ARTICLE IN PRESS

Nature Communications

https://doi.org/10.1038/s41467-026-68859-6

Article in Press

Branched-chain amino acid specialization drove diversification within *Calditenuaceae* (*Caldarchaeia*) and enables their cultivation

Received: 15 February 2025

Accepted: 16 January 2026

Cite this article as: Lai, D., Mosier, D., Palmer, M. *et al.* Branched-chain amino acid specialization drove diversification within *Calditenuaceae* (*Caldarchaeia*) and enables their cultivation. *Nat Commun* (2026). https://doi.org/10.1038/s41467-026-68859-6

Dengxun Lai, Damon Mosier, Marike Palmer, Xavier Mayali, Juliet Johnston, Walter Saldivar, Jonathan K. Covington, Jian-Yu Jiao, Ranjani Murali, Cale O. Seymour, Lan Liu, Zheng-Shuang Hua, Wen-Jun Li, Peter K. Weber, Jennifer Pett-Ridge, Daniel R. Colman, Eric S. Boyd, Takuro Nunoura, Jeremy A. Dodsworth & Brian P. Hedlund

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

ARTICLE IN PRESS

Branched-chain amino acid specialization drove diversification within *Calditenuaceae* (*Caldarchaeia*) and enables their cultivation

Dengxun Lai[1,2*], Damon Mosier[3,4], Marike Palmer[1,5], Xavier Mayali[6], Juliet Johnston[6], Walter Saldivar[3], Jonathan K. Covington[1], Jian-Yu Jiao[7], Ranjani Murali[1], Cale O. Seymour[1], Lan Liu[7], Zheng-Shuang Hua[8], Wen-Jun Li[7,9,10], Peter K. Weber[6], Jennifer Pett-Ridge[6], Daniel R. Colman[11], Eric S. Boyd[11], Takuro Nunoura[12], Jeremy A. Dodsworth[3*] & Brian P. Hedlund[1,13*]

[1]School of Life Sciences, University of Nevada, Las Vegas, NV, USA
[2]Department of Agroecology, Aarhus University, Slagelse, Denmark
[3]Department of Biology, California State University, San Bernardino, CA, USA
[4]Department of Microbiology, University of Calgary, Calgary, AB, Canada
[5]Department of Microbiology, University of Manitoba, Winnipeg, MB, Canada
[6]Physical and Life Sciences Directorate, Lawrence Livermore National Lab, Livermore, CA, USA
[7]State Key Lab of Biocontrol, Guangdong Provincial Key Lab of Plant Resources, School of Life Sciences, Sun Yat-Sen University, Guangzhou, PR China
[8]Department of Environmental Science and Engineering, University of Science and Technology of China, Hefei, PR China
[9]Southern Marine Science and Engineering Guangdong Lab (Zhuhai), School of Life Sciences, Sun Yat-Sen University, Guangzhou, PR China
[10]State Key Lab of Desert and Oasis Ecology, Key Lab of Ecological Safety and Sustainable Development in Arid Lands, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, PR China
[11]Department of Microbiology and Cell Biology, Montana State University, Bozeman, MT, USA
[12]Research Center for Bioscience and Nanoscience (CeBN), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan
[13]Nevada Institute of Personalized Medicine, University of Nevada Las Vegas, Las Vegas, NV, USA

*Correspondence to Dengxun Lai (laid.omics@gmail.com), Jeremy A. Dodsworth (jdodsworth@csusb.edu), and Brian P. Hedlund (brian.hedlund@unlv.edu)

ARTICLE IN PRESS

ARTICLE IN PRESS

Abstract

Many thermophiles that are abundant in high-temperature geothermal systems have never been cultivated and are poorly understood, including deeply branching members of the archaeal phylum *Thermoproteota*. Here, we describe the genome-guided cultivation of one such organism, *Calditenuis ramacidaminiphagus*, and show that it has evolved a heterotrophic metabolism focused on branched-chain amino acids (BCAAs). Initially, fluorescence in situ hybridization and nanoscale secondary ion mass spectrometry (FISH-nanoSIMS) showed that *Cal. ramacidaminiphagus* assimilated amino acids rapidly in casamino acid-amended enrichment cultures. Metagenome and metaproteome analyses showed a high abundance and expression of BCAA transporter genes, suggesting a BCAA-focused metabolism. This inference was supported by the subsequent enrichment of *Cal. ramacidaminiphagus* in BCAA-fed cultures, reaching 2.66×10^6 cells/mL and 48.7% of the community, whereas it was outcompeted when polar amino acids were included. Metabolic reconstruction and metaproteomics suggest that BCAAs are channeled into the mevalonate pathway for lipid biosynthesis and fuel ATP production through the TCA cycle coupled with aerobic respiration and through production of branched-chain organic acids by overflow metabolism. Ancestral state reconstructions and phylogenetic analyses of 62 *Caldarchaeales* genomes revealed multiple horizontal transfers of BCAA transporters to the ancestor of the genus *Calditenuis*. Our study highlights the crucial role of BCAAs in the early evolution and niche of this genus, and suggests a high degree of resource partitioning even within low-diversity thermophilic communities.

ARTICLE IN PRESS

ARTICLE IN PRESS

Introduction

Geothermal springs offer an exceptional window into the evolution, adaptation, and survival strategies of microorganisms[1,2]. The unique physicochemical characteristics of high-temperature geothermal springs promote the formation of microbial communities that are often simplified and include thermophiles and polyextremophiles[3,4,5,6]. The vast majority of thermophiles remain uncultivated[7,8]. While significant progress has been made to understand these uncultivated microorganisms through metagenomic studies, a critical gap remains in linking genomic predictions to physiological characteristics.

Among uncultivated thermophiles is the class *Nitrososphaeria*_A in the Genome Taxonomy Database (GTDB)[9], which was originally proposed as the phylum "Aigarchaeota" based on a metagenome-assembled genome (MAG)[10]. This class is a monophyletic sister taxon to *Nitrososphaeria sensu stricto*, the latter encompassing the globally widespread ammonia-oxidizing archaea (AOA). Previous evolutionary analyses indicated that the *Nitrososphaeria*/*Nitrososphaeria*_A clade originated in thermal environments[11,12,13,14] or transiently inhabited them before AOA diversified into moderate-temperature and cold environments[15], whereas *Nitrososphaeria*_A remained in geothermal environments, including terrestrial hot springs and fumaroles[8,12,16,17], subsurface geothermal systems[10,18], and deep-sea hydrothermal vent ecosystems[19]. A meta-analysis of 16S rRNA genes revealed at least nine genus-level groups within *Nitrososphaeria*_A[8], and a subsequent study assigned all *Nitrososphaeria*_A genomes to three families, *Wolframiiraptoraceae*, "*Caldarchaeaceae*", and *Calditenuaceae*[19]. The latter study reported tungsten-dependent growth of one member of the *Wolframiiraptoraceae*, *Wolframiiraptor gerlachensis*, on xylose and showed that tungsten transporters and tungsten-dependent aldehyde ferredoxin oxidoreductases were acquired during the evolution of the *Wolframiiraptoraceae*, but not its two sister families. *Wolframiiraptoraceae*, comprising strict anaerobes, is thought to have originated approximately 1.48 billion years ago (Ga)[11], and is the most ancient family within the *Nitrososphaeria*_A. In contrast, the *Calditenuaceae* are estimated to have diverged around 0.7 Ga, after the Neoproterozoic Oxygenation Event (NOE)[11]. Consistent with this inference, research showed that *Calditenuis aerorheumatis*[20] (syn. "*Candidatus* Calditenuis aerorheumensis"), a species in the *Calditenuaceae* from Yellowstone National Park, is aerobic, as evidenced by genes encoding a heme copper oxidase complex and a high number of

ARTICLE IN PRESS

transcripts mapping to this complex in a metatranscriptome[21]. *Calditenuis* is widely distributed[8,17] and *Calditenuis* species from hot springs in Tengchong, China (JZ_bin_19 in ref.[12]), have been predicted to oxidize CO aerobically to obtain energy and fix carbon via the reverse oxidative TCA (roTCA) pathway[22]. While systematic investigation of *Wolframiiraptoraceae* genomes has been conducted[19], genomes belonging to the *Calditenuaceae* are poorly represented in current databases and our knowledge about this family is limited.

The United States' Great Basin, located in the western part of the country, is characterized by rapid crustal thinning during the Miocene and increasing aridification, giving rise to widely distributed and poorly connected geothermal ecosystems[23]. One such ecosystem is the Great Boiling Spring (GBS) geothermal system in northwestern Nevada. This circumneutral Na–Cl spring system has temperatures reaching 95.5°C (boiling) and is home to diverse thermophiles[24,25], most of which remain uncultivated. *Calditenuis* is the most abundant archaeon in high-temperature GBS sediments based on 16S rRNA gene surveys[24] and metagenomic read mapping[26]. Previous studies have suggested that members of *Calditenuis* have the potential to assimilate fatty acids and amino acids based on genomic and transcriptomic data[12,16]. However, the validity of these predictions has yet to be experimentally confirmed, and other physiological characteristics of *Calditenuis* remain unknown.

In this study, we report the genome-guided enrichment and stable lab growth of a novel member of *Calditenuis*, designated *Calditenuis ramacidaminiphagus* LZ^{Ts}, the most abundant archaeal species in high-temperature GBS sediments. Although initial fluorescence in situ hybridization coupled with nanoscale secondary ion mass spectrometry (FISH-nanoSIMS) experiments in GBS sediments showed very low assimilatory activity by LZ^{Ts} compared to other cells, FISH-nanoSIMS with stable lab cultures of *Cal. ramacidaminiphagus* showed that LZ^{Ts} incorporated a large fraction of their carbon from amino acids. Analysis of MAGs of *Cal. ramacidaminiphagus* revealed a high abundance of branched-chain amino acid (BCAA) transporter genes, and metaproteomics confirmed the high expression of these genes in LZ^{Ts}. Based on these findings, we hypothesized that LZ^{Ts} could grow with BCAAs. To test this, we selectively enriched *Cal. ramacidaminiphagus* when BCAAs were supplied as the sole source of carbon. Metabolic reconstruction and metaproteomics suggested that BCAAs are utilized for isoprenoid biosynthesis

ARTICLE IN PRESS

and contribute to ATP production via the TCA cycle and aerobic respiration and through overflow metabolism, producing branched-chain organic acids. Ancestral state reconstructions and phylogenetic analyses revealed that BCAA transporter genes were acquired through several horizontal gene transfers from multiple sources to an ancestor of *Calditenuis*, underscoring the importance of BCAAs in the ecology and evolution of the genus. Given the importance of BCAAs, we propose the species epithet *ram.a.cid.a.mi.ni.pha'gus*, meaning "branched-chain amino acid eater".

Results

Predominance of *Cal. ramacidaminiphagus* in GBS

Shotgun metagenome sequencing was previously performed on sediment samples collected from GBS (Fig. 1a) at three sites: site A (85°C), site B (70°C), and site C (60°C)[24] (Fig. 1b). The relative abundances of archaea and bacteria based on metagenome read mapping correlated with temperature, with archaea being more abundant at higher temperatures (Fig. 1b), consistent with previous studies[24,25,27]. A single species, *Cal. ramacidaminiphagus*, was the dominant archaeon in the clay-rich sediments at 70°C and 85°C (Fig. 1c and Supplementary Data 1), comprising 83.7% and 24.5% of archaea at sites A and B, respectively. Given the high abundance of *Cal. ramacidaminiphagus* in the hot spring, we inferred it is ecologically important and that cultivating it could provide deeper insights into the ecology of high-temperature environments.

The cultivation of many archaea has proven to be challenging[28], so to better understand the physiology of this microorganism, we first used a FISH-nanoSIMS approach[29] to investigate the metabolism of single cells of *Cal. ramacidaminiphagus* within the complex community. We supplemented freshly collected GBS sediment with various ^{13}C-labeled substrates (algal amino acid mixture [algal AA], butyrate + propionate + acetate [BPA], starch, bicarbonate) to track the assimilation of different carbon sources, and with ^{15}N-labeled ammonium as a proxy for overall assimilatory activity (Supplementary Fig. 1 and Supplementary Data 2). Incubations were conducted under both aerobic and anaerobic conditions for four hours and 24 hours. *Cal. ramacidaminiphagus* cells were identified using catalyzed reporter deposition-FISH (CARD-FISH) and subsequently subjected to nanoSIMS analysis (Supplementary Fig. 2). Under aerobic conditions, *Cal. ramacidaminiphagus* cells collected after four hours in algal AA-amended

ARTICLE IN PRESS

incubations were enriched in ^{15}N compared to unlabeled controls, suggesting stimulation of *Cal. ramacidaminiphagus* activity (two-sided Mann-Whitney *U* test, p =0.040; Supplementary Data 3), yet the assimilation of carbon from amino acids by *Cal. ramacidaminiphagus* only became evident after 24 hours (two-sided Mann-Whitney *U* test, $p = 0.080$; Supplementary Data 3). The median ^{13}C enrichment of *Cal. ramacidaminiphagus* cells incubated with labeled algal AA after 4 hours or 24 hours was less than 0.2 atom percent excess (APE) (Supplementary Data 4), and there were no significant differences in ^{13}C and ^{15}N enrichment between *Cal. ramacidaminiphagus* cells and other cells after 4 hours of incubation. After 24 hours of incubation, however, other cells showed higher ^{13}C and ^{15}N enrichment than *Cal. ramacidaminiphagus* cells (two-sided Mann-Whitney *U* test, p = 4.63E-11, Supplementary Fig. 2 and Supplementary Data 3). Consequently, we could not distinguish: (i) slow amino acid assimilation by *Cal. ramacidaminiphagus*, (ii) catabolic use of amino acids while incorporating only a small amount of carbon into biomass, or (iii) isotope enrichment from cross-feeding[30] on metabolites derived from amino acids. We did not detect isotope incorporation by *Cal. ramacidaminiphagus* cells under anaerobic conditions, consistent with previous research suggesting that *Calditenuis* is aerobic[12,16,21].

Amino acids are preferred substrates for *Cal. ramacidaminiphagus* in stable lab cultures

Since in situ labeling experiments were not definitive, we aimed to cultivate *Cal. ramacidaminiphagus* in the lab, which would enable more incisive experimentation. Samples were taken from 85°C GBS sediments and inoculated into GBS salts medium[31] with 22 different combinations of carbon sources, H_2 concentrations, and terminal electron acceptors, including varying O_2 concentrations[26]. The primary cultures were sampled after 2 weeks and a secondary culture was sampled after 4 weeks, and community composition was evaluated by 16S rRNA gene sequencing (Supplementary Fig. 3; Supplementary Data 5). Of the 22 growth conditions tested, only two supported populations of *Cal. ramacidaminiphagus* that comprised >10% relative abundance of the community 16S rRNA genes. The highest relative abundance of *Cal. ramacidaminiphagus* was 25.3%, observed in culture G-10, where 0.02% casamino acids served as the sole carbon source, 10% v/v O_2 (2 atm pressure) was provided as the sole terminal electron acceptor, and the incubation temperature was 80°C[26]. The other enrichment was culture G-5, which was supplemented with 10% (v/v) of GBS hot spring mat extract and 2% (v/v) O_2, with 16.1% of

ARTICLE IN PRESS

the culture being *Cal. ramacidaminiphagus*[26]. *Cal. ramacidaminiphagus* was stably maintained in the G-10 culture in the lab, with transfers every three weeks over the course of 8 years.

Efforts to isolate *Cal. ramacidaminiphagus* were unsuccessful. Plating culture G-10 onto Gelrite-solidified medium yielded no colonies, and single-cell sorting via optical tweezers targeting rod-shaped cells isolated only a *Pyrobaculum* species. FISH-nanoSIMS was repeated with culture G-10 to investigate the metabolism of *Cal. ramacidaminiphagus* in the absence of an axenic culture (Supplementary Fig. 1 and Supplementary Data 6). We initially screened a few cells from the 6-hour treatments and detected only minor isotope enrichment, so we focused on 24-hour incubations. After 24 hours of incubation, both *Cal. ramacidaminiphagus* and other cells were enriched in ^{13}C from a variety of ^{13}C-labeled substrates and ^{15}N from ^{15}N-labeled ammonium compared to unlabeled controls (two-sided Mann-Whitney *U* test, $p < 0.01$; Fig. 2, Supplementary Data 3, and Supplementary Data 7), indicating active metabolism. The median ^{13}C enrichment of *Cal. ramacidaminiphagus* cells incubated with substrates was as follows: 17.0 APE with algal AA, 5.17 APE with propionate, 2.12 APE with acetate, 1.43 APE with xylose, and 1.0 APE with butyrate (Fig. 2 and Supplementary Data 7). Enrichment with labeled ribose, bicarbonate, or formate was less than 1.0 APE. Notably, *Cal. ramacidaminiphagus* assimilated a high amount of ^{13}C atoms from algal AA, although labeling was not higher than other cells (Fig. 2). *Cal. ramacidaminiphagus* cells did incorporate more ^{13}C and ^{15}N atoms than other cells in butyrate cultures, suggesting direct utilization of butyrate by *Cal. ramacidaminiphagus* (two-sided Mann-Whitney *U* test, $p < 0.01$; Fig. 2, Supplementary Fig. 2, Supplementary Data 3, and Supplementary Data 7). *Cal. ramacidaminiphagus* also assimilated considerable ^{13}C atoms from propionate, yet without higher labeling relative to other cells (Fig. 2, Supplementary Fig. 4). Additional interpretations about butyrate and propionate metabolisms are included in Supplementary Note 1.

