

Isolation, Identification, Whole-Genome Sequencing, and Annotation of Four *Bacillus* Species, *B. anthracis* RIT375, *B. circulans* RIT379, *B. altitudinis* RIT380, and *B. megaterium* RIT381, from Internal Stem Tissue of the Insulin Plant *Costus igneus*

Steven J. Polter,^a Alexander A. Caraballo,^b Yin P. Lee,^{c,d} Wilhelm W. H. Eng,^{c,d} Han M. Gan,^{c,d} Matthew S. Wheatley,^a Michael A. Savka,^a Bolaji N. Thomas,^b André O. Hudson^a

Thomas H. Gosnell School of Life Sciences, Rochester Institute of Technology, Rochester, New York, USA^a; College of Health Science and Technology, Rochester Institute of Technology, Rochester, New York, USA^b; Monash University Malaysia Genomics Facility, Monash University Malaysia, Selangor, Malaysia^c; School of Science, Monash University Malaysia, Selangor, Malaysia^d

S.J.P. and A.A.C. contributed equally to this work.

Here, we report the isolation, identification, whole-genome sequencing, and annotation of four *Bacillus* species from internal stem tissue of the insulin plant *Costus igneus*, grown in Puerto Rico. The plant is of medicinal importance, as extracts from its leaves have been shown to lower blood sugar levels of hyperglycemic rats.

Received 22 June 2015 Accepted 22 June 2015 Published 30 July 2015

Citation Polter SJ, Caraballo AA, Lee YP, Eng WWH, Gan HM, Wheatley MS, Savka MA, Thomas BN, Hudson AO. 2015. Isolation, identification, whole-genome sequencing, and annotation of four *Bacillus* species, *B. anthracis* RIT375, *B. circulans* RIT379, *B. altitudinis* RIT380, and *B. megaterium* RIT381, from internal stem tissue of the insulin plant *Costus igneus*. *Genome Announc* 3(4):e00847-15. doi:10.1128/genomeA.00847-15.

Copyright © 2015 Polter et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to André O. Hudson, aohsbi@rit.edu.

The pantropical plant *Costus igneus*, commonly referred to as the insulin plant, belongs to the *Costaceae* family and is considered a medicinal plant due to its antidiabetic properties. Recent studies have shown that extracts from the leaves of *C. igneus* were able to lower the blood sugar levels of hyperglycemic rats (1). In addition to its medicinal properties, phytochemical screenings of extracts from *C. igneus* revealed that iron, ascorbic acid, α -tocopherol, and β -carotene are in high abundance (2). We were interested in the isolation and identification of endophytic bacteria that associate with *C. igneus*, to gain some insights into the plant-bacteria symbiotic relationships. This study was facilitated by isolation, initial identification, and whole-genome sequencing and annotation of bacteria that associate with the plant. Briefly, internal tissue from a surface-sterilized stem, obtained from a farm in Puerto Rico, was used to inoculate tryptic soy broth (TSB) medium. Cultures were then serially diluted and plated on several media (tryptic soy agar, Luria agar, nutrient agar, potato dextrose agar, and R2A), which subsequently led to the isolation of several bacteria. Four Gram-positive bacteria were chosen for further analysis based on colony morphology and mi-

croscopic examination following Gram staining. The bacteria were identified as *B. anthracis*, *B. circulans*, *B. altitudinis*, and *B. megaterium*, based on nucleotide sequence analysis of the variable 3 (V3) region from the 16S rRNA gene amplified using the PCR primers V3-forward (5'-ACTCCTACGGGAGGCAGCAG-3') and V3-reverse (5'-ATTACCGCGGCTGCTGG-3') (3).

Genomic DNA was isolated from the endophytes using the GenElute bacterial genomic kit (Sigma-Aldrich, St. Louis, MO, USA) and prepared for whole-genome sequencing using the Nextera XT library preparation kit (Illumina, San Diego, CA, USA). Whole-genome sequencing was performed on the Illumina MiSeq (2 × 150-bp paired-end reads) located at the Monash University Malaysia Genomics Facility. The reads were error corrected and assembled *de novo* using SPAdes version 3.5 (4). Scaffolding of the contigs and *in silico* gap closing of the resulting scaffolds were performed using SSPACE and GapFiller, respectively (5, 6). Genome annotation was performed using the NCBI Prokaryotic Genomes Annotation Pipeline. The key attributes for each of the four genomes and annotations are summarized in Table 1.

