Extremophiles (2018) 22:983–991
https://doi.org/10.1007/s00792-018-1055-2

ORIGINAL PAPER

Thermus sediminis sp. nov., a thiosulfate-oxidizing and arsenate-reducing organism isolated from Little Hot Creek in the Long Valley Caldera, California

En-Min Zhou[1,2] · Wen-Dong Xian[1] · Chrisabelle C. Mefferd[2] · Scott C. Thomas[2] · Arinola L. Adegboruwa[2] · Nathan Williams[2,3] · Senthil K. Murugapiran[2] · Jeremy A. Dodsworth[4] · Rakesh Ganji[2] · Meng-Meng Li[1] · Yi-Ping Ding[1] · Lan Liu[1] · Tanja Woyke[5,6] · Wen-Jun Li[1] · Brian P. Hedlund[2,7]

Received: 19 May 2018 / Accepted: 6 September 2018 / Published online: 15 September 2018

Abstract

Thermus species are widespread in natural and artificial thermal environments. Two new yellow-pigmented strains, L198^T and L423, isolated from Little Hot Creek, a geothermal spring in eastern California, were identified as novel organisms belonging to the genus *Thermus*. Cells are Gram-negative, rod-shaped, and non-motile. Growth was observed at temperatures from 45 to 75 °C and at salinities of 0–2.0% added NaCl. Both strains grow heterotrophically or chemolithotrophically by oxidation of thiosulfate to sulfate. L198^T and L423 grow by aerobic respiration or anaerobic respiration with arsenate as the terminal electron acceptor. Values for 16S rRNA gene identity (≤97.01%), digital DNA–DNA hybridization (≤32.7%), OrthoANI (≤87.5%), and genome-to-genome distance (0.13) values to all *Thermus* genomes were less than established criteria for microbial species. The predominant respiratory quinone was menaquinone-8 and the major cellular fatty acids were iso-$C_{15:0}$, iso-$C_{17:0}$ and anteiso-$C_{15:0}$. One unidentified phospholipid (PL1) and one unidentified glycolipid (GL1) dominated the polar lipid pattern. The new strains could be differentiated from related taxa by β-galactosidase and β-glucosidase activity and the presence of hydroxy fatty acids. Based on phylogenetic, genomic, phenotypic, and chemotaxonomic evidence, the novel species *Thermus sediminis* sp. nov. is proposed, with the type strain L198^T (=CGMCC 1.13590^T=KCTC XXX).

Keywords *Thermus sediminis* sp. nov. · Thermophile · Polyphasic taxonomy · Genome sequencing · Geothermal springs · Little Hot Creek

Abbreviations

ANI	Average nucleotide identity
dDDH	Digital DNA–DNA hybridization
GGDC	Genome-to-genome distance calculator
DOE	Department of Energy
JGI	Joint Genome Institute
IMG	Integrated microbial genomes
CTAB	Cetyl trimethyl ammonium bromide
CRISPRs	Clustered regularly interspaced short palindromic repeats
COGs	Clusters of orthologous groups

Communicated by A. Oren.

En-Min Zhou, Wen-Dong Xian and Chrisabelle C. Mefferd contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00792-018-1055-2) contains supplementary material, which is available to authorized users.

✉ Wen-Jun Li
liwenjun3@mail.sysu.edu.cn

✉ Brian P. Hedlund
brian.hedlund@unlv.edu

Extended author information available on the last page of the article

Introduction

Brock and Freeze (1969) described *Thermus aquaticus*, which was isolated from the hot springs of Yellowstone National Park in the USA. Thereafter, many species of the genus *Thermus* have been isolated and described from various thermal environments worldwide. *Thermus* inhabits both natural and artificial thermal environments, including terrestrial geothermal areas, hot water taps, self-heating compost manure, and deep mines (Albuquerque et al. 2018). The genus *Thermus* currently includes 15 species with validly

published names, with the biotechnologically exalted *T. aquaticus* as the type species. *Thermus* species are of special interest with regard to biotechnology, horizontal gene transfer, and denitrification, and have been studied extensively (Blesa et al. 2018; Hedlund et al. 2011; Liu et al. 2015; Zhou et al. 2016). Their enzymes exhibit thermostability and resistance to other physical and chemical factors, which is important for biotechnological applications (Pantazaki et al. 2002; Vieille and Zeikus 2001). In addition, the high growth yield under laboratory conditions and genetic tractability of some strains has led to the use of some *Thermus* strains as models for basic research, particularly strains of *T. thermophilus* (Sazanov and Hinchliffe 2006; Yusupov et al. 2001).

Little Hot Creek (LHC), located on the eastern edge of the resurgent dome of the Long Valley Caldera near Mammoth Lakes, CA, USA, consists of several small circumneutral springs with different temperatures and discharge rates. Compared to other hydrothermal systems of the United States, such as Yellowstone National Park, the microbiology of LHC and other hot springs in the endorheic Great Basin have received little attention (Connon et al. 2008; Costa et al. 2009). A previous cultivation-independent 16S rRNA gene survey revealed moderately diverse, but highly novel microbial communities in the three distinct sources of LHC, including apparently abundant populations of *Aquificae*, *Thermodesulfobacteria*, and *Chloroflexi* (Vick et al. 2010). Phylotypes affiliated with phylum *Thermus–Deinococcus* represented 3% of the total bacterial sequences, and half of those were most closely related to *T. aquaticus* (Vick et al. 2010). We here describe the isolation and characterization of *Thermus* strains from LHC sediments, including novel strains L198^T and L423.