Genomic, metaproteomic, and proteomic insights suggest BCAA specialization

Since *Cal. ramacidaminiphagus* was maintained in the G-10 culture that contained casamino acids and assimilated large amounts of ^{13}C atoms from amino acids, we evaluated metabolic pathways related to amino acid metabolism, beginning with amino acid transporters. A high-quality *Cal. ramacidaminiphagus* MAG obtained from the G-10 culture designated LZ^{Ts} (estimated completeness 98.07%, estimated contamination 0%), 18 other high- and medium-quality MAGs,

ARTICLE IN PRESS

and one single-amplified genome (SAG) assigned to the genus *Calditenuis* by GTDB-Tk version 2.3.2[32] each encoded complete ABC transport systems for BCAAs and peptides. However, the assembled genomes did not encode transporters for PAAs and lacked genes encoding substrate-binding and ATP-binding proteins for neutral amino acids, suggesting specialization in BCAA utilization. To probe this idea further, differential abundance analysis using DESeq2[33] was conducted between the *Calditenuis* MAGs and other MAGs assigned to *Calditenuaceae*, revealing that BCAA transporter genes were highly enriched in *Calditenuis* (Fig. 3a). Specifically, the *Cal. ramacidaminiphagus* MAG, LZ^{Ts}, contains 9 copies of *livK* (encoding the substrate-binding protein (SBP)), 12 copies of *livM* (encoding the permease protein), 10 copies of *livH* (encoding another permease protein), 8 copies of *livF* (encoding the ATP-binding protein), and 11 copies of *livG* (encoding another ATP-binding protein). These BCAA transporter genes are clustered in the *Cal. ramacidaminiphagus* genome, and each gene encodes key features required for proper function: signal peptides and substrate binding domains for the SBPs (LivK-K01999), transmembrane helices for the permeases (LivM-K01999, LivH-K01997), and AAA domain and P-loop domains for the ATP-binding proteins (LivF-K01996, LivG-K01995) (Fig. 3b, Supplementary Data 8, and Supplementary Data 9). Several collocated genes likely involved in BCAA metabolism were also observed, including genes encoding acyl-CoA dehydrogenase (*acadM*-K00249), as well as genes involved in isoprenoid biosynthesis, such as geranylgeranyl diphosphate synthase (*idsA*-K01799) and hydroxymethylglutaryl-CoA reductase (*hmgcR*-K00021) (Fig. 3b). Notably, two long-chain acyl-CoA synthetase genes (*acsL*-K01897) were located beside BCAA transporters, suggesting that AcsL may activate BCAA-derived acyl intermediates or indicate a connection between BCAA metabolism and lipid biosynthesis.

To further probe the specificity of amino acid metabolism in the G-10 culture, we examined the expression of BCAA transporter genes in a G-10 soluble fraction metaproteome. The BCAA transporter SBP (LivK) and ATP-binding proteins (LivF and LivG) were detected, but the permeases were not observed (Supplementary Fig. 5 and Supplementary Data 10). Permeases are challenging to detect in proteomes due to their poor solubility as integral membrane proteins. The enrichment of ABC transport systems for BCAAs, their expression in the G-10 culture, and the absence of ABC transporters for other types of amino acids suggest that *Cal. ramacidaminiphagus*

ARTICLE IN PRESS

is adapted to specialize in BCAA metabolism. Thus, we reasoned that BCAAs are critical for the physiology of *Cal. ramacidaminiphagus*.

Given the predicted specialization of *Cal. ramacidaminiphagus* in BCAA metabolism, we investigated whether this specialization was reflected in the number and catalytic types of peptidases encoded in the *Calditenuaceae* genomes and proteomes. This analysis revealed that *Calditenuis* has a higher number of peptidases than other genera of *Calditenuaceae* (two-sided Mann-Whitney *U* test, $p = 4.61E-6$; Fig. 3c and Supplementary Data 11). Closer inspection showed that peptidase families M103, M20X, and T01X predominate in *Calditenuis* (Supplementary Data 12), suggesting a potential role in digesting proteins and peptides to release BCAAs. For instance, the TldD peptidase (M103.001) preferentially cleaves peptides near valines (https://www.ebi.ac.uk/merops/cgi-bin/pepsum?id=M103.001), and *Calditenuis* proteomes are enriched in valine relative to other amino acids (two-sided Mann-Whitney *U* test, $p = 7.31E-9$; Supplementary Fig. 5 and Supplementary Data 13). However, none of the identified *Cal. ramacidaminiphagus* peptidases possess classical signal peptides (Supplementary Data 14) and nearly all are globular (soluble), except for one M28B peptidase, which is a transmembrane protein with an extracellular peptidase domain. Given the predicted oligopeptide ABC transporter in *Calditenuis* (Supplementary Data 15), *Calditenuis* may supplement transported BCAAs with BCAAs from transported oligopeptides liberated by cytoplasmic peptidases. The near-absence of extracellular proteases could reflect the instability of extracellular proteins at high temperatures or a reliance on other microbes for extracellular proteolysis. Additional interpretations about peptidases and amino acid stoichiometry are included in Supplementary Note 2.

BCAAs channel into lipid synthesis, anabolism, and ATP production

The likely metabolism of all three BCAAs was reconstructed based on the *Cal. ramacidaminiphagus* LZ^{Ts} genome and from the G-10 metaproteome (Fig. 4). Complete anabolic and catabolic pathways for leucine, isoleucine, and valine were present in our model. The first step of the common pathway is catalyzed by branched-chain amino acid aminotransferase (IlvE), which transfers the amino group from BCAAs to α-ketoglutarate, forming branched-chain α-keto acids (syn. 2-oxoacids). The second step is catalyzed by 2-oxoisovalerate dehydrogenase (BCKDHA, BCKDHB, DBT), which decarboxylates the branched-chain α-keto acids to their corresponding

ARTICLE IN PRESS

acyl-CoA derivatives. The subsequent steps vary for leucine, isoleucine, and valine, although they share some common enzymes (Fig. 4). The metabolic products of leucine could be integrated into lipids since the key enzymes to channel leucine into mevalonate are present in *Calditenuis* and most of these enzymes were expressed in *Cal. ramacidaminiphagus* (Fig. 4 and Supplementary Data 10). A direct precursor to mevalonate, 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA), is a crucial intermediate in the biosynthesis of isoprenoids. *Cal. ramacidaminiphagus* likely utilizes the alternative mevalonate pathway for isoprenoid biosynthesis found in archaea[34,35], because it lacks the phosphomevalonate kinase (EC 2.7.4.2) used in the classical mevalonate pathway. Instead, it encodes phosphomevalonate decarboxylase (Pmd) and isopentenyl phosphate kinase (Ipk) to form isopentenyl diphosphate (Supplementary Fig. 7). Additionally, *Cal. ramacidaminiphagus* has the complete C10-C20 isoprenoid biosynthesis pathway, suggesting its capability to synthesize isoprenoids with carbon chain lengths ranging from C10 to C20 (Supplementary Fig. 7). Consequently, *Cal. ramacidaminiphagus* likely contributes to the in-situ production of isoprenoidal glycerol dialkyl glycerol tetraethers (GDGTs), which are abundant in GBS[36]. Indeed, three radical SAM proteins involved in the synthesis of GDGTs — tetraether synthase (Tes)[37], GDGT ring synthase A (GrsA), and GDGT ring synthase B (GrsB)[38] — were identified (Supplementary Data 16). All three BCAAs can likely be metabolized into acetyl-CoA and subsequently enter a complete TCA cycle (Supplementary Fig. 8), suggesting NADH may serve as a key electron carrier linking the TCA cycle to aerobic respiration. This could enable ATP synthesis when coupled with aerobic respiration using a heme copper oxidase complex (Cox) or the generation of intermediates for anabolism.

In addition to lipid synthesis and the TCA cycle, *Cal. ramacidaminiphagus* may produce branched-chain organic acids from BCAAs through overflow metabolism, which has been observed in both *Saccharolobus solfataricus*[39,40] and *Haloarcula marismortui*[41] under aerobic conditions. This metabolism would involve three steps (Fig. 4): (i) the oxidative deamination of amino acids to the α-keto acids 2-oxoisocaproate, 3-methyl-2-oxopentanoate, and 2-oxoisovalerate, catalyzed by IlvE; (ii) the oxidative decarboxylation of 2-oxoacids to their corresponding acyl-CoAs, isovaleryl-CoA (syn. 3-methylbutanoyl-CoA), 2-methylbutanoyl-CoA, and isobutyryl-CoA, catalyzed by 2-oxoacid:ferredoxin oxidoreductase (OFOR); and (iii) the hydrolysis of acyl-CoA coupled with substrate-level phosphorylation, catalyzed by acetate-CoA ligase (ACD; EC 6.2.1.13). Two key

ARTICLE IN PRESS

enzymes were identified in *Cal. ramacidaminiphagus*. The first is a putatively oxygen-resistant OFOR[39,42] and the second is ACD (Supplementary Data 17 and Supplementary Data 18). Although OFOR is typically sensitive to O_2[43,44], some *Sulfolobus* strains use an O_2-resistant variant that is active under aerobic conditions[39,43]. Protein alignments between OFOR from *Cal. ramacidaminiphagus* (herein CrOFOR) and the oxygen-resistant StOFOR2 from *Sulfolobus tokodaii* shows a high degree of similarity. Similar to StOFOR2[43], CrOFOR consists of an α-subunit with three fused domains, III (1a-233a), I (233a-518a), and II (518a-650a), and a single domain β-subunit, domain VI (1b-327b) of the β-subunit (Supplementary Fig. 9). The a or b in parentheses indicates the corresponding subunit. We further constructed computed structure models of CrOFOR and StOFOR2 using AlphaFold 3[45]. The computed topology of the domains of CrOFOR is highly similar to StOFOR2, and superimposition of the CrOFOR and StOFOR2 structures revealed similar folds and nearly identical active sites, suggesting that CrOFOR is O_2-tolerant and has broad 2-oxoacid specificity similar to StOFOR2[43] (Supplementary Fig. 10 and Supplementary Fig. 11). Additionally, OforA and OforB were detected in the G-10 metaproteome (Supplementary Data 10). ACD has previously been isolated from archaea, including *Pyrococcus furiosus* and *Thermococcus kodakarensis*, and is capable of catalyzing the formation of organic acids and ATP[42,46,47]. ACD was also detected in the G-10 metaproteome (Supplementary Data 10). We note that carboxylic acids can act as protonophores and therefore inhibit growth[48]. Given the high expression of these enzymes in the G-10 metaproteome, we suggest that consumption of carboxylic acids by other thermophiles may be a metabolic interdependency that could limit attempts to isolate *Cal. ramacidaminiphagus*.

BCAAs specifically enrich for *Cal. ramacidaminiphagus*

In light of our results, we hypothesized that *Cal. ramacidaminiphagus* could grow with BCAAs as carbon sources and electron donors. To test this hypothesis, lab cultures were established by transferring the G-10 culture into GBS salts medium containing only BCAAs, PAAs, or both BCAAs and PAAs as sole carbon sources, and the cultures were incubated for 7 days. DNA was extracted, and 16S rRNA genes were PCR-amplified and sequenced. Since *Calditenuis* is not present in the SILVA database, a 16S rRNA gene phylogenetic tree was constructed with a dataset of sequences from this group[8] to identify amplicon sequence variants (ASVs) belonging to *Calditenuis* (Supplementary Fig. 12). The results showed that BCAAs led to a strong enrichment

ARTICLE IN PRESS

of *Cal. ramacidaminiphagus* (48.7% of the population) and *Pyrobaculum*, while PAAs enriched for *Fervidibacter sacchari*, and the combination of BCAAs and PAAs enriched for *Thermocrinis jamiesonii* (Fig. 5a and Supplementary Data 19). To distinguish net growth from differential survival in these cultures, we performed qPCR using *Calditenuis*-specific primers to measure the absolute abundance of *Calditenuis*. This showed that the absolute abundance of *Calditenuis* in BCAA-amended cultures was 2.66×10^6 cells/mL, which is 278-fold higher than in the other cultures (Fig. 5b and Supplementary Data 19). Notably, the absolute abundance of *Calditenuis* in cultures amended with both PAAs and BCAAs was low, suggesting that PAAs or thermogenic products of PAAs may confer a competitive disadvantage or be toxic to *Calditenuis* at the tested concentrations. PAA toxicity has been observed in bacteria[49].

Phylogenomics, taxonomy, and inferred ecophysiology of *Calditenuaceae*

To further investigate the genomic potential and metabolic capacity of the *Calditenuaceae*, we collected and annotated 62 high- and medium-quality MAGs and one SAG belonging to the *Caldarchaeales*. Among these, 41 were newly obtained for this study. These 62 MAGs were exclusively derived from high-temperature, circumneutral pH, terrestrial geothermal springs in the U.S. Great Basin (Nevada), Yellowstone National Park (Wyoming), Malaysia, and China, and terrestrial fumaroles in Hawaii (Supplementary Data 20). The classification of the genomes assigned to the family *Calditenuaceae* was then determined based on monophyly observed in phylogenomic analyses using conserved gene/protein sets, relative evolutionary divergence, and pairwise genome distance calculations. Phylogenomic analyses of these 62 MAGs and 17 other *Caldarchaeales* MAGs available in public databases (Supplementary Data 20) provided robust evidence for the monophyly of the family *Calditenuaceae*, as supported by four marker gene/protein sets – 53 archaeal marker proteins (ar53)[9], 56 universally conserved single-copy proteins (uni56)[50], ribosomal proteins (Rprot)[51], and up-to-date archaeal core genes (UACG)[52] – (Fig. 6 and Supplementary Fig. 13). *Calditenuaceae* MAGs were then assigned to five genus-level groups based on relative evolutionary divergence[9] and average amino acid identity (AAI)[53] and 12 species based on average nucleotide identity (ANI)[54]. These genera include the previously recognized[8] genera G1 (JGI-OTU-1 in GTDB, *Calditenuis*[21]), G2 (AWOC01 in GTDB, herein *Ardentivivens*), G2 (HRBIN02 in GTDB, herein *Candentenecus*), JANXDH01 (not in ref[8]; JANXDH01 in GTDB, herein *Caldaricola*), and a novel group without a previous representative

ARTICLE IN PRESS

genome in public databases, G3 (herein *Kaftiomonas*) (Fig. 6, Supplementary Fig. 13, Supplementary Fig. 14, and Supplementary Data 20). G2 was originally designated based on 16S rRNA gene phylogeny[8] but contains two genera, *Ardentivivens* and *Candentenecus*, according to our phylogenomic analyses. A superimposed plot of ultrametric gene trees of 34 shared marker proteins further confirmed the coherence and independence of the five genera (Supplementary Fig. 15). Given the absence of representative genomes for genera G6 and G8 (Supplementary Fig. 14), we inferred that there are at least 12 genera within the still-undersampled class *Nitrososphaeria*_A (herein *Caldarchaeia*). Further sequencing efforts in shallow vents in Papua New Guinea, Yellowstone National Park, and the U.S. Great Basin, where 16S rRNA gene sequences of G6 and G8 were found, would enhance our understanding of this group. Nomenclatural type genomes were selected for the 12 species based on genome quality and the presence of rRNA genes. The proposed names have been registered and preliminarily endorsed under Path 1 of the SeqCode[55,56] (Supplementary Note 3). Additional interpretations about taxonomy, metabolism, and evolution of *Calditenuaceae* are included in Supplementary Note 4.

BCAA specialization as a key trait for the genus *Calditenuis*

As noted previously, genes encoding BCAA transporter proteins are enriched in *Calditenuis* genomes relative to other *Calditenuaceae* genomes (Fig. 3a and Supplementary Data 15), averaging 8-12 copies per gene per genome. To better understand the evolutionary relationships and origins of these BCAA transporters, BCAA transporter proteins from *Caldarchaeia* genomes and closely related proteins from UniProt[57] and the NCBI NR database were collected. Maximum-likelihood phylogenies of each component of the BCAA ABC transport system suggest a highly dynamic evolution of BCAA transporters. The major evolutionary patterns were the acquisition of BCAA transporters by the common ancestor of the genus *Calditenuis* via horizontal gene transfer from other archaea followed primarily by vertical evolution within the genus, as indicated by the observation that several *Calditenuis* BCAA transport proteins are monophyletic but not closely related (Fig. 7a and Supplementary Fig. 16).

