

## Complete genome sequence of *Intrasporangium calvum* type strain (7 KIP<sup>T</sup>)

Tijana Glavina Del Rio<sup>1</sup>, Olga Chertkov<sup>1,2</sup>, Montri Yasawong<sup>3</sup>, Susan Lucas<sup>1</sup>, Shweta Deshpande<sup>1</sup>, Jan-Fang Cheng<sup>1</sup>, Chris Detter<sup>1,2</sup>, Roxanne Tapia<sup>1,2</sup>, Cliff Han<sup>1,2</sup>, Lynne Goodwin<sup>1,2</sup>, Sam Pitluck<sup>1</sup>, Konstantinos Liolios<sup>1</sup>, Natalia Ivanova<sup>1</sup>, Konstantinos Mavromatis<sup>1</sup>, Amrita Pati<sup>1</sup>, Amy Chen<sup>4</sup>, Krishna Palaniappan<sup>4</sup>, Miriam Land<sup>1,5</sup>, Loren Hauser<sup>1,5</sup>, Yun-Juan Chang<sup>1,5</sup>, Cynthia D. Jeffries<sup>1,5</sup>, Manfred Rohde<sup>3</sup>, Rüdiger Pukall<sup>6</sup>, Johannes Sikorski<sup>6</sup>, Markus Göker<sup>6</sup>, Tanja Woyke<sup>1</sup>, James Bristow<sup>1</sup>, Jonathan A. Eisen<sup>1,7</sup>, Victor Markowitz<sup>4</sup>, Philip Hugenholtz<sup>1</sup>, Nikos C. Kyrpides<sup>1</sup>, Hans-Peter Klenk<sup>6</sup>, and Alla Lapidus<sup>\*</sup>

<sup>1</sup> DOE Joint Genome Institute, Walnut Creek, California, USA

<sup>2</sup> Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico, USA

<sup>3</sup> HZI – Helmholtz Centre for Infection Research, Braunschweig, Germany

<sup>4</sup> Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, California, USA

<sup>5</sup> Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

<sup>6</sup> DSMZ - German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig, Germany

<sup>7</sup> University of California Davis Genome Center, Davis, California, USA

\*Corresponding author: Alla Lapidus

**Keywords:** airborne, Gram-positive, non-motile, intercalary vesicles, nocardioform, *Actinobacteria*, *Intrasporangiaceae*, GEBA

---

*Intrasporangium calvum* Kalakoutskii *et al.* 1967 is the type species of the genus *Intrasporangium*, which belongs to the actinobacterial family *Intrasporangiaceae*. The species is a Gram-positive bacterium that forms a branching mycelium, which tends to break into irregular fragments. The mycelium of this strain may bear intercalary vesicles but does not contain spores. The strain described in this study is an airborne organism that was isolated from a school dining room in 1967. One particularly interesting feature of *I. calvum* is that the type of its menaquinone is different from all other representatives of the family *Intrasporangiaceae*. This is the first completed genome sequence from a member of the genus *Intrasporangium* and also the first sequence from the family *Intrasporangiaceae*. The 4,024,382 bp long genome with its 3,653 protein-coding and 57 RNA genes is a part of the *Genomic Encyclopedia of Bacteria and Archaea* project.

---

## Introduction

Strain 7 KIP<sup>T</sup> (= DSM 43043 = ATCC 23552 = JCM 3097) is the type strain of the species *Intrasporangium calvum*, which is the type species of its genus *Intrasporangium* [1,2]. The generic name derived from the Latin word *intra* meaning *within* and the Greek word *spora* meaning *a seed*. The name *Intrasporangium*, was selected to emphasize the possibility of intercalary formation of sporangia in mycelial filaments [3]. *Intrasporangium* is the type genus of the family *Intrasporangiaceae* and one out of currently nineteen genera in the family *Intrasporangiaceae* [4-6]. Strain 7 KIP<sup>T</sup> was first described in 1967 by Kalakoutskii *et al.* as an airborne organ-

