

The Assembled and Annotated Genome of the Fairy-Ring Fungus *Marasmius oreades*

Markus Hiltunen ^{*}, Sandra Lorena Ament-Velásquez, and Hanna Johannesson

Department of Organismal Biology, Uppsala University, Sweden

^{*}Corresponding author: E-mail: markus.hiltunen@ebc.uu.se.

Accepted: 27 May 2021

Abstract

Marasmius oreades is a basidiomycete fungus that grows in so called “fairy rings,” which are circular, underground mycelia common in lawns across temperate areas of the world. Fairy rings can be thought of as natural, long-term evolutionary experiments. As each ring has a common origin and expands radially outwards over many years, different sectors will independently accumulate mutations during growth. The genotype can be followed to the next generation, as mushrooms producing the sexual spores are formed seasonally at the edge of the ring. Here, we present new genomic data from 95 single-spore isolates of the species, which we used to construct a genetic linkage map and an updated version of the genome assembly. The 44-Mb assembly was anchored to 11 linkage groups, producing chromosome-length scaffolds. Gene annotation revealed 13,891 genes, 55% of which contained a pfam domain. The repetitive fraction of the genome was 22%, and dominated by retrotransposons and DNA elements of the KDZ and *Plavaka* groups. The level of assembly contiguity we present is so far rare in mushroom-forming fungi, and we expect studies of genomics, transposons, phylogenetics, and evolution to be facilitated by the data we present here of the iconic fairy-ring mushroom.

Significance

Genome assemblies of mushroom-forming fungi with complete chromosome sequences are currently rare, and the large clade of *Marasmius* is particularly undersampled, obstructing studies of genome evolution in this part of the fungal tree of life. Here, the nearly gapless chromosome and mitochondrion sequences of *Marasmius oreades*, commonly known as the Scotch bonnet or fairy-ring mushroom, are presented, alongside a high-quality gene annotation and repeat library. These data resources open up the opportunity to answer detailed questions about mushroom evolution.

Key words: mushroom, repeat library, transposon, linkage map, genomics, Agaricomycetes.

Introduction

Marasmius oreades is a member of the ubiquitous and important Marasmiaceae, a family of mainly litter-decaying basidiomycete fungi. The species, commonly known as the edible Scotch bonnet, is known for forming fairy rings; long-lived circular, underground mycelia that are commonly found in lawns and grasslands (fig. 1A). This characteristic makes the species a suitable model for evolutionary studies, as individual

mycelia are clearly discernable and the age of the rings can be inferred from the diameter (Burnett and Evans 1966). Fairy rings can be thought of as natural evolution experiments, where different sectors of the rings are subject to evolutionary forces such as mutation and selection during an extensive vegetative growth phase (Hiltunen et al. 2019). Mushrooms develop seasonally along the edge of the ring, where the sexual progeny in the form of basidiospores is produced, enabling the study of inheritance of variants present in the ring.