Materials and methods

Sample collection and *Thermus* strains isolation

Water and sediment samples were collected aseptically from LHC1 hot spring (GPS location N 37°41.436′, W 118°50.664′), in the Long Valley Caldera, located near Mammoth Lakes, CA, USA (Vick et al. 2010). The source temperature of LHC1 is about 80 °C and the water flows quickly into an outflow channel and dissipates into a marsh. One sediment sample, LHC1-D, was taken from the spring source. Another sediment, LHC1-H, was taken from the outflow channel of LHC1 at approximately 65 °C. The samples were transported to the laboratory in the dark and without temperature control. *Thermus* strains were obtained by standard serial dilution plating on *Thermus* medium (Castenholz 1969), and incubation at 65 °C, and presumptive *Thermus* colonies were restreaked > 3 times to obtain axenic cultures. All strains were routinely cultivated on *Thermus* medium at 65 °C unless otherwise stated. The strains were maintained as glycerol suspensions (20%, w/v) at − 80 °C. Strain L198^T was deposited in China General Microbiological Culture Collection Center (CGMCC) and Korean Collection for Type Cultures (KCTC) with the number CGMCC 1.13590^T and KCTC XXX, respectively.

16S rRNA gene analysis and genome sequencing, assembly, annotation and comparation

Genomic DNA for 16S rRNA gene PCR was extracted and the near-complete 16S rRNA genes were amplified and sequenced using methods described previously (Li et al. 2007). The EzBioCloud database was used to identify closest related taxa and to determine an approximate phylogenetic affiliation (Yoon et al. 2017). The 16S rRNA gene sequences of strains L198^T and L423 were deposited in GenBank under accession number LC382246 and LC382245, respectively. CLUSTAL_X was used to align the 16S rRNA gene sequences of strains L198^T, L423 and type strains of species of the genus *Thermus* (Thompson et al. 1997). Gaps at the 5′ and 3′ ends of the alignment were omitted from the analysis. Phylogenetic analyses were performed using the MEGA software package version 7.0 (Kumar et al. 2016) with neighbor-joining (Saitou and Nei 1987), maximum-likelihood (Felsenstein 1981) and maximum-parsimony (Fitch 1971) algorithms. The topology of the phylogenetic trees was evaluated by the bootstrap resamplings method with 1000 resamplings (Felsenstein 1985).

DNA for genome sequencing was isolated using the Joint Genome Institute's CTAB bacterial genomic DNA isolation protocol (http://my.jgi.doe.gov/general). The draft genome of strain L198^T was generated at the DOE's JGI using Pacific Biosciences sequencing technology (Eid et al. 2009). All raw reads were assembled using HGAP version 2.3.0 (Chin et al. 2013). The final draft assembly produced four contigs on four scaffolds, totaling 2,160,271 bp in size and with a mol% G + C of 68.2. The genome was annotated using the JGI microbial genome annotation pipeline (Mavromatis et al. 2009). Genes were identified using Prodigal (Hyatt et al. 2010) and manually curated with the JGI GenePRIMP pipeline (Pati et al. 2010). The predicted coding sequences were translated and used to search against the integrated microbial genomes (IMG) non-redundant database, Pfam, KEGG, COG, and InterPro databases to annotate predicted protein-coding genes. rRNA genes were predicted using hmmsearch tool from HMMER 3.0 (Eddy 2011); tRNA genes were identified using tRNAscan-SE 1.3.1 (Lowe and Eddy 1997); other non-coding RNAs and regulatory RNA features were found by searching the genome for the corresponding Rfam profiles using INFERNAL 1.0.2 (Nawrocki et al. 2009).

The draft genome sequence has been deposited at DDBJ/ENA/GenBank under the accession QURO00000000. The version described in this paper is version QURO01000000. Average nucleotide identity (ANI) and genome-to-genome distance (GGD) were calculated using the OrthoANI calculator (Lee et al. 2016). Digital DNA–DNA hybridization analysis was performed on the DSMZ genome-to-genome distance calculator platform (Meier-Kolthoff et al. 2013; http://ggdc.dsmz.de/distcalc2.php). To assess relationships between the sequenced *Thermus* strains, we performed a phylogenomic analysis and constructed the evolutionary tree based on single-copy marker genes using ezTree software with the default parameters (Wu et al. 2018). The bacterial pan genome analysis (BPGA) tool was employed to conduct an additional core- and pan genome analysis for comparative purposes with default settings (Chaudhari et al. 2016).

Morphological, physiological and chemotaxonomic characteristics

To determine the optimal growth conditions, strains L198^T and L423 were also cultured on R2A and T5 agar medium (Yu et al. 2013). Growth at different temperatures (30–80 °C), pHs [4–10, using the buffer system described by Xu et al. (2005)], and salt concentrations (0–5%, w/v NaCl), were tested with *Thermus* medium. All tests were performed at 65 °C, except when testing for the optimal growth temperature and temperature range. Gram staining was performed using the standard Gram reaction and was confirmed using the KOH lysis test method (Cerny 1978). Cell morphology of strains L198^T and L423 was examined by phase-contrast microscopy (BX51, Olympus) and scanning electron microscopy (JSM-6700F, JEOL) after growth in liquid *Thermus* medium for 48 h at 65 °C. Nitrate, ferric nitrilotriacetate (NTA), ferric citrate, and sulfur were evaluated as possible terminal electron acceptors in *Thermus* medium under anaerobic conditions as described by Kieft et al. (1999). The oxidation of thiosulfate was determined as described by Skirnisdottir et al. (2001). Arsenite oxidation and arsenate respiration were evaluated as described by Gihring and Banfield (2001). Catalase activity was determined by the formation of bubbles in 3% (v/v) H_2O_2 and oxidase activity was detected by the oxidation of tetramethyl-*p*-phenylenediamine. Single-carbon-source utilization tests were performed in modified minimal *Thermus* 162 medium without yeast extract and tryptone, and containing 0.05% NH_4Cl (Bjornsdottir et al. 2009; Vajna et al. 2012). The filter-sterilized (0.22 µm, Millipore) carbon sources were added in concentrations of 0.2% (w/v). Hydrolysis of starch, carboxymethyl cellulose, casein, and Tweens 20, 40, 60 and 80 was tested as described by Yu et al. (2013). The API ZYM, API 20NE (BioMérieux, France), and GEN III microplates (Biolog Inc., Hayward, CA, USA) were used to evaluate other enzyme activities and to determine additional physiological characters according to the manufacturer's instructions and Ming et al. (2014). Extraction and analysis of polar lipids was performed according to the method described by Minnikin et al. (1979) and Collins and Jones (1980). Quinones were extracted and purified as described by Collins et al. (1977) and analyzed by HPLC (Groth et al. 1997). The biomass for fatty acid analysis was harvested after 24 h of growth under standard conditions and analysis following the standard protocol of the microbial identification system (Agilent Technologies 7890A gas chromatograph; Sherlock Version 6.1; MIDI database: TSBA6).

