Nuclear and mitochondrial genome assemblies for the endangered wood-decaying fungus *Somion occarium*

*Rowena Hill[1], *Jamie McGowan[1], Vendula Brabcová[2], Seanna McTaggart[1], Naomi Irish[1], Tom Barker[1], Vanda Knitlhoffer[1], Sacha Lucchini[1], Kendall Baker[1], Leah Catchpole[1], Chris Watkins[1], Karim Gharbi[1], Gemy Kaithakottil[1], Alan Tracey[3], Jonathan MD Wood[3], Michal Tomšovský[4], Petr Baldrian[2], David Swarbreck[1], Neil Hall[1,5]

* equal contribution

Corresponding author – Rowena Hill, rowena.hill@earlham.ac.uk

[1] Earlham Institute, Norwich, UK

[2] Laboratory of Environmental Microbiology, Institute of Microbiology of the Czech Academy of Sciences, Prague, Czech Republic

[3] Tree of Life, Wellcome Sanger Institute, Hinxton, UK

[4] Department of Forest Protection and Wildlife Management, Faculty of Forestry and Wood Technology, Mendel University in Brno, Czech Republic

[5] School of Biological Sciences, University of East Anglia, Norwich, UK

Key words: saprotroph, white-rot, *Cerrenaceae*, *Polyporales*, fungal conservation

ABSTRACT

Somion occarium is a wood-decaying bracket fungus belonging to an order known to be rich in useful chemical compounds. Despite its widespread distribution, *S. occarium* has been assessed as endangered on at least one national Red List, presumably due to loss of old-growth forest habitat. Here, we present a near-complete, annotated nuclear genome assembly for *S. occarium* consisting of 31 Mbp arranged in 11 pseudochromosomes – 9 of which are telomere-to-telomere – as well as a complete mitochondrial genome assembly of 112.9 Kbp. We additionally performed phylogenomic analysis and annotated carbohydrate-active enzymes (CAZymes) to compare gene and CAZyme content across closely related species. This genome was sequenced as the representative for Kingdom *Fungi* in the European Reference Genome Atlas (ERGA) Pilot Project.

SIGNIFICANCE

Wood-decaying fungi are not only the foundation of nutrient cycling in our forests, but also known to produce many medically relevant chemical compounds. As *Somion occarium* is also endangered in at least one country, it is doubly important to produce high-quality genomic resources to facilitate study of this species. The sole fungal representative sequenced as part of the European Reference Genome Atlas (ERGA) Pilot Project, this genome assembly leads the way for future sequencing efforts of fungi within ERGA.

INTRODUCTION

The species *Somion occarium* (*Cerrenaceae*, *Polyporales*, *Basidiomycota*) is a bracket fungus and polypore, both of which are polyphyletic groupings based on growth form rather than ancestry. Usually found on dead wood of hardwood trees – including oaks (*Quercus* spp.) and beech (*Fagus* spp.) (GBIF 2024) – *S. occarium* is predominantly a white-rot decayer, or saprotroph. Its occasional presence on living trees suggests it is also capable of facultative parasitic growth, similar to other

species in the *Cerrenaceae* (Hallenberg et al. 2008; Justo et al. 2017). In addition to wood, the fungus has also occasionally been recorded from soil DNA metabarcoding (Větrovský et al. 2020). Such wood-decaying fungi are essential for carbon and nutrient cycling and the formation and stabilisation of soil aggregates in forest ecosystems (Miller & Lodge 2007).

S. occarium belongs to the order *Polyporales*, which contains many, but not all, of the polypore species. Polypores are iconic in forest ecosystems, and have long been culturally valued, including a rich history of their use in folk medicine (Grienke et al. 2014). Fittingly, the *Polyporales* are some of the highest known producers of bioactive compounds within fungi, including antibacterials, antifungals and drugs or drug leads (Prescott et al. 2023). Generating genomic data for species in the order provides the groundwork to support further discovery of potentially useful compounds.

