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

S. M. Thomas[1,2], P. Convey[2,5], H. M. Gan[3,4], D. A. Pearce[2,5,6] and G. Y. A. Tan[1,2,*]

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

Actinobacterium strain S63^T 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 Antarctic[1]. 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 Antarctica[2,3]. It is clear that various Antarctic microbial communities harbour new and previously undescribed species[4,5]. Here we report the draft genome sequence of strain S63^T, 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 bacteria[6], one of which is Actinobacterium strain S63^T. This strain was isolated from a sample collected from Spindrift Col on Signy Island, Antarctica. It was isolated on Starch Casein Nitrate agar[7] supplemented with 2% NaCl and incubated at 15°C for up to 4 months.

Strain S63^T is a Gram-positive coccoid bacterium. The strain is catalase positive, aerobic, non-motile and non-spore 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 test[8]. 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 depth[9]. Genome assembly was performed using Spades version 3.8.1 with the 'careful' option enabled[10]. The draft assembly was subsequently improved using *in silico* scaffolding and gap-closing[11,12]. 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 N_{50} of 194,740 bp.

*For correspondence. (e-mail: gyatan@um.edu.my)

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^T, the only validly described member of that genus. Four predicted proteins in the genome of strain S63^T 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^T 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 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 a part of the gene complex for the biosynthesis of trehalose. Production of trehalose may also increase survival of cold stress, as shown in an *Arthrobacter* sp. isolated from an Alpine permafrost sample, which was found to accumulate trehalose upon cold shock[14].

The strain has been deposited in the DSMZ culture collection as DSM 29435^T. 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.

1. Smith, R. I. L., Signy Island as a paradigm of biological and environmental change in Antarctic terrestrial ecosystems. In *Antarctic Ecosystems. Ecological Change and Conservation* (eds Kerry, K. R. and Hempel, G.), Springer, Berlin, 1990, pp. 32–50.
2. Cavicchioli, R., Microbial ecology of Antarctic aquatic systems. *Nat. Rev. Microbiol.*, 2015, **13**, 691–706.
3. Cowan, D. A. and Tow, L. A., Endangered Antarctic environments. *Annu. Rev. Microbiol*., 2004, **58**, 649–690.
4. Babalola, O. O., Kirby, B. M., Le Roes-Hill, M., Cook, A. E., Cary, S. C., Burton, S. G. and Cowan, D. A., Phylogenetic analysis of actinobacterial populations associated with Antarctic dry valley mineral soils. *Environ. Microbiol.*, 2009, **11**, 566–576.
5. Chong, C. W., Goh, Y. S., Convey, P., Pearce, D. A. and Tan, I. K. P., Spatial pattern in Antarctica: what can we learn from Antarctic bacterial isolates? *Extremophiles*, 2013, **17**, 733–745.
6. Pan, S. Y., Tan, G. Y. A., Convey, P., Pearce, D. A. and Tan, I. K. P., Diversity and bioactivity of actinomycetes from Signy Island terrestrial soils, maritime Antarctic. *Adv. Polar Sci*., 2013, **24**, 208–212.
7. Küster, E. and Williams, S., Selection of media for isolation of streptomycetes. *Nature*, 1964, **202**, 928–929.
8. Thomas, S., *Biodiversity and Bioactivity of Antarctic Actinobacteria*, M Sc thesis, University of Malaya, 2017.
9. Gan, H. M., Lee, Y. P. and Austin, C. M., Nanopore long-read guided complete genome assembly of *Hydrogenophaga intermedia* and genomic insights into 4-aminobenzenesulfonate, *p*-aminobenzoic acid and hydrogen metabolism in the genus *Hydrogenophaga*. *Front. Microbiol*., 2017, **8**, 1880.
10. Nurk, S. *et al.*, Assembling genomes and mini-metagenomes from highly chimeric reads. In Proceedings of the 17th Annual International Conference on Research in Computational Molecular Biology (eds Deng, M., Jiang, R., Sun, F. and Zhang, X.), 2013, Springer, Berlin, pp. 158–170.
11. Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D. and Pirovano, W., Scaffolding pre-assembled contigs using SSPACE. In *Bioinformatics*, Oxford, England, 2011, vol. 27, pp. 578–579.
12. Boetzer, M. and Pirovano, W., Toward almost closed genomes with GapFiller. *Genome Biol.*, 2012, **13**, 1.
13. Everatt, M. J., Convey, P., Bale, J. S., Worland, M. R. and Hayward, S. A. L., Responses of invertebrates to temperature and water stress: a polar perspective. *J. Therm. Biol.*, 2014, **54**, 118–132.
14. Chen, X. M. *et al.*, Regulation of expression of trehalose-6-phosphate synthase during cold shock in *Arthrobacter* strain A3. *Extremophiles*, 2011, **15**, 499–508.

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.

doi: 10.18520/cs/v115/i9/1695-1696