Results and discussion

Isolation and diversity of *Thermus* strains in Little Hot Creek

In total, 19 *Thermus* isolates were obtained from LHC1-D and LHC1-H. Preliminary analysis based on the 16S rRNA gene sequence comparison indicated 15 isolates were most closely related to *T. aquaticus* (97.89–97.93% 16S rRNA gene identity to *T. aquaticus* YT-1^T); two isolates were related to *T. oshimai* (99.87% 16S rRNA gene identity with *T. oshimai* DSM 12092^T); the other two isolates, L198^T from LHC1-H and L423 from LHC1-D, were almost identical (99.9% 16S rRNA gene identity), and shared relatively low 16S rRNA gene identity to existing *Thermus* type strains (97.01% with *T. islandicus* DSM 21543^T, 96.79% with *T. igniterrae* RF-4^T, 96.73% with *T. aquaticus* YT-1^T, 96.66% with *T. arciformis* CGMCC 1.6992^T, 96.11% with *T. composti* DSM 21686^T; < 96% for all other type strains). The phylogeny based on the maximum-likelihood method (Fig. 1) showed clustering of strains L198^T and L423 on a branch with type strains of *T. islandicus* and *T. composti*. Although bootstrap support was low for each method, the phylogenetic position of both strains was consistent in both neighbor-joining and maximum-parsimony phylogenies (supplementary Figure S1 and Figure S2). This suggested that strains L198^T and L423 are close relatives of *T. islandicus* DSM 21543^T and *T. composti* DSM 21686^T. To further investigate the taxonomic position and describe the novel strains, *T. islandicus* DSM 21543^T and *T. composti* DSM 21686^T were obtained from DSMZ and cultured under the same conditions as appropriate for specific comparative tests.

Phenotypic characters and chemotaxonomy of novel strains

Strains L198^T and L423 could growth on *Thermus* medium, R2A medium, and T5 agar. Both strains formed yellow

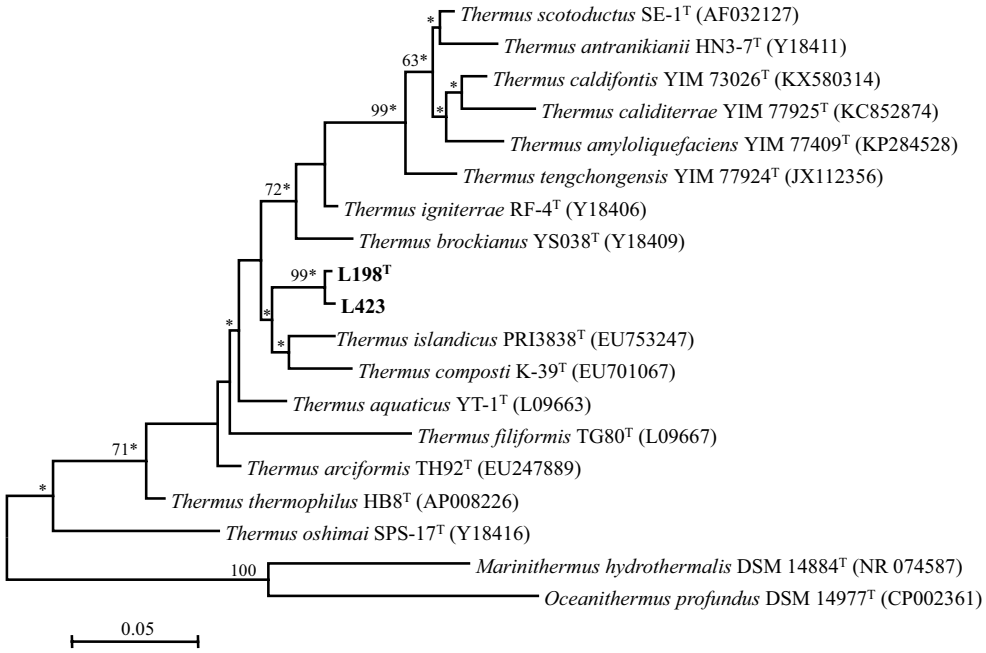

Fig. 1 Maximum-likelihood phylogenetic tree based on 16S rRNA gene sequences showing the position of strains L198^T and L423-2 within the genus *Thermus*. Sequences of *Marinithermus hydrothermalis* DSM 14884^T (NR 074587) and *Oceanithermus profundus* DSM 14977^T (CP002361) were used as outgroups. Bootstrap values (expressed as a percentage of 1000 replications) greater than 50% are given at the respective branching points. Bar: five substitutions per 100 nucleotide position. Asterisks denote nodes that were also recovered using the neighbor-joining and maximum-parsimony methods