The Earth BioGenome Project (EBP) is an ambitious initiative to sequence reference genomes for all eukaryotic species (Lewin et al. 2018). The European Reference Genome Atlas (ERGA) is the European node of EBP, and has coordinated a pilot project to establish and test a decentralised infrastructure to deliver the EBP's aspirations (Mc Cartney et al. 2024), for which *S. occarium* was selected as a representative for fungi. *S. occarium* was previously known as *Spongipellis delectans* (or the synonym *Hydnum occarium*), but recent division of *Spongipellis sensu lato* into several genera following DNA-based and morphological revision saw the restoration of the suppressed older genus name *Somion* for the *Sp. delectans* complex (Miettinen et al. 2023). North American *Sp. delectans* is now *Somion delectans*, whilst European *Sp. delectans* is *S. occarium*. There is one existing assembly for the *Cerrenaceae* family which *S. occarium* belongs to, for the species *Cerrena unicolor* (https://mycocosm.jgi.doe.gov/Cerun2/Cerun2.home.html), although the *S. occarium* ERGA assembly reported here represents the first EBP-level reference genome for the family.

While *S. occarium* is found across the northern temperate zone in Europe and northernmost Africa (Miettinen et al. 2023), the species is classed as endangered on the national Red List for the Czech Republic (Zíbarová et al. 2024), the locality from which the strain sequenced here was collected. Wood-decaying fungi are predominantly dependent on deadwood, which is more available in old-growth forest compared to younger or managed forests/plantations, and so the ongoing loss of old-growth forest habitat means that wood-decaying fungi are increasingly threatened by niche reduction (Lonsdale et al. 2008; Jönsson et al. 2008). Incidentally, this also means that certain fungal species can be used as indicator species for old-growth forest (Halme et al. 2009). Recognition of fungi in conservation discourse is relatively new (May et al. 2018), but progress is being made thanks to the enduring efforts of mycologists to raise their profile in Red Lists (Mueller et al. 2014; Dahlberg & Mueller 2011), as well as growing public and political awareness on the importance of fungi to ecosystems (IUCN 2021). Producing high-quality genomic resources for endangered species such as *S. occarium* is foundational to understanding and protecting them.

RESULTS AND DISCUSSION

Near-complete Genome Assembly for *Somion occarium*

We assembled the *S. occarium* genome using hifiasm (Cheng et al. 2021) from approximately 100-fold coverage of PacBio HiFi reads, with a mean read length of 18.6 Kbp. The final, manually curated assembly (table 1) exceeded the minimum reference standard defined by EBP assembly quality metrics (EBP 2023). Note that this final version of the assembly is a primary assembly and is highly heterozygous (1.69%; supplementary fig. S1).

We predicted telomeres on both ends of 9 out of the total 11 pseudo -chromosomes, with the remaining 2 contigs (SUPER_9 and SUPER_11) having a telomere at one end (fig. 1B). A total of 11 chromosomes was slightly lower than the number that has been previously reported for other

Polyporales species assembled to chromosome-level, which ranges from 12–14 (Wright et al. 2022; Ma et al. 2024; Chen et al. 2012, 2022), although this was explained by three chromosomal 'fusions' which were flagged when we examined synteny relative to the *Laetiporus sulphureus* gfLaeSulp1 assembly (fig. 1D). Each of these fusion events were supported following curation and manual checking of read alignments.

In addition to the nuclear genome assembly, we assembled and annotated the mitochondrial genome (fig. 1C), which was a single circular contig 112.9 Kbp in length and included all 14 core protein-coding genes typically present in fungal mitogenomes (Sandor et al. 2018).

Gene and CAZyme Content

Gene models were generated using REAT (Robust Eukaryotic Annotation Toolkit; https://github.com/EI-CoreBioinformatics/reat) and functionally annotated with AHRD (https://github.com/groupschoof/AHRD; supplementary file S1). We compared gene content across species with available genome annotations within the *Polyporales* 'residual' clade as defined by Justo et al. (2017), including *Laetiporus sulphureus* as an outgroup (supplementary table S1). This first involved inference of a species tree from orthologous genes using OrthoFinder (Emms & Kelly 2019), where we found that the divergence of families in the phylogenomic tree corresponded with previous multilocus phylogenetic analysis within the order (Justo et al. 2017). We additionally predicted carbohydrate-active enzyme (CAZyme) genes using run_dbcan (Zheng et al. 2023; supplementary file S2), as CAZymes are a major component of the gene repertoire necessary for white-rot (Hage et al. 2021). The total number of genes and CAZymes was slightly lower than either the closest available relative within *Cerrenaceae*, *Cerrena unicolor*, or *Panus rudis* in the sister family of *Panaceae* (fig. 1E). The proportions of CAZymes known to act on various substrates was similar across all taxa, however.

