The Complete Genome Sequence of *Streptomyces cf. griseus* (XyelbKG-1), an Ambrosia Beetle-Associated Actinomycete.

Kirk J. Grubbs, Peter H.W. Biedermann, Garret Suen, Sandra M. Adams, Joseph A. Moeller, Jonathan L. Klassen, Lynne A. Goodwin, Tanja Woyke, A. Christine Munk, David Bruce, Chris Detter, Roxanne Tapia, Cliff S. Han and Cameron R. Currie

Department of Bacteriology, University of Wisconsin-Madison, Madison, Wisconsin, Cellular and Molecular Pathology Graduate Program, University of Wisconsin-Madison, Madison, Wisconsin, Institute of Ecology and Evolution, Division Behavioral Ecology, University of Berne, Bern, Switzerland, DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, Wisconsin, DOE Joint Genome Institute, Walnut Creek, California, Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico

ABSTRACT

*Streptomyces cf. griseus* (XylebKG-1) is an insect-associated strain of the well-studied Actinobacterial species *S. griseus*. Here we present the genome of XylebKG-1 and discuss its similarity to the genome of *S. griseus* subsp. griseus NBRC13350. XylebKG-1 was isolated from the fungus cultivating *Xyleborinus saxesenii* system. Given its similarity to free-living *S. griseus* subsp. griseus NBRC13350, comparative genomics will elucidate critical components of bacterial interactions with insects.

*Corresponding author. Mailing address: DOE Great Lakes Bioenergy Research Center and Department of Bacteriology, 6155 MSB, 1550 Linden Drive, University of Wisconsin-Madison, Madison, WI 53706. Phone: (608)265-8034. Fax: (608)262-9865. E-mail: currie@bact.wisc.edu.
Streptomyces griseus is a soil bacterium known for its production of secondary metabolites including streptomycin, the first effective antibiotic for tuberculosis (19). Here we present the genome sequence of Streptomyces cf. griseus (XylebKG-1), which, to our knowledge, is the first strain of S. griseus associated with an insect. XylebKG-1 was isolated from the ambrosia beetle Xyleborinus saxeseni, which cultivates a fungus for food (2). Actinobacteria-specific isolations from both beetles and their fungal galleries resulted in isolation of XylebKG-1.

DNA from pure isolates was extracted using a bead-beating protocol (10) and the genome was sequenced at the DOE Joint Genome Institute (JGI). A non-contiguous finished genome of XylebKG-1 was generated using a shotgun approach employing a combination of Illumina (3) and 454 sequencing technologies (15). An Illumina shotgun library (69,927,062 reads totaling ~5.3 Gbp) and two 454 GS (FLX Titanium) shotgun libraries (688,595 standard reads and 172,566 20 Kbp paired-end reads totaling ~352 Mbp) were sequenced and assembled. The 454 data was assembled using Newbler, version 2.3 (Roche) and Illumina sequencing data was assembled with VELVET, version 0.7.63 (20). All assemblies were integrated using parallel phrap, version SPS D 4.24 (High Performance Software, LLC). Illumina data was used to correct potential base errors and increase consensus quality using the software Polisher developed at the JGI (Alla Lapidus, unpublished). Possible misassembled regions were corrected using gapResolution (Cliff Han, unpublished), Dupfinisher (11), or by sequencing cloned bridging PCR fragments with subcloning. Gaps between contigs were closed by editing in Consed (6-8), by PCR and by Bubble PCR (J-F Cheng, unpublished) primer walks.

The total size of the genome is 8,727,768 bp and the final assembly is based on 352.4 Mbp of 454 sequence (38X coverage) and 5.3 Gbp of Illumina sequence (257x coverage).

The overall genome G+C content is 72.1%. Generation (http://compbio.ornl.gov/generation/), Glimmer (5) and Critica (v1.05) (1) were used to predict a total of 7,265 candidate protein-encoding gene models. RNAmmer (13) annotated six 16S
rRNAs and six 23S rRNAs. A tRNAscan-SE (14) search revealed 66 tRNAs corresponding to all 20 standard amino acids. Additional PRIAM (4), KEGG (12) and COG (18) analyses were completed and can be accessed at http://genome.ornl.gov/microbial/streACT1.

Evidence for XylebKG-1 as a strain of S. griseus is based on similarity between the genomes of XylebKG-1 and the type strain S. griseus subsp. griseus NBRC13350 (16). Both have similar genome sizes (8.7 Mbp to 8.5 Mbp), G+C content (72.1% to 72.2%), six rRNA operons and 66 tRNAs. An average nucleotide identity analysis conducted between both genomes using Jspecies (v1.2.1) (17) revealed a 98.98 ANIb value (91.68% genome alignment) and a 98.95 ANIm value (94.13% genome alignment), indicating a species level degree of similarity (9). Given this similarity, XylebKG-1 represents a unique opportunity to study genetic elements involved in Actinobacteria/Insect associations.

Nucleotide sequence accession number: The genome sequence of Streptomyces cf. griseus XylebKG-1 is deposited in GenBank under accession ADFC00000000.

Acknowledgements. This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494) supporting G.S., S.M.A., J.A.M. and C.R.C. This work was also funded by the National Institute of Food and Agriculture, United States Department of Agriculture, under ID number WISO1321 supporting K.J.G. P.H.W.B. was partly supported by a DOC-fellowship of the Austrian Academy of Sciences at the Department of Behavioral Ecology, University of Bern. The work conducted by the US Department of Energy Joint Genome Institute is supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231. Additional support was provided by the National Science Foundation (MCB-0702025). Support for J.L.K. comes from an NSERC postdoctoral fellowship.
69

References


2. Batra, L. R. 1967. Ambrosia Fungi: A Taxonomic Revision, and Nutritional Studies of


Sequencer Traces UsingPhred. I. Accuracy Assessment. Genome Research 8:175-
185.


and J. M. Tiedje. 2007. DNA-DNA hybridization values and their relationship to whole-


11. Han, C. and P. Chain. 2006. Finishing Repetitive Regions Automatically with
Dupfinisher. In H. Valafar (ed.), Proceedings of the 2006 International Conference on
Bioinformatics Computational Biology, BIOCOMP’06. CSREA Press, Las Vegas,
Nevada, USA.

for representation and analysis of molecular networks involving diseases and drugs.
Nucleic Acids Research 38:D355-D360.