A summary of the secondary metabolite analysis of the four

TABLE 1 Sequencing and annotation details of the four *Bacillus* species isolated from *Costus igneus*

Strain	BioProject no.	BioSample no.	Accession no.	Organism	Genome coverage (×)	Genome size (bp)	No. of contigs	No. of ORFs ^a	No. of tRNAs	No. of rRNAs
RIT375	PRJNA285407	SAMN03753508	LDPG00000000	<i>Bacillus anthracis</i> RIT375	66	5,677,918	62	5,666	98	15
RIT379	PRJNA285407	SAMN03753509	LDPH00000000	<i>Bacillus circulans</i> RIT379	71	5,454,271	103	4,950	85	22
RIT380	PRJNA285407	SAMN03753582	LDPI00000000	<i>Bacillus altitudinis</i> RIT380	97	3,972,159	85	3,985	87	9
RIT381	PRJNA285407	SAMN03753583	LDPJ00000000	<i>Bacillus megaterium</i> RIT381	41	5,864,214	240	5,616	151	41

^a ORFs, open reading frames.

species using the antibiotics Secondary Metabolite Analysis Shell (antiSMASH) tool revealed interesting results regarding the identification of gene clusters that are involved in the synthesis of secondary metabolites. For example, the genome of *Bacillus anthracis* RIT375 resulted in the identification of gene clusters for the synthesis of a lasso peptide in contig 1 and a petrobactin in contig 3, and *Bacillus circulans* RIT379 possess gene clusters for lantipeptide and cepacian in contigs 7 and 11. *Bacillus altitudinis* RIT380 possess gene clusters for bacilysin, lichenysin, and sporulation killing factor in contigs 8, 9, and 13, respectively. *Bacillus megaterium* RIT381 possessed gene clusters for surfactin and carotenoid in contigs 1 and 3 (7, 8).

Nucleotide sequence accession numbers. The nucleotide sequences have been deposited at DDBJ/EMBL/GenBank under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

A.O.H. and M.A.S. thank the College of Science (COS) at the Rochester Institute of Technology (RIT) for ongoing support. A.A.C. is supported by the Louis Stokes Alliance for Minority Participation (LSAMP) through the National Science Foundation (HRD-1202480). This work was also supported by the Monash University Malaysia Tropical Medicine and Biology Multidisciplinary Platform.

REFERENCES

- Shetty AJ, Choudhury D, Rejeesh Nair V, Kuruvilla M, Kotian S, Kotian S. 2010. Effect of the insulin plant (*Costus igneus*) leaves on dexamethasone-induced hyperglycemia. *Int J Ayurveda Res* 1: 100–102.
- Hegde PK, Rao HA, Rao PN. 2014. A review on insulin plant (*Costus igneus* NaK). *Pharmacogn Rev* 8:67–72. <http://dx.doi.org/10.4103/0973-7847.125536>.
- Hudson AO, Ahmad NH, Van Buren R, Savka MA. 2010. Sugarcane and grapevine endophytic bacteria: isolation, detection of quorum sensing signals and identification by 16S v3 rDNA sequence analysis, p 801–806. *In* Mendez-Vilas A (ed), *Current research, technology and education topics in applied microbiology and microbial biotechnology*, vol 2. Formatex Research Center, Badajoz, Spain.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <http://dx.doi.org/10.1089/cmb.2012.0021>.
- Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579. <http://dx.doi.org/10.1093/bioinformatics/btq683>.
- Boetzer M, Pirovano W. 2012. Toward almost closed genomes with Gap-Filler. *Genome Biol* 13:R56. <http://dx.doi.org/10.1186/gb-2012-13-6-r56>.
- Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R. 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* 39:W339–W346. <http://dx.doi.org/10.1093/nar/gkr466>.
- Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T. 2013. antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res* 41:W204–W221. <http://dx.doi.org/10.1093/nar/gkt449>.