CONCLUSION

Here we present the first near-complete, annotated genome assembly for the wood-decaying bracket fungus *Somion occarium* (prev. *Spongipellis delectans*). The first representative for Kingdom *Fungi* under the ERGA initiative, this new high-quality genome resource will enable further exploration of the genetic basis of saprotrophy in an ecologically and chemically important lineage of fungi, the *Polyporales*.

MATERIALS AND METHODS

Sample Collection and Isolation

The *S. occarium* strain sequenced here was collected as a fresh basidiome growing on dead wood of *Quercus cerris*, on 20th October 2005 at the Rendez-vous nature monument in Czech Republic (Valtice, Břeclav, South Moravian region, 48.7499006N, 16.7939872E). This is the same locality from which the epitype of *S. occarium* was collected (Miettinen et al. 2023). The basidiome was placed on a petri dish of 2% malt extract agar, and the resulting culture was deposited in the culture collection of MENDELU (accession MUcc 838) and simultaneously kept in the Culture Collection of Basidiomycetes of Institute of Microbiology, Prague (accession CCBAS136). Its identity was confirmed using phylogenetic analysis by Tomšovský (2012), under the previous name of *Spongipellis delectans*.

Nucleic Acid Isolation and Sequencing

DNA and RNA extractions

Fresh mycelia were obtained after 2 weeks of stationary cultivation in 2% malt extract liquid media (24ºC, dark), washed in deionised water, hand squeezed and stored at -80ºC prior to DNA or RNA extraction. High molecular weight (HMW) DNA was isolated in 8 aliquots, each consisting of up to

200mg of frozen biomass, using a modified phenol-chloroform extraction (Sagova-Mareckova et al. 2008). Frozen biomass was homogenised in liquid nitrogen to a fine powder. All vortexing steps in the workflow were replaced by 50x repeated inversion of tubes to prevent DNA fragmentation. DNA was eluted in 10mM Tris-Cl, pH 8.5, and stored at -80ºC. RNA was extracted in 6 aliquots of at least 50mg of mycelia using the Quick RNA Fungal/Bacterial Miniprep kit (Zymo Research) according to the manufacturer's protocol excepting DNAse treatment, which was carried out separately using the Turbo DNA-free kit (Invitrogen) and stored at -80ºC. Quality of the isolated DNA/RNA was checked with gel electrophoresis, where samples with HMW DNA or RNA showing three separate bands for RNA subunits were considered acceptable.

PacBio HiFi genome sequencing

Four HWM DNA extractions were combined to construct a PacBio HiFi library at the Earlham Institute, Norwich, UK using the SMRTbell Express Template Prep Kit 2.0 (PacBio, P/N 100-983-900). In total, 18µg of HWM DNA was manually sheared with the Megaruptor 3 instrument (Diagenode, P/N B06010003) according to the operations manual. After shearing, the sample underwent AMPure PB bead (PacBio, P/N 100-265-900) purification and concentration before undergoing library preparation using the SMRTbell Express Template Prep Kit 2.0 (PacBio, P/N 100-983-900). The library was prepared according to the HiFi protocol v03 (PacBio, P/N 101-853-100) and the final library was size fractionated using the SageELF system (Sage Science, P/N ELF0001) and a 0.75% cassette (Sage Science, P/N ELD7510). The library was quantified by fluorescence (Invitrogen Qubit 3.0, P/N Q33216) and the size of the library fractions were estimated from a smear analysis performed on the FEMTO Pulse System (Agilent, P/N M5330AA). The loading calculations for sequencing were completed using the PacBio SMRTLink Binding Calculator 10.1. Sequencing primer v2 was annealed to the adapter sequence of the HiFi library. The library was bound to the sequencing polymerase with the Sequel II Binding Kit v2.0 (PacBio, P/N 101-842-900). Calculations

for primer and polymerase binding ratios were kept at default values for the library type. Sequel II DNA internal control 1.0 was spiked into the library at the standard concentration prior to sequencing. The sequencing chemistry used was Sequel II Sequencing Plate 2.0 (PacBio, P/N 101-820-200) and the Instrument Control Software v10.1.0.125432. The library was sequenced on the Sequel IIe on one Sequel II SMRTcell 8M. The parameters for sequencing were diffusion loading, 30-hour movie, 4-hour pre-extension time, 2-hour immobilisation, and 70pM on plate loading concentration.

