Draft genome sequence of a novel actinobacterium from the family *Intrasporangiaceae* isolated from Signy Island, Antarctica

S. M. Thomas¹,², P. Convey²,⁵, H. M. Gan³,⁴, D. A. Pearce²,⁵,⁶ and G. Y. A. Tan¹,²,*

¹Institute of Biological Sciences, and
²National Antarctic Research Centre, University of Malaya, 50603 Kuala Lumpur, Malaysia
³Genomics Facility, Tropical and Medicine Biology Platform, Monash University Malaysia, Jalan Lagoon Selatan, 47500 Bandar Sunway, Selangor, Malaysia
⁴Centre for Integrative Ecology, School of Life and Environmental Sciences, Deakin University, Geelong 3220, Australia
⁵British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge CB3 0ET, United Kingdom
⁶Department of Applied Sciences, Faculty of Health and Life Sciences, Northumbria University at Newcastle, Newcastle-upon-Tyne, NE1 8ST, United Kingdom

Actinobacterium strain S63T isolated from a soil sample collected from Spindrift Col on Signy Island (South Orkney Islands, maritime Antarctic) is a new species of the *Intrasporangiaceae* family. Here we report a draft genome sequence with an approximate size of 5 Mbp contained in 54 contigs (69.33% GC content). Preliminary analysis revealed the presence of cold active protein coding sequences, which may indicate an adaptation to the harsh polar environment from which the strain was isolated.

Keywords: Actinobacteria, Antarctic soil, *Humibacillus*.

SIGNY Island is located in the South Orkney Islands archipelago in the Atlantic sector of the Southern Ocean. The island is well known for its seals, penguins, skuas and other birds, as well as being a paradigmatic example of the terrestrial vegetation and ecosystems of the maritime Antarctic¹. Within its maximum dimensions (6.5 × 5 km), most terrestrial microbial communities here have a strong marine influence, receiving nutritional and water input through sea-spray and the movement of marine vertebrates. Culture-dependent and culture-independent studies increasingly indicate that diverse microbial life thrives, despite the environmental pressures of Antarctica²,³. It is clear that various Antarctic microbial communities harbour new and previously undescribed species⁴,⁵. Here we report the draft genome sequence of strain S63T, a novel species of the family *Intrasporangiaceae*, isolated from a soil sample collected from Signy Island.

Soils from Signy Island have yielded a diversity of bacteria⁶, one of which is Actinobacterium strain S63T. This strain was isolated from a sample collected from Spindrift Col on Signy Island, Antarctica. It was isolated on Starch Casein Nitrate agar⁷ supplemented with 2% NaCl and incubated at 15°C for up to 4 months.

Strain S63T is a Gram-positive coccoid bacterium. The strain is catalase positive, aerobic, non-motile and non-spor forming. The diaminopimelic acid isomer in the cell wall peptidoglycan is LL-DAP. This strain also produces amylase and lipase at 15°C, but does not produce biosurfactants in the drop-collapsing test⁸. Additional phenotypic information supporting the novelty of the strain is being documented for a future publication.

Total genomic DNA extraction was carried out using the MG™ DNA purification kit (Macrogen Inc., Korea) following the manufacturer’s instructions. Genomic DNA was fragmented using adaptive focused acoustic technology (AFA, Covaris) to a target size of 250 bp. The fragmented DNA was then end-repaired, ligated to Truseq adapters, and PCR-enriched using Truseq PCR master mix and a primer cocktail following the manufacturer’s protocol (TruSeq DNA sample preparation kit, Illumina, CA, USA). The final sequencing library was quantified using qPCR according to the qPCR quantification protocol guide (KAPA Library Quantification kits for illumina sequencing platforms) and the library size was confirmed using the Agilent 1000 DNA kit on an Agilent 2100 Bioanalyzer (Agilent Technologies, Böblingen, Germany). The resulting library was sequenced using a standard flow cell, and 200 cycles (100 paired end) on the illumina HiSeq2000 platform.

A total of 47,882,286 paired-end reads (4.79 Gbp) were generated on the HiSeq, of which only 6 million reads were subsequently subsampled for de novo assembly to reduce the generation of spurious contigs due to overly high sequencing depth⁹. Genome assembly was performed using Spades version 3.8.1 with the ‘careful’ option enabled¹⁰. The draft assembly was subsequently improved using in silico scaffolding and gap-closing¹¹,¹². The final draft genome has an accumulated length of 5,023,713 bp (69.33% GC, 120 × genome coverage) contained in 54 contigs with an N₅₀ of 194,740 bp.
Annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline, leading to the prediction of 4529 coding sequences. Among these, 2671 had annotated functions and 1858 were hypothetical proteins. One copy of each ribosomal RNA (5S, 16S and 23S) was identified in the draft genome, along with 47 tRNAs and 3 ncRNAs.

Based on 16S rRNA gene sequence analysis and BLAST results from the NCBI database, the strain was found to share <97% similarity with *Humibacillus xanthopallidus* KV-663\(^1\), the only validly described member of that genus. Four predicted proteins in the genome of strain S63\(^7\) were annotated as cold-shock proteins (Locus tags: BA895_07505, BA895_14940, BA895_17625 and BA895_20550). Subsequent protein domain analysis using InterProScan confirmed the presence of protein domains commonly found in cold-shock proteins such as IPR011129, IPR002059 and IPR012156, indicating potential capability of the strain to resist the extreme environment stresses of Antarctica.

Strain S63\(^1\) also possesses genes coding for enzymes involved in the conversion of maltose to trehalose. Trehalose is a source of energy as well as a cryoprotectant and is involved in anhydrobiosis in many groups of organisms.\(^13\) It could, in this case, assist this bacterium to withstand prolonged desiccation. The genome also contained alpha-amylase and glucoamylase genes, which are involved in the conversion of maltose to trehalose. Trehalose may also increase survival and is a source of energy as well as a cryoprotectant and is involved in anhydrobiosis in many groups of organisms.\(^7\)

The strain has been deposited in the DSMZ culture collection as DSM 29435\(^6\). The partial 16S rRNA gene sequence (1366 bp) has been deposited in the NCBI GenBank database under the accession number KJ547654. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the NCBI Reference Sequence: NZ_MAST00000000.1.

### Acknowledgements

We acknowledge the University of Malaya Research Grant (RP002E-13SUS & RP007-2012D) and the Ministry of Science, Technology and Innovation Flagship Grant (FP0712E012). P.C. is supported by Natural Environment Research Council core funding to the British Antarctic Survey’s ‘Biodiversity, Evolution and Adaptation’ Team. P.C. and D.A.P. are also supported by Visiting Professorships to the University of Malaya.

---


\(^8\) Thomas, S., Biodiversity and Bioactivity of Antarctic Actinobacteria, MSc thesis, University of Malaya, 2017.


\(^12\) Boetzer, M. and Pirovano, W., Toward almost closed genomes with GapFiller. *Genome Biol.*, 2012, 13, 1.
