

# IMA genome#F17

## Journal Article

**Author(s):**

Wingfield, Brenda D.; Berger, Dave K.; Coetzee, Martin P.A.; Duong, Tuan A.; Martin, Anke; Pham, Nam Q.; van den Berg, Noelani; Wilken, P. Markus; Arun-Chinnappa, Kiruba Shankari; Barnes, Irene; Buthelezi, Sikelela; Dahanayaka, Buddhika Amarasinghe; Durán, Alvaro; Engelbrecht, Juanita; Feurtey, Alice; Fourie, Arista; Fourie, Gerda; Hartley, Jesse; Kabwe, Eugene N.K.; Maphosa, Mkhululi; et al.

**Publication date:**

2022

**Permanent link:**

<https://doi.org/10.3929/ethz-b-000591920>

**Rights / license:**

[Creative Commons Attribution 4.0 International](#)

**Originally published in:**

IMA Fungus 13(1), <https://doi.org/10.1186/s43008-022-00104-3>



# IMA genome-F17

Draft genome sequences of an *Armillaria* species from Zimbabwe, *Ceratocystis colombiana*, *Elsinoë necatrix*, *Rosellinia necatrix*, two genomes of *Sclerotinia minor*, short-read genome assemblies and annotations of four *Pyrenophora teres* isolates from barley grass, and a long-read genome assembly of *Cercospora zeina*

Brenda D. Wingfield<sup>1\*</sup>, Dave K. Berger<sup>2</sup>, Martin P. A. Coetzee<sup>1</sup>, Tuan A. Duong<sup>1</sup>, Anke Martin<sup>3</sup>, Nam Q. Pham<sup>2</sup>, Noelani van den Berg<sup>1</sup>, P. Markus Wilken<sup>1</sup>, Kiruba Shankari Arun-Chinnappa<sup>3,4</sup>, Irene Barnes<sup>1</sup>, Sikelela Buthelezi<sup>1</sup>, Buddhika Amarasinghe Dahanayaka<sup>3</sup>, Alvaro Durán<sup>5</sup>, Juanita Engelbrecht<sup>1</sup>, Alice Feurtey<sup>6,7</sup>, Arista Fourie<sup>1</sup>, Gerda Fourie<sup>1</sup>, Jesse Hartley<sup>1</sup>, Eugene N. K. Kabwe<sup>8</sup>, Mkhululi Maphosa<sup>1</sup>, Deborah L. Narh Mensah<sup>1,10</sup>, David L. Nsibo<sup>2</sup>, Lizel Potgieter<sup>6,7</sup>, Barsha Poudel<sup>3</sup>, Eva H. Stukenbrock<sup>6,7</sup>, Chanel Thomas<sup>1</sup>, Nilofar Vaghefi<sup>3,9</sup>, Tanya Welgemoed<sup>8</sup> and Michael J. Wingfield<sup>2</sup>

## IMA GENOME-F 17A

Draft genome sequence of an *Armillaria* species from Zimbabwe

### Introduction

The genus *Armillaria* includes at least 38 species, most of which are facultative necrotrophs (Gregory and Rishbeth 1991). Pathogenicity of these organisms can result in *Armillaria* root and stem rot and what is referred to as shoestring root rot (Morrison 1991). This disease can bring about massive devastation to woody plants grown for horticulture, agriculture, as well as natural and managed forests across the various continents (Baumgartner and Rizzo 2001, 2002; Guillaumin et al. 1993; Labbé et al. 2015). The saprophytic nature of some *Armillaria* spp. results in enhancement of forest ecosystems through the breakdown of woody material, resulting in carbon and mineral cycling (Baumgartner et al. 2011; Heinzelmann et al. 2019). Transition from a saprophytic to a pathogenic lifestyle, and vice versa, can occur due to

intra-species variation, forest management systems, the state of the host (e.g. stressed or healthy), as well as environmental factors (e.g. elevation) (Legrand et al. 1996; Prospero et al. 2004; Tsykun et al. 2012).