© The Author(s) 2021. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

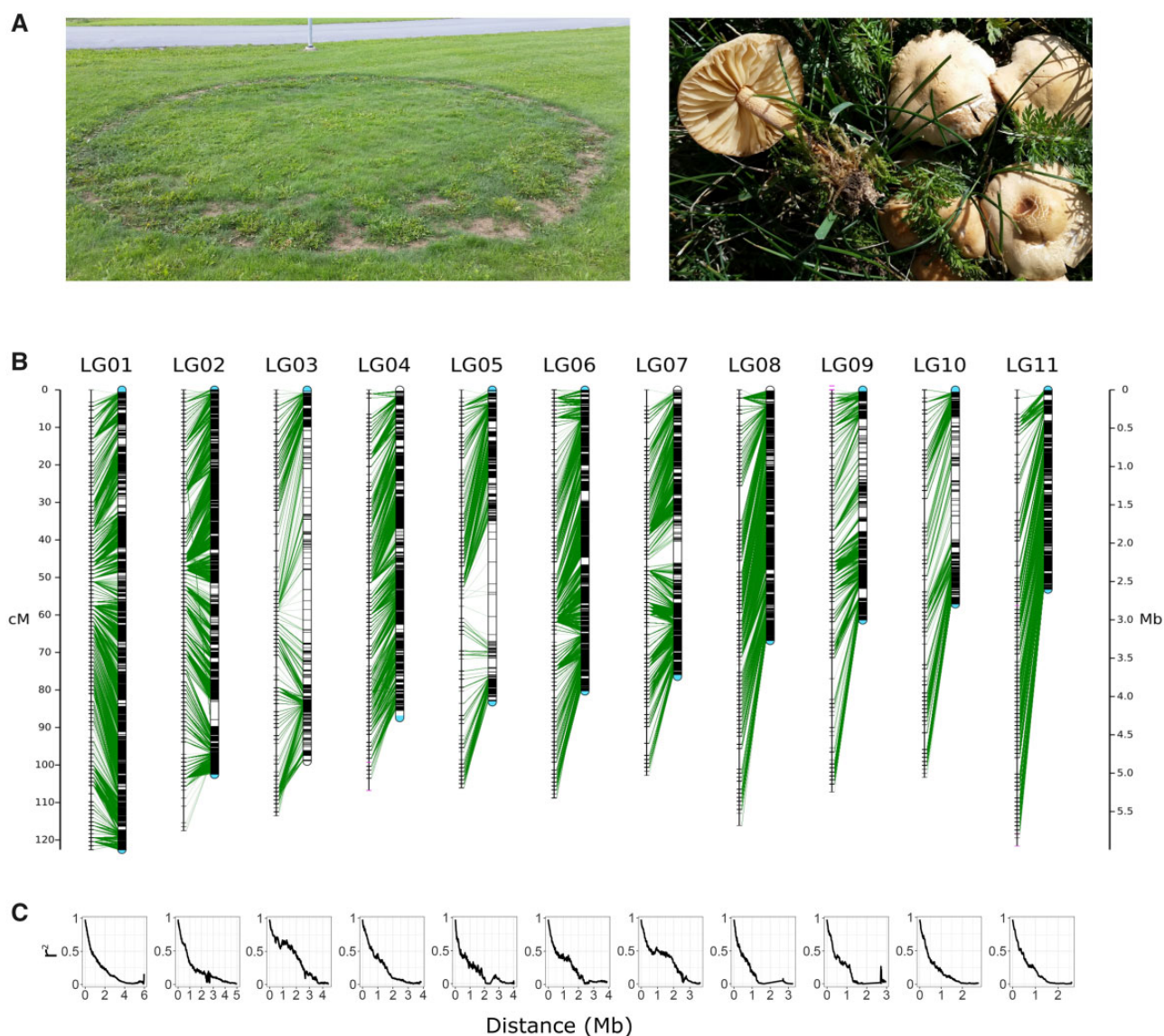


FIG. 1.—Linkage in the *Marasmius oreades* genome. (A) Photographs of a *M. oreades* fairy ring (° 6.7 m) and fruiting bodies (photos by M. Hiltunen). (B) The 11 linkage groups (LGs) and corresponding putative chromosome sequences of the *M. oreades* genome. The LGs are shown as lines to the left, with marker positions in cM distance drawn as horizontal lines across the LGs. The corresponding chromosome sequences are shown to the right, with green lines connecting the marker genetic and physical coordinates. Markers mapping to another chromosome than the expected one are shown in magenta on the LGs (e.g., at the end of LG04). Telomere repeats are indicated in blue on the chromosome sequences (not drawn to scale). (C) Linkage disequilibrium (r^2) decay between markers with increasing physical distance for each linkage group.

The genome assembly of *M. oreades* was recently presented and used to study primarily single-basepair mutations accumulating during growth (Hiltunen et al. 2019). However, mutations affecting larger chromosomal regions, including movement of transposable elements (TEs), are increasingly being recognized as important additional sources of variability in natural populations (Barrón et al. 2014; Stuart et al. 2016; Oggenfuss et al. 2020). An improved version of the *M. oreades* genome assembly, with nearly completely assembled chromosomes, is expected to enable studies of such mutations in this system. Fungi typically have small genome sizes

compared with other eukaryotes. Hence, whole-genome sequencing of haploid offspring can be used cost-effectively to produce linkage data of different chromosomal regions (Labbé et al. 2008; Foulongne-Oriol 2012; Fierst 2015).

Results and Discussion

In this study, we present genomic data of 95 single-spore isolates of *M. oreades* (supplementary table S1, Supplementary Material online). These data are a valuable resource to study heritage and meiotic processes, and were here

