

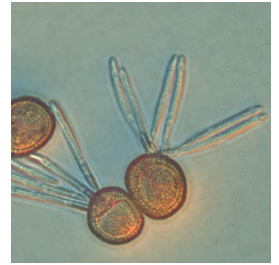


Sep 01, 2025

Version 1

Draft genome of the switchgrass head smut pathogen *Tilletia maclaganii* V.1

DOI

<https://dx.doi.org/10.17504/protocols.io.dm6gpmr55gzp/v1>

Gian Maria Niccolò Benucci^{1,2}, Kevin L. Myers³, Chathu Wijewardana¹, Gary C. Bergstrom³, Stephen Mondo⁴, Robert Riley⁴, Anna Lipzen⁴, Jonathan Del Rosario⁴, Mi Yan⁴, Vivian Ng⁴, Igor V. Grigoriev^{4,5}, Acer VanWallendael^{2,6,7,8,9}, David B. Lowry^{2,6,7}

¹Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing MI, 48824 USA;

²Great Lakes Bioenergy Research Center (GLBRC), 48824 East Lansing, MI USA;

³School of Integrative Plant Science, Plant Pathology and Plant-Microbe Biology Section, Cornell University, Ithaca, NY 14853-5904, USA;

⁴U.S. Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA;

⁵Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, CA, 94720, USA;

⁶Department of Plant Biology, Michigan State University, East Lansing MI, 48824 USA;

⁷Plant Resilience Institute, Michigan State University, East Lansing MI, 48824 USA;

⁸Department of Horticultural Science, North Carolina State University, Raleigh, NC, 27695;

⁹Department of Crop and Soil Science, North Carolina State University, Raleigh, NC, 27695

**gian benucci**

Michigan State University

Create & collaborate more with a free account

Edit and publish protocols, collaborate in communities, share insights through comments, and track progress with run records.

[Create free account](#)



DOI: <https://dx.doi.org/10.17504/protocols.io.dm6gpmr55gzp/v1>

Protocol Citation: Gian Maria Niccolò Benucci, Kevin L. Myers, Chathu Wijewardana, Gary C. Bergstrom, Stephen Mondo, Robert Riley, Anna Lipzen, Jonathan Del Rosario, Mi Yan, Vivian Ng, Igor V. Grigoriev, Acer VanWallendael, David B. Lowry 2025. Draft genome of the switchgrass head smut pathogen *Tilletia maclaganii*. **protocols.io**
<https://dx.doi.org/10.17504/protocols.io.dm6gpmr55gzp/v1>

License: This is an open access protocol distributed under the terms of the **[Creative Commons Attribution License](#)**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Protocol status: Working

We use this protocol and it's working

Created: July 16, 2025

Last Modified: September 01, 2025

Protocol Integer ID: 222661

Keywords: Smut, pathogen, genome, bioenergy, switchgrass, switchgrass head smut pathogen tilletia, significant pathogen of the bioenergy crop switchgrass, annotation of the first reference genome, first reference genome, draft genome, bioenergy crop switchgrass, significant pathogen, tilletia maclaganii

Funders Acknowledgements:

Department of Energy

Grant ID: DE-SC0018409

Abstract

The head smut (*Tilletia maclaganii*) is a significant pathogen of the bioenergy crop switchgrass. *T. maclaganii* typically is more prevalent in older stands of switchgrass and can contribute to significant biomass loss. Here, we outline the methods for the sequencing, assembly, and annotation of the first reference genome for *Tilletia maclaganii*.

Image Attribution

Photo by Christine Layton

Guidelines

NA



Materials

DNA and RNA isolation:

- Macherey-Nagel Nucleobond HMW DNA kit
- New England Biolabs Monarch Total RNA Miniprep Kit (New England Biolabs, MA, USA)
- Thermo Scientific Nanodrop (Thermo Scientific, MA, USA)
- Agilent 5200 Fragment Analyzer (Agilent Technologies, Santa Clara, CA, USA)

Genome sequencing:

- PacBio Sequel IIe sequencing platform (Pacific Biosciences, CA, USA)
- Megaruptor 3 (Diagenode, Seraing, Belgium)
- SMRTbell Barcoded Adapter Plate 3.0 (Pacific Biosciences, CA, USA)
- SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences, CA, USA)
- AMPure PB Beads (Pacific Biosciences, CA, USA)
- BluePippin (Sage Science, MA, USA)
- SMRT Link 10.2, 8M v1 SMRT cells, Version 2.0 sequencing chemistry

Transcriptome sequencing:

- PerkinElmer Sciclone NGS robotic liquid handling system
- Illumina TruSeq Stranded mRNA HT sample prep kit
- KAPA Biosystems next-generation sequencing library qPCR kit (KAPA Biosystem, MA)
- Roche LightCycler 480 real-time PCR instrument (Roche, MA, USA)
- Illumina NovaSeq 6000 sequencing platform
- NovaSeq XP v1.5 reagent kits, S4 flow cell