colonies and were rod-shaped (2.0–5.0 µm in length and 0.2–0.5 µm in width, supplementary Figure S3) and non-motile. The growth of both strains occurred at salinities of 0–1.5% (w/v) added NaCl. Weak growth was also observed at salinities of 2.0% (w/v) added NaCl for strain L423. The new isolates could grow aerobically or with arsenate (1 mM) as a terminal electron acceptor under anaerobic conditions. Anaerobic respiration of arsenate is consistent with relatively high As concentrations (> 9 µM of total As) in LHC and nearby Hot Creek (Vick et al. 2010; Wilkie and Hering 1998) and high rates of As cycling in these streams (Wilkie and Hering 1998). The detailed physiological and biochemical characteristics of strain L198^T and L423 are given in the species description and Table 1. The chemotaxonomic features of strain L198^T were similar to that of other species of *Thermus* (Da Costa et al. 2006). The predominant respiratory quinone was menaquinone MK-8 (98%), with menaquinone MK-7 (2%) as the minor fraction. Major cellular fatty acids of strain L198^T were iso-$C_{15:0}$, iso-$C_{17:0}$, and anteiso-$C_{15:0}$. The overall fatty acid composition of strains L198^T and L423 were very similar (supplementary Table S1), and broadly similar to that of *T. islandicus* DSM 21543^T; however, iso-$C_{15:0}$ was more abundant in strains L198^T and L423 than in *T. islandicus* DSM 21543^T, and the opposite was true for anteiso-$C_{15:0}$. The fatty acid composition of *T. composti* DSM 21686^T was distinct from the other three strains, especially with respect to higher relative abundance of iso-$C_{17:0}$ and anteiso-$C_{17:0}$. Hydroxy fatty acids were detected as minor fractions only in strains L198^T and L423, but not in strains DSM 21543^T and DSM 21686^T. Low levels of branched chain 3-hydroxy fatty acids were also present in some strains of *T. aquaticus*, *T. filiformis*, and *T. tengchongensis* (Carreto et al. 1996; Hudson et al. 1987; Yu et al. 2013). The predominant polar lipids of strain L198^T were one unidentified phospholipid (PL1) and one unidentified glycolipid (GL1), which is similar with *T. islandicus* DSM 21543^T and *T. composti* DSM 21686^T, and consistent with other members of genus *Thermus*. Strain L198^T was distinct from DSM 21543^T and DSM 21686^T by the presence of an unidentified phospholipid (PL2). The complete polar lipid profiles of all strains of this study are given in supplementary Figure S4.

General genome properties and comparative genomics

The strain L198^T high-quality draft genome is 2.16 Mbp long with a 68.21% G + C content. Strain L198^T has a chromosome (scaffold Ga0070539_14) and at least two likely plasmids (scaffold Ga0070539_12 and Ga0070539_13). A total of 2308 genes were predicted, comprising 2251 protein-coding and 57 RNA genes. Three CRISPR arrays were

Table 1 Phenotypic characteristics differentiate strain L198^T from related *Thermus* strains

Characteristic	L198^T	L423	*T. islandicus* DSM 21543^T	*T. composti* DSM 21686^T	*T. aquaticus* YT-1Ta
Isolation source	Hot spring in USA	Hot spring in USA	Hot spring in Iceland	Thermophilic phase of compost in Hungary	Hot spring in USA
Pigmentation	Yellow	Yellow	Yellow	White	Yellow
Temperature range for growth (°C)	45–75	45–75	45–79	40–80	40–79
pH range for growth	5.0–9.0	5.0–9.0	5.5–10.5	5.0–9.0	6.0–9.5
Salinity for growth (NaCl %, w/v)	0–1.5	0–2.0	0–0.5	0–1.5	0–1.0
Catalase	+	+	−	+	+
β-galactosidase[b]	+	+	−	−	−
β-glucosidase[b]	+	+	−	−	+
Hydrolysis of					
Starch	−	−	+	−	+
Aesculin	+	+	+	+	−
Tween 20	−	−	+	+	+
Tween 40	+	−	+	+	+
Tween 60	+	−	+	+	+
Tween 80	−	−	−	+	−
Utilization of					
D-glucose	−	−	−	+	+
L-arabinose	+	+	−	−	−
D-fructose	+	+	−	+	+
D-galactose	+	+	−	−	−
L-rhamnose	+	+	−	−	−
Sucrose	w	+	−	−	+
Maltose	w	+	+	+	+
Major fatty acids (> 10%)	Iso-$C_{15:0}$, iso-$C_{17:0}$, anteiso-$C_{15:0}$	Iso-$C_{15:0}$, iso-$C_{17:0}$, anteiso-$C_{15:0}$	Iso-$C_{15:0}$, iso-$C_{17:0}$, anteiso-$C_{15:0}$, anteiso-$C_{17:0}$	Iso-$C_{17:0}$, anteiso-$C_{17:0}$, iso-$C_{16:0}$	Iso-$C_{17:0}$, iso-$C_{15:0}$, $C_{16:0}$, iso-$C_{16:0}$
Presence of 3-OH fatty acids	+	+	−	−	+

Data were confirmed in this work unless otherwise specified. Tests are scored as positive (+), negative (−) or weakly positive (w)

[a]Data for *T. aquaticus* YT-1^T was taken from Brock and Freeze (1969), and Chung et al. (2000)

[b]Data for enzyme activities of β-galactosidase and β-glucosidase were obtained with API ZYM strip in this work

identified. The coding regions accounted for 97.53% of the whole genome and 1815 genes were assigned to a putative function with the remaining annotated as hypothetical proteins. A total of 1618 genes (70.10%) were assigned into clusters of orthologous groups (COGs). The most abundant COG category was amino acid transport and metabolism (10.8%, COG category E), followed closely by general function prediction only (10.6%, COG category R), and translation, ribosomal structure, and biogenesis (10.2%, COG category J). The properties of the genomes and the distribution of genes into COGs are presented in supplementary Table S2. The genome sequence similarity between strains L198^T and genome sequences of other *Thermus* strains with sequenced genomes are shown in supplementary Table S3. The highest digital DDH, orthoANIb, and GGDC values were 32.7%, 87.47%, and 0.13 for *T. aquaticus* YT-1^T, which are below the standard DDH (70%), ANI (95–96%), and GGDC (0.258) values used for delineation of bacterial species (Goris et al. 2007; Kim et al. 2014; Richter and Rosselló-Móra 2009; Stackebrandt and Goebel 1994; Wayne et al. 1987). 442 single-copy marker genes were identified from the genomes sequences of 15 *Thermus* species. The maximum-likelihood phylogenetic tree based on the concatenation of single-copy marker genes revealed strain L198^T represents a close relative of *T. aquaticus* YT-1^T, which is in agreement with the highest overall genome related index. Strain L198^T and *T. aquaticus* YT-1^T belong to a highly supported branch (100% bootstrap value) along with *T. islandicus* DSM 21543^T, *T. arciformis* CGMCC 1.6992^T and *T. thermophilus* HB8^T (supplementary Figure S5).