used to construct a genetic linkage map of the species. The linkage map consisted of 11 high-confidence linkage groups (LGs), spanning a total distance of 1,226 cM and a mean of 111 cM per LG (fig. 1B). We used the linkage map as a basis for a new genome assembly by reassembling previously sequenced PacBio, Oxford Nanopore MiniON (ONT) long reads, and 10X Genomics (10XG) Chromium linked reads (supplementary table S1, Supplementary Material online) (data released in Hiltunen et al. [2019]). We were able to anchor 98% of assembled sequence (43,635,136 out of 44,372,355 bp) to the 11 LGs of the map, producing scaffolds of putative chromosome length (supplementary table S2, Supplementary Material online). The assembly size of *M. oreades* is smaller than that of the related *Marasmius fiardii* (59,447,912 bp) (Miyachi et al. 2020). This difference may be explained by a true difference in genome size between the species and/or that the *M. fiardii* assembly is partly heterozygous as a result of sequencing a dikaryotic culture. As expected, linkage disequilibrium in the *M. oreades* genome degrades over physical distance between markers on the chromosomes (fig. 1C). Twenty-two gaps remain in the genome (lengths unknown). Of the assembled scaffolds, seven were found to start and end in telomere repeat sequence (TAAC[3-8], or the reverse complement), with the remaining four containing either the 5' or 3' telomere repeat (fig. 1C). Linked read data do not support linkage between the scaffolds with only one assembled telomere, causing us to believe that the telomeres in those cases are in reality close to the scaffold ends but were not anchored or assembled. Coverage of the spore isolates, and fractions of shared 10X Chromium barcodes between adjacent regions are mostly even across the scaffolds (supplementary fig. S1, Supplementary Material online). The recombination rate appears to increase towards the subtelomeres (fig. 2A). Although the exact determination of centromere regions would require additional experimental data, the combination of low recombination rate, low gene, and high repeat density, in addition to a drop in GC content may indicate such regions in our assembly (e.g., at ~2.5 Mb on Chr1; fig. 2A). The mitochondrial genome was assembled to a 66,224 bp contig; a typical size for Agaricomycetes (Araújo et al. 2021). Twenty-one scaffolds (737,219 bp) remained unanchored, one of which contained a telomere repeat. No bacterial contamination was detected (supplementary fig. S2, Supplementary Material online). Almost all expected single-copy conserved orthologs (BUSCOs) could be found in the genome (98.5% complete + fragmented orthologs; basidiomycota_odb9, $n = 1,335$), confirming its high base-level accuracy (Simão et al. 2015). The level of assembly contiguity presented here is crucial for the study of, for example, TEs and chromosome evolution (Thomma et al. 2016), and yet, is rare among mushroom-forming fungi: less than 8% of available Agaricomycotina genomes on JGI Mycocosm ($n = 467$; <https://mycocosm.jgi.doe.gov/mycocosm/home>; accessed January 26, 2021) have L95 values of 20 or lower (most

studied fungi in the Agaricomycotina have between 10 and 20 chromosomes).

Our *M. oreades* genome assembly was subjected to annotation of genes and repetitive sequences. The analysis predicted 13,891 genes (15,053 mRNAs), covering 55.5% (24,614,704 bp) of the genome, with an average gene length of 1,772 bp and six exons per mRNA (supplementary table S2, Supplementary Material online); estimates comparable to other Agaricomycetes (Ohm et al. 2010; Li et al. 2018; Miyachi et al. 2020). The predicted protein set showed high concordance with the basidiomycete conserved ortholog database (98.5% complete + fragmented BUSCOs; basidiomycota_odb9, $n = 1,335$). Out of the predicted mRNAs, 8,222 (55%) contained a pfam domain (supplementary table S2, Supplementary Material online). Genes are generally concentrated towards the center of chromosomes and decrease in density towards the telomeres (fig. 2A). For confident TE calling, we constructed a library of repetitive sequences de novo, and manually curated the sequences therein (supplementary table S3, Supplementary Material online). Repeat annotation revealed 21.8% (9,685,063 bp) of the genome to be repetitive, and repeats tend to form regional clusters on the chromosomes (fig. 2A), similarly to in other fungi (Stajich et al. 2010; Foulongne-Oriol et al. 2013). Long-terminal repeat (LTR) elements are the most abundant mobile elements, covering 14% of the genome (fig. 2B and supplementary table S2, Supplementary Material online), which reflects findings in other basidiomycetes (Labbé et al. 2012; Amselem et al. 2015; Castanera et al. 2016). DNA elements constitute 6% of the genome, about half of which belong to the KDZ or *Plavaka* groups (Iyer et al. 2014). We characterized the terminal inverted repeat (TIR) motif of these understudied elements and summarized the information in supplementary figure S3, Supplementary Material online.

Taken together, the data presented here enable the use of *M. oreades* as a model for detailed studies of, for example, structural mutations, TE mobilization, and gene family evolution.

