

1 **Genome Sequences from Five Type Strain Members of the**
2 **Archaeal Family *Sulfolobaceae*: *Acidianus ambivalens*,**
3 ***Acidianus infernus*, *Stygiolobus azoricus*, *Sulfuracidifex***
4 ***metallicus*, and *Sulfurisphaera ohwakuensis***

5
6 James A. Counts, Nicholas P. Vitko, and Robert M. Kelly*

7
8 *Department of Chemical and Biomolecular Engineering,*
9 *North Carolina State University, Raleigh, NC, USA*

10
11
12
13 **ASM Microbiology Resource Announcements (MRA)**

14
15
16 **Running Title:** Genomes of Extremely Thermoacidophilic Archaea

17
18
19 **Keywords:** *Sulfolobaceae, extreme thermoacidophiles, chemolithoautotrophy, sulfur reduction,*
20 *sulfur oxidation, metal mobilization, biooxidation*

21
22
23
24
25 *Address correspondence to: **Robert M. Kelly**
26 Dept. of Chemical and Biomolecular Engineering
27 North Carolina State University
28 EB-1, 911 Partners Way
29 Raleigh, NC 27695-7905
30 Phone: (919) 515-6396
31 Fax: (919) 515-3465
32 Email: rmkelly@ncsu.edu

Abstract: Presented are five genomes from the polyextremophilic ($T_{\text{opt}} > 65\text{ }^{\circ}\text{C}$, $\text{pH}_{\text{opt}} < 3.5$) archaeal family *Sulfolobaceae*, greatly expanding order-wide genomic diversity. Included are the only obligate anaerobic species, several facultative sulfur-utilizers, two metal-mobilizers, one facultative chemolithoautotroph with robust metabolic versatility, and some of the most thermophilic thermoacidophiles to-date.

Extremely thermoacidophilic *Sulfolobales* (*Archaea*, *Crenarchaeota*, *Thermoprotei*, *Sulfolobales*, *Sulfolobaceae*) are ubiquitous in highly acidic geothermal environments around the world. These organisms subsist in inorganic-rich (sulfur-rich, metal-rich, organic carbon-limited) environments, ranging from calderas to mud pits to sulfur pools, acting as microbial transformers of inorganic, energetic materials percolating from subsurface terrestrial activities. The genomes presented here come from the species *Acidianus ambivalens* DSM-3772, *Acidianus infernus* DSM-3191, *Stygiolobus azoricus* DSM-6296, *Sulfuracidifex metallicus* DSM-6482, and *Sulfurisphaera ohwakuensis* DSM-12421, which have diverse metabolic capabilities, as well as being globally distributed (**Table 1**). This work represents a second stage in sequencing that complements previous work, focused on metal-biooxidizers of the genera *Metallosphaera* and *Acidianus* (1). Additionally, this research reports three previously unsequenced type-strains, and provides more complete information for two recent genome sequencing projects associated with the other two type strains.

All of the species presented here were obtained from isolates deposited at the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ). All of the cultures were grown as recommended by DSMZ; further, the *Acidianus* spp. and *S. azoricus* were cultured anaerobically with a hydrogen/carbon dioxide (20/80) headspace. Organic solvent-based DNA extraction was performed with phenol, chloroform, and isoamyl alcohol, and an isopropanol precipitation, similarly to the method described by Geslin *et al.* (2).

Single Molecule Real Time (SMRT) sequencing was performed with either a single SMRT cell on the RSII system or multiplexed (Barcoded adapter kit 8A) on a SMRT cell with the Sequel system (Pacific Biosciences, Menlo Park, CA, USA). Both library configurations were size-selected with BluePippin (15 or 3 kb) and prepared with the Template Prep Kit 1.0-SPv3. Additional short-read data were prepared in either single-end 150 bp or paired-end 250 bp reads using the TruSeq Nano DNA v2 library preparation kit and run on an Illumina (San Diego, CA, USA) on either MiSeq or NextSeq 500 sequencers. Assemblies were performed with a recent repeat graph assembler designed for long error-prone reads, Flye, with default parameters and a 2.5 Mb genome size estimate (3) [v2.4.2], using raw consensus long-read data (**Table 1**). Following the initial assembly, short read data were mapped to all of the constructs using Bowtie2 (4) [v2.3.2], with medium-sensitivity, in order to correct for sequencing errors from long-read data. These short reads were curated with Trimmomatic (5) [v0.38], with the criteria of phred33 scores greater than 30 for head and tail and 28 within a sliding window of four bases, removing reads under 145 bp (single-end) and 240 bp (paired-end). Finally, genomes were deposited in the NCBI genomes database where they were annotated via the PGAAP algorithm (6).

In summation, five genome sequences are presented representing diverse members from family *Sulfolobaceae*. While the *S. metallicus* genome remains unclosed, the draft genome provided here greatly improves a previous short-read assembly that produced 167 contigs ([BBBY01](#)); and the *A. ambivalens* genome (previously 67 contigs)(7) is now closed. These data provide new information to assist researchers in fields ranging from fundamental microbiology and ecology to biotechnology and metabolic engineering.

Acknowledgements. This work was supported in part by grants from the U.S. Air Force Office of Sponsored Research (FA9550-17-1-0268) and the U.S. National Science Foundation (CBET-1264052). JA Counts acknowledges support from a U.S. NIH Biotechnology Traineeship (T32 GM008776-16).

86 **Data Availability.** All relevant bioproject, biosample, and accession numbers are presented in
87 **Table 1** and hyperlinked to the respective repository sites from the NCBI database.