Various groups have conducted omics-based research on *Armillaria* species (Akulova et al. 2020; Anderson et al. 2018; Caballero et al. 2022; Collins et al. 2017; Collins et al. 2013; Heinzelmann et al. 2020; Kolesnikova et al. 2019; Linnakoski et al. 2021; Misiek et al. 2011; Misiek and Hoffmeister 2012; Sipos et al. 2017; Sonnenbichler et al. 1997; Sun et al. 2020; Wingfield et al. 2016a, b; Zhan et al. 2020). These genomics, proteomics and metabolomics studies were done to gain insight into the molecular mechanisms and biochemical properties that drive the pathogenicity and virulence of *Armillaria* spp. This information would eventually help to develop efficient strategies for identifying these fungi, containing their spread, and minimising damage to forest ecosystems.

The previously determined nuclear and mitochondrial genomes of various *Armillaria* species are providing invaluable resources for genome-based research (Table 1). Studies using these genomes have broadened our understanding of the biology of the *Armillaria* species and the evolution of their genomes. For

\*Correspondence: [brenda.wingfield@fabi.up.ac.za](mailto:brenda.wingfield@fabi.up.ac.za)

<sup>1</sup> Department of Biochemistry, Genetics and Microbiology, Faculty of Natural and Agricultural Sciences, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa  
Full list of author information is available at the end of the article



**Table 1** Genome information for the published *Armillaria* species in comparison to *Armillaria* African Clade B isolate CMW4456

Species	Strain	Number of scaffolds	Assembly size (Mbp)	Number of predicted protein coding genes	GC content (%)	Origin	References
<i>Armillaria altimontana</i> (NABS X)	837–10	100	73.74	19,326	47.8	USA	Caballero et al. (2022)
<i>Armillaria borealis</i>	AB13-TR4-IP16	44,365	66.59	21,969	N/A	Russia	Akulova et al. (2020)
<i>Armillaria cepistipes</i>	B5	182	75.52	23,461	47.6	Italy	Sipos et al. (2017)
<i>Armillaria fuscipes</i>	CMW2740	24,403	52.98	14,515	N/A	South Africa	Wingfield et al. (2016a, b)
<i>Armillaria gallica</i>	Ar21-2	319	85.34	25,704	47.5	USA	Sipos et al. (2017)
<i>Armillaria gallica</i>	012 m	63	87.31	26,261	47.4	China	Zhan et al. (2020)
<i>Armillaria mellea</i>	DSM 3731	4,377	58.36	14,473	49.1	France	Collins et al. (2013)
<i>Armillaria ostoyae</i>	C18/9	106	60.11	22,705	48.3	Switzerland	Sipos et al. (2017)
<i>Armillaria solidipes</i>	C28-4	229	58.01	20,811	48.4	USA	Sipos et al. (2017)
<i>Armillaria solidipes</i> (form <i>A. ostoyae</i> )	ID001	72	55.74	16,357	48.3	USA	Caballero et al. (2022)
<i>Armillaria</i> African Clade B sp.	CMW4456	840	54.95	13,600	46.5	Zimbabwe	Described here

N/A Not available

example, Sipos et al. (2017) showed that genome evolution in the genus was predominantly caused by gene family expansion. Kolesnikova et al. (2019) assembled the complete mitochondrial genomes of *A. borealis*, *A. gallica*, *A. sinapina*, and *A. solidipes* and found a high degree of variation in size, gene content and genomic organization among these phylogenetically closely related species. Recently, the first chromosome-level *Armillaria* genome assembly became available, revealing genome-wide recombination in the genome of *A. ostoyae* (Heinzmann et al. 2020).

Sequenced *Armillaria* species originate primarily in the Northern Hemisphere. The genome of only one species from the Southern Hemisphere, *A. fuscipes* from South Africa, has so far been published (Wingfield et al. 2016a, b). It is known that species from the Northern Hemisphere, Australasia together with Southern America, and Africa, respectively, reside in distinct monophyletic lineages (Coetzee et al. 2011; Koch et al. 2017). Genomes of species in these geographic locations may, therefore, have followed very different evolutionary pathways. Within the African clade, Coetzee et al. (2005) identified two lineages, referred to as *A. fuscipes* and African Group B. Here, we report the genome of an *Armillaria* isolate belonging to African Group B (sensu Coetzee et al. 2005), sequenced using both long- and short-read technologies. This genome expands the sequence resources for *Armillaria* species from the Southern Hemisphere and Africa.