Materials and Methods

Here, we generated genomic data of 95 single-spore isolates of *M. oreades*. Spore prints were collected from four fruiting bodies from one fairy ring located in a lawn in Uppsala, Sweden (59°51'27.3"N, 17°34'19.2"E). Spores were germinated on culture plates containing potato dextrose agar, and newly germinated single spores were transferred to fresh plates. Successful isolation of single-spore cultures was verified by inspecting cultures for absence of clamp connections, which is diagnostic of the monokaryotic (haploid) condition in this and many other species (Mallett and Harrison 1988). Isolates were transferred and grown in 2% liquid malt extract for several days before DNA purification with the Zymo Fungal/Bacterial Miniprep kit. Paired-end libraries (2×151 bp)

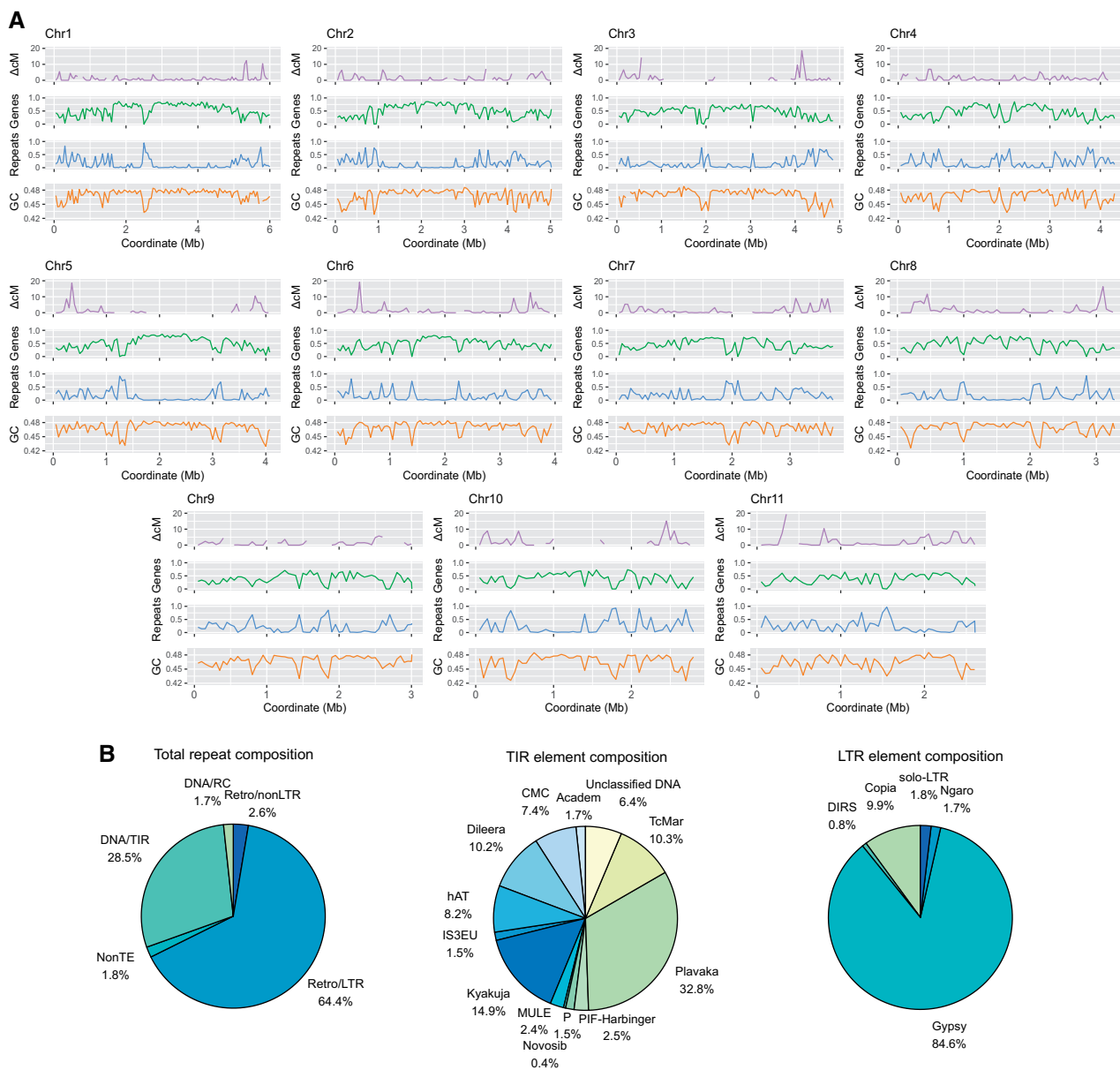


FIG. 2.—Genome characteristics. (A) Panels from top to bottom: recombination rate as difference in mean cM distance between windows; gene density as fraction of occupied base pairs per window; repeat density as fraction of occupied base pairs per window; GC content as fraction of bases per window. In all tracks, 50-kb nonoverlapping windows were used. (B) Repeat content in the genome as total repeat composition, and composition of the two most abundant classes: TIR and LTR elements.

were prepared from 1 ng DNA with the Nextera XT kit with dual indexes (Illumina) for each isolate separately, before pooling equal volumes of all 95 libraries and sequencing on one lane of Illumina HiSeqX (HiSeq X SBS chemistry).