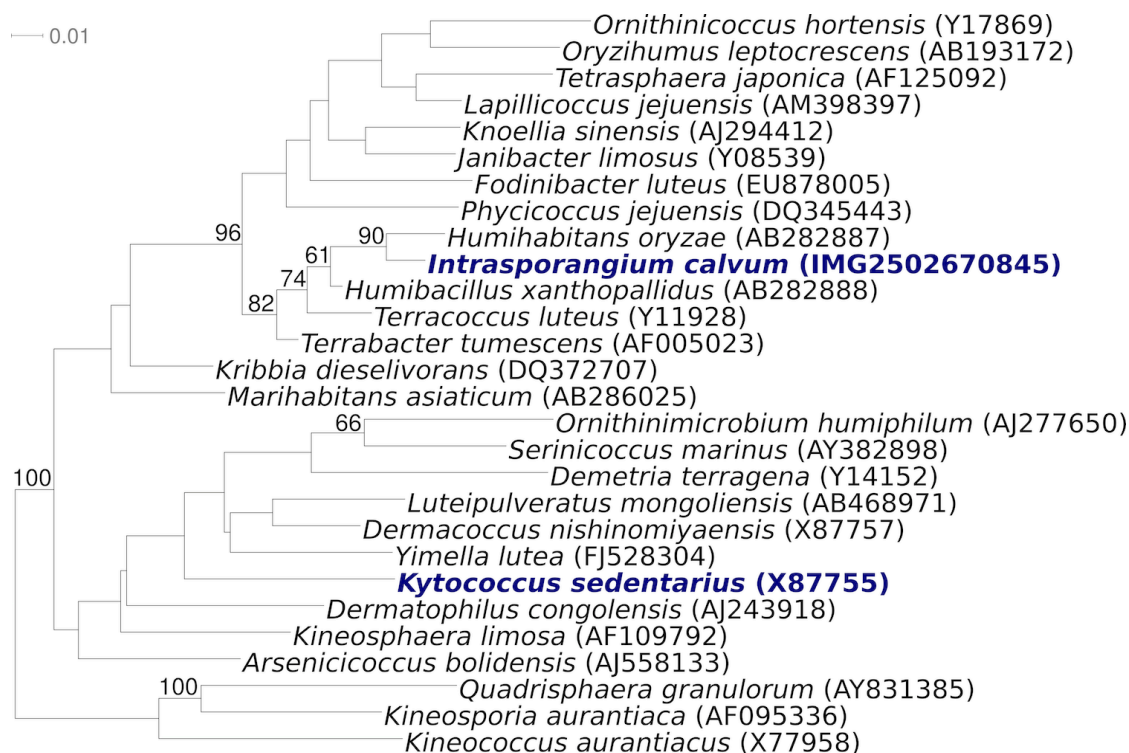
ism, which was isolated under nonselective conditions on plates of meat-peptone agar exposed to the atmosphere of a school dining room [1,7,8]. *I. calvum* is of particular interest because the type of its menaquinones is different from all other representatives of the family *Intrasporangiaceae* [8]. Here we present a summary classification and a set of features for *I. calvum* 7 KIP<sup>T</sup>, together with the description of the complete genomic sequencing and annotation.

## Classification and features

The 16S rRNA gene of strain 7 KIP<sup>T</sup> shares 92.6–98.7% sequence identity with the sequences of the type strains from the other members of the family *Intrasporangiaceae* [9], with *Humihabitans oryzae* as the closest relative. The 16S rRNA gene sequence of 7 KIP<sup>T</sup> is 99% identical to the uncultured *Intrasporangiaceae* clone HT06Ba24, isolated from soil of a former coal gasification site in Gliwice, Poland [10,11] and AKAU4164, isolated from uranium contaminated soil in Oak Ridge, USA [10,12]. The environmental samples database (env\_nt) contains the marine metagenome clone 1096626841081 (AA-CY020552144) from surface water (92% sequence identity with 7 KIP<sup>T</sup>). The genomic survey sequences database (gss) contains the metagenomic clone 1061002660518 from Floreana island in Punta Cormorant, Ecuador [10], which shares 93% sequence identity with 7 KIP<sup>T</sup> (as of July 2010). One of the 16S rRNA sequences of strain 7 KIP<sup>T</sup> was compared using NCBI BLAST under default values (e.g., considering only the best 250 hits) with the

most recent release of the Greengenes database [13] and the relative frequencies, weighted by BLAST scores, of taxa and keywords, weighted by BLAST scores, were determined. The five most frequent genera were *Janibacter* (29.6%), *Terrabacter* (19.8%), *Sanguibacter* (8.4%), *Dermacoccus* (7.7%) and *Tetrasphaera* (6.2%). The five most frequent keywords within the labels of environmental samples which yielded hits were 'skin' (9.1%), 'human' (4.7%), 'microbiome/temporal/topographical' (4.5%), 'sludge' (4.4%) and 'heel/plantar' (3.1%). The single most frequent keyword within the labels of environmental samples which yielded hits of a higher score than the highest scoring species was 'contaminated/soil/uranium' (33.3%).