Next, proteins involved in BCAA metabolism in *Caldarchaeia* genomes were screened, with a focus on *Calditenuaceae* (Fig. 4 and Fig. 7b). Notably, all members of *Calditenuaceae* have OforA, OforB, and ACD, indicating that the capability to produce branched-chain organic acids is a

ARTICLE IN PRESS

common and ancient feature within this family. This finding is consistent with previous studies showing that this pathway may be widely distributed in archaea[39]. Only members of *Calditenuis* have genes encoding BCKDHA, BCKDHB, and DBT, providing an alternative mechanism to convert 2-oxoacids derived from BCAAs to their CoA conjugates and suggesting that these genes were likely acquired by the ancestor of *Calditenuis*. Phylogenetic analyses show that the BCKDHA and BCKDHB sequences from *Calditenuis* each form two monophyletic clusters (Supplementary Fig. 17), indicating early gene-duplication events within the genus followed by vertical retention. DBT sequences from *Calditenuis* form a single monophyletic cluster, reflecting vertical inheritance (Supplementary Fig. 18). Notably, the BCKDHA, BCKDHB, and DBT sequences from *Calditenuis* are phylogenetically adjacent to homologous sequences from *Thermus*, suggesting a potential evolutionary connection between these thermophilic taxa. All species of *Calditenuis* encode most or all proteins necessary for BCAA metabolism, except for malonate-semialdehyde dehydrogenase (MmsA), which is absent in *Ca. aerorheumatis* and *Calditenuis fumarioli* (Fig. 7b), suggesting that the valine metabolism pathway is incomplete in these species. Additionally, several genes encoding proteins necessary to channel isoleucine and valine into the TCA cycle are absent in *Caldaricola*, *Kaftiomonas, Ardentivivens*, and *Candentenecus*, suggesting they may only generate ATP from isoleucine and valine via production of branched-chain organic acids. *Kaftiomonas* also lacks genes encoding HMGCR (Fig. 7b), which plays a crucial role in the conversion of BCAA degradation products into intermediates necessary for isoprenoid synthesis. The absence of HMGCR suggests that BCAAs may not be direct precursors for isoprenoid synthesis in *Kaftiomonas*.

Ancestral reconstruction of BCAA transporter genes

To better understand how *Calditenuis* acquired so many copies of BCAA transporter genes, we conducted ancestral reconstruction using the Amalgamated Likelihood Estimation (ALE) method for BCAA transporter genes, with PAA transporter genes as a comparison. This gene tree-aware approach, based on probabilistic models, includes parameters for gene originations (gene birth or HGTs from outside the sampled genomes), duplications, losses, and transfers, and has previously been used for tree reconciliation[58]. The evolution of PAA transporter genes in *Calditenuaceae* was simple, with horizontal transfer to the common ancestor of *Kaftiomonas*, *Ardentivivens*, and *Candentenecus* followed by vertical evolution (Fig. 8 and Supplementary Fig. 19). In contrast, the

ARTICLE IN PRESS

gene gain and loss patterns and phylogenetic analyses of BCAA transporters were more complex, with each gene having been acquired and lost multiple times. We summarize these dynamics as net gains and losses to focus on the cumulative evolutionary outcomes (Fig. 8). Closer inspection reveals the following conclusions: (i) complete BCAA transport complexes are ubiquitous across *Caldarchaeia*, suggesting conserved BCAA metabolism throughout the class; (ii) gene gain events began in the last common ancestor of extant *Calditenuaceae* (node 1), suggesting positive selection for BCAA metabolism; (iii) extensive BCAA transporter gene evolution occurred within the common ancestor of *Calditenuis* (node 2) and within the species of *Calditenuis*, especially *Cal. ramacidaminiphagus*; for example, *Cal. ramacidaminiphagus* experienced net gains of 8 *livK*, 10 *livH*, 12 *livM*, 11 *livG*, and 8 *livF* gene copies; and (iv) a parallel gain of BCAA transporters in *Caldarchaeaceae* suggesting a common adaptation towards BCAA metabolism among aerobic members of the *Caldarchaeia*. Our observations suggest that BCAA-dependent metabolism is crucial in the evolution of ecophysiology of *Calditenuis*.

Discussion

Despite over a century focused mainly on studies of pure cultures, most of the microbial diversity on Earth remains uncultivated[59]. This is particularly true for archaea, which have been less studied compared to bacteria[59,60]. Our study combined a variety of approaches and expanded our knowledge about the genomic diversity, metabolism, ecological functions, and evolution of an uncultivated and poorly understood class within the phylum *Thermoproteota*, which is often dominant in high-temperature geothermal systems.

We demonstrated BCAA specialization in *Cal. ramacidaminiphagus* LZ^{Ts}, which is the dominant archaeon in GBS. While research on BCAA metabolism in microorganisms remains limited, the presence of pathways for both complete (TCA cycle, aerobic respiration) and partial (branched-chain organic acid production) oxidation of BCAAs in LZ^{Ts} may be driven by the metabolic demands for anabolism and catabolism under different conditions. We speculate that planktonic autotrophs in the water column, mainly *Thermocrinis*, *Pyrobaculum*, and their viruses[25,61,62], and sediments could serve as a source of necromass, releasing peptides and amino acids, including BCAAs, that become available for heterotrophic metabolism[63]. A high genome copy number for

ARTICLE IN PRESS

BCAA transporters and peptidases could allow them to differentiate and/or respond quickly to changes in amino acid flux, which would provide *Cal. ramacidaminiphagus* with a competitive advantage in the low dissolved organic carbon spring[5]. Furthermore, our study suggests that BCAA specialization is a conserved feature in *Calditenuis*, while the other genera in *Calditenuaceae* could also utilize different amino acids. Consistent with these results, Beam et al., 2015[21] showed that BCAA transporter genes are highly transcribed in *Cal. aerorheumatis* within a hot spring in Yellowstone National Park, USA. In particular, abundant transcripts were detected for 8 homologs of the BCAA substrate-binding protein, 12 homologs of the BCAA permease protein, and 5 copies of the BCAA ATP-binding protein, strongly suggesting active BCAA metabolism by *Ca. aerorheumatis* in situ. Other transcripts were consistent with complete oxidation of BCAAs via the TCA cycle coupled with aerobic respiration and the predicted overflow metabolism reactions (i.e., IlvE, OFOR, and ACD). In contrast, a separate study showed that *Candentenecus silaniferventis* and *Candentenecus fervidifontanae* assimilated the PAA aspartate in a hot spring in Yunnan, China[6]. Understanding the specialization of AA metabolism could facilitate the cultivation of *Calditenuis* species and other thermophiles.

Our research underscores a recent report of strong resource partitioning among dissolved organic carbon pools in a similar low-diversity thermophilic community[6]. Such partitioning allows for the coexistence of multiple taxa and minimizes direct competition for resources. Supporting this hypothesis, *W. gerlachensis* from GBS preferentially assimilates xylose[19] over other sugars, while *F. sacchari* specializes in polysaccharide degradation[26]. Our findings reveal strong ecological partitioning according to amino acid types, representing the first study, to our knowledge, to integrate evolutionary genomics and ecophysiology experiments to demonstrate this pattern. These results suggest that even simple ecosystems, such as high-temperature sediments in GBS, have highly specialized metabolic niches.

Methods

Microbial community composition in GBS sediments

Metagenomes derived from sediment samples from three locations within GBS (N40°39′41″ W119°21′58″), identified as site A (85°C), site B (70°C), and site C (60°C) (Fig. 1a, b) were

ARTICLE IN PRESS

described previously[24] and are available on the Integrated Microbial Genomes and Microbiomes platform (IMG/M; https://img.jgi.doe.gov/m/) under Gold Analysis Project IDs Ga0197142, Ga0197143, and Ga0197144[24]. Taxonomic assignment of the protein sequences was conducted using Kraken2[64]. A custom reference database was established by combining RefSeq complete archaeal and bacterial genomes with high-quality representative MAGs or isolates from GBS, including *Cal. ramacidaminiphagus* (GCA_050589285.1), *Ardentivivens gerlachensis* (GCA_054084685.1), *Wolframiiraptor gerlachensis* (GCA_021323375.2), *Fervidibacter sacchari* PD1^T (GCA_030520105.1), and *Thermoflexus hugenholtzii* JAD2^T (GCF_900187885.1). Low-complexity regions in protein sequences[65] were masked using segmasker provided as part of NCBI-BLAST suite. For novel genomes not yet represented in NCBI and thus lacking NCBI taxonomy IDs, existing identifiers for taxa absent from the sample was used as placeholder identifiers (Supplementary Data 1).

Metagenome sequencing, assembly, binning, and annotation

Previously we described a matrix of 22 enrichment cultures that were inoculated with GBS sediments and tracked for two transfers[26]. All cultures led to a decrease in the relative abundance of *Calditenuis*, but one culture, G-10 (0.02% casamino acids (mass vol^{-1}), 10% O_2, 80 ºC), had observable growth after the second transfer and maintained *Calditenuis* at a relative abundance of 25.3% (Supplementary Fig. 3; Supplementary Data 5). *Calditenuis* was stable in this culture for over 8 years in the lab. For the G-10 lab enrichment culture[26], DNA was sequenced using the Illumina MiSeq platform (2×250 bp) and Oxford Nanopore Technology. For long read sequencing, high-molecular weight DNA was extracted using the JGI CTAB protocol with modifications, including three freeze-thaw cycles following lysozyme treatment. RNase A was used to denature ribonucleic acids, followed by an additional chloroform/isoamyl extraction step. The sequencing library was prepared with the EXP-NBD104 Native Barcoding Expansion Kit (Oxford Nanopore Technologies, Oxford, UK) and the SQK-LSK109 Ligation Sequencing Kit, and sequenced on a MinION Mk1B device (Oxford Nanopore Technologies, Oxford, UK) with a FLO-MIN106 flow cell. Long-read data were assembled using metaFLYE v. 2.8.272[66] and this assembly was used as untrusted contigs for assembling Illumina reads using metaSPAdes v. 3.12.173[67]. Detailed information is available in ref[26].

ARTICLE IN PRESS

A total of seven high- and medium-quality *Calditenuaceae* MAGs and one SAG were obtained from the GTDB database (Supplementary Data 20). Three others from GBS were derived from existing assemblies of publicly available metagenomes on the Integrated Microbial Genomes and Microbiomes platform (IMG/M; https://img.jgi.doe.gov/m/)[68]. Samples for metagenomes were collected from Tengchong, China and Yellowstone National Park, USA and sequenced and binned according to methods in the original publications describing those metagenomes[2,17]. Briefly, metagenomic DNA extracted from geothermal springs in Tengchong, China and Yellowstone National Park, USA, was sequenced using the Illumina NovaSeq, HiSeq3000, or HiSeq4000 platforms, employing paired-end or single-read configurations. All MAGs were subjected to quality checks with CheckM v1.0.12[69] and CheckM2 v1.0.1[70], and classification with GTDB Toolkit v.2.3.2[32]. This was followed by open reading frame (ORF) prediction using Prodigal v.2.6.3[71]. Protein sequences in the MAGs were annotated using eggnog-mapper v.2[72], and the results against the KEGG, COG, and arCOG databases were collected (Supplementary Data 15, Supplementary Data 21 and Supplementary Data 22). The NCBI Protein Basic Local Alignment Search Tool (BLASTp)[73] was used to identify enzymes involved in BCAA metabolism, including OFOR (Supplementary Data 17), ACD (Supplementary Data 18), 3-methylcrotonyl-CoA carboxylase (MCCC) (Supplementary Data 23), hydroxymethylglutaryl-CoA lyase (HMGCL) (Supplementary Data 24), and methylmalonyl-CoA epimerase (MCEE) (Supplementary Data 25). Transporters were also confirmed using the Transporter Classification Database[74] (Supplementary Data 26). Peptidases were identified using METABOLIC[75] using the MEROPS peptidase database (Supplementary Data 12)[76]. Ribosomal RNA sequences and tRNA sequences were identified (Supplementary Data 27 and Supplementary Data 28) with barnap v.0.9[77] and tRNAscan-SE v.2.0.9[78], respectively. Signal peptides were predicted using SignalP versions 6.0[79] and 5.0[80], while transmembrane domains were identified with DeepTMHMM[81].

Phylogenetic analyses of genes

Protein sequences from the class *Caldarchaeia* annotated as BCAA or PAA transporters were collected. The top five best-matched sequences were retrieved from the NR and Uniprot[57] databases using BLASTp[73] with an E-value threshold of 1E-10 and a minimum sequence coverage of 30%. The related sequences were clustered at 90% amino acid similarity using CD-HIT[82], then combined with the *Caldarchaeia* protein sequences and dereplicated. Sequences were aligned

ARTICLE IN PRESS

using MAFFT v.7.3[83] with the L-INS-I algorithm and trimmed with TrimAI v.1.4[84] in automated mode. Phylogenetic trees were constructed using IQ-TREE v.2.2.6[85] with the optimal substitution model identified by ModelFinder[86]. Branch support was assessed using Shimodaira-Hasegawa approximate likelihood ratio tests (SH-aLRT)[87] and ultrafast bootstrapping (UFBoot)[88], with each method performed using 1,000 replicates. The same approach was used to construct phylogenetic trees for other gene sequences, with appropriate sequence variations. For the 16S rRNA gene tree, sequences from *Caldarchaeia* longer than 940 bp were retrieved, and representative sequences were selected from reference[8]. For the phylogenetic tree used to identify *Calditenuis*, ASVs from BCAA experiments were included.

Phylogenomic analyses and ancestral character state reconstructions

All MAGs and SAGs classified as belonging to the family *Calditenuaceae*, 11 type genomes from the *Wolframiiraptoraceae*[19], and other publicly available genomes in the class *Caldarchaeia* were used for the construction of a species tree. As reference sequences, genomes for members of the *Nitrososphaerales* (syn. *Thaumarchaeota*) were included in the taxon set (Supplementary Data 20). Four marker gene/protein sets were used, including ar53[9], uni56[50], Rprot[51], and UACG[52]. Individual marker alignments were extracted and subjected to model testing using ProtTest v3.4.2[89] to identify the optimal evolutionary models. The alignments were then concatenated using FASconCAT-G v3.4.2[90], resulting in four concatenated, partitioned data matrices. Each partition (i.e., marker) within the data matrices was assigned the optimal evolutionary model (Supplementary Data 29-32). The concatenated data matrices were subjected to maximum-likelihood analysis using RAxML v. 8.2.12[91] or IQ-TREE v.2.2.6[85], with each partition having independently estimated model parameters and evolutionary rates. For RAxML, branch support was inferred from 1000 bootstrap pseudoreplicates. For IQ-TREE, branch support was determined through SH-aLRT[87] and UFBoot[88], with each method performed using 1,000 replicates.

Given similar topologies of the species trees constructed using ar53, uni56, and Rprot, a cladogram based on the ar53 species tree was selected as the foundation for reconstructing ancestral character states. To ensure accurate ancestral reconstruction, only high-quality genomes with $\geqslant 90\%$ completeness and $\leqslant 5\%$ contamination were selected. For analyses targeting genes related to BCAA and PAA transporters, a gene-tree-aware approach was employed using the

ARTICLE IN PRESS

ALEml_undated algorithm[58]. In contrast, for analyses focusing on the entire orthologous families, gain and loss events were estimated using a Wagner parsimony approach within the COUNT package[92].

Lab cultivation amended with BCAA, PAA, and BPAA

0.5 mL of G-10 culture (~25% *Calditenuis* relative abundance)[26] was used to inoculate 50 mL cultures amended with a final concentration of 2 mM BCAA, 2 mM PAA, or 2 mM each for both, and incubated at 80ºC for one week. DNA was extracted using the FastDNA Spin Kit for Soil (MP Biomedicals, Solon, OH, USA). DNA was shipped to Molecular Research DNA Lab (Texas, USA) for sequencing with the updated Earth Microbiome Project (EMP) 16S rRNA primer set[93,94]: 515F 5′-GTGYCAGCMGCCGCGGTAA-3′ and 806R 5′-GGACTACNVGGGTWTCTAAT-3′. Raw Illumina fastq files were processed using QIIME2 (version 2023.2)[95]. ASV tables were rarefied and normalized to the lowest sequencing depth.