For the genome assembly and annotation, we utilized published PacBio, ONT, and 10XG genome sequence reads, and Illumina HiSeq transcriptome reads (supplementary table S1, Supplementary Material online) (Hiltunen et al. 2019). In short, these data were collected the following way: a

monokaryotic single spore of the species was isolated (strain O3SP1), grown in liquid malt extract, and DNA was obtained by the Qiagen Genomic Tip. A linked-read library was prepared from this DNA (10X Genomics Chromium Genome) and sequenced (half a lane of HiSeqX). For PacBio and Nanopore sequencing, the DNA was further purified using the MoBio PowerClean kit. For PacBio, a 10-kb insert size library was prepared and sequenced with RSII (eight SMRT cells, C4 chemistry, P6 polymerase). Two Nanopore libraries

were prepared, one with the LSK108 kit and sequenced on a R9.4 flowcell, and another with the LSK308 kit, sequenced on a R9.5 flowcell. A cDNA library was prepared, using the same culturing method as above, but immediately freezing harvested tissue in liquid nitrogen, extracting total RNA with the RNeasy Plant Mini kit (QIAGEN, Germany), preparing the library with NEBNext Ultra Directional RNA Library Prep kit for Illumina (New England Biolabs). Sequencing was performed in one lane of HiSeq2500 (v4 chemistry).

We newly assembled the data presented by Hiltunen et al. (2019) as follows. Raw PacBio and Nanopore reads were assembled using Canu v1.7 (Koren et al. 2017). This backbone assembly was split at misassembly positions as identified by Tigmint v1.1.2 (Jackman et al. 2018) from a BWA mem 0.7.17-r1188 (Li and Durbin 2009) mapping of the 10XG reads. Scaffolds were built by ARBitR v0.1 (Hiltunen et al. 2020) using the 10XG reads mapped by EMA v0.6.2 (Shajii et al. 2018) to the broken assembly. Scaffold gaps were filled using LR_Gapcloser (Xu et al. 2019) and error-corrected long reads from Canu. Next, reasoning that the gap filling step may have introduced some sequence that was already present in contigs that could not be integrated into scaffolds, the assembly was scanned for such redundant contigs, by splitting the assembly into long and short scaffolds (threshold: 100 kb) and aligning the short ones to the long ones with the nucmer program of MUMmer v4.0.0beta2 (Kurtz et al. 2004). Short contigs with >97% sequence identity and >95% alignment coverage to a long scaffold were removed. The resulting assembly was subjected to another round of ARBitR scaffolding, gap filling, and removal of redundant contigs as above. Then the next step was to create a genetic linkage map to anchor and order the assembled scaffolds into long pseudomolecule sequences.

After adapter and primer trimming of the 95 single-spore isolate sequencing reads using Trimmomatic v0.32 (Bolger et al. 2014), reads were mapped to the scaffolded assembly using BWA mem. We used Lep-MAP3 v0.2 (Rastas 2017) to generate a linkage map from these data, by calling variants using Samtools mpileup (Li et al. 2009) and piping into Lep-MAP3. Variants were filtered with the Lep-MAP3 module Filtering2; allowing for a heterozygote rate of 0.1 (in theory, no sites can be heterozygous in the haploid spore isolates, but we included this parameter due to index hopping; van der Valk et al. 2020), a maximum of ten individuals with missing data per site, and noninformative sites removed. After filtering, 141,011 sites remained. A linkage map was constructed using a LOD score of 20 (Lep-MAP3 modules SeparateChromosomes2 and OrderMarkers2). Resulting LGs with less than 100 markers were removed, resulting in 11 long, high-confidence LGs.

Next, we anchored the assembly scaffolds to the linkage map to create pseudomolecule sequences using Allmaps v0.9.14 (Tang et al. 2015). Again, gaps were filled with LR_Gapcloser and redundant scaffolds removed as above.