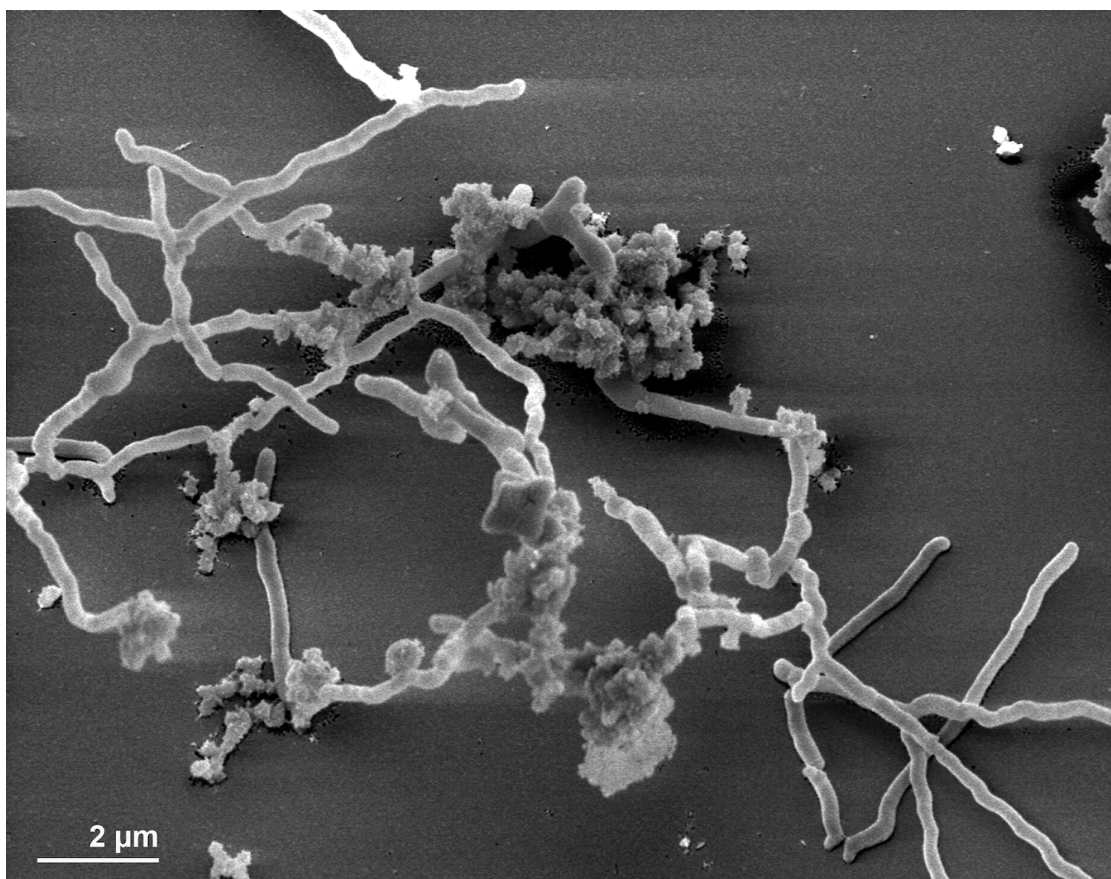
Figure 1 shows the phylogenetic neighborhood of *I. calvum* 7 KIP<sup>T</sup> in a 16S rRNA based tree. The sequences of the two 16S rRNA gene copies in the genome are differ by only one nucleotide from each other and by up to one nucleotide from the previously published sequence generated from DSM 43043 (AJ566282).



**Figure 1.** Phylogenetic tree highlighting the position of *I. calvum* 7 KIP<sup>T</sup> relative to the type strains of the genera within the family *Intrasporangiaceae*. The trees were inferred from 1,406 aligned characters [14,15] of the 16S rRNA gene sequence under the maximum likelihood criterion [16] and rooted with the type strains of the genera within the family *Kineosporiaceae* [17]. The branches are scaled in terms of the expected number of substitutions per site. Numbers above branches are support values from 650 bootstrap replicates [18] if larger than 60%. Lineages with type strain genome sequencing projects registered in GOLD [19] are shown in blue, published genomes in bold [20].

Strain 7 KIPT<sup>T</sup> forms a branching mycelium, which tends to break into irregular fragments, *i.e.*, typically nocardioform [1,8]. The mycelium may bear intercalary vesicles that do not contain spores [Table 1, Figure 2, 7,26]. The vesicles of strain 7 KIPT<sup>T</sup> are ovoid and lemon-shaped (5-15 µm in diameter) [1,7]. Several round or oval bodies (1.2-1.5 µm in diameter) may be observed in the vesicles of older cultures [1,7]. The oval bodies in the vesicles of strain 7 KIPT<sup>T</sup> are nonmotile but may undergo a Brownian movement (in mature vesicles) [1,7]. There was no aerial mycelium observed from the strain 7 KIPT<sup>T</sup> [1,7,8]. The mycelial filaments penetrate the agar and form compact, small colonies (1-5 mm of diameter) [1]. These colonies are round, glistening and whitish (cream-whitish in old colonies) when the cells are grown on meat-extract peptone agar [1]. Strain 7 KIPT<sup>T</sup> is aerobic and Gram-positive (Gram-variable in old cultures) and not acid-fast [1]. Strain 7 KIPT<sup>T</sup> is rather fastidious in nutritional requirements [1]. Growth is seemingly dependent on some unidentified substances present in the peptone used in the growth medium [1]. The strain prefers complex me-

dia for growth, especially containing peptone and yeast extract [1,7]. As such, the growth characteristics on a variety of media such as meat-extract peptone, blood serum broth, oatmeal agar, Sauton medium agar and other media, also in combination of different atmospheric gases and their concentrations, have been studied in detail [1]. Strain 7 KIPT<sup>T</sup> is able to grow between 28°C and 37°C, however, the cells grow faster at 37°C than 28°C, but it does not at 45°C [1]. It grows slowly on meat-extract peptone medium and the first signs of macroscopic growth will appear after 3-5 days when incubated at 28°C [1]. Strain 7 KIPT<sup>T</sup> does not grow on the majority of synthetic mineral media that are routinely used for actinomycetes [1,7]. Strain 7 KIPT<sup>T</sup> is able to reduce nitrate to nitrite when KNO<sub>3</sub> is added to the growth medium (meat-extract peptone broth) [1]. The liquefaction of gelatin does not occur when the strain 7 KIPT<sup>T</sup> was grown on meat-extract peptone gelatine [1]. Strain 7 KIPT<sup>T</sup> has no antibiotic activity against *Micrococcus luteus*, *Staphylococcus aureus*, *Escherichia coli*, *Bacillus subtilis*, *Candida albicans* and *Mycobacterium* sp. v-5 [1].



**Figure 2.** Scanning electron micrograph of *I. calvum* 7 KIPT<sup>T</sup>

**Table 1.** Classification and general features of *I. calvum* 7 KIP<sup>T</sup> according to the MIGS recommendations [21].

