna lacicola* gen. nov., sp. nov., a planktonic freshwater bacterium with stream-lined genome. Int J Syst Evol Microbiol. 2014;64:3254–63.
60. Suzuki S, Kuenen JG, Schipper K, van der Velde S, Ishii SI, Wu A, et al. Physiological and genomic features of highly alkaliphilic hydrogen-utilizing *Betaproteobacteria* from a continental serpentinizing site. Nat Commun. 2014;5:3900.
61. Fones EM, Colman DR, Kraus EA, Nothaft DB, Poudel S, Rempfert KR, et al. Physiological adaptations to serpentinization in the Samail Ophiolite, Oman. ISME J. 2019;13:1750–62.
62. Mendler K, Chen H, Parks DH, Lobb B, Hug LA, Doxey AC. AnnoTree: visualization and exploration of a functionally annotated microbial tree of life. Nucleic Acids Res. 2019;47:4442–8.
63. Kato S, Yumoto I, Kamagata Y. Isolation of acetogenic bacteria that induce biocorrosion by utilizing metallic iron as the sole electron donor. Appl Environ Microbiol. 2015;81:67–73.
64. Morrill PL, Brazelton WJ, Kohl L, Rietze A, Miles SM, Kavanagh H, et al. Investigations of potential microbial methanogenic and carbon monoxide utilization pathways in ultra-basic reducing springs associated with present-day continental serpentinization: the Tablelands, NL, CAN. Front Microbiol. 2014;5:613.
65. Chivian D, Brodie EL, Alm EJ, Culley DE, Dehal PS, DeSantis TZ, et al. Environmental genomics reveals a single-species ecosystem deep within Earth. Science. 2008;322:275–8.
66. Lin L-H, Wang P-L, Rumble D, Lippmann-Pipke J, Boice E, Pratt LM, et al. Long-term sustainability of a high-energy, low-diversity crustal biome. Science 2006;314:479.
67. Karnachuk OV, Frank YA, Lukina AP, Kadnikov VV, Beletsky AV, Mardanov AV, et al. Domestication of previously uncultivated *Candidatus* Desulforudis audaxviator from a deep aquifer in Siberia sheds light on its physiology and evolution. ISME J. 2019;13:1947–59.
68. Heuer V, Elvert M, Tille S, Krummen M, Mollar XP, Hmelo LR, et al. Online $\delta^{13}C$ analysis of volatile fatty acids in sediment/porewater systems by liquid chromatography-isotope ratio mass spectrometry. Limnol Oceanogr Methods. 2006;4:346–57.
69. Ijiri A, Harada N, Hirota A, Tsunogai U, Ogawa NO, Itaki T, et al. Biogeochemical processes involving acetate in sub-seafloor sediments from the Bering Sea shelf break. Org Geochem. 2012;48:47–55.
70. Hanselmann KW. Microbial energetics applied to waste repositories. Experientia. 1991;47:645–87.
71. Wang G, Spivack AJ, D'Hondt S. Gibbs energies of reaction and microbial mutualism in anaerobic deep subseafloor sediments of ODP Site 1226. Geochim Cosmochim Acta. 2010;74:3938–47.
72. Mason OU, Hazen TC, Borglin S, Chain PS, Dubinsky EA, Fortney JL, et al. Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater Horizon oil spill. ISME J. 2012;6:1715–27.
73. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
74. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77.
75. Wu Y-W, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics. 2016;32:605–7.
76. Wu YW, Tang YH, Tringe SG, Simmons BA, Singer SW. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. Microbiome. 2014;2:26.
77. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25:1043–55.
78. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.
79. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, et al. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res. 2016;44:D286–D93.
80. Katoh K, Kuma K-I, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 2005;33:511–8.
81. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3.
82. Criscuolo A, Gribaldo S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol. 2010;10:210.
83. Imachi H, Nobu MK, Nakahara N, Morono Y, Ogawara M, Takaki Y, et al. Isolation of an archaeon at the prokaryote-eukaryote interface. Nature. 2020;577:519–25.
84. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. A complete domain-to-species taxonomy for bacteria and archaea. Nat Biotechnol. 2020;38:1079–86.

85. Le SQ, Gascuel O. An improved general amino acid replacement matrix. Mol Biol Evol. 2008;25:1307–20.
86. Schrempf D, Lartillot N, Szöllősi G. Scalable empirical mixture models that account for across-site compositional heterogeneity. Mol Biol Evol. 2020;37:3616–31.
87. Williams TA, Schrempf D, Szöllősi GJ, Cox CJ, Foster PG, Embley TM. Inferring the deep past from molecular data. Genome Biol Evol. 2021;13:evab067.
88. Lemoine F, Domelevo Entfellner JB, Wilkinson E, Correia D, Dávila Felipe M, De Oliveira T, et al. Renewing Felsenstein's phylogenetic bootstrap in the era of big data. Nature. 2018;556:452–6.
89. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 2012;28:3150–2.
90. Lin H, Chen W, Ding H. AcalPred: a sequence-based tool for discriminating between acidic and alkaline enzymes. PLoS One. 2013;8:e75726.

ACKNOWLEDGEMENTS

The authors greatly appreciate Mr. Sejima and others of Happo-one Development Co., Ltd. for their kind permission and cooperation in conducting field studies at Hakuba Happo hot springs. This work was supported by the JSPS KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas, "Hadean Bioscience" project (nos. JP26106006 to K. Kurokawa and JP26106004 to Y. Kamagata). This work was partly funded by the JSPS KAKENHI Grant-in-aid for Scientific Research nos. JP26710012 and JP18H02426 to H. Tamaki, JP15H05620 to R. Nakai, and JP18H03367 to M.K. Nobu.

AUTHOR CONTRIBUTIONS

MKN and RN designed and performed metagenomic, phylogenetic, and thermodynamic analyses, interpreted the data, and wrote the manuscript. ST performed sampling, DNA extraction, and cultivation. HM, AT, and KK performed metagenomic sequencing and metagenomic analysis. AI performed chemical measurements. SS supported chemical data collection. HT and YK designed the project, supervised sampling and cultivation, and supported manuscript preparation.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41396-022-01197-9.

Correspondence and requests for materials should be addressed to Masaru Konishi Nobu or Hideyuki Tamaki.