The assembly was polished first using HyPo (downloaded February 23, 2020) (Kundu et al. 2019), with 10XG reads mapped by EMA and raw PacBio and ONT reads with Minimap2 v 2.17-r941 (Li 2018, 2). A second round of polishing was performed using EMA mappings of the 10XG reads with Pilon (Walker et al. 2014).

The mitochondrial genome was identified in the assembly by GC content and read coverage. The circular chromosome had been assembled into one contig, with a long overhang aligning back to itself. The ribosomal RNA large subunit (*rnl*) was identified in this contig by BLAST, using the *rnl* gene from the related species *Moriliophthora perniciosa* (NCBI GenBank: AY376688.1) as query. The contig representing the mitochondrial genome was broken manually within Geneious v10.2.4 (<http://www.geneious.com>), at a coordinate just before the *rnl* BLAST hit. The two resulting sequences were reassembled with the de novo assembly program in Geneious where the overlap was collapsed.

Genome completeness was assessed with BUSCO v2.0.1 with the basidiomycota_odb9 database (Simão et al. 2015) and telomeres were located using a custom script (available at https://github.com/markhilt/genome_analysis_tools/blob/master/find_telomeres.py). An assessment of bacterial contamination was performed using Blobtools v1.1.1 (Laetsch and Blaxter 2017), with the raw assembly as output by Canu, 10XG mapped reads (10X Genomics Longranger), and BlastN results to the nt database (1e-25) as input data.

From the single-spore sequencing data set, we calculated linkage disequilibrium between variants the following way. Reads were remapped to the finished assembly by BWA mem (Li and Durbin 2009) and variants were called using Platypus v0.8.1.1 (Rimmer et al. 2014). Heterozygous calls were disregarded and the resulting vcf file was recoded into a phased format (genotype calls are per definition phased as the spore data set is haploid). Linkage disequilibrium between variants on the same chromosome was calculated using the hap-r2 module of VCFtools v0.1.17 (Danecek et al. 2011).

Genome annotation was performed using Funannotate v1.8.3 (Palmer and Stajich 2021). We followed the Funannotate manual (<https://funannotate.readthedocs.io>; last accessed March 2021) closely during most steps, with the following exceptions. RepeatMasker v. open-4.0.9 (<http://www.repeatmasker.org/RepeatMasker/>) was run standalone to softmask the genome prior to annotation, using a custom, manually curated repeat library (described below). GeneMark-ET v4.62_lic (Lukashin and Borodovsky 1998) was included during structural gene predictions. Functional information about gene predictions was collected with eggNOG mapper online (<http://eggno-mapper.embl.de/>; February 2, 2021), standalone runs of InterProScan v5.30-69.0 (Jones et al. 2014), and the fungal version of antiSMASH v5.1.1 (Blin et al. 2019). Phobius (Käll et al. 2004) was run using the wrapper from Funannotate.

We compiled a repeat library for the new assembly the following way. RepeatModeler v2.0.1 (Zeng et al. 2018) with the `-LTRStruct` option was initially used to mine the genome for repetitive sequences. Individual copies with flanking regions of the sequences from RepeatModeler were collected from the genome using BlastN, aligned with MAFFT v7.407 (Katoh and Standley 2013), manually inspected, and curated. We focused primarily on three points during manual curation: extending flanks to capture the whole element, creating a new consensus, and classifying the consensus sequence. Boundaries of each element were defined based on target site duplications (TSD), and for the new consensus, all sequence between the TSDs was included (but not the TSDs themselves). Structural information (e.g., LTR and TIR regions) was collected by self-alignments using MAFFT and used for classification, in addition to protein domains as identified by CDD (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>; last accessed February 15, 2021). We based our classifications on the system proposed by Wicker et al. (2007), following the structural information, TSD and protein domain architecture provided therein for defining *Gypsy*, *Copia*, and most DNA TEs. Additionally, *Plavaka* autonomous elements were defined by the pfam18759 domain, and *Dileera* and *Kyakuja* elements by protein homology of the transposase to previously identified families (Iyer et al. 2014). *Helitrons* were classified by having the pfam14214 and pfam05970 domains (Castanera et al. 2014). For LINEs we used homology to sequences in RepBase for classification, as identified by Censor (Bao et al. 2015). Nonautonomous elements were defined by having TSDs and structural features but lacking the protein domains required for transposition, also in most cases spanning less than 1,000 bp. We found only nonautonomous DNA elements, which we classified based on terminal sequence motifs and TSD length (Kojima 2020b; Storer et al. 2021). In cases where we had already identified autonomous families of certain elements, we used their TIR and TSD characteristics for classification of nonautonomous families. We noticed widespread insertions of putative nonautonomous MULE and *Academ* elements (Kojima 2020a). Reasoning that these superfamilies likely have autonomous families in the genome that had escaped capture by RepeatModeler, we used hmmer v3.3.2 (hmmer.org) to search for them. We used the hmm profile for pfam10551 for MULE, and for *Academ* we obtained an hmm profile from Muszewska et al. (2017). Resulting hits were curated the same way as above. We also searched for previously unidentified *Sola*, *Dada*, *Zator*, and *Kolobok* elements using hmm profiles from Muszewska et al. (2017) but without finding multiple well-supported hits.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