MIGS ID	Property	Term	Evidence code
		Domain <i>Bacteria</i>	TAS [22]
		Phylum <i>Actinobacteria</i>	TAS [23]
		Class <i>Actinobacteria</i>	TAS [4,24]
		Subclass <i>Actinobacteridae</i>	TAS [4,6]
	Current classification	Order <i>Actinomycetales</i>	TAS [2,4,6,25]
		Suborder <i>Micrococccineae</i>	TAS [4,6]
		Family <i>Intrasporangiaceae</i>	TAS [4-6]
		Genus <i>Intrasporangium</i>	TAS [1,2]
		Species <i>Intrasporangium calvum</i>	TAS [1,2]
		Type strain 7 KIP	TAS [1]
	Gram stain	positive	TAS [1]
	Cell shape	branching mycelium, which tends to break into irregular fragments	TAS [1,8]
	Motility	none	TAS [1,26]
	Sporulation	none	TAS [26]
	Temperature range	28°C–37°C	TAS [1]
	Optimum temperature	37°C	NAS
	Salinity	not reported	
MIGS-22	Oxygen requirement	aerobic	TAS [1,7]
	Carbon source	carbohydrates	TAS [1]
	Energy source	chemoorganotroph	TAS [1,7]
MIGS-6	Habitat	air	TAS [1]
MIGS-15	Biotic relationship	free-living	NAS
MIGS-14	Pathogenicity	none	NAS
	Biosafety level	1	TAS [27]
	Isolation	air in a school dining room	TAS [1]
MIGS-4	Geographic location	Russia	NAS
MIGS-5	Sample collection time	1967	TAS [1]
MIGS-4.1	Latitude	not reported	
MIGS-4.2	Longitude	not reported	
MIGS-4.3	Depth	not reported	
MIGS-4.4	Altitude	not reported	

Evidence codes - IDA: Inferred from Direct Assay (first time in publication); TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from of the Gene Ontology project [28]. If the evidence code is IDA, then the property was directly observed by one of the authors or an expert mentioned in the acknowledgements.

## Chemotaxonomy

Strain 7 KIP<sup>T</sup> contains LL-diaminopimelic acid (LL-A<sub>2</sub>pm) in the cell wall and possesses the A3 $\gamma$ -type of peptidoglycan [29,30]. The amino acid at position 1 of the peptide subunit is L-alanine [30]. The cell wall structure of strain 7 KIP<sup>T</sup> is characterized by the cross-linkage of the A3 $\gamma$ -type peptidoglycan via a triglycine-interpeptide bridge and by a glycine residue bound to the  $\alpha$ -carboxyl group of the D-glutamic acid position 2 of the peptide subunit [29,30]. Strain 7 KIP<sup>T</sup> possesses a totally unsaturated menaquinone with eight isoprene units

(MK-8) instead of a partially saturated menaquinone with two of eight isoprene units hydrogenated (MK-8(H<sub>4</sub>)) which is the characteristic menaquinone of all other representatives of the family *Intrasporangiaceae* [8,30]. Cells of strain 7 KIP<sup>T</sup> contain glucosamine-containing phospholipids (phospholipids type 4) [7]. Polar lipids of the strain are phosphatidyl-inositol, phosphatidylinositol mannosides, phosphatidylglycerol and diphosphatidyl-glycerol [30]. The major cellular fatty acids are saturated branched-chain acids: iso-

C<sub>15:0</sub> (37.8%), anteiso-C<sub>15:0</sub> (12.6%), iso-C<sub>16:0</sub> (12.3%), iso-C<sub>14:0</sub> (5.0%), anteiso-C<sub>17:0</sub> (3.9%), iso-C<sub>17:1</sub> (3.7%), iso-C<sub>15:1</sub> (3.5%), iso-C<sub>16:1</sub> (3.1%) and straight chain acid C<sub>15:0</sub> (2.7%) [30]. Polyamine contents (μmol per g dry wt) of strain 7 KIP<sup>T</sup> are putrescine (2.02), spermidine (1.03), spermine (0.31), cadaverine (0.30), 1,3-diaminopropane (0.17), *sym*-homospermidine (0.05) and tyramine (0.17) [29].

## Genome sequencing and annotation

### Genome project history

This organism was selected for sequencing on the basis of its phylogenetic position [31], and is part of the *Genomic Encyclopedia of Bacteria and Archaea* project [32]. The genome project is deposited in the Genome OnLine Database [19] and the complete genome sequence is deposited in GenBank. Sequencing, finishing and annotation were performed by the DOE Joint Genome Institute (JGI). A summary of the project information is shown in Table 2.

**Table 2.** Genome sequencing project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	Finished
MIGS-28	Libraries used	Three genomic libraries: one standard and one paired ended 454 pyrosequence library and one standard Illumina library
MIGS-29	Sequencing platforms	454 GS FLX Titanium, Illumina GAii
MIGS-31.2	Sequencing coverage	59.7 × pyrosequence: 95.2 × Illumina
MIGS-30	Assemblers	Newbler version 2.0.0-PostRelease-11/04/2008, phrap
MIGS-32	Gene calling method	Prodigal 1.4, GenePRIMP
	INSDC ID	CP002343
	Genbank Date of Release	December 29, 2010
	GOLD ID	Gc01572
	NCBI project ID	43527
	Database: IMG-GEBA	2503538011
MIGS-13	Source material identifier	DSM 43043
	Project relevance	Tree of Life, GEBA

### Growth conditions and DNA isolation

*I. calvum* 7 KIP<sup>T</sup> was grown in medium 65 (GYM Streptomyces medium) supplemented with one third of BHI (medium 215) [33] at 28°C. DNA was isolated from 0.5-1 g of cell paste using Qiagen Genomic 500 DNA Kit (Qiagen, Hilden, Germany) following the standard protocol as recommended by the manufacturer, with modification st/LALMP for cell lysis as described by Wu *et al.* [32].