Bioinformatic analysis:

- bbmap suite (<http://sourceforge.net/projects/bbmap/>)
- BBtools version 38.79 [kmercountexact.sh default; bbnorm.sh pigz passes=1 bits=16 target=9999999 min=174; bbduk.sh maxgc=0.4].
- Flye version 2.9-b1768 [-g 100k --asm-coverage 100 --pacbio-hifi]
- Prodigal version 2.6.3 [-p meta]
- HMMER version 3.1b2 [--domtblout]
- BBDuk [bbduk.sh k=25 mm=f kmask=N]
- BBTools [k=25 mm=f mkf=0.03 ordered ow]
- RACON version 1.4.13 [-u -t 36]
- Trinity version 2.11.0

Safety warnings

 NA

Ethics statement

NA

Before start

NA

DNA and RNA isolation

- 1 *Tilletia maclaganii* was isolated from switchgrass cultivar Shawnee in Ithaca NY, USA (42.4654130 N, -76.4590049 W) in 2021. Mycelium from a single basidiospore (isolate Tm001-NY21) was grown in Malt Extract Peptone Broth (MPB) medium for 14 days at 24°C. DNA was extracted with a Macherey-Nagel Nucleobond HMW DNA kit. For RNA isolation, mycelium of *Tilletia maclaganii* Tm001-NY21 was grown in 100 ml MPB medium for 14 days with shaking (50 rpm) at 24°C. Three hours prior to harvest, the fungus was challenged with pentachloronitrobenzene (PCNB) fungicide at a final concentration of 1 ppm. After incubation with PCNB, the mycelium was harvested and filtered, flash-frozen in liquid nitrogen, and disrupted with mortar and pestle. Total RNA was then extracted using a New England Biolabs Monarch Total RNA Miniprep Kit (New England Biolabs, MA, USA) according to the manufacturer's instructions. Final RNA quality was assessed with a Thermo Scientific Nanodrop (Thermo Scientific, MA, USA) and Agilent 5200 Fragment Analyzer (Agilent Technologies, Santa Clara, CA, USA).

DNA sequencing

- 2 DNA was sequenced using the PacBio Sequel IIe sequencing platform (Pacific Biosciences, CA, USA). An input of 1.5 µg of genomic DNA was sheared around 10 kb using the Megaruptor 3 (Diagenode, Seraing, Belgium). The sheared DNA underwent exonuclease treatment to remove single-stranded ends. Subsequently, it was subjected to a DNA damage repair enzyme mix and an end-repair/A-tailing mix. The DNA was then ligated with the SMRTbell Barcoded Adapter Plate 3.0 (Pacific Biosciences, CA, USA), using barcoded overhang adapters from the SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences, CA, USA). Lastly, the DNA was purified using AMPure PB Beads (Pacific Biosciences, CA, USA). Libraries were pooled based on the required amount of reads per sample, ensuring each sample received adequate coverage for analysis. Pooled libraries are then size-selected using the 0.75% agarose gel cassettes with Marker S1 and High Pass protocol on the BluePippin (Sage Science, MA, USA). PacBio Sequencing primer was then annealed to the SMRTbell template library and sequencing polymerase was bound to them using Sequel II Binding kit 2.0. The prepared SMRTbell template libraries were then sequenced on a PacBio Sequel IIe sequencer using SMRT Link 10.2, 8M v1 SMRT cells, and Version 2.0 sequencing chemistry with 1×1800 movie run times.

RNA sequencing

- 3 RNA was sequenced using the Illumina NovaSeq 6000 sequencing platform. Plate-based RNA sample preparation was performed with the PerkinElmer Sciclone NGS robotic liquid handling system, utilizing Illumina's TruSeq Stranded mRNA HT sample prep kit. Poly-A selection of mRNA was carried out following the protocol outlined by Illumina in their

user guide (https://support.illumina.com/sequencing/sequencing_kits/truseq-stranded-mrna.html). The PCR conditions were as follows: 1 µg of total RNA starting material per sample and 8 cycles of PCR were used for library amplification. The prepared library was then quantified using KAPA Biosystems next-generation sequencing library qPCR kit (KAPA Biosystem, MA) and run on a Roche LightCycler 480 real-time PCR instrument (Roche, MA, USA). Once quantified, the library was multiplexed with other libraries, and the pooled libraries sequenced on an Illumina NovaSeq 6000 sequencing platform using NovaSeq XP v1.5 reagent kits, S4 flow cell, following a 2×150 indexed run recipe.