Real-time quantitative PCR (qPCR)

Quantification of *Calditenuis* 16S rRNA genes was performed by qPCR using a CFX96 real-time PCR system (Bio-Rad, CA, USA). *Calditenuis*-specific primers targeting the V5 regions (*Calditenuis*_new_F 5′-TGCGGGCTAGCTGTCGGGG-3′ and *Calditenuis*_new_R 5′-TCGTCAGGTAGGGTCGTCAG-3′) were designed, and their specificity was verified using the NCBI Primer-BLAST tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and 16S rRNA gene sequences from *Calditenuaceae* genomes (Supplementary Data 27). These primers amplify a ~180 bp product and are specific to *Calditenuis* from GBS but not to *Calditenuis* from other locations. The standard curve was established using triplicate reactions of 10^3-10^9 copies/µL dilutions of the purified plasmid G04b_L4_A095[5], which contains a near-full-length *Calditenuis* 16S rRNA gene clone. Each 25 µL qPCR reaction mixture consisted of 12.5 µL of Power SYBR Green PCR Master Mix (Applied Biosystems, CA, USA), 1 µL of each primer (100 µM), 1 µL of template DNA, and 9.5 µL of water. The qPCR procedure was as follows: initial denaturation at 95°C for 15 minutes; followed by 40 cycles of 95°C for 15 seconds, 64°C for 30 seconds, and 72°C for 45 seconds. Fluorescence was detected during the 72°C step of each cycle. This was followed by a melt curve analysis from 60-95°C in 0.3°C increments (5 seconds each). Each sample was analyzed in triplicate.

ARTICLE IN PRESS

Stable isotope probing and nanoSIMS

To infer the metabolic activities of *Cal. ramacidaminiphagus*, stable isotope probing was conducted both in a natural setting at GBS in Nevada and in a lab setting on the G-10 cultures, followed by sample fixation in 1% paraformaldehyde for 1 hour on ice. Samples were spotted on indium tin oxide-coated slides (ITO SuperEpoxy 2, ArrayIt Corp) for nanoSIMS and CARD-FISH[19] with horseradish peroxidase labeled probe AigG1-1012 (5'-AGGTAGGGTCGTCAGCCCGA-3') and 40% formamide. These samples were sent to Lawrence Livermore National Lab (LLNL) for nanoSIMS analysis. ITO-coated slides were cut using a diamond saw to fit onto metal sample holders and analyzed with a CAMECA nanoSIMS 50 at Lawrence Livermore National Lab. A charge-coupled device (CCD) camera was used to locate fiducial markers and identify FISH-positive cells by capturing their X-Y coordinates with the real-time imaging unit. The primary $^{133}Cs^+$ ion beam was set to 1.5 pA, producing a beam diameter of approximately 100 nm at 16 keV. Rastering was conducted over 20 × 20 μm analysis areas with a dwell time of 1 ms per pixel for 19–30 scans (cycles), generating images at 256 × 256-pixel resolution. To ensure intracellular isotopic analysis without penetrating through the ITO coating, sputtering equilibrium was achieved by pre-sputtering with a 90-pA beam to a depth of ~10 nm. The secondary ion mass spectrometer was optimized for a mass resolving power of ~7000, and secondary electron images, along with quantitative secondary ion images, were simultaneously acquired for $^{12}C_2^-$, $^{13}C^{12}C^-$, $^{12}C^{14}N^-$, and $^{12}C^{15}N^-$ using individual electron multipliers in pulse counting mode. NanoSIMS datasets from two independent experimental runs were processed using L'Image (L. Nittler, Carnegie Institution of Washington, Washington, DC, USA) to apply deadtime and image shift corrections before generating $^{13}C^{12}C/^{12}C_2$ and $^{12}C^{14}N/^{12}C^{15}N$ ratio images, reflecting ^{13}C and ^{15}N incorporation into biomass. Regions of interest (ROIs) for isotopic ratio quantification were manually delineated around individual cells. A cell was considered isotopically enriched and thus anabolically active if its isotopic composition exceeded three standard deviations above the mean isotopic composition of unlabeled (non-*Calditenuis*) cells. While the exact substrates metabolized by *Cal. ramacidaminiphagus* are unknown, predictions from genomic data were used to inform lab experiment[12,21]. Therefore, a variety of isotopically labeled substrates were added to both lab and natural samples (Supplementary Fig. 1; Supplementary Data 2 and Supplementary Data 6). The ^{15}N-ammonium served as positive controls

ARTICLE IN PRESS

for detecting activity, while the substrates labeled with ^{13}C were used to analyze the metabolic capabilities of *Cal. ramacidaminiphagus*.

For natural samples, sediment and water slurry were collected close to the "B site" at GBS (Fig. 1b), a location with temperature and pH ranges of 77.5-79.4°C and 7.54-7.59, respectively. Two mL of the slurry was added to 8 mL of GBS spring water (previously sampled, filtered, autoclaved, and N_2-sparged) dosed with labeled substrates (Supplementary Data 2 and Supplementary Fig. 1) and incubated in 60 mL amber serum vials at 75°C. Samples were harvested at 4- and 24-hour time points over a three-day period from July 18-20, 2017. Post-incubation, cells were fixed in preparation for FISH and nanoSIMS analysis. Sediment-containing samples were centrifuged with Nycodenz to separate out larger particles, which interfere with nanoSIMS analysis. The treatment steps were as follows: 200 µL of fixed samples were gently layered on top of 200 µL of 80% Nycodenz in a 1.5 mL microcentrifuge tube, then centrifuged at 8,367 x g for 10 minutes. The upper sample layer and the sample-Nycodenz interface (approximately 230 µL) were harvested and transferred to a new 1.5 mL tube, mixed with DEPC water to a total volume of 500 µL, and centrifuged again at 8,367 x g for 5 minutes. The supernatant was removed and discarded without disturbing the pellet. The pellet was then washed twice with 500 µL 1x PBS, each time centrifuging at 8,367 x g for 5 minutes. Following the final wash, the pellet was resuspended in 50 µL of 50% ethanol in water and stored at -20°C for later use in CARD-FISH and nanoSIMS experiments.

Lab samples were derived from the existing G-10 cultures (Supplementary Fig. 3 and Supplementary Data 5)[26]. After the standard 3-week incubation period, labeled substrates were added and incubation continued for an additional 24 hours. Cultures were harvested at 6- and 24-hour time points, and cells were fixed in preparation for FISH and nanoSIMS analysis (Supplementary Data 6).

Computed structure models (CSMs) of CrOFORA and CrOFORB

AlphaFold 3 was used to generate a homodimeric computed structure model (CSM) of the *Cal. ramacidaminiphagus* CrOFOR using the AlphaFold 3 webserver (https://golgi.sandbox.google.com/)[45]. For comparison and control, a homodimeric CSM of *S. tokodaii* StOFOR2 was also included[43]. These structures, along with their corresponding predicted

ARTICLE IN PRESS

aligned error (PAE) plots, were visualized and edited using ChimeraX v.1.8[96]. To identify potential key active site residues, multiple sequence alignments of OforA and OforB with the StOFOR2 α-subunit and β-subunit, respectively, were created using the MAFFT online server v. 7 with the MAFFT-DASH L-INS-i algorithm (https://mafft.cbrc.jp/alignment/server/)[97,98]. Known active site residues of StOFOR2 were compared to the aligned residues from CrOFOR using BioEdit v. 7.7.1[99]. Additionally, the amino acid identity (AAI) between the α- or β-subunits of *Cal. ramacidaminiphagus* CrOFOR and *S. tokodaii* StOFOR2 was calculated using BLASTp[73].

Metaproteomics

Metaproteomics was conducted as in ref[26]. Briefly, the G-10 culture containing *Cal. ramacidaminiphagus* was grown in "spring water medium" with 0.02% m/v casamino acids and 10% v/v O_2 headspace for three weeks at 80°C. Cells were pelleted by centrifugation, resuspended in SDS lysis buffer, homogenized, incubated, and centrifuged. Proteins in the supernatant were precipitated with acetone, pelleted, washed, and stored before being sent to the UC Davis proteomics center. Protein concentrates were dried, solubilized, and subjected to clean-up, reduction, alkylation, and tryptic proteolysis via filter-aided sample preparation on suspension-trap devices. Data-independent analysis (DIA) was performed using UHPLC coupled to a timsTOF Pro in parallel accumulation-serial fragmentation mode. Liquid chromatography was conducted on a UHPLC nano-flow Easy nLC with a PepSep C18 column and a 45-minute gradient. The mass spectrometry acquisition included four 25 m/z precursor windows per 100 ms TIMS scan, covering 400 - 1200 m/z, with 16 TIMS scans creating 64 total windows. Raw files were processed with Spectronaut v. 18 using DirectDIA analysis mode. Peptides were matched against a custom database of high-quality genomes derived from the G-10 culture metagenome and isolate genomes from GBS, along with common lab contaminants. Analysis parameters included allowance for two missed cleavages, minimum peptide length of 7 amino acids, maximum peptide mass of 4,600 Da, fixed carbamidomethylation of cysteine residues, and variable modifications for methionine oxidation and protein N-terminal acetylation. Filtering criteria included a decoy FDR of less than 1% for peptide spectrum matches and protein group identifications. Decoy hits, contaminants, and single-site identifications were excluded from further analysis.

ARTICLE IN PRESS

Data availability

Genomes were submitted to eLMSG and/or IMG, with nomenclatural type genomes submitted to GenBank, under accession numbers listed in Supplementary Data 20. Raw sequence data for metagenomes from which type genomes were derived are available in the SRA under the run accessions SRX27892530 [https://www.ncbi.nlm.nih.gov/search/all/?term=SRX27892530], SRX27857028 [https://www.ncbi.nlm.nih.gov/search/all/?term=SRX27857028], SRX25997117 [https://www.ncbi.nlm.nih.gov/search/all/?term=SRX25997117], SRX29285230 [https://www.ncbi.nlm.nih.gov/search/all/?term=SRX29285230], SRX29285228 [https://www.ncbi.nlm.nih.gov/search/all/?term=SRX29285228], and SRX25997118 [https://www.ncbi.nlm.nih.gov/search/all/?term=SRX25997118]. Raw files for 16S rRNA gene sequencing are available in the SRA under BioProject accession number PRJNA1219692 [https://www.ncbi.nlm.nih.gov/search/all/?term=PRJNA1219692]. Metaproteomics (MSV000093641 [https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?task=c4115bc81d584c77a217ef174c8fa6ce]) data are available from the Center for Computational Mass Spectrometry's MassIVE repository via the ProteomeXchange Consortium. Source data are provided with this paper.

Code availability

Custom code is available on GitHub (https://github.com/alexximalayaunlv/Calditenuis_paper.git) and Zenodo with the access code DOI: 10.5281/zenodo.18097983. The Python scripts in this repository were used to: (1) process ALE results and (2) select genomes with specific marker proteins for genealogical concordance analysis using Densitree.

References

1. Amend JP, Shock EL. Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and Bacteria. *FEMS Microbiol Rev* **25**, 175-243 (2001).
2. Qi Y-L, *et al.* Analysis of nearly 3000 archaeal genomes from terrestrial geothermal springs sheds light on interconnected biogeochemical processes. *Nat Commun* **15**, 4066 (2024).
3. Guo L, *et al*. Hydrogeochemical constraints shape hot spring microbial community compositions: evidence from acidic, moderate-temperature springs and alkaline, high-temperature springs, southwestern Yunnan geothermal areas, China. *J Geophys Res Biogeosci* **126**, e2020JG005868 (2021).
4. Chiriac CM, *et al.* Differences in temperature and water chemistry shape distinct diversity patterns in thermophilic microbial communities. *Appl Environ Microbiol* **83**, e01363-17 (2017).
5. Costa KC, *et al.* Microbiology and geochemistry of great boiling and mud hot springs in the United States Great Basin. *Extremophiles* **13**, 447-459 (2009).
6. Lai D, *et al.* Resource partitioning and amino acid assimilation in a terrestrial geothermal spring. *ISME J* **17**, 2112-2122 (2023).

ARTICLE IN PRESS

7. Steen AD, *et al.* High proportions of bacteria and archaea across most biomes remain uncultured. *ISME J* **13**, 3126-3130 (2019).
8. Hedlund BP, *et al.* Uncultivated thermophiles: current status and spotlight on 'Aigarchaeota'. *Curr Opin Microbiol* **25**, 136-145 (2015).
9. Parks DH, *et al.* GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* **50**, D785-D794 (2022).
10. Nunoura T, *et al.* Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res* **39**, 3204-3223 (2011).
11. Yang Y, *et al.* The evolution pathway of ammonia-oxidizing archaea shaped by major geological events. *Mol Biol Evol* **38**, 3637-3648 (2021).
12. Hua Z-S, *et al.* Genomic inference of the metabolism and evolution of the archaeal phylum Aigarchaeota. *Nat Commun* **9**, 2832 (2018).
13. Abby SS, Kerou M, Schleper C. Ancestral reconstructions decipher major adaptations of ammonia-oxidizing archaea upon radiation into moderate terrestrial and marine environments. *mBio* **11**, e02371-20 (2020).
14. Luo Z-H, *et al.* Temperature, pH, and oxygen availability contributed to the functional differentiation of ancient Nitrososphaeria. *ISME J* **18**, wrad031 (2024).
15. Sheridan PO, *et al.* Gene duplication drives genome expansion in a major lineage of Thaumarchaeota. *Nat Commun* **11**, 5494 (2020).
16. Balbay MG, *et al.* Metabolic versatility of Caldarchaeales from geothermal features of Hawai'i and Chile as revealed by five metagenome-assembled genomes. *Front Microbiol* **14**, 1216591 (2023).
17. Colman DR, *et al.* Covariation of hot spring geochemistry with microbial genomic diversity, function, and evolution. *Nat Commun* **15**, 7506 (2024).
18. Nunoura T, *et al.* Genetic and functional properties of uncultivated thermophilic crenarchaeotes from a subsurface gold mine as revealed by analysis of genome fragments. *Environ Microbiol* **7**, 1967-1984 (2005).
19. Buessecker S, *et al.* An essential role for tungsten in the ecology and evolution of a previously uncultivated lineage of anaerobic, thermophilic Archaea. *Nat Commun* **13**, 3773 (2022).
20. Oren A, *et al.* Lists of names of prokaryotic Candidatus taxa.). Microbiology Society (2020).
21. Beam JP, *et al.* Ecophysiology of an uncultivated lineage of Aigarchaeota from an oxic, hot spring filamentous 'streamer'community. *ISME J* **10**, 210-224 (2016).
22. Mall A, *et al.* Reversibility of citrate synthase allows autotrophic growth of a thermophilic bacterium. *Science* **359**, 563-567 (2018).
23. Faulds JE, *et al.* Characterizing structural controls of geothermal fields in the northwestern Great Basin: A progress report. *Geothermal Resources Council Transactions* **30**, 69-76 (2006).
24. Thomas SC, *et al.* Position-specific metabolic probing and metagenomics of microbial communities reveal conserved central carbon metabolic network activities at high temperatures. *Front Microbiol* **10**, 1427 (2019).
25. Cole JK, *et al.* Sediment microbial communities in Great Boiling Spring are controlled by temperature and distinct from water communities. *ISME J* **7**, 718-729 (2013).
26. Nou NO, *et al.* Genome-guided isolation of the hyperthermophilic aerobe *Fervidibacter sacchari* reveals conserved polysaccharide metabolism in the Armatimonadota. *Nat Commun* **15**, 9534 (2024).
27. Colman DR, *et al.* Geobiological feedbacks and the evolution of thermoacidophiles. *ISME J* **12**, 225-236 (2018).
28. Sun Y, *et al.* Perspectives on cultivation strategies of archaea. *Microb Ecol* **79**, 770-784 (2020).
29. Behrens S, *et al.* Linking microbial phylogeny to metabolic activity at the single-cell level by using enhanced element labeling-catalyzed reporter deposition fluorescence in situ hybridization (EL-FISH) and NanoSIMS. *Appl Environ Microbiol* **74**, 3143-3150 (2008).
30. Seth EC, Taga ME. Nutrient cross-feeding in the microbial world. *Front Microbiol* **5**, 350 (2014).
31. Dodsworth JA, *et al. Thermoflexus hugenholtzii* gen. nov., sp. nov., a thermophilic, microaerophilic, filamentous bacterium representing a novel class in the *Chloroflexi*, *Thermoflexia* classis nov., and description of *Thermoflexaceae* fam. nov. and *Thermoflexales* ord. nov. *Int J Syst Evol Microbiol* **64**, 2119-2127 (2014).
32. Chaumeil P-A, *et al.* GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* **38**, 5315-5316 (2022).
33. Love M, Anders S, Huber W. Differential analysis of count data–the DESeq2 package. *Genome Biol* **15**, 550 (2014).
34. Lombard J, Moreira D. Origins and early evolution of the mevalonate pathway of isoprenoid biosynthesis in the three domains of life. *Mol Biol Evol* **28**, 87-99 (2011).
35. Nishimura H, *et al.* Biochemical evidence supporting the presence of the classical mevalonate pathway in the thermoacidophilic archaeon *Sulfolobus solfataricus*. *J Biochem* **153**, 415-420 (2013).