### Genome sequencing and assembly

The genome was sequenced using a combination of Illumina and 454 sequencing platforms. All general aspects of library construction and sequencing can be found at the JGI website [34]. Pyrosequencing reads were assembled using the Newbler assembler version 2.0.0-PostRelease-11/04/2008 (Roche). The initial Newbler assembly consisted of 28 contigs in two scaffolds and

was converted into a phrap assembly by making fake reads from the consensus, collecting the read pairs in the 454 paired end library. Illumina GAii sequencing data (309MB) was assembled with Velvet [35] and the consensus sequences were shredded into 1.5 kb overlapped fake reads and assembled together with the 454 data. The 454 draft assembly was based on 226.2 Mb 454 draft data and all of the 454 paired end data. Newbler parameters are -consed -a 50 -l 350 -g -m -ml 20. The Phred/Phrap/Consed software package [36] was used for sequence assembly and quality assessment in the following finishing process. After the shotgun stage, reads were assembled with parallel phrap (High Performance Software, LLC). Possible mis-assemblies were corrected with gapResolution [34], Dupfinisher, or sequencing cloned bridging PCR fragments with subcloning or

transposon bombing (Epicentre Biotechnologies, Madison, WI) [20]. Gaps between contigs were closed by editing in Consed, by PCR and by Bubble PCR primer walks (J.-F.Chang, unpublished). A total of 139 additional reactions were necessary to close gaps and to raise the quality of the finished sequence. Illumina reads were also used to correct potential base errors and increase consensus quality using a software Polisher developed at JGI [37]. The error rate of the completed genome sequence is less than one error in 100,000. Together, the combination of the Illumina and 454 sequencing platforms provided  $154.9 \times$  coverage of the genome. The final assembly contains 847,906 pyrosequencing and 11,758,818 Illumina reads.

### Genome annotation

Genes were identified using Prodigal [38] as part of the Oak Ridge National Laboratory genome annotation pipeline, followed by a round of manual curation using the JGI GenePRIMP pipeline [39].

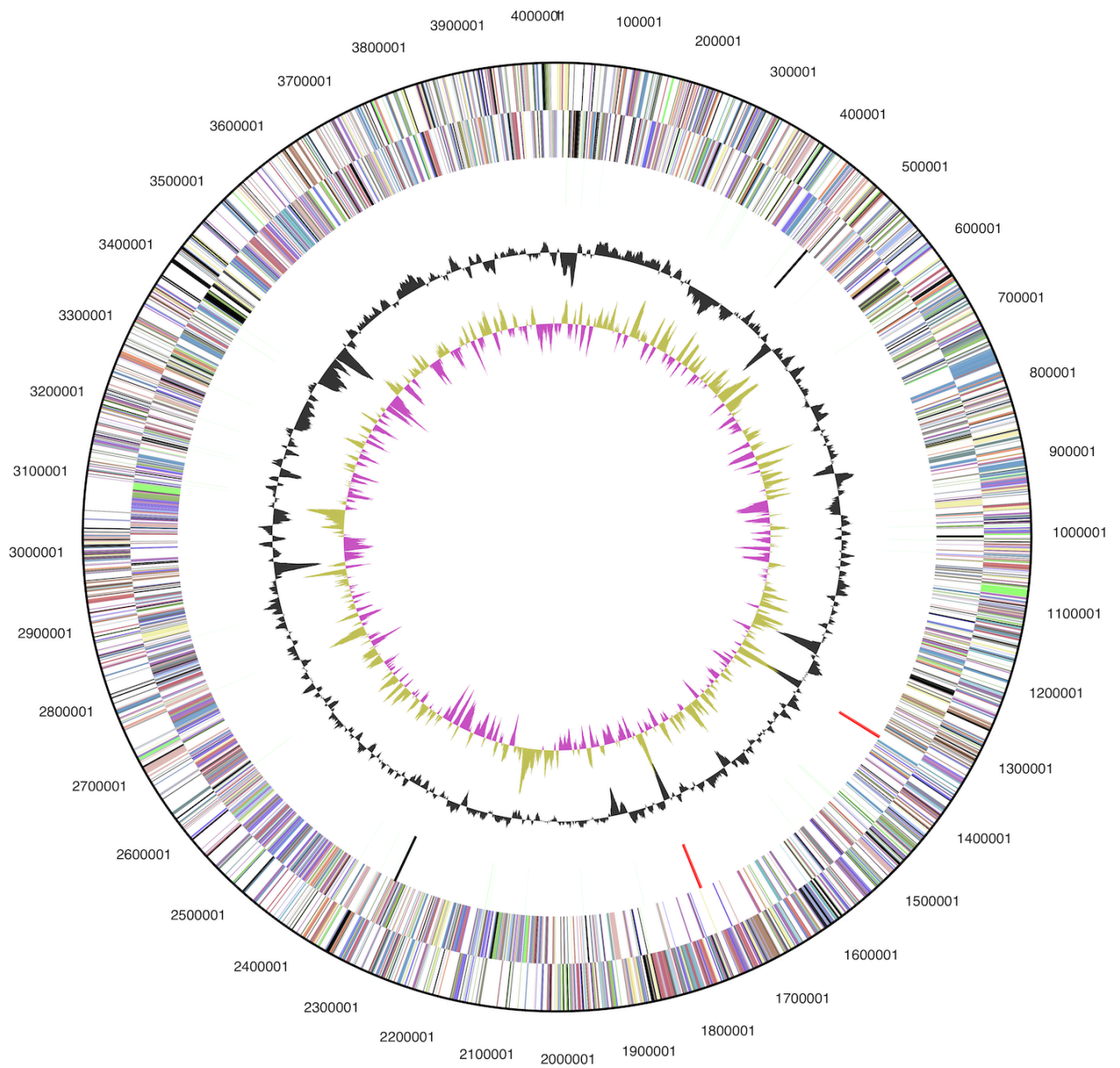
The predicted CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant database, UniProt, TIGRFam, Pfam, PRIAM, KEGG, COG, and InterPro databases. Additional gene prediction analysis and functional annotation was performed within the Integrated Microbial Genomes - Expert Review (IMG-ER) platform [40].