The pan genome for the 15 genome *Thermus* genome set included 5193 genes, including 1295 core genes (supplementary Table S4). Among the 1295 core genes, ~ 89.4% could be assigned to COG categories with general functional prediction only (COG category R). The numbers of species-specific genes ranged from 51 to 230 genes in the genomes, with 92 in the genome of strain L198^T.

Analysis of the genome sequence

Analysis of the genome of strain L198^T was generally consistent with the phenotypic data. Genes for flagellar assembly and L- and P rings were absent from the genome. Genes for nitrogen oxyanion and nitrogen oxide respiration were also absent from the genome, also in agreement with the cultivation experiments. The L198^T genome has a *sox* gene cluster (*soxABCDXYZ*, IMG gene ID 2617069958–2617069969) predicted to encode a thiosulfate oxidation system (Friedrich et al. 2005), which is consistent with the ability to oxidize thiosulfate. A high affinity phosphate transport system (Pts, IMG gene ID 2617070823) was identified in L198^T genome, likely allowing arsenate to enter the cells due to the structural similarity of arsenate and phosphate (Muller et al. 2007). Genes for both respiratory arsenate reduction (arsenate reductase, ArsC, IMG gene ID 2617070204) and an arsenite efflux pump protein (ArsB, IMG gene ID 2617070395) were present, consistent with the arsenate respiration phenotype in strain L198^T. A search for carbohydrate-active enzymes (CAZymes) (Lombard et al. 2013; Yin et al. 2012) revealed a total of 58 CAZymes, of which 15 are glycoside hydrolases (GHs), including GHs that probably are involved in depolymerization of chitin (GH23) and starch (GH4, GH42, GH57, GH77, four genes belonging to GH13 family, two genes belonging to GH31 family, and two genes belonging to GH2 family). The strain also encodes for GH36, which may aid in removal of terminal galactosidase from glycolipids or glycoproteins. The L198^T genome encodes 44 peptidases as identified by the MEROPS database (Rawlings et al. 2016). Both CAZymes and peptidases are generally compatible with the abilities of L198^T to utilize various carbohydrates or amino acids as carbon and energy sources, and these characteristics are common to other *Thermus* species (Albuquerque et al. 2018).

Phylogenetic and genomic analysis clearly place the new isolates within the genus *Thermus*, as supported by a number of phenotypic traits, including negative Gram stain, rod shape, lack of motility, use of menaquinone 8 (MK-8) as the predominant respiratory quinone, and phospholipids and glycolipids as the major polar lipids. However, they differ from closely related species based on their genomic distinctness (16S rRNA gene, dDDH, ANI, and GGDC) and on their colony color, enzyme activities, and different carbon substrate oxidation ability (Table 1). Strains L198^T and L423 formed yellow colonies distinct from *T. composti* DSM 21686^T, which are white. Strain L198^T has a wider NaCl tolerance range (0–1.5%) than *T. islandicus* DSM 21543^T (0–0.5%) and *T. aquaticus* YT-1^T (0–1.0%). Similarly to *T. islandicus* DSM 21543^T, strains L198^T also exhibits mixotrophic growth by oxidation of thiosulfate and has a *sox* gene cluster predicted to encode a thiosulfate oxidation system. However, the activity of thiosulfate oxidation is negative for *T. aquaticus* YT-1^T. The activities of β-galactosidase and β-glucosidase, and the ability to oxidize L-arabinose, D-galactose, and L-rhamnose distinguish strain L198^T and closely related type strains *T. composti* DSM 21686^T and *T. islandicus* DSM 21543^T. Deviations were also observed concerning the cellular fatty acid profiles, which mainly differed in the amount of iso-$C_{15:0}$, iso-$C_{16:0}$, iso-$C_{17:0}$, anteiso-$C_{15:0}$, and anteiso-$C_{17:0}$, and the presence of hydroxy fatty acids in strains L198^T and L423 (supplementary Table S1).

In conclusion, on the basis of phylogenetic, genomic, and phenotypic analyses, we suggest that strains L198^T and L423 represent a novel species of the genus *Thermus*. We, therefore, propose the name *Thermus sediminis* sp. nov. with the type strain L198^T.

Description of *Thermus sediminis* sp. nov.

Thermus sediminis (se.di′mi.nis. L. gen. n. *sediminis* of a sediment, referring to the source from which the type strain was isolated).

Cells are Gram stain-negative and rod-shaped and 2.0–5.0 µm in length and 0.2–0.5 µm in width. Non-motile. Good growth on *Thermus* medium, R2A medium, and T5 agar. Form circular, convex, yellow-pigmented colonies on *Thermus* medium. Thermophilic, with an optimum growth temperature of 60–70 °C (range 45–75 °C). Growth occurs from pH of 5.0–9.0 (optimum 7.0). Tolerant salinities of 0–2% (w/v) added NaCl. Aerobic; does not grow anaerobically with nitrate (4.5 mM), ferric nitrilotriacetate (NTA-Fe, 10 mM), ferric citrate (10 mM), or sulfur as terminal electron acceptors and does not ferment glucose. Grows mixotrophically by oxidation of thiosulfate to sulfate. Oxidase and catalase were positive. Positive for hydrolysis of aesculin, gelatin, and galactosidase, but negative for hydrolysis of starch, carboxymethyl cellulose, casein, arginine and Tweens 20 and 80, nitrate reduction, urease, indole production, and glucose fermentation. Hydrolysis of Tweens 40 and 60 are variable. D-fructose, D-galactose, L-rhamnose, and L-arabinose can be utilized as sole carbon sources, but D-glucose, D-lactose, D-cellobiose, D-trehalose, D-raffinose, D-melibiose, myo-inositol, and malate are not utilized. Weak growth is exhibited on sucrose, maltose, and glycerol. Utilization of pyruvate is variable. Results from API ZYM test strips are positive for alkaline phosphatase, esterase C4, esterase lipase C8, leucine arylamidase,