The authors thank Aaron A. Vogan and Alexander Sang-Jae Suh for advice regarding curation of repetitive elements. We acknowledge support from the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council, and SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure. The authors acknowledge funding from ‘Stiftelsen Lars Hiertas minne’ to MH for methodological expenses including sequencing and reagents.

Data Availability

Newly generated sequencing reads from 95 single-spore isolates are available at the NCBI Sequence Read Archive (BioProject PRJNA702659). The new genome version Maror2 is available at the NCBI Genome database (JAFLEJ000000000; BioProject PRJNA525964). Previously generated reads used here for assembly and annotation are available under the BioProject PRJNA525964. The repeat library and genetic map are available on Dryad (<https://doi.org/10.5061/dryad.000000034>).

Literature Cited

- Amselem J, Lebrun M-H, Quesneville H. 2015. Whole genome comparative analysis of transposable elements provides new insight into mechanisms of their inactivation in fungal genomes. *BMC Genomics* 16:141.
- Araújo DS, et al. 2021. Comparative mitogenomics of Agaricomycetes: diversity, abundance, impact and coding potential of putative open-reading frames. *Mitochondrion* 58:1–13.
- Bao W, Kojima KK, Kohany O. 2015. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob DNA*. 6:11–11.
- Barrón MG, Fiston-Lavier A-S, Petrov DA, González J. 2014. Population genomics of transposable elements in *Drosophila*. *Annu Rev Genet*. 48:561–581.
- Blin K, et al. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res*. 47(W1):W81–W87.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120.
- Burnett JH, Evans EJ. 1966. Genetical homogeneity and the stability of the mating-type factors of ‘Fairy Rings’ of *Marasmius oreades*. *Nature* 210(5043):1368–1369.
- Castanera R, et al. 2014. Highly expressed captured genes and cross-kingdom domains present in *Helitrons* create novel diversity in *Pleurotus ostreatus* and other fungi. *BMC Genomics* 15:1071.
- Castanera R, et al. 2016. Transposable elements versus the fungal genome: impact on whole-genome architecture and transcriptional profiles. *PLoS Genet*. 12(6):e1006108.
- Danecek P, et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27(15):2156–2158.
- Fierst JL. 2015. Using linkage maps to correct and scaffold de novo genome assemblies: methods, challenges, and computational tools. *Front Genet*. 6:220.