ARTICLE IN PRESS

ARTICLE IN PRESS

36. Paraiso JJ, *et al.* The distribution and abundance of archaeal tetraether lipids in US Great Basin hot springs. *Front Microbiol* **4**, 247 (2013).
37. Zeng Z, *et al.* Identification of a protein responsible for the synthesis of archaeal membrane-spanning GDGT lipids. *Nat Commun* **13**, 1545 (2022).
38. Zeng Z, *et al.* GDGT cyclization proteins identify the dominant archaeal sources of tetraether lipids in the ocean. *Proc Natl Acad Sci* **116**, 22505-22511 (2019).
39. Stark H, *et al.* Oxidative Stickland reactions in an obligate aerobic organism–amino acid catabolism in the Crenarchaeon *Sulfolobus solfataricus*. *FEBS J* **284**, 2078-2095 (2017).
40. Lewis AM, *et al.* The biology of thermoacidophilic archaea from the order *Sulfolobales*. *FEMS Microbiol Rev* **45**, fuaa063 (2021).
41. Bräsen C, Schönheit P. Regulation of acetate and acetyl-CoA converting enzymes during growth on acetate and/or glucose in the halophilic archaeon *Haloarcula marismortui*. *FEMS Microbiol Lett* **241**, 21-26 (2004).
42. Yokooji Y, *et al.* Genetic examination of initial amino acid oxidation and glutamate catabolism in the hyperthermophilic archaeon *Thermococcus kodakarensis*. *J Bacteriol* **195**, 1940-1948 (2013).
43. Yan Z, *et al.* Crystal structures of archaeal 2-oxoacid: ferredoxin oxidoreductases from *Sulfolobus tokodaii*. *Sci Rep* **6**, 33061 (2016).
44. Eram MS, *et al.* Molecular and biochemical characterization of bifunctional pyruvate decarboxylases and pyruvate ferredoxin oxidoreductases from *Thermotoga maritima* and *Thermotoga hypogea*. *J Biochem* **158**, 459-466 (2015).
45. Abramson J, *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024).
46. Mai X, Adams M. Purification and characterization of two reversible and ADP-dependent acetyl coenzyme A synthetases from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bacteriol* **178**, 5897-5903 (1996).
47. Awano T, *et al.* Characterization of two members among the five ADP-forming acyl coenzyme A (acyl-CoA) synthetases reveals the presence of a 2-(imidazol-4-yl) acetyl-CoA synthetase in *Thermococcus kodakarensis*. *J Bacteriol* **196**, 140-147 (2014).
48. Baker-Austin C, Dopson M. Life in acid: pH homeostasis in acidophiles. *Trends Microbiol* **15**, 165-171 (2007).
49. Rodrigues TA, *et al.* Copper II-polar amino acid complexes: toxicity to bacteria and larvae of *Aedes aegypti*. *An Acad Bras Cienc* **89**, 2273-2280 (2017).
50. Eloe-Fadrosh EA, *et al.* Global metagenomic survey reveals a new bacterial candidate phylum in geothermal springs. *Nat Commun* **7**, 10476 (2016).
51. Hug LA, *et al.* A new view of the tree of life. *Nat Microbiol* **1**, 1-6 (2016).
52. Na S-I, *et al.* UACG: Up-to-Date Archaeal Core Genes and Software for Phylogenomic Tree Reconstruction. *J Microbiol* **61**, 683-692 (2023).
53. Konstantinidis KT, Tiedje JM. Towards a genome-based taxonomy for prokaryotes. *J Bacteriol* **187**, 6258-6264 (2005).
54. Jain C, *et al.* High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* **9**, 5114 (2018).
55. Hedlund BP, *et al.* SeqCode: a nomenclatural code for prokaryotes described from sequence data. *Nat Microbiol* **7**, 1702-1708 (2022).
56. Whitman WB, *et al.* Development of the SeqCode: a proposed nomenclatural code for uncultivated prokaryotes with DNA sequences as type. *Syst Appl Microbiol* **45**, 126305 (2022).
57. Consortium U. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* **47**, D506-D515 (2019).
58. Szöllősi GJ, *et al.* Efficient exploration of the space of reconciled gene trees. *Syst Biol* **62**, 901-912 (2013).
59. Rodríguez del Río Á, *et al.* Functional and evolutionary significance of unknown genes from uncultivated taxa. *Nature* **626**, 377-384 (2024).
60. Baker BJ, Dick GJ. Omic approaches in microbial ecology: charting the unknown. *Microbe* **8**, 353-359 (2013).
61. Murphy CN, *et al.* Community microrespirometry and molecular analyses reveal a diverse energy economy in great boiling spring and sandy's spring west in the US great basin. *Appl Environ Microbiol* **79**, 3306-3310 (2013).
62. Palmer M, *et al.* Diversity and distribution of a novel genus of hyperthermophilic *aquificae* viruses encoding a proof-reading family-A DNA polymerase. *Front Microbiol* **11**, 583361 (2020).
63. Fernandes-Martins MC, *et al.* Ecological dichotomies arise in microbial communities due to mixing of deep hydrothermal waters and atmospheric gas in a circumneutral hot spring. *Appl Environ Microbiol* **87**, e01598-21 (2021).
64. Lu J, *et al.* Metagenome analysis using the Kraken software suite. *Nat Protoc* **17**, 2815-2839 (2022).

ARTICLE IN PRESS

ARTICLE IN PRESS

65. Wootton JC, Federhen S. Statistics of local complexity in amino acid sequences and sequence databases. *Comput Chem* **17**, 149-163 (1993).
66. Kolmogorov M, *et al.* metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat Methods* **17**, 1103-1110 (2020).
67. Nurk S, *et al.* metaSPAdes: a new versatile metagenomic assembler. *Genome Res* **27**, 824-834 (2017).
68. Chen I-MA, *et al.* IMG/M v. 5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res* **47**, D666-D677 (2019).
69. Parks DH, *et al.* CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**, 1043-1055 (2015).
70. Chklovski A, et al. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. Nat Methods 20, 1203–1212 (2023).
71. Hyatt D, *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform* **11**, 1-11 (2010).
72. Huerta-Cepas J, *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**, D309-D314 (2019).
73. Boratyn GM, *et al.* BLAST: a more efficient report with usability improvements. *Nucleic Acids Res* **41**, W29-W33 (2013).
74. Saier Jr MH, *et al.* The transporter classification database (TCDB): 2021 update. *Nucleic Acids Res* **49**, D461-D467 (2021).
75. Zhou Z, *et al.* METABOLIC: high-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. *Microbiome* **10**, 33 (2022).
76. Rawlings ND, *et al.* MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* **42**, D503-D509 (2014).
77. Seemann T. barrnap 0.9: rapid ribosomal RNA prediction. Available online at https://github.com/tseemann/barrnap (2013)
78. Chan PP, *et al.* tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res* **49**, 9077-9096 (2021).
79. Teufel F, *et al.* SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol* **40**, 1023-1025 (2022).
80. Almagro Armenteros JJ, *et al.* SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**, 420-423 (2019).
81. Hallgren J, *et al.* DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *BioRxiv*, 2022.2004. 2008.487609 (2022).
82. Fu L, *et al.* CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150-3152 (2012).
83. Katoh K, *et al.* MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* **33**, 511-518 (2005).
84. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973 (2009).
85. Nguyen L-T, *et al.* IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**, 268-274 (2014).
86. Kalyaanamoorthy S, *et al.* ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**, 587-589 (2017).
87. Guindon S, *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307-321 (2010).
88. Minh BQ, Nguyen MAT, Von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* **30**, 1188-1195 (2013).
89. Darriba D, *et al.* ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**, 1164–1165 (2011).
90. Kück P, Longo GC. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Front Zool* **11**, 1-8 (2014).
91. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313 (2014).
92. Csűös M. Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. *Bioinformatics* **26**, 1910-1912 (2010).

ARTICLE IN PRESS

ARTICLE IN PRESS

93. Parada AE, Needham DM, Fuhrman JA. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol* **18**, 1403-1414 (2016).
94. Apprill A, *et al.* Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat Microb Ecol* **75**, 129-137 (2015).
95. Hall M, Beiko RG. 16S rRNA gene analysis with QIIME2. *Microbiome analysis: methods and protocols*, 113-129 (2018).
96. Pettersen EF, *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **30**, 70-82 (2021).
97. Rozewicki J, *et al.* MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res* **47**, W5-W10 (2019).
98. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* **20**, 1160-1166 (2019).
99. Hall T, Biosciences I, Carlsbad C. BioEdit: an important software for molecular biology. *GERF Bull Biosci* **2**, 60-61 (2011).

Acknowledgments

We thank Dave and Sandy Jamieson for access to GBS and acknowledge that GBS is on the historical lands of the Cui Ui Ticutta (Cui-ui eaters) band of Northern Paiutes. We also thank the operators of Gongxiaoshe resort and hotel and the operators of Yunnan Tengchong Volcano and Spa Tourist Attraction Development Corporation. Part of this work was carried out at Lawrence Livermore National Lab (LLNL) under Contract DE-AC52-07NA2734 (X.M., J.J., P.W., J.P.-R). Funding was also provided by the U.S. National Science Foundation (DEB 1557042 and 2038420, D.L., D.M., M.P., X.M., J.J., J.K.C, C.O.S., J.A.D., W.S., B.P.H.), NASA (80NNSC17KO548 and 80NSSC25M0046, D.L., D.M., M.P., X.M., J.J., J.K.C, C.O.S., J.A.D., W.S., B.P.H.), the National Natural Science Foundation of China (Nos. 92251302, 32200007, and 32370011), and the Novo Nordisk Foundation (NNF24OC0089849). Data from Yellowstone National Park were generated with funding from a NASA grant to D.R.C. and E.S.B. (80NSSC19M0150), and with sample collection under the National Park Service permit YELL-SCI-5544.

Contributions

B.P.H., J.A.D., and D.L. conceived the study. D.M., J.A.D., W.S., and B.P.H. conducted field sampling. D.M., J.A.D., D.L., and B.P.H. conducted lab experiments. D.M., J.A.D., X.M., J.J., J. P.-R., P.K.W., and D.L. analyzed samples with nanoSIMS and performed nanoSIMS data analysis and interpretation. M.P., D.L., T.N., and B.P.H. conducted nomenclature and taxonomic work. J.-Y.J., L.L., D.R.C., M.P., D.L., E.S.B., Z.-S.H., and W.-J.L. were involved in metagenomic sampling, sequencing, assembly, and binning. J.K.C. conducted computed structure modeling. R.M conducted nitric oxide reductase screening. D.L. and C.O.S. conducted bioinformatics

ARTICLE IN PRESS

analyses. D.L. wrote the first drafts of the manuscript. B.P.H contributed to the writing and revision of manuscript. All authors contributed to the final version of the paper.

Competing interests

The authors have no competing interests to declare.

ARTICLE IN PRESS

ARTICLE IN PRESS

FIGURE LEGENDS:

Figure 1. Research location and in situ abundance of *Cal. ramacidaminiphagus*. a, Map showing the location of Great Boiling Spring (GBS) in the northwestern U.S. Great Basin. Map data sourced from Natural Earth (https://www.naturalearthdata.com/). **b,** Three sampling sites with indicated temperatures and the relative abundance of archaea and bacteria. **c,** Microbial community composition derived from sediment metagenomes at site A, site B, and site C ($n = 1$ per site). Source data are provided as a Source Data file.

Figure 2. ^{13}C and ^{15}N incorporation in single cells from GBS cultured with isotopically labeled substrates, as measured by nanoSIMS. a, Hybrid violin/box plots showing isotopic enrichments of *Cal. ramacidaminiphagus* (red) and other cells (grey). The violin plot shows the density distribution of all data points, whereas the box plot shows the median (central line) and 25th and 75th percentiles (box edges). The whiskers extend to the minimum and maximum values that are not classified as outliers. Outliers were identified using the Tukey method, defined as values beyond 1.5 × the interquartile range (IQR; the difference between the 25th and 75th percentiles). Each point reflects the enrichment of cellular ^{13}C or ^{15}N in a single cell. The number of independent cells in each group is denoted by *n*. Statistical significance was calculated using the two-sided Mann-Whitney U-test (*, $p < 0.01$). Asterisks indicate significant differences between unlabeled controls (asterisks below violins) and *Cal. ramacidaminiphagus* or other cells (asterisks above brackets) (red, $p < 0.01$; grey, $p < 0.05$). *Cal. ramacidaminiphagus* and other cells exhibited significantly higher ^{13}C and ^{15}N values compared to unlabeled cells in the G-10 culture. Exact *p*-values are provided in Supplementary Data 3. Isotope labeling and nanoSIMS experiments were conducted once. **b,** CARD-FISH imaging (upper left) of *Cal. ramacidaminiphagus* (Cy3, marked with a yellow circle), scanning electron microscope (SEM) image (lower left), and nanoSIMS ion ratio images reflecting ^{13}C assimilation (upper right) and ^{15}N assimilation (lower right) of representative *Cal. ramacidaminiphagus* cells and other cells for an amino acid experiment. DAPI, 4′,6-diamidino-2-phenylindole. Images are representative of 140 *Cal. ramacidaminiphagus* cells that were analyzed by nanoSIMS. Source data are provided as a Source Data file.

Figure 3. BCAA transporter gene enrichment, phylogeny, genomic contigs, and metabolism in *Calditenuaceae*. a, Volcano plot illustrating differentially abundant genes between *Calditenuis* and other genera in *Calditenuaceae*, determined by DEseq2 based on gene copy numbers. The horizontal dash-line corresponds to -log10 (adjusted $p = 0.05$). Significantly enriched genes are color-coded: BCAA transporter genes in orange and other genes in blue. **b,** BCAA transporter gene clusters in *Cal. ramacidaminiphagus*. For BCAA transporter proteins, signal peptides, transmembrane domains, and enzyme localizations were predicted using DeepTMHMM. Other

ARTICLE IN PRESS

protein domains within the BCAA transporter proteins were identified using HMMscan against the entire protein family (Pfam) database, setting a full-sequence E-value reporting threshold of 10^{-3} and a per-domain independent E-value threshold of 0.1. The Greek letters correspond to those in Fig. 7a. Abbreviations: *acsL*: long-chain acyl-CoA synthetase (K01897); *acadM*: acyl-CoA dehydrogenase (K00249); *alA*: alanine dehydrogenase (K19244); *idsA*: geranylgeranyl diphosphate synthase (K13787); maiA: maleate isomerase (K01799); *hmgcR*: hydroxymethylglutaryl-CoA reductase (K00021). **c,** Violin box plot showing the total number of peptidases in *Calditenuis* high-quality genomes (*n*=9) and other genera in *Calditenuaceae* (*n*=31). The violin outlines the distribution of values, the box bounds indicate the interquartile range (25^{th}–75^{th} percentiles), and the whiskers extend to the minimum and maximum values that are not classified as outliers. Statistical significance was calculated using the two-sided Mann-Whitney U-test ($p < 4.6E-06$). Source data are provided as a Source Data file.

Figure 4. BCAA metabolism in *Cal. ramacidaminiphagus* LZ^{Ts}. BCAA metabolism proposed based on KOs (black), BLASTp (green), and metaproteomics (red), with priority given as follows: metaproteomics > KOs and BLASTp. Leucine, isoleucine, and valine metabolism are colored pink, green, and yellow, respectively. Key enzymes involved in the production of branched-chain organic acids through overflow metabolism are represented by blue dots. Compounds with dashed box represent branched-chain organic acids.

Figure 5. BCAA specialization in *Calditenuis*. a, Mixed lab cultures were established with the addition of branched-chain amino acids (BCAA; orange), polar amino acids (PAA; purple), or both BCAA and PAA (brown), with each condition having three biological replicates (*n*=3), for seven days. The relative abundance of amplicon sequence variants (ASVs) after incubation is shown using boxplots, where the box bounds indicate the interquartile range (25^{th}–75^{th} percentiles) and the whiskers extend to the minimum and maximum values. **b,** Absolute abundance of *Calditenuis* 16S rRNA genes measured by qPCR), using the same cultures used in (a; three biological replicates (*n*=3). The grey color indicates the initial abundance (Time=0). Boxplot central marks show the median, with edges at the 25^{th} and 75^{th} percentiles, and error bars indicate the standard deviation of three technical replicates. Conditions sharing the same lowercase letter indicate no significant differences, as determined by one-way ANOVA followed by Tukey's post-hoc test ($p > 0.05$). Exact *p*-values are provided in Supplementary Data 19. These enrichment experiments were conducted once. Source data are provided as a Source Data file.

Figure 6. Genome relatedness indices (a), and phylogenomic reconstruction with geographical source (b) in the family *Calditenuaceae*. a, Average nucleotide identity (ANI) and amino acid identity (AAI) matrix for designated nomenclatural type genomes for each

ARTICLE IN PRESS

Calditenuaceae species. **b,** A maximum-likelihood phylogenomic reconstruction was based on the concatenation of 53 conserved archaeal marker proteins (ar53). The alignment was partitioned, and the appropriate substitution model was applied to each partition. Branch support for the phylogeny was determined using traditional bootstrapping with 1000 pseudoreplicates. Supported branches above 90% are indicated by filled circles at the nodes. Taxonomic groups at the species and genus levels were determined based on phylogenomic placement, ANI, AAI, and relative evolutionary divergence. Genome IDs in bold and marked with stars are proposed nomenclatural types. The genome ID in gray indicates a species that lacks a high-quality representative genome. The outer ring indicates the geographical locations. Source data are provided as a Source Data file.