PacBio Iso-Seq

One PacBio Iso-Seq library was constructed starting from 315ng of total RNA according to the guidelines laid out in the Iso-Seq protocol v02 (PacBio, 101-763-800), using SMRTbell express template prep kit 2.0 (PacBio, 102-088-900). Please see the Supplementary Material for full details. The sequencing chemistry used was Sequel II Sequencing Plate 2.0 (PacBio, 101-820-200) and the Instrument Control Software v10.1.0.119549. The Iso-Seq library was sequenced on the Sequel IIe instrument with one Sequel II SMRTcell 8M cell. The parameters for sequencing were diffusion loading, 30-hour movie, 2-hour immobilisation time, 2-hour pre-extension time, and 80pM on plate loading concentration.

Illumina RNA-Seq

One RNA-Seq library was constructed using the NEBNext Ultra II RNA Library prep for Illumina kit (NEB#E7760L), NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB#7490) and NEBNext Multiplex Oligos for Illumina (E6440S) at a concentration of 10μM. Please see the Supplementary Material for full details. The library was loaded onto a v1.5 NovaSeq SP flow cell using the NovaSeq Xp Flow Cell Dock. The flow cell was then loaded onto the NovaSeq 6000 along with an NovaSeq 6000 v1.5 SP cluster cartridge, buffer cartridge, and 300 cycle SBS cartridge (Illumina). The

NovaSeq was run with NVCS v1.7.5 and RTA v3.4.4, and was set up to sequence 150bp paired-end reads. The data was demultiplexed and converted to fastq using bcl2fastq2.

Genome Assembly

PacBio HiFi reads were randomly subsampled to approximately 100x coverage and assembled using hifiasm v0.16.1 (Cheng et al. 2021). Alternative haplotypes were removed using purge_dups v1.0.1 (Guan et al. 2020). The assembly was manually curated by loading into the Gap5 sequence editor (Bonfield & Whitwham 2010) and each contig's integrity was assessed. Four breaks were made due to erroneously assembled reads at contig ends and one join was made between two contigs based on repeat analysis and genome synteny with the *Trametes hirsuta* reference genome (accession GCA_001302255.2).

Assembly completeness was estimated using BUSCO v5.4.7 with the Polyporales_odb10 dataset (Manni et al. 2021). The assembly consensus quality value (QV) was estimated using Merqury v1.3 (Rhie et al. 2020). Telomeric repeats (TTAGGG) at the ends of fragments were identified using tidk v0.2.31 (Brown 2023), with at least five repeats required for positive telomere prediction. The mitochondrial genome was assembled separately using IPA (https://github.com/PacificBiosciences/pbipa).

Genome Annotation

Annotation of repetitive elements was performed using the EI-Repeat pipeline v1.1.0 (https://github.com/EI-CoreBioinformatics/eirepeat), which masked the genome assembly using RepeatMasker v4.0.7 (Smit et al. 2015) with a repeat library from RepBase and a *de novo* library of repeats constructed using RepeatModeler v1.0.11 (Smit & Hubley 2015).

Gene models were annotated using REAT – Robust Eukaryotic Annotation Toolkit (https://github.com/EI-CoreBioinformatics/reat) and Minos (https://github.com/EI-CoreBioinformatics/minos), incorporating transcript assemblies from Illumina RNA-Seq and PacBio Iso-Seq data, alignment of protein sequences from related species, and evidence guided gene prediction with AUGUSTUS (Stanke and Morgenstern 2005). Please see the Supplementary Material for full details. Gene models were functionally annotated using EI-FunAnnot pipeline v1.3 (https://github.com/EI-CoreBioinformatics/eifunannot) utilising AHRD v3.3.3 (https://github.com/groupschoof/AHRD) with hits against fungal proteins from both Swiss-Prot and TrEMBL (downloaded 15/06/2022; The UniProt Consortium 2021) generated with BLAST v2.6.0 (Camacho et al. 2009) as well as results from InterProScan v5.22.61 (Jones et al., 2014).

