



Genome Announcement

Draft genome sequence of *Streptomyces* sp. strain F1, a potential source for glycoside hydrolases isolated from Brazilian soil



Ricardo Rodrigues de Melo^{a,b,1}, Gabriela Felix Persinoti^{a,1},
Douglas Antonio Alvaredo Paixão^a, Fábio Márcio Squina^a,
Roberto Ruller^{a,*}, Helia Harumi Sato^{b,*}

^a Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Campinas, São Paulo, Brazil

^b Universidade Estadual de Campinas (UNICAMP), Faculdade de Engenharia de Alimentos, Departamento de Ciência de Alimentos, Campinas, São Paulo, Brazil

ARTICLE INFO

Article history:

Received 26 September 2016

Accepted 22 November 2016

Available online 10 February 2017

Associate Editor: John McCulloch

Keywords:

Actinobacteria

Glycoside hydrolases

Streptomyces

Draft genome sequence

Soil bacteria

ABSTRACT

Here, we show the draft genome sequence of *Streptomyces* sp. F1, a strain isolated from soil with great potential for secretion of hydrolytic enzymes used to deconstruct cellulosic biomass. The draft genome assembly of *Streptomyces* sp. strain F1 has 69 contigs with a total genome size of 8,142,296 bp and G + C 72.65%. Preliminary genome analysis identified 175 proteins as Carbohydrate-Active Enzymes, being 85 glycoside hydrolases organized in 33 distinct families. This draft genome information provides new insights on the key genes encoding hydrolytic enzymes involved in biomass deconstruction employed by soil bacteria.

© 2017 Sociedade Brasileira de Microbiologia. Published by Elsevier Editora Ltda. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Streptomyces species are aerobic Gram-positive bacteria best known industrially as producers of natural antibiotics,¹ but they are also recognized for their capacity to utilize cellulosic biomass.² Phylogenetically, *Streptomyces* is the largest genus

of the Actinobacteria phylum. During their lifetime, these soil bacteria are able to differentiate, produce aerial mycelia and a wide variety of secondary metabolites.³ Although a large number of *Streptomyces* species can grow on plant biomass, understanding of key genes encoding hydrolytic enzymes involved in biomass degrading by *Streptomyces* is currently limited to a few soil-isolates.^{2,4–7} *Streptomyces* sp. strain F1

* Corresponding authors.

E-mails: roberto.ruller@bioetanol.org.br (R. Ruller), heliah@fea.unicamp.br (H.H. Sato).

¹ These authors contributed equally to this work.

<http://dx.doi.org/10.1016/j.bjm.2016.11.010>

1517-8382/© 2017 Sociedade Brasileira de Microbiologia. Published by Elsevier Editora Ltda. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Table 1 – Summary of genome features of *Streptomyces* sp F1.

Features	
Source	Soil
Genome size, Mb	8.1
GC content %	0.73
tRNA	90
rRNA	3
Protein coding sequences	7262

was isolated from soil containing decomposing organic matter collected in Campinas, São Paulo, Brazil. This isolated strain showed ability to grow in culture medium containing cellulose or hemicellulose as sole carbon source, and to secrete extracellular enzymes belonging to the glycoside hydrolases (GHs) families. Glycoside hydrolases are a group of enzymes that play an important role in the conversion of lignocellulosic biomass into small chemical building blocks, which can then be used to produce biofuels and other important intermediary molecules.⁸ Here, we show the draft genome sequence of *Streptomyces* sp. F1, to identify GHs family members and to improve understanding of natural biomass utilization by soil bacteria.

Genomic DNA extraction from *Streptomyces* sp. F1 was carried out using FastDNA SPIN Kit for soil (MP Biomedicals, Irvine, CA) according to the manufacturer's instructions. The genome was sequenced by whole genome shotgun sequencing using the Illumina HiSeq 2500 System at CTBE Sequencing and Robotics NGS facility, generating 8,147,881 paired end reads (2× 100 bp). Reads were preprocessed with Trimmomatic,⁹ to remove low-quality and adapter sequences and were assembled using Spades version 3.6.¹⁰ The genome size was estimated to be 8,205,272, with approximately 100× coverage. The draft genome assembly of *Streptomyces* sp. F1 has 69 contigs, 8,142,296 bp in length with G+C content of 72.65% (Table 1), an N50 of 296,926 bp, and the largest contig was 760,841 bp. Genome completeness was evaluated using CheckM,¹¹ which revealed that the assembly is 100% complete, considering 460 marker genes from Streptomycetaceae family.