Figure 7. BCAA metabolism and evolutionary history in *Calditenuaceae.* **a,** Phylogeny of *Calditenuaceae* BCAA substrate-binding proteins. Protein sequences annotated as BCAA substrate-binding proteins, along with sequences from the NCBI NR database and UniProt database, were clustered, aligned, and trimmed. A maximum-likelihood phylogenetic tree was constructed using IQ-tree based on the best-fit model (LG+F+R5) with UFBoot and SH-aLRT support values inferred from 1000 replicates each (see methods for details). BCAA SBP homologs in *Cal. ramacidaminiphagus* are shown using Greek letters corresponding to Fig. 3b. The inset table shows the closest reviewed UniProt hit to the *Cal. ramacidaminiphagus* LZ^{Ts} sequences (Q58662, *Methanocaldococcus jannaschii*) with alignment statistics (identity, positives, and E-value). **b,** Circles represent the presence or absence of specific genes in each species. Larger circles indicate that more than 50% of the genomes for a species contain the gene, while smaller circles indicate that 50% or fewer of the genomes contain the gene. 'Genome size' refers to the average estimated complete genome size of the species. Leucine, isoleucine, and valine metabolism are colored pink, green, and yellow, respectively. Source data are provided as a Source Data file.

Figure 8. Gene-tree-aware ancestral state reconstruction of BCAA and PAA transporter genes in the family *Calditenuaceae*. The phylogenetic tree, derived from the ar53 analysis, illustrates gene content evolution within the *Calditenuaceae*. Circles along the branches represent net changes in gene content, where brown to yellow hues indicate net gene losses, and pink to purple hues indicate net gene gains. Color intensity corresponds to the magnitude of change. Net change is calculated as the number of gene gains minus gene losses along each branch. Positive values reflect net gains, while negative values reflect net losses. Numbers inside the circles denote specific genes of interest, arranged as in the accompanying dot matrix. To the right, gene profiles indicate the average copy number per species, with darker shades representing higher copy numbers of genes encoding transporters for branched-chain amino acids (BCAA; orange) or polar amino acids (PAA; purple). Source data are provided as a Source Data file.

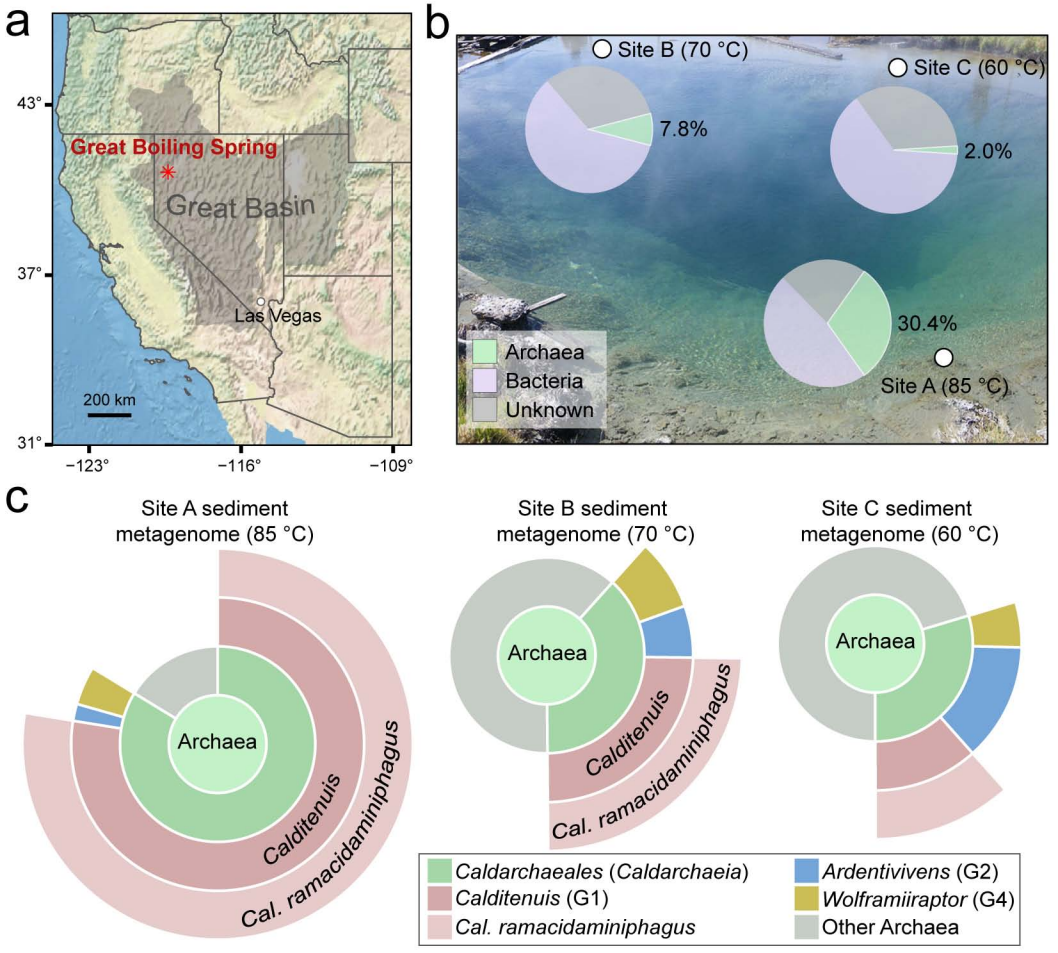
ARTICLE IN PRESS

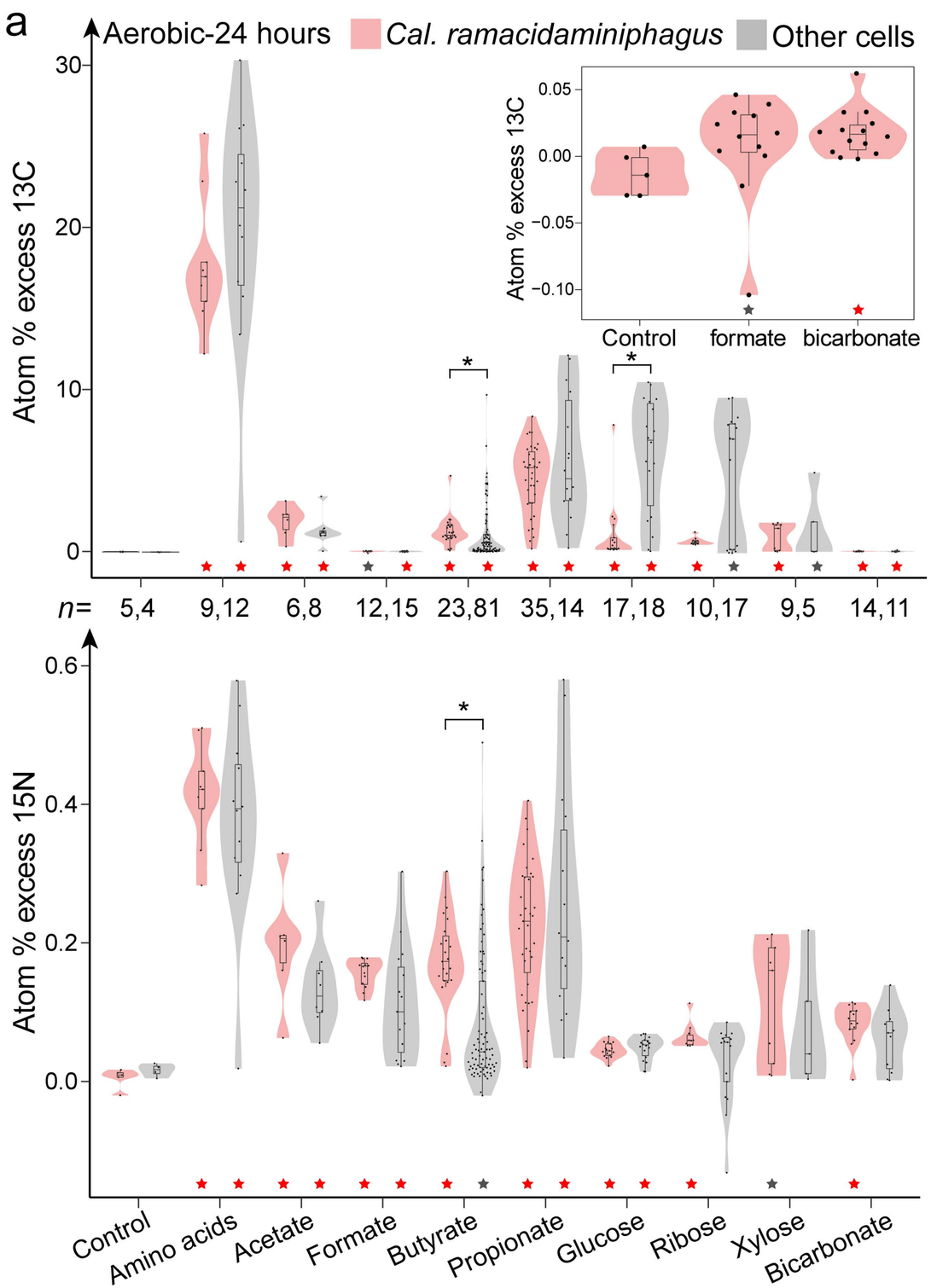
Editor's Summary

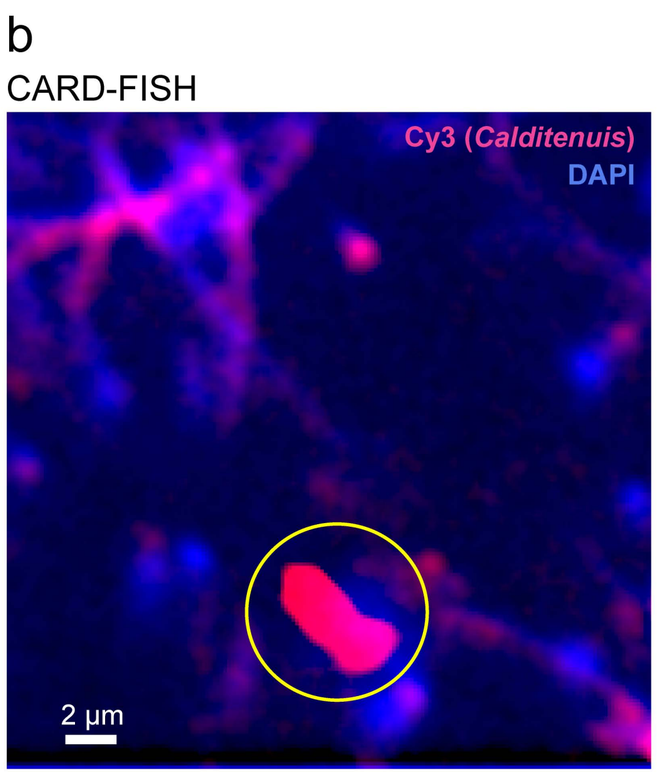
Many thermophiles that are abundant in geothermal systems have never been cultivated and are poorly understood. Here, Lai et al. describe the cultivation of one such organism, a deeply branching member of the archaeal phylum *Thermoproteota*, and provide evidence that it has evolved to specialize in branched-chain amino acid metabolism.

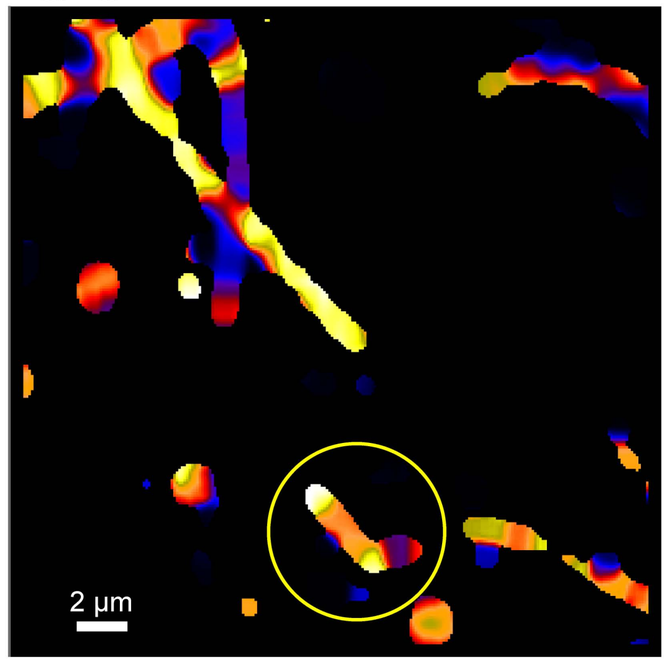
Peer Review Information: *Nature Communications* thanks Melina Kerou and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.

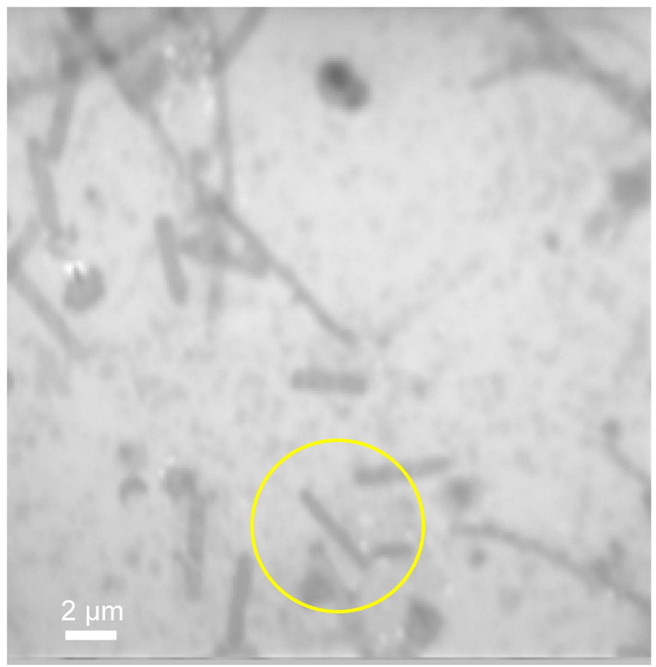
ARTICLE IN PRESS

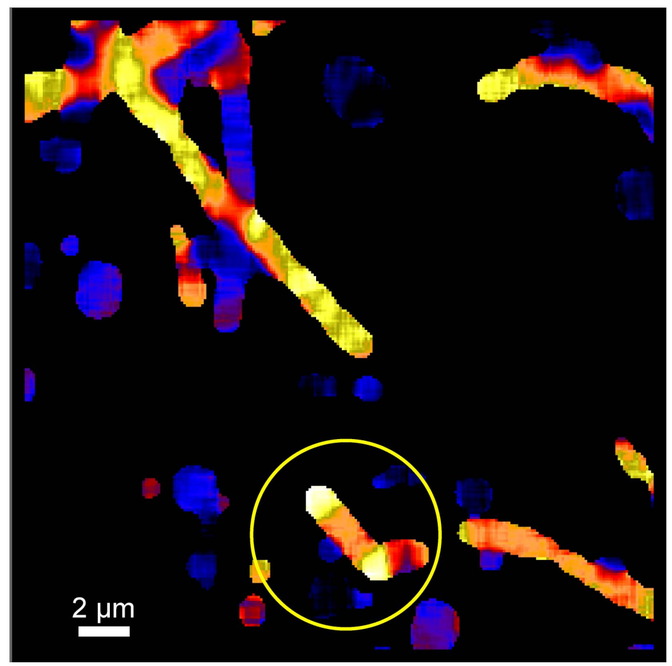

a
43°
37°
31°
−123°
−116°
−109°
Great Boiling Spring
Great Basin
Las Vegas
200 km
b
Site B (70 °C)
7.8%
Site C (60 °C)
2.0%
30.4%
Site A (85 °C)
Archaea
Bacteria
Unknown
c
Site A sediment metagenome (85 °C)
Site B sediment metagenome (70 °C)
Site C sediment metagenome (60 °C)
Archaea
Calditenuis
Cal. ramacidaminiphagus
Caldarchaeales (Caldarchaeia)
Calditenuis (G1)
Cal. ramacidaminiphagus
Ardentivivens (G2)
Wolframiiraptor (G4)
Other Archaea