The mitochondrial genome was annotated using MFannot v1.35 (Lang et al. 2023) and visualised using OGDRAW (Greiner et al. 2019).

Assembly and annotation data were submitted to the European Nucleotide Archive using COPO (Shaw et al. 2020).

Phylogenomics and Comparative Genomics

Assessment of syntenic conservation and collinearity with the chromosome-scale genome assembly of *Laetiporus sulphureus* gfLaeSulp1 (Wright et al. 2022) was performed using MCScanX v1.1 (Wang et al. 2012) and visualised using SynVisio (Bandi & Gutwin 2020).

For phylogenomic reconstruction of the *Polyporales* 'residual' clade *sensu* Justo et al. (2017), we inferred orthologous gene families by running OrthoFinder v2.5.4 (Emms & Kelly 2019) on the *S. occarium* proteome alongside proteomes from other residual *Polyporales* taxa with available data and *Laetiporus sulphureus* as an outgroup (supplementary table S1). CAZyme genes for all taxa

were predicted using run_dbcan v4 (Zheng et al. 2023; Zhang et al. 2018; Huang et al. 2018) and visualised alongside the STAG species tree produced by OrthoFinder.

ACKNOWLEDGMENTS

We thank the essential work of the ERGA Pilot Project co-ordinators, Ann McCartney, Giulio Formenti and Alice Mouton, in making this work possible with their coordination effort as part of the ERGA Pilot Project daily activities. Many thanks to Felix Shaw, Debby Ku and Aaliyah Providence for their assistance with the COPO data upload. We thank Steven Singer for permitting the use of unpublished data for *Cerrena unicolor* from Mycocosm. The authors acknowledge the work delivered via the Laboratory Managers and Research Computing Groups at Earlham Institute who manage and deliver High Performance Computing at EI. We thank Thomas Brabbs and James Lipscombe for assistance and advice with laboratory automation, and Neil Shearer for LIMS support.

FUNDING

This work was supported by funding from the Biotechnology and Biological Sciences Research Council (BBSRC), part of UK Research and Innovation, Core Capability Grant (BB/CCG1720/1). Part of this work was delivered via the BBSRC National Capabilities in Genomics and Single Cell Analysis (BBS/E/T/000PR9816) by members of the Technical Genomics and the Core Bioinformatics Groups, and in e-Infrastructure (BBS/E/T/000PR9814) by members of the e-Infrastructure group at the Earlham Institute. Part of this work was supported by the Earlham Institute Strategic Programme Grant Decoding Biodiversity (BBX011089/1) and its constituent work package – Decode WP2 Genome Enabled Analysis of Diversity to Identify Gene Function, Biosynthetic Pathways, and Variation in Agri/Aquacultural Traits (BBS/E/ER/230002B). PB and VB were supported by the Czech Science Foundation (21-09334J) and by the Ministry of Education, Youth and Sports of the Czech

Republic (Talking microbes - understanding microbial interactions within One Health framework; CZ.02.01.01/00/22_008/0004597).

DATA AVAILABILITY

Genome assembly and annotation data are available in the European Nucleotide Archive under BioProject PRJEB75241 and accession GCA_964035595.1.

CRediT Author Contributions

RH: formal analysis, investigation, methodology, visualisation, writing - original draft, writing - review & editing

JM: formal analysis, investigation, methodology, visualisation, writing - original draft, writing - review & editing

VB: methodology, investigation, writing - original draft, writing - review & editing

SM: project administration, writing - review & editing

NI: methodology, investigation, writing - original draft

TB: investigation, writing - original draft

VK: investigation

SL: methodology

KB: project administration

LC: supervision

CW: project administration, supervision

KG: resources, supervision, writing - review & editing

GK: formal analysis, resources, software, writing - review & editing

AT: data curation

JMDW: data curation, supervision

MT: resources, writing - original draft, writing - review & editing

PB: resources, writing - original draft, writing - review & editing

DS: resources, software, supervision, writing - review & editing

NH: resources, supervision, writing - review & editing

REFERENCES

Bandi V, Gutwin C. 2020. Interactive Exploration of Genomic Conservation. In: Proceedings of the 46th Graphics Interface Conference. Canadian Human-Computer Communications Society: University of Toronto. doi: 10.20380/GI2020.09.