naphthol-AS-BI-phosphohydrolase, β-galactosidase, α-glucosidase, and β-glucosidase, negative for valine arylamidase, cystine arylamidase, trypsin, α-chymotrypsin, β-glucuronidase, α-mannosidase, β-fucosidase, or *N*-acetyl-β-glucosaminidase, and variable for lipase C14, acid phosphatase, and α-galactosidase. The Biolog GEN III MicroPlate system results indicated oxidation of limited carbon sources: D-turanose, D-fructose, D-galactose, 3-methyl glucose, L-rhamnose, D-glucose-6-PO_4, D-fructose-6-PO_4, D-galacturonic acid, L-galactonic acid, D-glucuronic acid, glucuronamide, α-keto-glutaric acid, and acetoacetic acid. Weak results were observed for D-maltose, sucrose, D-mannose, glycerol, *N*-acetyl-D-glucosamine, and L-fucose. The predominant respiratory quinone is menaquinone MK-8. The major fatty acids are iso-$C_{15:0}$, iso-$C_{17:0}$ and anteiso-$C_{15:0}$. The polar lipids include two unidentified phospholipids (PL1–PL2), an unidentified aminophospholipid (PLN), and three unidentified glycolipids (GL1–GL3).

The type strain L198^T (= CGMCC 1.13590^T = KCTC XXX), was isolated from LHC1 hot spring in the Long Valley Caldera, California, USA. The DNA G + C content is 68.21% deduced from the genomic data. Strain L423 is an additional strain of this species.

Acknowledgements We thank the National Forest Service (Inyo National Forest, Mammoth Lakes Office) for permission to sample Little Hot Creek. This work was supported by the National Natural Science Foundation of China (nos. 31600103 and 31470139), China Postdoctoral Science Foundation (2016M602569), and National Science Foundation grants (OISE-0968421 and DBI-1005223). Whole genome sequencing conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract no. DE-AC02-05CH11231. Wen-Jun Li was also supported by Guangdong Province Higher Vocational Colleges and Schools Pearl River Scholar Funded Scheme (2014). B.P. Hedlund was also funded by a gift from Greg Fullmer through the UNLV Foundation.

Compliance with ethical standards

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Albuquerque L, Rainey FA, Da Costa MS (2018) *Thermus*. In: Whitman WB et al (eds) Bergey's manual of systematics of archaea and bacteria. https://doi.org/10.1002/9781118960608.gbm00477.pub2

Bjornsdottir SH, Petursdottir SK, Hreggvidsson GO, Skirnisdottir S, Hjorleifsdottir S, Arnfinnsson J, Kristjansson JK (2009) *Thermus islandicus* sp. nov., a mixotrophic sulfur-oxidizing bacterium isolated from the Torfajokull geothermal area. Int J Syst Evol Microbiol 59:2962–2966. https://doi.org/10.1099/ijs.0.007013-0

Blesa A, Averhoff B, Berenguer J (2018) Horizontal gene transfer in *Thermus* spp. Curr Issues Mol Biol 29:23–36. https://doi.org/10.21775/cimb.029.023

Brock TD, Freeze H (1969) *Thermus aquaticus* gen. n. and sp. n., a nonsporulating extreme thermophile. J Bacteriol 98:289–297

Carreto L, Wait R, Nobre M, Da Costa MS (1996) Determination of the structure of a novel glycolipid from *Thermus aquaticus* 15004 and demonstration that hydroxy fatty acids are amide linked to glycolipids in *Thermus* spp. J Bacteriol 178:6479–6486. https://doi.org/10.1128/jb.178.22.6479-6486.1996

Castenholz RW (1969) Thermophilic blue-green algae and the thermal environment. Bacteriol Rev 33:476–504

Cerny G (1978) Studies on the aminopeptidase test for the distinction of Gram-negative from Gram-positive bacteria. Eur J Appl Microbiol Biotechnol 5:113–122. https://doi.org/10.1007/BF00498805

Chaudhari NM, Gupta VK, Dutta C (2016) BPGA—an ultra-fast pan-genome analysis pipeline. Sci Rep 6:24373

Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. https://doi.org/10.1038/nmeth.2474

Chung AP, Rainey FA, Valente M, Nobre MF, Da Costa MS (2000) *Thermus igniterrae* sp. nov. and *Thermus antranikianii* sp. nov., two new species from Iceland. Int J Syst Evol Microbiol 50:209–217. https://doi.org/10.1099/00207713-50-1-209

Collins M, Jones D (1980) Lipids in the classification and identification of coryneform bacteria containing peptidoglycans based on 2, 4-diaminobutyric acid. J Appl Microbiol 48:459–470. https://doi.org/10.1111/j.1365-2672.1980.tb01036.x

Collins M, Pirouz T, Goodfellow M, Minnikin D (1977) Distribution of menaquinones in actinomycetes and corynebacteria. J Gen Microbiol 100:221–230. https://doi.org/10.1099/00221287-100-2-221

Connon SA, Koski AK, Neal AL, Wood SA, Magnuson TS (2008) Ecophysiology and geochemistry of microbial arsenic oxidation within a high arsenic, circumneutral hot spring system of the Alvord Desert. FEMS Microbiol Ecol 64:117–128. https://doi.org/10.1111/j.1574-6941.2008.00456.x