88

89

REFERENCES CITED

1. Counts JA, Vitko NP, Kelly RM. 2018. Complete genome sequences of extremely thermoacidophilic metal-mobilizing type strain members of the Archaeal family Sulfolobaceae, *Acidianus brierleyi* DSM-1651, *Acidianus sulfidivorans* DSM-18786, and *Metallosphaera hakonensis* DSM-7519. Microbiol Resour Announc 7:e00831-18.
2. Geslin C, Le Romancer M, Erauso G, Gaillard M, Perrot G, Prieur D. 2003. PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, "*Pyrococcus abyssi*". J Bacteriol 185:3888-94.
3. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540-546.
4. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357-9.
5. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-20.
6. Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2008. Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation. OMICS 12:137-41.
7. Bertran E, Ward LM, Johnston DT. 2020. Draft genome sequence of *Acidianus ambivalens* DSM 3772, an aerobic thermoacidophilic sulfur disproportionator. Microbiol Resour Announc 9.

- 112 8. Zillig W, Yeats S, Holz I, Böck A, Rettenberger M, Gropp F, Simon G. 1986. *Desulfurolobus*
113 *ambivalens*, gen. nov., sp. nov., an autotrophic archaeobacterium facultatively oxidizing or
114 reducing sulfur. Syst Appl Microbiol 8:197-203.
- 115 9. Segerer A, Neuner A, Kristjansson JK, Stetter KO. 1986. *Acidianus infernus*, gen. nov., sp.
116 nov. and *Acidianus brierleyi* comb. nov. : facultatively aerobic, extremely acidophilic
117 thermophilic sulfur-metabolizing archaeobacteria. Int J of Syst Bacteriol 36:559-564.
- 118 10. Segerer AH, Trincone A, Gahrtz M, Stetter KO. 1991. *Stygiolobus azoricus* gen. nov., sp.
119 nov. represents a novel genus of anaerobic, extremely thermoacidophilic archaeobacteria
120 of the order *Sulfolobales*. Int J of Syst Bacteriol 41:495-501.
- 121 11. Huber G, Stetter KO. 1991. *Sulfolobus metallicus*, sp. nov., a novel strictly
122 chemolithoautotrophic thermophilic Archaeal species of metal-mobilizers. Syst Appl
123 Microbiol 14:372-378.
- 124 12. Kurosawa N, Itoh YH, Iwai T, Sugai A, Uda I, Kimura N, Horiuchi T, Itoh T. 1998.
125 *Sulfurisphaera ohwakuensis* gen. nov., sp. nov., a novel extremely thermophilic acidophile
126 of the order *Sulfolobales*. Int J Syst Bacteriol 48:451-456.
- 127

Table 1. Sequenced Species' Collection, Sequencing, Assembly and Meta-Data					
Species (ref)	<i>Acidianus ambivalens</i> (8)	<i>Acidianus infernus</i> (9)	<i>Stygiolobus azoricus</i> (10)	<i>Sulfuracidifex metallicus</i> (11)	<i>Sulfurisphaera ohwakuensis</i> (12)
Collection IDs	DSM3772 ^T JCM9191 ^T	DSM3191 ^T JCM8955 ^T	DSM6296 ^T JCM9021 ^T	DSM6482 ^T JCM9184 ^T	DSM12421 ^T JCM9065 ^T
Isolation Site	Leirhnjúkur, Mývatn Iceland	Pisciarelli Solfatara, Naples, Italy	São Miguel Island, Azores	Krafla, Mývatn, Iceland	Hot spring in Ohwaku Valley, Hakone, Japan
Metabolism	Chemolithoautotroph	Chemolithoautotroph	Chemolithoautotroph	Chemolithoautotroph	Chemolithoautotroph
Oxygen Metabolism	Facultative	Facultative	Anaerobe	Aerobe	Facultative
T _{opt} (°C) pH _{opt}	80 2.0	90 2.0	80 2.5	65 2.0	85 2.0
Long Reads	299,403 ¹ (661X)	104,250 ^{1,‡} (438X)	185,364 ¹ (482X)	62,361 ^{2,‡} (308X)	151,460 ¹ (362X)
Short Reads	3,961,682 ³ (260X)	4,171,032 ⁴ (469X)	4,830,704 ⁴ (605X)	5,477,710 ⁴ (620X)	5,400,166 ⁴ (481X)
Genome Size, bp	2,252,027 [†]	2,220,671	1,987,069 [†]	2,199,731	2,803,915 [†]
# of Contigs (Largest)		4 (2,184,866)		5 (1,839,470)	
G+C (mol%)	34.2	34.4	37.6	38.6	32.7
Bioproject	PRJNA488459	PRJNA488459	PRJNA488459	PRJNA463410	PRJNA488459
BioSample	SAMN09933089	SAMN09933090	SAMN09933091	SAMN09933021	SAMN09933092
Genome Accession	CP045482	WFIY00000000	CP045483	WGGD00000000	CP045484
Read Accessions	SRX7690445 (PacBio) SRX7690446 (Illumina)	SRX7690447 (PacBio) SRX7690448 (Illumina)	SRX7690449 (PacBio) SRX7690450 (Illumina)	SRR11038957 (PacBio) SRR11038956 (Illumina)	SRX7690451 (PacBio) SRX7690452 (Illumina)
(† = Closed; ‡ = Trimmomatic Filtered or Long-Read Filtered; >5K <i>A. infernus</i> ; >8K <i>S. metallicus</i> ; Avg = AverageChemistries/Platforms: 1) PacBio Sequel/3kb Selection, 2) PacBio RSII/15kb Selection, 3) Illumina NextSeq 500/SE-150bp, 4) Illumina MiSeq/PE-250bp					