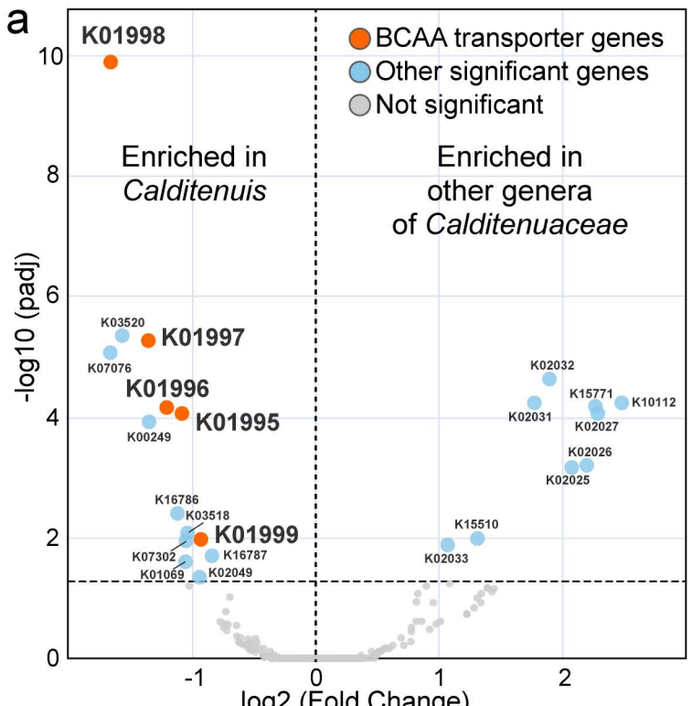

a
Aerobic-24 hours
Cal. ramacidaminiphagus
Other cells
Atom % excess 13C
30
20
10
0
Atom % excess 13C
0.05
0.00
−0.05
−0.10
Control
formate
bicarbonate
*
*
n= 5,4 9,12 6,8 12,15 23,81 35,14 17,18 10,17 9,5 14,11
Atom % excess 15N
0.6
0.4
0.2
0.0
*
Control
Amino acids
Acetate
Formate
Butyrate
Propionate
Glucose
Ribose
Xylose
Bicarbonate

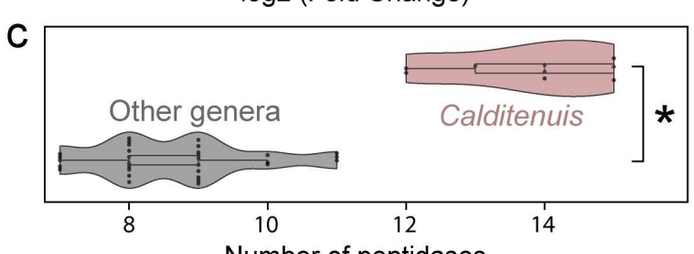

b
CARD-FISH
Cy3 (Calditenuis)
DAPI
2 µm

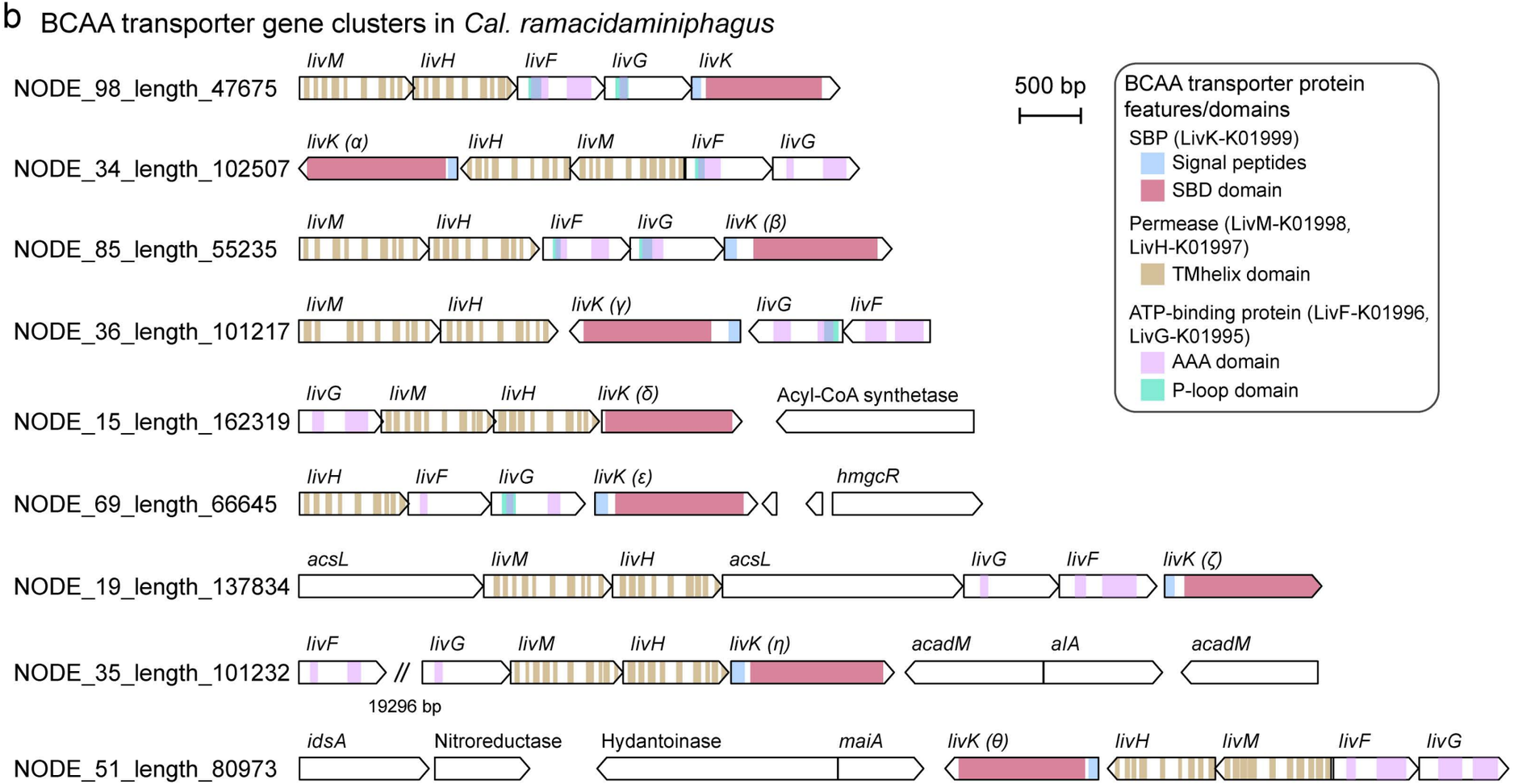

13C/12C
2 µm
40000
34285
28571
22857
17142
11428
5714
0

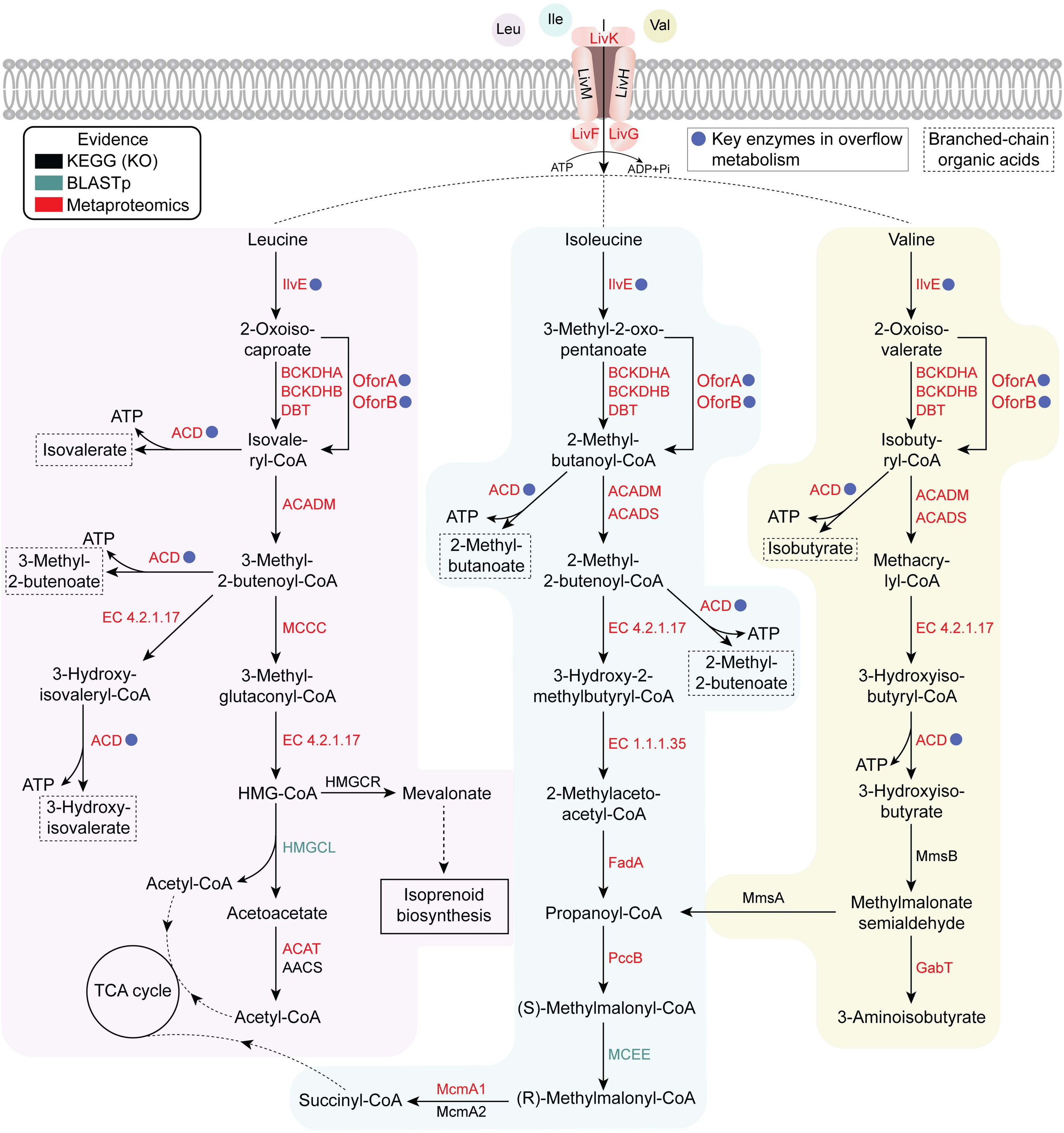

SEM
2 µm

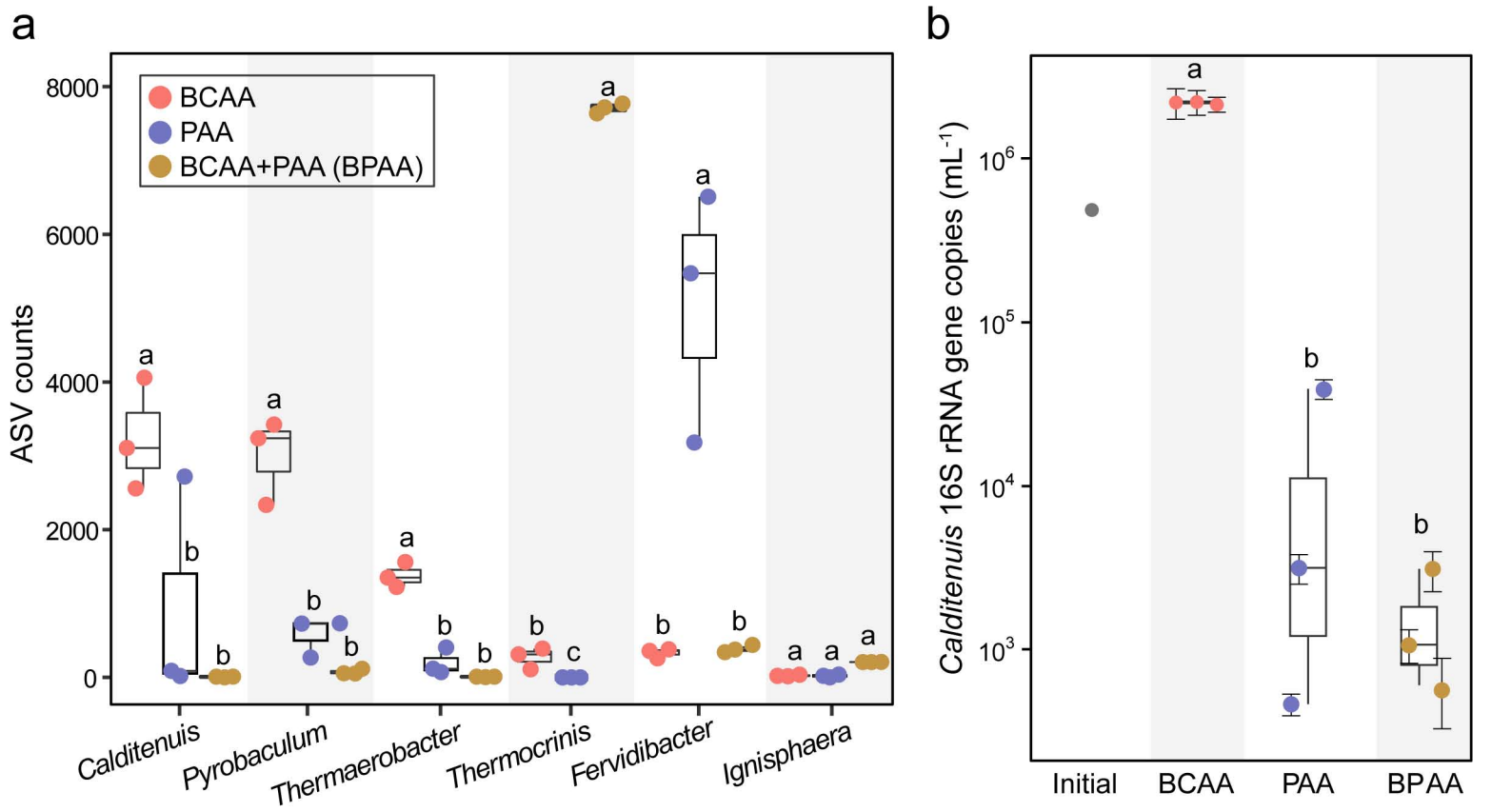

15N/14N
2 µm
2000
1714
1428
1142
857
571
285
0


a
K01998
BCAA transporter genes
Other significant genes
Not significant
Enriched in
Calditenuis
Enriched in
other genera
of Calditenuaceae
K03520
K01997
K07076
K01996
K01995
K00249
K16786
K03518
K01999
K07302
K16787
K01069
K02049
K02032
K15771
K10112
K02031
K02027
K02026
K02025
K15510
K02033
-log10 (padj)
log2 (Fold Change)

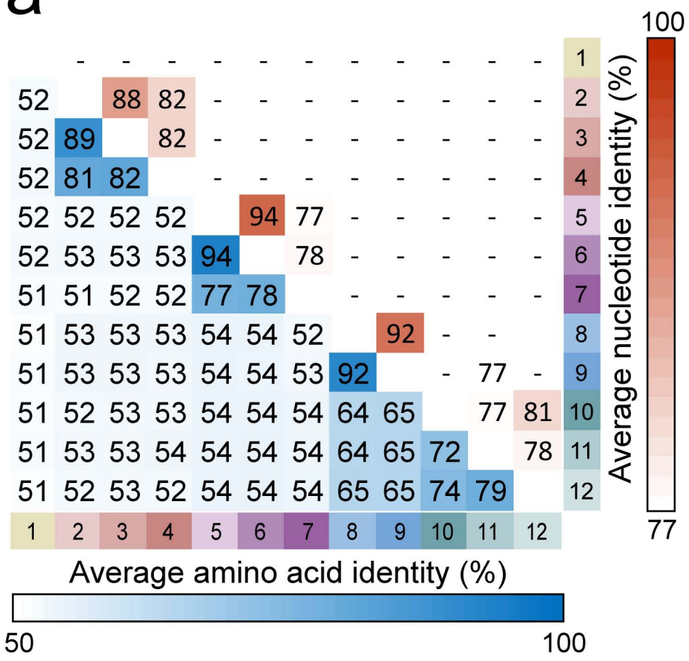

c
Other genera
Calditenuis
*
Number of peptidases


b BCAA transporter gene clusters in Cal. ramacidaminiphagus
NODE_98_length_47675
livM livH livF livG livK
NODE_34_length_102507
livK (α) livH livM livF livG
NODE_85_length_55235
livM livH livF livG livK (β)
NODE_36_length_101217
livM livH livK (γ) livG livF
NODE_15_length_162319
livG livM livH livK (δ) Acyl-CoA synthetase
NODE_69_length_66645
livH livF livG livK (ε) hmgcR
NODE_19_length_137834
acsL livM livH acsL livG livF livK (ζ)
NODE_35_length_101232
livF // livG livM livH livK (η) acadM aIA acadM
19296 bp
NODE_51_length_80973
idsA Nitroreductase Hydantoinase maiA livK (θ) livH livM livF livG
500 bp
BCAA transporter protein features/domains
SBP (LivK-K01999)
Signal peptides
SBD domain
Permease (LivM-K01998, LivH-K01997)
TMhelix domain
ATP-binding protein (LivF-K01996, LivG-K01995)
AAA domain
P-loop domain

