

Genome Sequence of *Fusarium* Isolate MYA-4552 from the Midgut of *Anoplophora glabripennis*, an Invasive, Wood-Boring Beetle

 Joshua R. Herr,^a Erin D. Scully,^b Scott M. Geib,^c Kelli Hoover,^d John E. Carlson,^e David M. Geiser^f

Department of Plant Pathology and Center for Plant Science Innovation, University of Nebraska, Lincoln, Nebraska, USA^a; Stored Product Insect and Engineering Research Unit, USDA, Agricultural Research Service, Center for Grain and Animal Health Research, Manhattan, Kansas, USA^b; Tropical Crop and Commodity Protection Research Unit, USDA-ARS Daniel K. Inouye Pacific Basin Agricultural Research Center, Hilo, Hawaii, USA^c; Department of Entomology and Center for Chemical Ecology, The Pennsylvania State University, University Park, Pennsylvania, USA^d; The Schatz Center for Tree Molecular Genetics, Department of Ecosystem Science and Management, The Pennsylvania State University, University Park, Pennsylvania, USA^e; Department of Plant Pathology & Environmental Microbiology, The Pennsylvania State University, University Park, Pennsylvania, USA^f

J.R.H. and E.D.S. contributed equally to this work.

The *Fusarium solani* species complex (FSSC) is a clade of environmentally ubiquitous fungi that includes plant, animal, and insect associates. Here, we report the draft genome sequence of the undescribed species FSSC 6 (isolate MYA-4552), housed in the gut of the wood-boring cerambycid beetle *Anoplophora glabripennis*.

Received 24 May 2016 Accepted 30 May 2016 Published 21 July 2016

Citation Herr JR, Scully ED, Geib SM, Hoover K, Carlson JE, Geiser DM. 2016. Genome sequence of *Fusarium* isolate MYA-4552 from the midgut of *Anoplophora glabripennis*, an invasive, wood-boring beetle. *Genome Announc* 4(4):e00544-16. doi:10.1128/genomeA.00544-16.

Copyright © 2016 Herr et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to David M. Geiser, dgeiser@psu.edu.

The *Fusarium solani* species complex (FSSC) is a ubiquitous and functionally diverse group of fungi consisting of plant pathogens (1), human pathogens (2), and insect associates, including nutritional symbionts (3–5). Of the symbionts inhabiting the Asian longhorned beetle (ALB) gut, FSSC 6 has been persistently vertically transmitted in a colony maintained in quarantine at Penn State University, PA, for 16 generations and has been detected in all field-collected specimens in the United States (5, 6). Previous analyses have demonstrated that FSSC 6 can degrade lignocellulose (7), synthesize several essential amino acids (6), and contribute to the synthesis of sterols and other essential nutrients (8). To more thoroughly understand the metabolic capacities of FSSC 6 and its potential roles in ALB digestive physiology, we sequenced and assembled the genome of isolate ATCC MYA-4522 derived from an ALB larva collected in New York (4, 5).

Genomic DNA was extracted from freeze-dried tissue using the QIAamp DNA minikit (Qiagen, Redwood City, CA). Following extraction, genomic DNA (gDNA) concentration and integrity were assessed via agarose gel electrophoresis and using a NanoDrop ND-8000 UV-Vis spectrophotometer (Thermo Fisher Scientific). Three 454 FLX+ pyrosequencing libraries (Roche, Branford, CT) were prepared using the GS FLX Titanium sequencing kit XLR70, according to the manufacturer's specifications. The three 454 FLX+ runs provided a total of 7.138 Gb of single-end sequencing data. Low-quality base calls and 454-specific sequencing adapters were removed using Trimmomatic (9). Sequencing errors, assessed as erroneous k-mers, and redundant data over 30× coverage were filtered using the khmer software package (10). A combined filtered data set provided initial assembly drafts using the SPAdes (11), SOAPdenovo (12), and Velvet (13) genome assembly programs, and contigs from each