Costa KC, Navarro JB, Shock EL, Zhang CL, Soukup D, Hedlund BP (2009) Microbiology and geochemistry of great boiling and mud hot springs in the United States Great Basin. Extremophiles 13:447–459. https://doi.org/10.1007/s00792-009-0230-x

Da Costa MS, Rainey FA, Nobre MF (2006) The genus *Thermus* and relatives. In: The prokaryotes. Springer, New York, pp 797–812. https://doi.org/10.1007/0-387-30747-8_32

Eddy SR (2011) Accelerated profile HMM searches. PLoS Comput Biol 7:e1002195. https://doi.org/10.1371/journal.pcbi.1002195

Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S (2009) Real-time DNA sequencing from single polymerase molecules. Science 323:133–138. https://doi.org/10.1126/science.1162986

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17:368–376. https://doi.org/10.1007/BF01734359

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x

Fitch WM (1971) Toward defining the course of evolution: minimum change for a specific tree topology. Syst Biol 20:406–416. https://doi.org/10.1093/sysbio/20.4.406

Friedrich CG, Bardischewsky F, Rother D, Quentmeier A, Fischer J (2005) Prokaryotic sulfur oxidation. Curr Opin Microbiol 8:253–259. https://doi.org/10.1016/j.mib.2005.04.005

Gihring TM, Banfield JF (2001) Arsenite oxidation and arsenate respiration by a new *Thermus* isolate. FEMS Microbiol Lett 204:335–340. https://doi.org/10.1111/j.1574-6968.2001.tb10907.x

Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM (2007) DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 57:81–91. https://doi.org/10.1099/ijs.0.64483-0

Groth I, Schumann P, Rainey F, Martin K, Schuetze B, Augsten K (1997) *Demetria terragena* gen. nov., sp. nov., a new genus of actinomycetes isolated from compost soil. Int J Syst Bacteriol 47:1129–1133. https://doi.org/10.1099/00207713-47-4-1129

Hedlund BP, McDonald A, Lam J, Dodsworth JA, Brown J, Hungate B (2011) Potential role of *Thermus thermophilus* and *T. oshimai* in high rates of nitrous oxide (N_2O) production in ~ 80 °C hot springs in the US Great Basin. Geobiology 9:471–480. https://doi.org/10.1111/j.1472-4669.2011.00295.x

Hudson JA, Morgan HW, Daniel RM (1987) *Thermus filiformis* sp. nov., a filamentous caldoactive bacterium. Int J Syst Bacteriol 37:431–436. https://doi.org/10.1099/00207713-37-4-431

Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform 11:119. https://doi.org/10.1186/1471-2105-11-119

Kieft T, Fredrickson JK, Onstott TC, Gorby YA, Kostandarithes HM, Bailey TJ, Kennedy DW, Li SW, Plymale AE, Spadoni CM, Gray MS (1999) Dissimilatory reduction of Fe(III) and other electron acceptors by a *Thermus* isolate. Appl Environ Microbiol 65:1214–1221

Kim M, Oh H-S, Park S-C, Chun J (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346–351. https://doi.org/10.1099/ijs.0.059774-0

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054

Lee I, Kim YO, Park S-C, Chun J (2016) OrthoANI: an improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 66:1100–1103. https://doi.org/10.1099/ijsem.0.000760

Li WJ, Xu P, Schumann P, Zhang YQ, Pukall R, Xu LH, Stackebrandt E, Jiang CL (2007) *Georgenia ruanii* sp. nov., a novel actinobacterium isolated from forest soil in Yunnan (China), and emended description of the genus *Georgenia*. Int J Syst Evol Microbiol 57:1424–1428. https://doi.org/10.1099/ijs.0.64749-0

Liu Y, Song J, Tan T, Liu L (2015) Production of fumaric acid from l-malic acid by solvent engineering using a recombinant thermostable fumarase from *Thermus thermophilus* HB8. Appl Biochem Biotechnol 175:2823–2831. https://doi.org/10.1007/s12010-014-1468-z

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2013) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:D490–D495. https://doi.org/10.7868/S0026898415060208

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:0955–0964. https://doi.org/10.1093/nar/25.5.0955

Mavromatis K, Ivanova NN, Chen I-MA, Szeto E, Markowitz VM, Kyrpides NC (2009) The DOE-JGI standard operating procedure for the annotations of microbial genomes. Stand Genom Sci 1:63. https://doi.org/10.4056/sigs.632

Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinform 14:60. https://doi.org/10.1186/1471-2105-14-60

Ming H, Yin YR, Li S, Nie GX, Yu TT, Zhou EM, Liu L, Dong L, Li WJ (2014) *Thermus caliditerrae* sp. nov., a novel thermophilic species isolated from a geothermal area. Int J Syst Evol Microbiol 64:650–656. https://doi.org/10.1099/ijs.0.056838-0

Minnikin D, Collins M, Goodfellow M (1979) Fatty acid and polar lipid composition in the classification of *Cellulomonas*, *Oerskovia* and related taxa. J Appl Microbiol 47:87–95. https://doi.org/10.1111/j.1365-2672.1979.tb01172.x

Muller D, Médigue C, Koechler S, Barbe V, Barakat M, Talla E, Bonnefoy V, Krin E, Arsène-Ploetze F, Carapito C, Chandler M, Cournoyer B, Cruveiller S, Dossat C, Duval S, Heymann M, Leize E, Lieutaud A, Lièvremont D, Makita Y, Mangenot S, Nitschke W, Ortet P, Perdrial N, Schoepp B, Siguier P, Simeonova DD, Rouy Z, Segurens B, Turlin E, Vallenet D, Van Dorsselaer A, Weiss S, Weissenbach J, Lett MC, Danchin A, Bertin PN (2007) A tale of two oxidation states: bacterial colonization of arsenic-rich environments. PLoS Genet 3:e53. https://doi.org/10.1371/journal.pgen.0030053