- Foulongne-Oriol M. 2012. Genetic linkage mapping in fungi: current state, applications, and future trends. *Appl Microbiol Biotechnol.* 95(4):891–904.
- Foulongne-Oriol M, Murat C, Castanera R, Ramirez L, Sonnenberg ASM. 2013. Genome-wide survey of repetitive DNA elements in the button mushroom *Agaricus bisporus*. *Fungal Genet Biol.* 55:6–21.
- Hiltunen M, Grudzinska-Sterno M, Wallerman O, Ryberg M, Johannesson H. 2019. Maintenance of high genome integrity over vegetative growth in the fairy-ring mushroom *Marasmius oreades*. *Curr Biol.* 29(16):2758–2765.e6.
- Hiltunen M, Ryberg M, Johannesson H. 2020. ARBitR: an overlap-aware genome assembly scaffold for linked reads. *Bioinformatics.* doi: 10.1093/bioinformatics/btaa975.
- Iyer LM, et al. 2014. Lineage-specific expansions of TET/JBP genes and a new class of DNA transposons shape fungal genomic and epigenetic landscapes. *Proc Natl Acad Sci U S A.* 111(5):1676–1683.
- Jackman SD, et al. 2018. Tigrint: correcting assembly errors using linked reads from large molecules. *BMC Bioinformatics* 19(1):393.
- Jones P, et al. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30(9):1236–1240.
- Käll L, Krogh A, Sonnhammer ELL. 2004. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol.* 338(5):1027–1036.
- Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Kojima KK. 2020a. *AcademH*, a lineage of *Academ* DNA transposons encoding helicase found in animals and fungi. *Mob DNA.* 11:15.
- Kojima KK. 2020b. Structural and sequence diversity of eukaryotic transposable elements. *Genes Genet Syst.* 94(6):233–252.
- Koren S, et al. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res.* 27(5):722–736.
- Kundu R, Casey J, Sung W-K. 2019. HyPo: super fast & accurate polisher for long read genome assemblies. *bioRxiv.* 2019.12.19.882506. doi: 10.1101/2019.12.19.882506.
- Kurtz S, et al. 2004. Versatile and open software for comparing large genomes. *Genome Biol.* 5(2):R12.
- Labbé J, et al. 2008. A genetic linkage map for the ectomycorrhizal fungus *Laccaria bicolor* and its alignment to the whole-genome sequence assemblies. *New Phytol.* 180(2):316–328.
- Labbé J, et al. 2012. Characterization of transposable elements in the ectomycorrhizal fungus *Laccaria bicolor*. *PLoS One* 7(8):e40197.
- Laetsch D, Blaxter M. 2017. BlobTools: interrogation of genome assemblies [version 1; peer review: 2 approved with reservations]. *F1000Res.* 6:1287.
- Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34(18):3094–3100.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25(14):1754–1760.
- Li H, et al. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078–2079.
- Li H, et al. 2018. The genome sequences of 90 mushrooms. *Sci Rep.* 8(1):9982.
- Lukashin AV, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 26(4):1107–1115.
- Mallett KI, Harrison LM. 1988. The mating system of the fairy ring fungus *Marasmius oreades* and the genetic relationship of fairy rings. *Can J Bot.* 66(6):1111–1116.
- Miyauchi S, et al. 2020. Large-scale genome sequencing of mycorrhizal fungi provides insights into the early evolution of symbiotic traits. *Nat Commun.* 11(1):5125.
- Muszewska A, Steczkiewicz K, Stepniewska-Dziubinska M, Ginalski K. 2017. Cut-and-paste transposons in fungi with diverse lifestyles. *Genome Biol Evol.* 9(12):3463–3477.
- Oggenfuss U, et al. 2020. A population-level invasion by transposable elements in a fungal pathogen. *bioRxiv.* 2020.02.11.944652. doi: 10.1101/2020.02.11.944652.
- Ohm RA, et al. 2010. Genome sequence of the model mushroom *Schizophyllum commune*. *Nat Biotechnol.* 28(9):957–963.
- Palmer J, Stajich J. 2021. Funannotate v1.8.3: eukaryotic genome annotation (Version 1.8.3). Zenodo. doi: 10.5281/zenodo.1134477.
- Rastas P. 2017. Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics* 33(23):3726–3732.
- Rimmer A, et al. 2014. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet.* 46(8):912–918.
- Shajii A, Numanagic I, Whelan C, Berger B. 2018. Statistical binning for barcoded reads improves downstream analyses. *Cell Syst.* 7(2):219–226.e5.
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31(19):3210–3212.
- Stajich JE, et al. 2010. Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom *Coprinopsis cinerea* (*Coprinus cinereus*). *Proc Natl Acad Sci U S A.* 107(26):11889–11894.
- Storer J, Hubley R, Rosen J, Wheeler TJ, Smit AF. 2021. The Dfam community resource of transposable element families, sequence models, and genome annotations. *Mob DNA.* 12(1):2.
- Stuart T, et al. 2016. Population scale mapping of transposable element diversity reveals links to gene regulation and epigenomic variation. *eLife* 5:e20777.
- Tang H, et al. 2015. ALLMAPS: robust scaffold ordering based on multiple maps. *Genome Biol.* 16:3.
- Thomma BPHJ, et al. 2016. Mind the gap; seven reasons to close fragmented genome assemblies. *Fungal Genet Biol.* 90:24–30.
- van der Valk T, Vezzi F, Ormestad M, Dalén L, Guschanski K. 2020. Index hopping on the Illumina HiSeqX platform and its consequences for ancient DNA studies. *Mol Ecol Resour.* 20(5):1171–1181.
- Walker BJ, et al. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9(11):e112963.
- Wicker T, et al. 2007. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet.* 8(12):973–982.
- Xu G-C, et al. 2019. LR_GapCloser: a tiling path-based gap closer that uses long reads to complete genome assembly. *Gigascience* 8(1):giy157.
- Zeng L, Kortschak RD, Raison JM, Bertozzi T, Adelson DL. 2018. Superior ab initio identification, annotation and characterisation of TEs and segmental duplications from genome assemblies. *PLoS One* 13(3):e0193588.

Associate editor: Li-Jun Ma