Streptomyces sp. F1 showed highest 16S rDNA sequence similarity with *Streptomyces misionensis* strain NRRL B-3230^T. In silico DNA–DNA hybridization (DDH)¹² and Average Nucleotide Identity/Alignment fraction (gANI/AF)¹³ values of *Streptomyces* sp. F1 compared to *Streptomyces misionensis* DSM 40306, were 94.2% and 99.4%/0.99, respectively, suggesting that strain F1 may be classified as *Streptomyces misionensis*.

Streptomyces sp. F1 genome was annotated using IMG-JGI Microbial Genome Annotation Pipeline (img.jgi.doe.gov). It has been predicted to include 7355 genes, being 7262 protein-coding genes, 3 rRNA (5S (1), 16S (1), 23S (1)), and 90 tRNA genes (Table 1). According to IMG functional annotation, 4453 genes were classified into COG categories, 5526 in PFAM protein families, 1542 in TIGRFAM families, and 714 in Transporter Classification. Further classification according to dbCAN showed that 175 proteins were classified as Carbohydrate-Active Enzymes, being 85 glycoside hydrolases organized in 33 distinct families. The current genome

assembly provides a preliminary landscape of the genomic and metabolic capabilities of *Streptomyces* sp. F1.

Nucleotide sequence accession number

The whole genome sequences of *Streptomyces* sp. F1 have been deposited at DDBJ/EMBL/GenBank under accession number FKJI03000000.

Conflict of interest

The authors declare no conflicts of interest.

Acknowledgements

The authors gratefully acknowledge the Brazilian National Council for Scientific and Technological Development (CNPq) for their financial support and fellowships, and CNPEM-CTBE for the use of Sequencing and Robotics NGS facility.

REFERENCES

- Barbe V, Bouzon M, Mangenot S, et al. Complete genome sequence of *Streptomyces cattleya* NRRL 8057, a producer of antibiotics and fluorometabolites. *J Bacteriol.* 2011;193(18):5055–5056.
- Takasuka TE, Book AJ, Lewin GR, Currie CR, Fox BG. Aerobic deconstruction of cellulosic biomass by an insect-associated *Streptomyces*. *Sci Rep.* 2013;3:1030.
- Chater KF, Biró S, Lee KJ, Palmer T, Schrepf H. The complex extracellular biology of *Streptomyces*. *FEMS Microbiol Rev.* 2010;34(2):171–198.
- Franco-Cirigliano MN, Rezende RDC, Gravina-Oliveira MP, et al. *Streptomyces misionensis* PESB-25 produces a thermoacidophilic endoglucanase using sugarcane bagasse and corn steep liquor as the sole organic substrates. *Biomed Res Int.* 2013;2013:1–10.
- Wibberg D, Al-Dilaimi A, Busche T, et al. Complete genome sequence of *Streptomyces reticuli*, an efficient degrader of crystalline cellulose. *J Biotechnol Elsevier BV.* 2016;222:13–14.
- Book AJ, Lewin GR, McDonald BR, et al. Cellulolytic *Streptomyces* strains associated with herbivorous insects share a phylogenetically linked capacity to degrade lignocellulose. *Appl Environ Microbiol.* 2014;80(15):4692–4701.
- Semêdo LTAS, Gomes RC, Linhares AA, et al. *Streptomyces drozdowiczii* sp. nov., a novel cellulolytic streptomycete from soil in Brazil. *Int J Syst Evol Microbiol.* 2004;54(4):1323–1328.
- Horn S, Vaaje-Kolstad G, Westereng B, Eijsink VG. Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels.* 2012;5(1):45.
- Bolger AM, Lohse M, Trimmomatic Usadel B. A flexible trimmer for Illumina Sequence Data. *Bioinforma Adv.* 2014;30:2114.
- Bankevich A, Nurk S, Antipov D, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012;19(5):455–477.

11. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW, Check M. Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 2015;25(7):1043-1055.
12. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics.* 2013;14:60.
13. Varghese NJ, Mukherjee S, Ivanova N, et al. Microbial species delineation using whole genome sequences. *Nucleic Acids Res.* 2015;43(14):6761-6771.