Nawrocki EP, Kolbe DL, Eddy SR (2009) Infernal 1.0: inference of RNA alignments. Bioinformatics 25:1335–1337. https://doi.org/10.1093/bioinformatics/btp157

Pantazaki A, Pritsa A, Kyriakidis D (2002) Biotechnologically relevant enzymes from *Thermus thermophilus*. Appl Microbiol Biotechnol 58:1–12. https://doi.org/10.1007/s00253-001-0843-1

Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC (2010) GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. Nat Methods 7:455–457. https://doi.org/10.1038/NMETH.1457

Rawlings ND, Barrett AJ, Finn R (2016) Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res 44:D343–D350. https://doi.org/10.1093/nar/gkv1118

Richter M, Rosselló-Móra R (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci 106:19126–19131. https://doi.org/10.1073/pnas.0906412106

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454

Sazanov LA, Hinchliffe P (2006) Structure of the hydrophilic domain of respiratory complex I from *Thermus thermophilus*. Science 311:1430–1436. https://doi.org/10.1126/science.1123809

Skirnisdottir S, Hreggvidsson GO, Holst O, Kristjansson JK (2001) Isolation and characterization of a mixotrophic sulfur-oxidizing *Thermus scotoductus*. Extremophiles 5:45–51. https://doi.org/10.1007/s007920000172

Stackebrandt E, Goebel BM (1994) Taxonomic note: a place for DNA–DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int J Syst Bacteriol 44:846–849. https://doi.org/10.1099/00207713-44-4-846

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882. https://doi.org/10.1093/nar/25.24.4876

Vajna B, Kanizsai S, Kéki Z, Márialigeti K, Schumann P, Tóth EM (2012) *Thermus composti* sp. nov., isolated from oyster mushroom compost. Int J Syst Evol Microbiol 62:1486–1490. https://doi.org/10.1099/ijs.0.030866-0

Vick T, Dodsworth JA, Costa K, Shock E, Hedlund BP (2010) Microbiology and geochemistry of Little Hot Creek, a hot spring environment in the Long Valley Caldera. Geobiology 8:140–154. https://doi.org/10.1111/j.1472-4669.2009.00228.x

Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol Rev 65:1–43. https://doi.org/10.1128/MMBR.65.1.1-43.2001

Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI (1987) Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. Int J Syst Evol Microbiol 37:463–464. https://doi.org/10.1099/00207713-37-4-463

Wilkie JA, Hering JG (1998) Rapid oxidation of geothermal arsenic (III) in streamwaters of the eastern Sierra Nevada. Environ Sci Technol 32:657–662. https://doi.org/10.1021/es970637r

Wu Y-W (2018) ezTree: an automated pipeline for identifying phylogenetic marker genes and inferring evolutionary relationships among uncultivated prokaryotic draft genomes. BMC Genom 19:921. https://doi.org/10.1186/s12864-017-4327-9

Xu P et al (2005) *Naxibacter alkalitolerans* gen. nov., sp. nov., a novel member of the family '*Oxalobacteraceae*' isolated from China. Int J Syst Evol Microbiol 55:1149–1153. https://doi.org/10.1099/ijs.0.63407-0

Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res 40:W445–W451. https://doi.org/10.1093/nar/gks479

Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, Seo H, Chun J (2017) Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol 67:1613–1617. https://doi.org/10.1099/ijsem.0.001755

Yu TT, Yao JC, Ming H, Yin YR, Zhou EM, Liu MJ, Tang SK, Li WJ (2013) *Thermus tengchongensis* sp. nov., isolated from a geothermally heated soil sample in Tengchong, Yunnan, southwest China. Antonie Van Leeuwenhoek 103:513–518. https://doi.org/10.1007/s10482-012-9833-9

Yusupov MM, Yusupova GZ, Baucom A, Lieberman K, Earnest TN, Cate J, Noller HF (2001) Crystal structure of the ribosome at 5.5 Å resolution. Science 292:883–896. https://doi.org/10.1126/science.1060089

Zhou EM, Murugapiran SK, Mefferd CC, Liu L, Xian WD, Yin YR, Ming H, Yu TT, Huntemann M, Clum A, Pillay M, Palaniappan K, Varghese N, Mikhailova N, Stamatis D, Reddy TBK, Ngan CY, Daum C, Shapiro N, Markowitz V, Ivanova N, Spunde A, Kyrpides N, Woyke T, Li WJ, Hedlund BP (2016) High-quality draft genome sequence of the *Thermus amyloliquefaciens* type strain YIM 77409^{T} with an incomplete denitrification pathway. Stand Genom Sci 11:1–9. https://doi.org/10.1186/s40793-016-0140-3

Affiliations

En-Min Zhou[1,2] · Wen-Dong Xian[1] · Chrisabelle C. Mefferd[2] · Scott C. Thomas[2] · Arinola L. Adegboruwa[2] · Nathan Williams[2,3] · Senthil K. Murugapiran[2] · Jeremy A. Dodsworth[4] · Rakesh Ganji[2] · Meng-Meng Li[1] · Yi-Ping Ding[1] · Lan Liu[1] · Tanja Woyke[5,6] · Wen-Jun Li[1] · Brian P. Hedlund[2,7]

[1] State Key Laboratory of Biocontrol, Guangdong Key Laboratory of Plant Resources, School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, People's Republic of China

[2] School of Life Sciences, University of Nevada, Las Vegas, Las Vegas, NV 89154, USA

[3] Las Vegas High School PAL Program, Clark County School District, Las Vegas, NV 89154, USA

[4] Department of Biology, California State University, San Bernardino, San Bernardino, CA, USA

[5] Department of Energy, Joint Genome Institute, Walnut Creek, CA 94598, USA

[6] Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[7] Nevada Institute of Personalized Medicine, University of Nevada, Las Vegas, Las Vegas, NV 89154, USA