### Genome properties

The genome consists of a 4,024,382 bp long chromosome with a 70.7% GC content (Table 3 and Figure 3). Of the 3,710 genes predicted, 3,653 were protein-coding genes, and 57 RNAs; ninety pseudogenes were also identified. The majority of the protein-coding genes (71.3%) were assigned with a putative function while the remaining ones were annotated as hypothetical proteins. The distribution of genes into COGs functional categories is presented in Table 4.

**Table 3.** Genome Statistics

Attribute	Value	% of Total
Genome size (bp)	4,024,382	100.00%
DNA coding region (bp)	3,618,708	89.92%
DNA G+C content (bp)	2,845,385	70.70%
Number of replicons	1	
Extrachromosomal elements	0	
Total genes	3,710	100.00%
RNA genes	57	1.54%
rRNA operons	2	
Protein-coding genes	3,653	98.46%
Pseudo genes	90	2.43%
Genes with function prediction	2,645	71.29%
Genes in paralog clusters	360	7.70%
Genes assigned to COGs	2,674	72.08%
Genes assigned Pfam domains	2,871	77.39%
Genes with signal peptides	1,101	29.68%
Genes with transmembrane helices	860	23.18%
CRISPR repeats	0	



**Figure 3.** Graphical circular map of the genome. From outside to the center: Genes on forward strand (color by COG categories), Genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew.

**Table 4.** Number of genes associated with the general COG functional categories

Code	value	%age	Description
J	160	5.4	Translation, ribosomal structure and biogenesis
A	1	0.0	RNA processing and modification
K	230	7.8	Transcription
L	181	6.1	Replication, recombination and repair
B	2	0.1	Chromatin structure and dynamics
D	38	1.3	Cell cycle control, cell division, chromosome partitioning
Y	0	0.0	Nuclear structure
V	46	1.6	Defense mechanisms
T	131	4.4	Signal transduction mechanisms
M	155	5.2	Cell wall/membrane/envelope biogenesis
N	2	0.1	Cell motility
Z	0	0.0	Cytoskeleton
W	0	0.0	Extracellular structures
U	34	1.2	Intracellular trafficking and secretion, and vesicular transport
O	98	3.3	Posttranslational modification, protein turnover, chaperones
C	223	7.5	Energy production and conversion
G	199	6.7	Carbohydrate transport and metabolism
E	305	10.3	Amino acid transport and metabolism
F	79	2.7	Nucleotide transport and metabolism
H	141	4.8	Coenzyme transport and metabolism
I	151	5.1	Lipid transport and metabolism
P	132	4.5	Inorganic ion transport and metabolism
Q	89	3.0	Secondary metabolites biosynthesis, transport and catabolism
R	358	12.1	General function prediction only
S	212	7.2	Function unknown
-	1,036	27.9	Not in COGs

## Acknowledgements

We would like to gratefully acknowledge the help of Gabriele Gehrlich-Schröter for growing *I. calvum* cultures and Susanne Schneider for DNA extraction and quality analysis (both at DSMZ). This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231, Lawrence Livermore

National Laboratory under Contract No. DE-AC52-07NA27344, and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396, UT-Battelle and Oak Ridge National Laboratory under contract DE-AC05-00OR22725, as well as German Research Foundation (DFG) INST 599/1-2 and SI 1352/1-2, and Thailand Research Fund Royal Golden Jubilee Ph.D. Program No. PHD/0019/2548' for MY.