Bonfield JK, Whitwham A. 2010. Gap5—editing the billion fragment sequence assembly. Bioinformatics. 26:1699–1703. doi: 10.1093/bioinformatics/btq268.

Brown M. 2023. A Telomere Identification toolKit (tidk). https://github.com/tolkit/telomeric-identifier.

Camacho C et al. 2009. BLAST+: architecture and applications. BMC Bioinformatics. 10:421. doi: 10.1186/1471-2105-10-421.

Chen C-L et al. 2022. Sexual Crossing, Chromosome-Level Genome Sequences, and Comparative Genomic Analyses for the Medicinal Mushroom *Taiwanofungus Camphoratus* (Syn. *Antrodia Cinnamomea* , *Antrodia Camphorata)*. Microbiol Spectr. 10:e02032-21. doi: 10.1128/spectrum.02032-21.

Chen S et al. 2012. Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. Nat Commun. 3:913. doi: 10.1038/ncomms1923.

Cheng H, Concepcion GT, Feng X, Zhang H, Li H. 2021. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat Methods. 18:170–175. doi: 10.1038/s41592-020-01056-5.

Dahlberg A, Mueller GM. 2011. Applying IUCN red-listing criteria for assessing and reporting on the conservation status of fungal species. Fungal ecol. 4:147–162. doi: 10.1016/j.funeco.2010.11.001.

EBP. 2023. *Report on Assembly Standards v5*. https://www.earthbiogenome.org/report-on-assembly-standards (Accessed April 9, 2024).

Emms DM, Kelly S. 2019. OrthoFinder: Phylogenetic orthology inference for comparative genomics. Genome Biol. 20:238. doi: 10.1101/466201.

GBIF. 2024. *Spongipellis delectans* (Peck) Murrill. https://www.gbif.org/species/2543646 (Accessed April 17, 2024).

Greiner S, Lehwark P, Bock R. 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 47:W59–W64. doi: 10.1093/nar/gkz238.

Grienke U, Zöll M, Peintner U, Rollinger JM. 2014. European medicinal polypores – A modern view on traditional uses. J Ethnopharmacol. 154:564–583. doi: 10.1016/j.jep.2014.04.030.

Guan D et al. 2020. Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics. 36:2896–2898. doi: 10.1093/bioinformatics/btaa025.

Hage H et al. 2021. Gene family expansions and transcriptome signatures uncover fungal adaptations to wood decay. Environ Microbiol. 23:5716–5732. doi: 10.1111/1462-2920.15423.

Hallenberg N, Ryberg M, Nilsson RH, Wood AR, Wu S-H. 2008. *Pseudolagarobasidium* (Basidiomycota): on the reinstatement of a genus of parasitic, saprophytic, and endophytic resupinate fungi. Botany. 86:1319–1325. doi: 10.1139/B08-088.

Halme P et al. 2009. Perennial polypores as indicators of annual and red-listed polypores. Ecol Indic. 9:256–266. doi: 10.1016/j.ecolind.2008.04.005.

Huang L et al. 2018. dbCAN-seq: a database of carbohydrate-active enzyme (CAZyme) sequence and annotation. Nucleic Acids Res. 46:D516–D521. doi: 10.1093/nar/gkx894.

IUCN. 2021. Re:wild and IUCN SSC become first global organizations to call for the recognition of fungi as one of three kingdoms of life critical to protecting and restoring Earth. https://www.iucn.org/news/species-survival-commission/202108/rewild-and-iucn-ssc-become-first-global-organizations-call-recognition-fungi-one-three-kingdoms-life-critical-protecting-and-restoring-earth (Accessed April 17, 2024).

Jones P et al. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics. 30:1236–1240. doi: 10.1093/bioinformatics/btu031.

Jönsson MT, Edman M, Jonsson BG. 2008. Colonization and extinction patterns of wood-decaying fungi in a boreal old-growth *Picea abies* forest. J Ecol. 96:1065–1075. doi: 10.1111/j.1365-2745.2008.01411.x.