assembler were combined into scaffolds using Mix (14) and the haploid *Nectria haematococca* (FSSC undescribed species [11]) genome assembly as a guide (15). The final assembly consisted of 103 contigs, with a G+C content of 51.84%, and with a total length of 46.9 Mbp; this length represents 99% of the estimated genome size from k-mer spectral analysis (10). The largest contig (prescaffolding) was 4,887,614 bases in length, and the assembly consisted of an N_{50} of 1,729,713 bp, with half of the genome represented by eight contigs. While approximately 91.9% of the genome assembly aligned to the FSSC 11 genome, 200 kbp of small contigs putatively represent supernumerary chromosomal regions. Total reads, as well as small contigs, were mapped using Bowtie 2 (16) to both the genome assembly and to the NCBI reference genome collection and were verified to be of fungal origin. Utilizing gene models optimized for fungi and gene features from previously sequenced *Fusarium* genomes (15, 17), the program MAKER (18) was used to annotate the FSSC 6 MYA-4552 genome. This annotation predicted 13,157 genes represented by open reading frames.

Nucleotide sequence accession number. The whole-genome shotgun sequencing project for the FSSC 6 isolate MYA-4552 has been deposited at DDBJ/EMBL/GenBank under the accession no. [LWBZ00000000](https://www.ncbi.nlm.nih.gov/nuccore/LWBZ00000000). The version described in this paper is the first version of the assembly.

ACKNOWLEDGMENTS

This research was supported by grant USDA-NRI-CRSEES, grant 2009-35302-05286, and Korean World Class University Project R31-2009-000-20025-0.

FUNDING INFORMATION

This work, including the efforts of Erin Scully, Scott Geib, and Kelli Hoover, was funded by USDA-NRI-CSREES (2009-35302-05286). This work, including the efforts of John Carlson, was funded by Korean World Class University (R31-2009-000-20025-0).

The U.S. Department of Agriculture, Agricultural Research Service, is an equal opportunity/affirmative action employer and all agency services are available without discrimination. Mention of commercial products and organizations in this manuscript is solely to provide specific information. It does not constitute endorsement by USDA-ARS over other products and organizations not mentioned.