## References

1. Kalakoutskii LV, Kirillova IP, Krassilnikov NA. A new genus of the *Actinomycetales* - *Intrasporangium* gen. nov. *J Gen Microbiol* 1967; **48**:79-85.
2. Skerman VBD, McGowan V, Sneath PHA. Approved lists of bacterial names. *Int J Syst Bacteriol* 1980; **30**:225-420. [doi:10.1099/00207713-30-1-225](https://doi.org/10.1099/00207713-30-1-225)



3. Kalakoutslis LV. Genus *Intrasporangium*. In: Bergey's Manual of Systematic Bacteriology. Vol. 2. PHA Sneath, NS Mair, ME Sharpe, JG Holt (eds). p. 1504-1506. Williams & Wilkins, 1986. Baltimore, USA
4. Stackebrandt E, Rainey FA, Ward NL. Proposal for a new hierarchic classification system, *Actinobacteria* classis nov. *Int J Syst Bacteriol* 1997; **47**:479-491. [doi:10.1099/00207713-47-2-479](https://doi.org/10.1099/00207713-47-2-479)
5. Stackebrandt E, Schumann P. Description of *Bogoriellaceae* fam. nov., *Dermacoccaceae* fam. nov., *Rarobacteraceae* fam. nov. and *Sanguibacteraceae* fam. nov. and emendation of some families of the suborder *Micrococccineae*. *Int J Syst Evol Microbiol* 2000; **50**:1279-1285. [PubMed](https://pubmed.ncbi.nlm.nih.gov/11284282/)
6. Zhi XY, Li WJ, Stackebrandt E. An update of the structure and 16S rRNA gene sequence-based definition of higher ranks of the class *Actinobacteria*, with the proposal of two new suborders and four new families and emended descriptions of the existing higher taxa. *Int J Syst Evol Microbiol* 2009; **59**:589-608. [PubMed](https://pubmed.ncbi.nlm.nih.gov/19111111/)  
[doi:10.1099/ijs.0.65780-0](https://doi.org/10.1099/ijs.0.65780-0)
7. Holt JG, Krieg NR, Sneath PHA, Staley JT, Williams ST. 1994. Bergey's Manual of Determinative Bacteriology, 9th ed. Williams & Wilkins, Baltimore.
8. Stackebrandt E, Schumann P. 2006. *Archaea. Bacteria: Firmicutes, Actinomycetes*, p. 297-321. In M Dworkin, S Falkow, E Rosenberg, K-H Schleifer, E Stackebrandt (eds), *The Prokaryotes: A Handbook on the Biology of Bacteria*, Vol. 3. Springer Science + Business Media LLC, New York, printed in Singapore
9. Chun J, Lee JH, Jung Y, Kim M, Kim S, Kim BK, Lim YW. EzTaxon: a web-based tool for the identification of prokaryotes based on 16S ribosomal RNA gene sequences. *Int J Syst Evol Microbiol* 2007; **57**:2259-2261. [PubMed](https://pubmed.ncbi.nlm.nih.gov/17311111/)  
[doi:10.1099/ijs.0.64915-0](https://doi.org/10.1099/ijs.0.64915-0)
10. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res* 2009; **37**:D26-D31. [PubMed](https://pubmed.ncbi.nlm.nih.gov/19111111/) [doi:10.1093/nar/gkn723](https://doi.org/10.1093/nar/gkn723)
11. Kunapuli U, Lueders T, Meckenstock RU. The use of stable isotope probing to identify key iron-reducing microorganisms involved in anaerobic benzene degradation. *ISME J* 2007; **1**:643-653. [PubMed](https://pubmed.ncbi.nlm.nih.gov/17311111/) [doi:10.1038/ismej.2007.73](https://doi.org/10.1038/ismej.2007.73)
12. Brodie EL, Desantis TZ, Joyner DC, Baek SM, Larsen JT, Andersen GL, Hazen TC, Richardson PM, Herman DJ, Tokunaga TK, et al. Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. *Appl Environ Microbiol* 2006; **72**:6288-6298. [PubMed](https://pubmed.ncbi.nlm.nih.gov/16811111/)  
[doi:10.1128/AEM.00246-06](https://doi.org/10.1128/AEM.00246-06)
13. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 2006; **72**:5069-5072. [PubMed](https://pubmed.ncbi.nlm.nih.gov/16811111/)  
[doi:10.1128/AEM.03006-05](https://doi.org/10.1128/AEM.03006-05)
14. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000; **17**:540-552. [PubMed](https://pubmed.ncbi.nlm.nih.gov/11284282/)
15. Lee C, Grasso C, Sharlow MF. Multiple sequence alignment using partial order graphs. *Bioinformatics* 2002; **18**:452-464. [PubMed](https://pubmed.ncbi.nlm.nih.gov/12311111/)  
[doi:10.1093/bioinformatics/18.3.452](https://doi.org/10.1093/bioinformatics/18.3.452)
16. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 2008; **57**:758-771. [PubMed](https://pubmed.ncbi.nlm.nih.gov/18311111/)  
[doi:10.1080/10635150802429642](https://doi.org/10.1080/10635150802429642)
17. Yarza P, Richter M, Peplies J, Euzéby J, Amann R, Schleifer KH, Ludwig W, Glöckner FO, Rosselló-Móra R. The All-Species Living Tree project: A 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst Appl Microbiol* 2008; **31**:241-250. [PubMed](https://pubmed.ncbi.nlm.nih.gov/18311111/)  
[doi:10.1016/j.syapm.2008.07.001](https://doi.org/10.1016/j.syapm.2008.07.001)
18. Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A. How many bootstrap replicates are necessary? *Lect Notes Comput Sci* 2009; **5541**:184-200. [doi:10.1007/978-3-642-02008-7\\_13](https://doi.org/10.1007/978-3-642-02008-7_13)
19. Liolios K, Chen IM, Mavromatis K, Tavernarakis N, Hugenholtz P, Markowitz VM, Kyrpides NC. The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res* 2009; **38**:D346-D354. [PubMed](https://pubmed.ncbi.nlm.nih.gov/19111111/)  
[doi:10.1093/nar/gkp848](https://doi.org/10.1093/nar/gkp848)
20. Sims D, Brettin T, Detter JC, Han C, Lapidus A, Copeland A, Glavina Del Rio T, Nolan M, Chen F, Lucas S, et al. Complete genome sequence of *Kytococcus sedentarius* type strain (541<sup>T</sup>). *Stand Genomic Sci* 2009; **1**:12-20. [doi:10.4056/sigs.761](https://doi.org/10.4056/sigs.761)
21. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, et al. The minimum information about a genome sequence (MIGS) specification. *Nat*