Justo A et al. 2017. A revised family-level classification of the *Polyporales* (*Basidiomycota*). Fungal Biology. 121:798–824. doi: 10.1016/j.funbio.2017.05.010.

Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 37:540–546. doi: 10.1038/s41587-019-0072-8.

Lewin HA et al. 2018. Earth BioGenome Project: Sequencing life for the future of life. PNAS. 115:4325–4333. doi: 10.1073/pnas.1720115115.

Lonsdale D, Pautasso M, Holdenrieder O. 2008. Wood-decaying fungi in the forest: conservation needs and management options. Eur J Forest Res. 127:1–22. doi: 10.1007/s10342-007-0182-6.

Lang BF et al. 2023. Mitochondrial genome annotation with MFannot: a critical analysis of gene identification and gene model prediction. Front Plant Sci. 14:1222186. doi: 10.3389/fpls.2023.1222186.

Ma J-X et al. 2024. Whole genome sequencing and annotation of *Daedaleopsis sinensis*, a wood-decaying fungus significantly degrading lignocellulose. Front Bioeng Biotechnol. 11:1325088. doi: 10.3389/fbioe.2023.1325088.

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Mol Biol Evol. 38:4647–4654. doi: 10.1093/molbev/msab199.

May TW et al. 2018. Recognition of the discipline of conservation mycology. Conserv Biol. doi: 10.1111/cobi.13228.

Mc Cartney AM et al. 2024. The European Reference Genome Atlas: piloting a decentralised approach to equitable biodiversity genomics. npj Biodiversity. 3:28. doi: 10.1038/s44185-024-00054-6.

Miettinen O et al. 2023. A revised genus-level classification for *Cerrenaceae* (*Polyporales, Agaricomycetes*). Fungal Systematics and Evolution. 12:271–322. doi: 10.3114/fuse.2023.12.14.

Miller R, Lodge D. 2007. Fungal Responses to Disturbance: Agriculture and Forestry. In: Environmental and Microbial Relationships. Kubicek, CP & Druzhinina, IS, editors. The Mycota, A Comprehensive Treatise on Fungi as Experimental Systems for Basic and Applied Research Vol. IV Springer-Verlag: Berlin pp. 50–52.

Mueller GM, Dahlberg A, Krikorev M. 2014. Bringing fungi into the conservation conversation: The Global Fungal Red List Initiative. Fungal Conservation. 4:12–16. doi: 10.1017/s0953756202226659.

Nurk S et al. 2020. HiCanu: accurate assembly of segmental duplications, satellites, and allelic variants from high-fidelity long reads. Genome Res. doi: 10.1101/gr.263566.120.

Prescott TAK, Hill R, Mas-Claret E, Gaya E, Burns E. 2023. Fungal Drug Discovery for Chronic Disease: History, New Discoveries and New Approaches. Biomolecules. 13:986. doi: 10.3390/biom13060986.

Rhie A, Walenz BP, Koren S, Phillippy AM. 2020. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. Genome Biol. 21:245. doi: 10.1186/s13059-020-02134-9.

Sagova-Mareckova M et al. 2008. Innovative Methods for Soil DNA Purification Tested in Soils with Widely Differing Characteristics. Appl Environ Microbiol. 74:2902–2907. doi: 10.1128/AEM.02161-07.

Sandor S, Zhang Y, Xu J. 2018. Fungal mitochondrial genomes and genetic polymorphisms. Appl Microbiol Biotechnol. 102:9433–9448. doi: 10.1007/s00253-018-9350-5.

Shaw F et al. 2020. COPO: a metadata platform for brokering FAIR data in the life sciences. F1000Res. 9:495. doi: 10.12688/f1000research.23889.1.

Smit A, Hubley R. 2015. RepeatModeler Open-1.0. http://www.repeatmasker.org.

Smit A, Hubley R, Green P. 2015. RepeatMasker Open-4.0. http://www.repeatmasker.org.

Stanke M, Morgenstern B. 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Research. 33:W465–W467. doi: 10.1093/nar/gki458.

The UniProt Consortium. 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research. 49:D480–D489. doi: 10.1093/nar/gkaa1100.