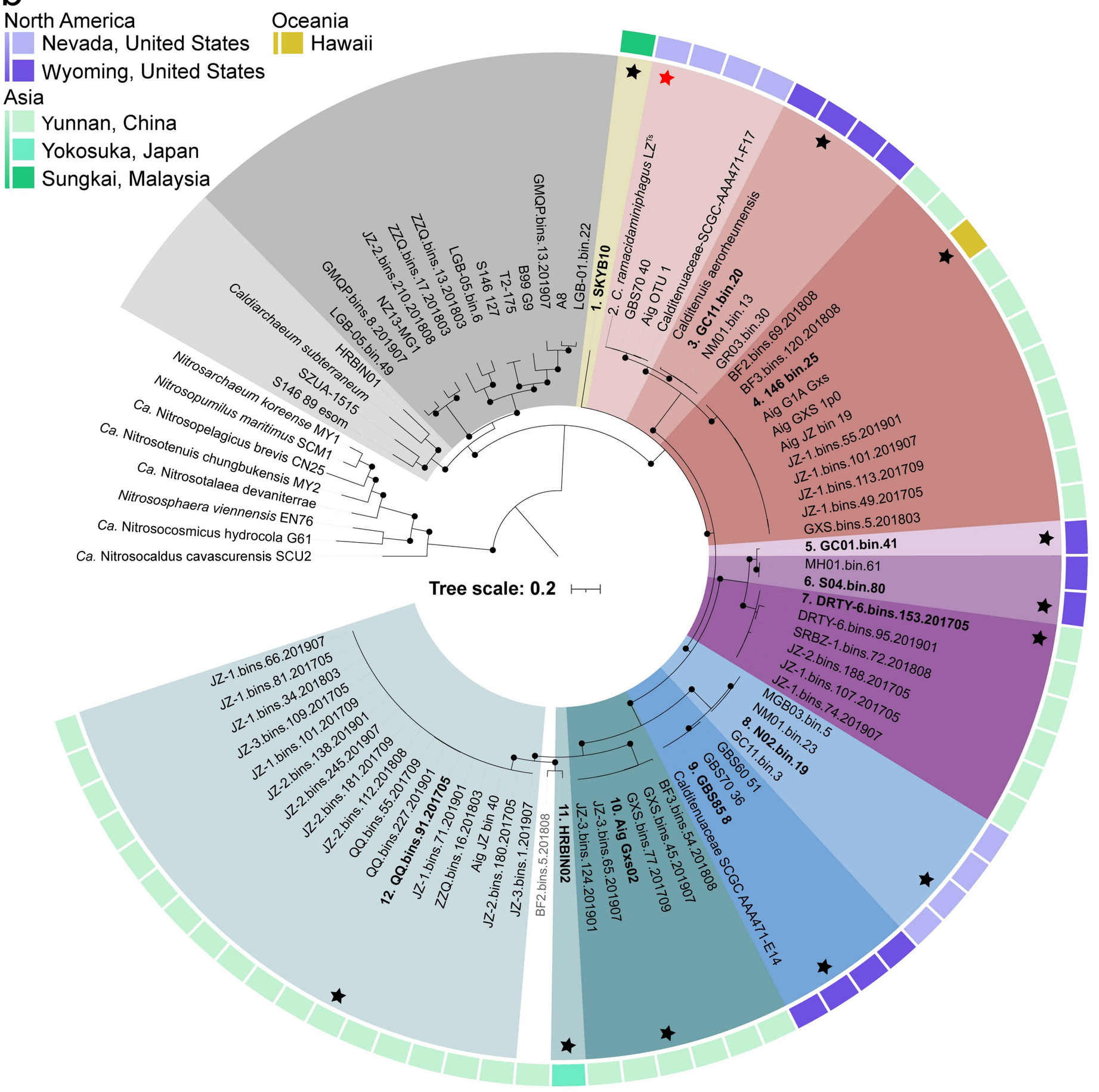

Leu
Ile
Val
LivK
LivM
LivH
LivF
LivG
ATP
ADP+Pi
Evidence
KEGG (KO)
BLASTp
Metaproteomics
Key enzymes in overflow metabolism
Branched-chain organic acids
Leucine
Isoleucine
Valine
IlvE
2-Oxoiso-caproate
3-Methyl-2-oxo-pentanoate
2-Oxoiso-valerate
BCKDHA
BCKDHB
DBT
OforA
OforB
ATP
ACD
Isovalerate
Isovale-ryl-CoA
2-Methyl-butanoyl-CoA
Isobuty-ryl-CoA
ACADM
ACADS
3-Methyl-2-butenoate
3-Methyl-2-butenoyl-CoA
2-Methyl-butanoate
2-Methyl-2-butenoyl-CoA
Isobutyrate
Methacry-lyl-CoA
EC 4.2.1.17
MCCC
2-Methyl-2-butenoate
3-Hydroxy-isovaleryl-CoA
3-Methyl-glutaconyl-CoA
3-Hydroxy-2-methylbutyryl-CoA
3-Hydroxyiso-butyryl-CoA
3-Hydroxy-isovalerate
HMG-CoA
HMGCR
Mevalonate
HMGCL
Acetyl-CoA
Acetoacetate
Isoprenoid biosynthesis
ACAT
AACS
TCA cycle
EC 1.1.1.35
2-Methylaceto-acetyl-CoA
FadA
Propanoyl-CoA
MmsA
MmsB
Methylmalonate semialdehyde
3-Hydroxyiso-butyrate
PccB
GabT
3-Aminoisobutyrate
(S)-Methylmalonyl-CoA
MCEE
(R)-Methylmalonyl-CoA
McmA1
McmA2
Succinyl-CoA


a
b
BCAA
PAA
BCAA+PAA (BPAA)
ASV counts
8000
6000
4000
2000
0
Calditenuis
Pyrobaculum
Thermaerobacter
Thermocrinis
Fervidibacter
Ignisphaera
Calditenuis 16S rRNA gene copies (mL-1)
10^6
10^5
10^4
10^3
Initial
BCAA
PAA
BPAA

a



b



Caldarchaeales (*Caldarchaeia*)

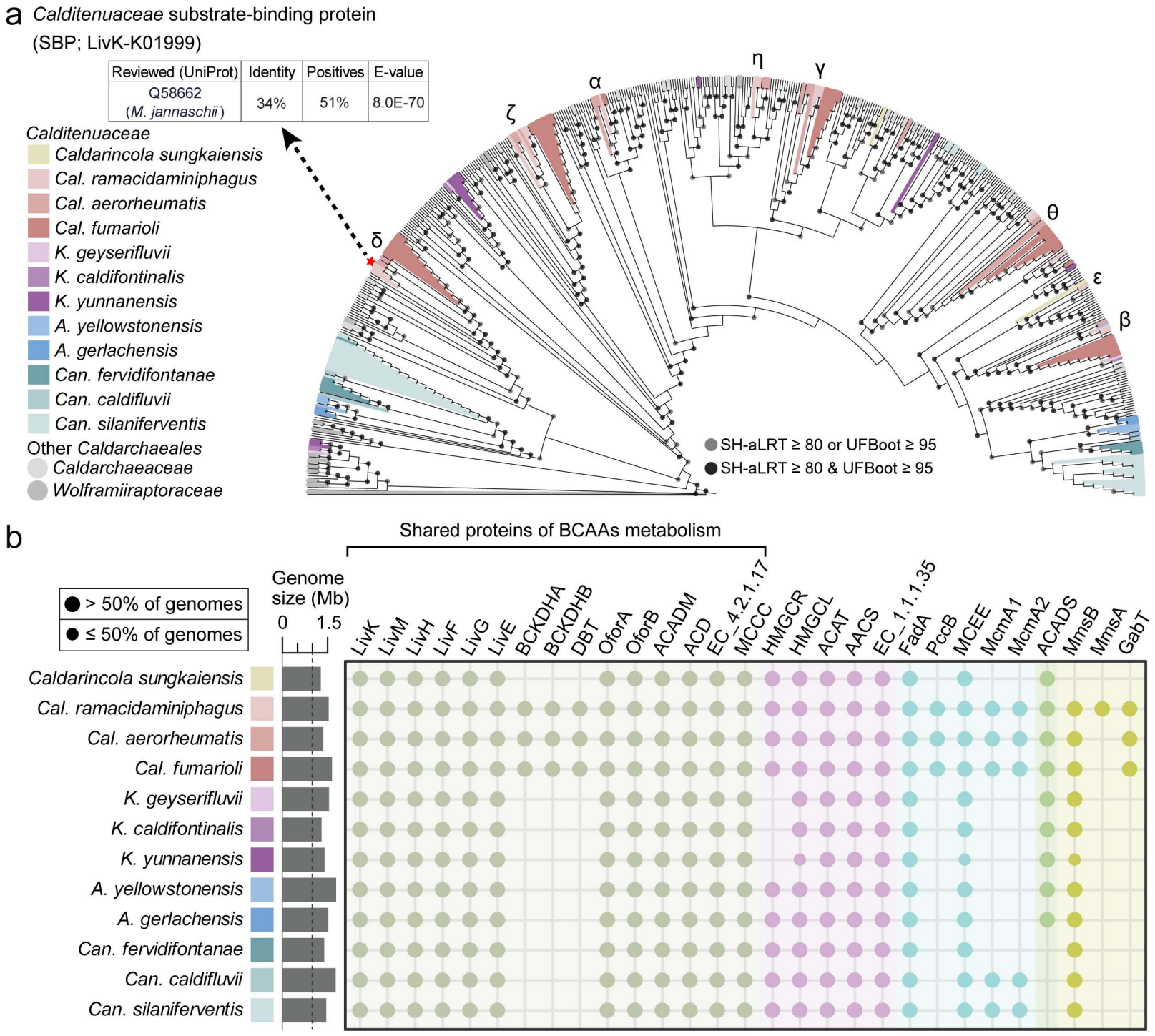
- *Caldarchaeaceae*
- *Wolframiiraptoraceae*

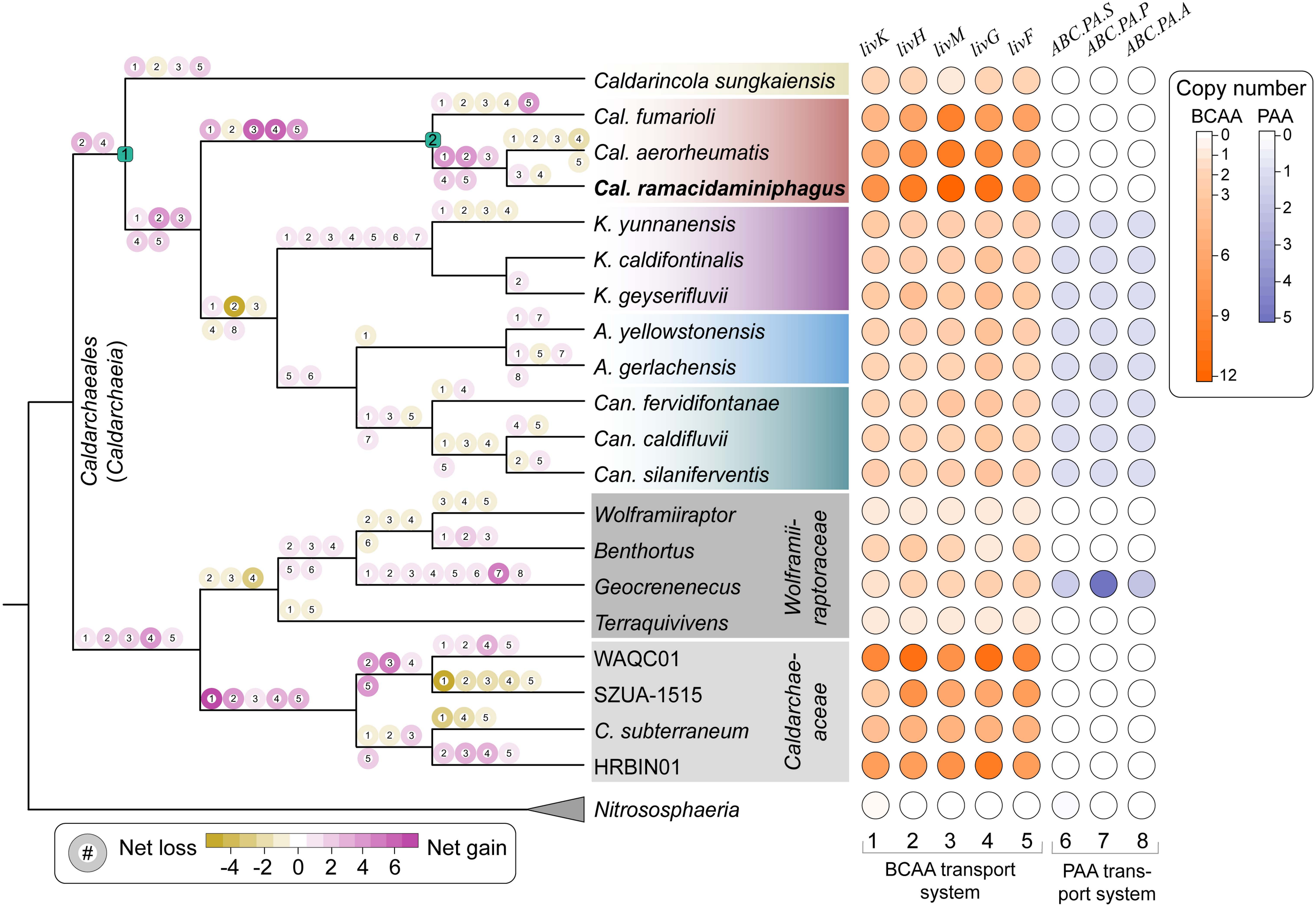
Calditenuaceae

1. *Caldaricola sungkaiensis*
2. *Calditenuis ramacidaminiphagus*
3. *Calditenuis aerorheumatis*
4. *Calditenuis fumarioli*
5. *Kaftiomonas geyserifluvii*
6. *Kaftiomonas caldifontinalis*
7. *Kaftiomonas yunnanensis*
8. *Ardentivivens yellowstonensis*
9. *Ardentivivens gerlachensis*
10. *Candentenecus fervidifontanae*
11. *Candentenecus caldifluvii*
12. *Candentenecus silaniferventis*

● Branch support > 90 %

★ Nomenclatural type genomes


a
Calditenuaceae substrate-binding protein
(SBP; LivK-K01999)
Reviewed (UniProt) | Identity | Positives | E-value
Q58662 (M. jannaschii) | 34% | 51% | 8.0E-70
Calditenuaceae
Caldarincola sungkaiensis
Cal. ramacidaminiphagus
Cal. aerorheumatis
Cal. fumarioli
K. geyserifluvii
K. caldifontinalis
K. yunnanensis
A. yellowstonensis
A. gerlachensis
Can. fervidifontanae
Can. caldifluvii
Can. silaniferventis
Other Caldarchaeales
Caldarchaeaceae
Wolframiiraptoraceae
α
β
γ
δ
ε
ζ
η
θ
SH-aLRT ≥ 80 or UFBoot ≥ 95
SH-aLRT ≥ 80 & UFBoot ≥ 95
b
Shared proteins of BCAAs metabolism
> 50% of genomes
≤ 50% of genomes
Genome size (Mb)
0
1.5
LivK
LivM
LivH
LivF
LivG
LivE
BCKDHA
BCKDHB
DBT
OforA
OforB
ACADM
ACD
EC_4.2.1.17
MCCC
HMGCR
HMGCL
ACAT
AACS
EC_1.1.1.35
FadA
PccB
MCEE
McmA1
McmA2
ACADS
MmsB
MmsA
GabT
Caldarincola sungkaiensis
Cal. ramacidaminiphagus
Cal. aerorheumatis
Cal. fumarioli
K. geyserifluvii
K. caldifontinalis
K. yunnanensis
A. yellowstonensis
A. gerlachensis
Can. fervidifontanae
Can. caldifluvii
Can. silaniferventis


livK
livH
livM
livG
livF
ABC.PA.S
ABC.PA.P
ABC.PA.A
Copy number
BCAA
PAA
0
1
2
3
6
9
12
0
1
2
3
4
5
Caldarincola sungkaiensis
Cal. fumarioli
Cal. aerorheumatis
Cal. ramacidaminiphagus
K. yunnanensis
K. caldifontinalis
K. geyserifluvii
A. yellowstonensis
A. gerlachensis
Can. fervidifontanae
Can. caldifluvii
Can. silaniferventis
Wolframiiraptor
Benthortus
Geocrenenecus
Terraquivivens
Wolframii-raptoraceae
WAQC01
SZUA-1515
C. subterraneum
HRBIN01
Caldarchae-aceae
Nitrososphaeria
Caldarchaeales (Caldarchaeia)
Net loss
-4 -2 0 2 4 6
Net gain
1 2 3 4 5
BCAA transport system
6 7 8
PAA trans-port system