- Biotechnol* 2008; **26**:541-547. [PubMed](#)  
[doi:10.1038/nbt1360](https://doi.org/10.1038/nbt1360)
22. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc Natl Acad Sci USA* 1990; **87**:4576-4579. [PubMed](#)  
[doi:10.1073/pnas.87.12.4576](https://doi.org/10.1073/pnas.87.12.4576)
  23. Garrity GM, Holt JG. The Road Map to the Manual. In: Garrity GM, Boone DR, Castenholz RW (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 1, Springer, New York, 2001, p. 119-169.
  24. Euzéby JP, Tindall BJ. Nomenclatural type of orders: corrections necessary according to Rules 15 and 21a of the Bacteriological Code (1990 Revision), and designation of appropriate nomenclatural types of classes and subclasses. Request for an opinion. *Int J Syst Evol Microbiol* 2001; **51**:725-727. [PubMed](#)
  25. Buchanan RE. Studies in the nomenclature and classification of bacteria. II. The primary subdivisions of the Schizomycetes. *J Bacteriol* 1917; **2**:155-164. [PubMed](#)
  26. Lechevalier H, Lechevalier MP. Ultramicroscopic structure of *Intrasporangium calvum* (*Actinomyce-tales*). *J Bacteriol* 1969; **100**:522-525. [PubMed](#)
  27. Classification of bacteria and archaea in risk groups. TRBA 466.
  28. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al.* Gene Ontology: tool for the unification of biology. *Nat Genet* 2000; **25**:25-29. [PubMed](#) [doi:10.1038/75556](https://doi.org/10.1038/75556)
  29. Busse HJ, Schumann P. Polyamine profiles within genera of the class *Actinobacteria* with LL-diaminopimelic acid in the peptidoglycan. *Int J Syst Bacteriol* 1999; **49**:179-184. [PubMed](#)  
[doi:10.1099/00207713-49-1-179](https://doi.org/10.1099/00207713-49-1-179)
  30. Schumann P, Prauser H, Rainey FA, Stackebrandt E, Hirsch P. *Friedmanniella antarctica* gen. nov., sp. nov., an LL-diaminopimelic acid-containing actinomycete from Antarctic sandstone. *Int J Syst Bacteriol* 1997; **47**:278-283. [PubMed](#)  
[doi:10.1099/00207713-47-2-278](https://doi.org/10.1099/00207713-47-2-278)
  31. Klenk HP, Goeker M. En route to a genome-based classification of *Archaea* and *Bacteria*? *Syst Appl Microbiol* 2010; **33**:175-182. [PubMed](#)  
[doi:10.1016/j.syapm.2010.03.003](https://doi.org/10.1016/j.syapm.2010.03.003)
  32. Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, Kunin V, Goodwin L, Wu M, Tindall BJ, *et al.* A phylogeny-driven genomic encyclopaedia of *Bacteria* and *Archaea*. *Nature* 2009; **462**:1056-1060. [PubMed](#)  
[doi:10.1038/nature08656](https://doi.org/10.1038/nature08656)
  33. List of growth media used at DSMZ:  
[http://www.dsmz.de/microorganisms/media\\_list.php](http://www.dsmz.de/microorganisms/media_list.php).
  34. DOE Joint Genome Institute.  
<http://www.jgi.doe.gov>
  35. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using deBruijn graphs. *Genome Res* 2008; **18**:821-829. [PubMed](#)  
[doi:10.1101/gr.074492.107](https://doi.org/10.1101/gr.074492.107)
  36. Phrap and Phred for Windows, MacOS, Linux, and Unix. <http://www.phrap.com>
  37. Lapidus A, LaButti K, Foster B, Lowry S, Trong S, Goltsman E. POLISHER: An effective tool for using ultra short reads in microbial genome assembly and finishing. AGBT, Marco Island, FL, 2008.
  38. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010; **11**:119. [PubMed](#) [doi:10.1186/1471-2105-11-119](https://doi.org/10.1186/1471-2105-11-119)
  39. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC. Gene-PRIMP: a gene prediction improvement pipeline for prokaryotic genomes. *Nat Methods* 2010; **7**:455-457. [PubMed](#) [doi:10.1038/nmeth.1457](https://doi.org/10.1038/nmeth.1457)
  40. Markowitz VM, Ivanova NN, Chen IMA, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 2009; **25**:2271-2278. [PubMed](#)  
[doi:10.1093/bioinformatics/btp393](https://doi.org/10.1093/bioinformatics/btp393)