Bioinformatic analysis

- 4 Circular consensus sequencing (CCS) reads were initially filtered using the BBMap suite (<http://sourceforge.net/projects/bbmap/>) to remove low-quality reads, sequencing errors, and other artifacts. The mitochondrial genome was assembled separately from the nuclear genome using CCS reads. Organelle-derived reads were isolated based on coverage and GC content, excluding nuclear sequences using a maximum coverage threshold of 1.5 times the k-mer coverage peak and a GC content cutoff of 0.40 using BBTools version 38.79 [kmercountexact.sh default; bbnorm.sh pigz passes=1 bits=16 target=9999999 min=174; bbduk.sh maxgc=0.4]. An initial mitochondrial assembly was generated with Flye [1] version 2.9-b1768 [-g 100k --asm-coverage 100 --pacbio-hifi], followed by gene prediction using Prodigal [2] version 2.6.3 [-p meta]. Predicted genes were searched against a curated mitochondrial HMM database using HMMER [3] version 3.1b2 [--domtblout]. Contigs with putative mitochondrial genes were masked for ribosomal loci using BBTools [bbduk.sh k=25 mmf=kmask=N] and a curated eukaryotic ribosomal database. These masked contigs were used to recruit additional CCS reads using BBTools [k=25 mmf=mkf=0.03 ordered ow], and a second Flye assembly was performed with the same parameters. An additional iteration of read recruitment and assembly followed. Contigs shorter than 1 kpb were removed from the assembly and polish was performed with two rounds of Racon [4] version 1.4.13 [racon -u -t 36] to produce the final mitochondrial assembly.
- 5 The mitochondrial assembly was used to remove organelle-derived reads from the CCS. The filtered nuclear reads were assembled using Flye version 2.9-b1768 [-t 32 --pacbio-hifi], followed by two additional rounds of Racon [4] polishing.
- 6 Genome and mitochondrial assemblies were annotated using the JGI Annotation Pipeline [5-6] with supporting Illumina transcriptome RNA-Seq data.
- 7 RNA-Seq reads were filtered using BBduk from the BBmap suite. Raw reads were evaluated for artifact sequences by kmer matching [kmer=25], allowing for 1 mismatch. Detected artifacts were trimmed from the 3' end of the reads. RNA spike-in reads, PhiX reads, and reads containing any N's were removed. Quality trimming was performed using the phred trimming method with a threshold of Q6. After trimming, reads shorter than the length threshold were discarded (minimum length 25 bases or 1/3 of the original

read length, whichever was longer). Finally, the trimmed and filtered reads were assembled into consensus sequences using Trinity [7] version 2.11.0.

References

- 8 1. Kolmogorov, M., Yuan, J., Lin, Y. et al. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37, 540–546 <https://doi.org/10.1038/s41587-019-0072-8>
2. Hyatt, D., Chen, GL., LoCascio, P.F. et al. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11, 119. <https://doi.org/10.1186/1471-2105-11-119>
3. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. 2013. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* 41(12):e121. <https://doi.org/10.1093/nar/gkt263>
4. Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* 27(5):737–746. <https://doi.org/10.1101/gr.214270.116>
5. Kuo A, Bushnell B, Grigoriev IV. 2014. Fungal genomics: sequencing and annotation, p 1–52. In Martin F (ed), *Fungi. Advances in botanical research*. Elsevier Academic Press, Cambridge, United Kingdom.
6. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otilar R, Riley R, Salamov A, Zhao X, Korzeniewski F, Smirnova T, Nordberg H, Dubchak I, Shabalov I. 2014. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res.* 42(Database issue):D699–704. <https://doi.org/10.1093/nar/gkt1183>
7. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 29(7):644–52. <https://doi.org/10.1038/nbt.1883>

Protocol references

References

1. Kolmogorov, M., Yuan, J., Lin, Y. et al. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37, 540–546 <https://doi.org/10.1038/s41587-019-0072-8>
2. Hyatt, D., Chen, G.L., LoCascio, P.F. et al. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11, 119. <https://doi.org/10.1186/1471-2105-11-119>
3. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. 2013. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* 41(12):e121. <https://doi.org/10.1093/nar/gkt263>
4. Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* 27(5):737–746. <https://doi.org/10.1101/gr.214270.116>
5. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 29(7):644–52. <https://doi.org/10.1038/nbt.1883>
6. Kuo A, Bushnell B, Grigoriev IV. 2014. Fungal genomics: sequencing and annotation, p 1–52. In Martin F (ed), *Fungi. Advances in botanical research*. Elsevier Academic Press, Cambridge, United Kingdom.
7. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otiillar R, Riley R, Salamov A, Zhao X, Korzeniewski F, Smirnova T, Nordberg H, Dubchak I, Shabalov I. 2014. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res.* 42(Database issue):D699–704. <https://doi.org/10.1093/nar/gkt1183>

Acknowledgements

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number DE-SC0018409. The work (proposal: 10.46936/10.25585/60000010) conducted by the U.S. Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231.