REFERENCES

- Zhang N, O'Donnell K, Sutton DA, Nalim FA, Summerbell RC, Padhye AA, Geiser DM. 2006. Members of the *Fusarium solani* species complex that cause infections in both humans and plants are common in the environment. *J Clin Microbiol* 44:2186–2190. <http://dx.doi.org/10.1128/jcm.00120-06>.
- Short DPG, O'Donnell K, Thrane U, Nielsen KF, Zhang N, Juba JH, Geiser DM. 2013. Phylogenetic relationships among members of the *Fusarium solani* species complex in human infections and the descriptions of *F. keratoplasticum* sp. nov. and *F. petroliphilum* stat. nov. *Fungal Genet Biol* 53:59–70 <http://dx.doi.org/10.1016/j.fgb.2013.01.004>.
- Kasson MT, O'Donnell K, Rooney AP, Sink S, Ploetz RC, Ploetz JN, Konkol JL, Carrillo D, Freeman A, Mendel Z, Smith JA, Black AW, Hulcr J, Bateman C, Stefkova K, Campbell PR, Geering ADW, Dann EK, Eskalen A, Mohotti K, Short DPG, Aoki T, Fenstermacher KA, Davis DD, Geiser DM. 2013. An inordinate fondness for *Fusarium*: phylogenetic diversity of fusaria cultivated by ambrosia beetles in the genus *Euwallacea* on avocado and other plant hosts. *Fungal Genet Biol* 56:147–157. <http://dx.doi.org/10.1016/j.fgb.2013.04.004>.
- Geib SM, Filley TR, Hatcher PG, Hoover K, Carlson JE, Jimenez-Gasco Mdel M, Nakagawa-Izumi A, Sleighter RL, Tien M. 2008. Lignin degradation in wood-feeding insects. *Proc Natl Acad Sci USA* 105:12932–12937. <http://dx.doi.org/10.1073/pnas.0805257105>.
- Geib SM, Scully ED, Jimenez-Gasco Mdel M, Carlson JE, Tien M, Hoover K. 2012. Phylogenetic analysis of *Fusarium solani* associated with the Asian longhorned beetle, *Anoplophora glabripennis*. *Insects* 3:141–160. <http://dx.doi.org/10.3390/insects3010141>.
- Ayayee PA, Larsen T, Rosa C, Felton GW, Ferry JG, Hoover K. 2016. Essential amino acid supplementation by gut microbes of a wood-feeding cerambycid. *Environ Entomol* 45:66–73. <http://dx.doi.org/10.1093/ee/nvv153>.
- Scully ED, Geib SM, Hoover K, Tien M, Tringe SG, Barry KW, Glavina del Rio T, Chovatia M, Herr JR, Carlson JE. 2013. Metagenomic profiling reveals lignocellulose degrading system in a microbial community associated with a wood-feeding beetle. *PLoS One* 8:e73827. <http://dx.doi.org/10.1371/journal.pone.0073827>.
- Scully ED, Geib SM, Carlson JE, Tien M, McKenna D, Hoover K. 2014. Functional genomics and microbiome profiling of the Asian longhorned beetle (*Anoplophora glabripennis*) reveal insights into the digestive physiology and nutritional ecology of wood feeding beetles. *BMC Genomics* 15:1096. <http://dx.doi.org/10.1186/1471-2164-15-1096>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <http://dx.doi.org/10.1093/bioinformatics/btu170>.
- Crusoe MR, Alameldin H, Awad S, Boyce RR, Bucher E, Caldwell A, Cartwright R, Constantinides B, Edverson G, Fay S, Fenton J, Fenzi T, Fish J, Garcia-Gutierrez L, Garland J, Gluck J, Gonzalez I, Guermont S, Guo J, Gupta A, Herr JR, Howe A, Hyer A, Härpfer A, Irber L, Kidd R, Wiencko HL, Lin D, Lippi J, Mansour T, McNulty P, McDonald E, Mizzi J, Nederbragt L, Ortiz-Zuazaga H, Ory J, Pell J, Pepe-Ranney C, Picett R, Russ ZN, Schwarz E, Scott C, Seaman J, Sievert S, Simpson J, Skennerton CT, Spencer J, Srinivasan RR, Standage D, Stapleton J, et al. 2015. The khmer software package: enabling efficient nucleotide sequence analysis. *F1000Res* 4:900.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu S-M, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam T-W, Wang J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaScience* 1:18. <http://dx.doi.org/10.1186/2047-217X-1-18>.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <http://dx.doi.org/10.1101/gr.074492.107>.
- Soueidan H, Maurier F, Groppi A, Sirand-Pugnet P, Tardy F, Citti C, Dupuy V, Nikolski M. 2013. Finishing bacterial genome assemblies with mix. *BMC Bioinformatics* 14:S16. <http://dx.doi.org/10.1186/1471-2105-14-s15-s16>.
- Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, Schmutz J, Taga M, White GJ, Zhou S, Schwartz DC, Freitag M, Ma L-J, Danchin EGJ, Henrissat B, Coutinho PM, Nelson DR, Straney D, Napoli CA, Barker BM, Gribskov M, Rep M, Kroken S, Molnár I, Rensing C, Kennell JC, Zamora J, Farman ML, Selker EU, Salamov A, Shapiro H, Pangilinan J, Lindquist E, Lamers C, Grigoriev IV, Geiser DM, Covert SF, Temporini E, VanEtten HD. 2009. The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. *PLoS Genet* 5:e1000618. <http://dx.doi.org/10.1371/journal.pgen.1000618>.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <http://dx.doi.org/10.1038/nmeth.1923>.
- Cuomo CA, Güldener U, Xu J-R, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma L-J, Baker SE, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang Y-L, DeCaprio D, Gale LR, Gnerre S, Goswami RS, Hammond-Kosack K, Harris LJ, Hilburn K, Kennell JC, Kroken S, Magnuson JK, Mannhaupt G, Mauceli E, Mewes H-W, Mitterbauer R, Muehlbauer G, Munsterkötter M, Nelson D, O'Donnell K, Ouellet T, Qi W, Quesneville H, Roncero MIG, Seong K-Y, Tetko IV, Urban M, Waalwijk C, Ward TJ, Yao J, Birren BW, Kistler HC. 2007. The *Fusarium graminearum* genome reveals a Link between localized polymorphism and pathogen specialization. *Science* 317:1400–1402. <http://dx.doi.org/10.1126/science.1143708>.
- Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, Holt C, Sánchez Alvarado A, Yandell M. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res* 18:188–196. <http://dx.doi.org/10.1101/gr.6743907>.