Tomšovský M. 2012. Delimitation of an almost forgotten species *Spongipellis litschaueri* (Polyporales, Basidiomycota) and its taxonomic position within the genus. Mycol Progress. 11:415–424. doi: 10.1007/s11557-011-0756-z.

Větrovský T et al. 2020. GlobalFungi, a global database of fungal occurrences from high-throughput-sequencing metabarcoding studies. Scientific Data. 7:228. doi: 10.1038/s41597-020-0567-7.

Wang Y et al. 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 40:e49. doi: 10.1093/nar/gkr1293.

Wright R et al. 2022. The genome sequence of the chicken of the woods fungus, *Laetiporus sulphureus* (Bull.) Murrill, 1920. Wellcome Open Res. 7:83. doi: 10.12688/wellcomeopenres.17750.1.

Zhang H et al. 2018. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 46:W95–W101. doi: 10.1093/nar/gky418.

Zheng J et al. 2023. dbCAN3: automated carbohydrate-active enzyme and substrate annotation. Nucleic Acids Res. 51:W115–W121. doi: 10.1093/nar/gkad328.

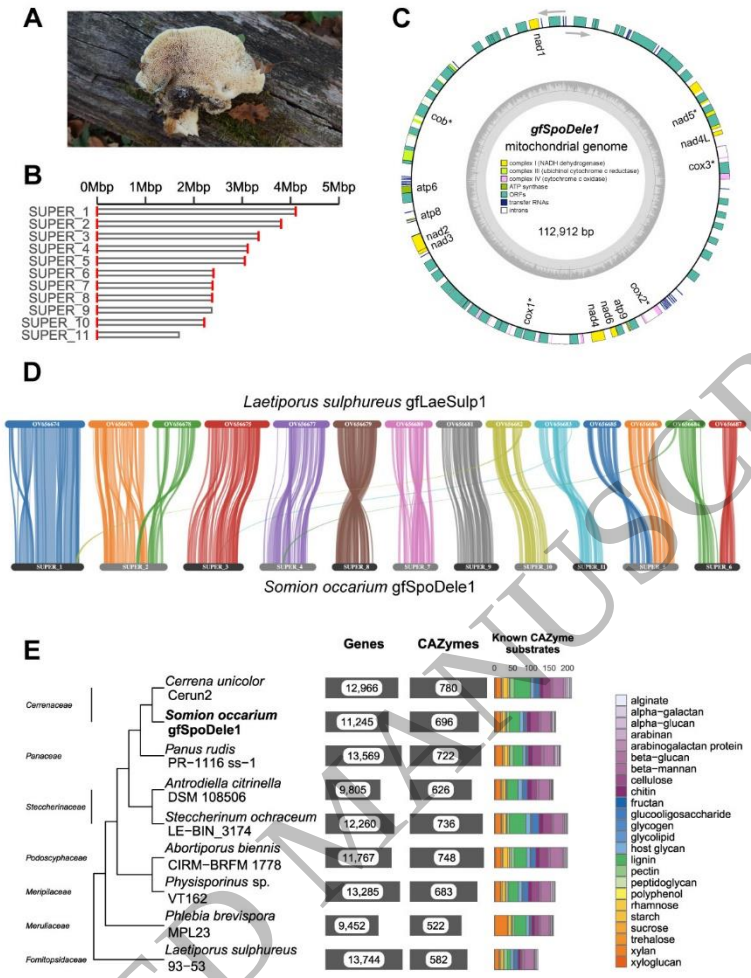
Zíbarová L et a. 2024. Red list of fungi (macromycetes) of the Czech Republic. Příroda 46:1–192.

1

Table 1. Summary statistics for the gfSpoDele1 assembly and annotation.

	Nuclear genome	Mitogenome
Number of contigs	11	1
Total size (bp)	30,870,404	112,912
N50 (Mbp)	3.05	-
Kmer quality value (QV)	66.53	-
Number of protein-coding genes (high confidence/low confidence)	11,245 (9,409/1,836)	14 (35 ORFs, 19 intronic ORFs)
BUSCO completeness (assembly)	C:94.4%[S:93.8%,D:0.6%], F:0.8%,M:4.8%,n:4464	-
BUSCO completeness (annotation)	C:98.6%[S:97.7%,D:0.9%], F:0.1%,M:1.3%,n:4464	-

2



1