#### Sequenced strains

**Zimbabwe:** Stapelford, Manicaland isolated from *Brachystegia utilis*, 2001, *E. Mwenje* (culture CMW4456; PREM 63337—dried culture).

#### Nucleotide sequence accession number

The Whole Genome Shotgun project of the *Armillaria* sp. genome has been deposited at DDBJ/ENA/GenBank under the accession JANDKJ000000000. The version described in this paper is version JANDKJ010000000.

#### Materials and methods

The culture of isolate CMW4456 were grown and maintained in MYA (1.5% Malt extract, 0.2% Yeast extract, 1.5% Agar) at 24 °C in the dark for 4 weeks. DNA was extracted from the harvested mycelia using the method described by Duong et al. (2013). PacBio sequencing was conducted on the Sequel IIE system using the circular consensus sequencing (CCS) mode at Inqaba Biotechnical Industries (Pty) Ltd. (Pretoria, South Africa).

For short read sequencing on the Illumina HiSeq platform, genomic DNA was extracted from cultures grown in MY broth (1.5% Malt extract, 0.2% Yeast extract) for six weeks at 24 °C in the dark. Harvested cultures were kept at – 80 °C, followed by lyophilization. DNA was extracted with the Qiagen DNEasy Plant Pro Kit (50) Cat. No. 69204 (Qiagen, Sandton, South Africa) following the manufacturer's instructions. Illumina paired-end library

preparation and whole-genome sequencing was done with an insert size of 350 bp and read-length of 150 bp at MacroGen.

Trimmomatic v. 0.38 (Bolger et al. 2014) was used to trim adapter sequences and low-quality ends of the Illumina reads (ILLUMINACLIP, TruSeq3-PE.fa:2:30:10:8; LEADING, 3; TRAILING, 3; MINLEN, 30).

The PacBio HiFi reads were assembled with CLC Genomics Workbench v 22.0.1 (QIAGEN, Aarhus). The assembly was subsequently polished with the trimmed Illumina HiSeq reads, using Pilon v. 1.23 (Walker et al. 2014). Genome completeness was evaluated with Benchmarking Universal Single-Copy Orthologs (BUSCO) v. 5.3.2, using the agaricales\_odb10 lineage dataset (Manni et al. 2021). AUGUSTUS v. 3.4.0 (Keller et al. 2011; Stanke et al. 2008; Stanke and Waack 2003) was used to predict protein coding genes, applying gene models of the closely related species, *Coprinus cinereus*. QUASt v 5.0.2 (Gurevich et al. 2013) was used to evaluate metrics, including contig number, total length, GC content, and N50 for the genome assemblies. BUSCO, AUGUSTUS and QUASt were run using the Galaxy platform (Afgan et al. 2018; The Galaxy Community 2022) (<https://usegalaxy.eu/>). *Armillaria ostoyae* strain C18/9 genome (Sipos et al. 2017) with accession number FUEG00000000.1 was used as the reference genome for genome quality evaluation.

The identity of the *Armillaria* African Clade B isolate CMW4456 for which a genome was sequenced was confirmed based on phylogenetic grouping with published DNA sequences. DNA sequences from the internal transcribed spacer region (ITS) and the translation elongation factor one alpha (*tef1- $\alpha$* ) were extracted from the genome. Since few *tef1- $\alpha$*  sequences from African isolates are available in databases, the *tef1- $\alpha$*  sequence was compared to sequences on GenBank using BLASTn. The ITS sequence was included in the data matrix of Coetzee et al. (2005) and aligned using the online version of MAFFT v. 7. (Katoh et al. 2019). The TrN+G nucleotide substitution model was determined as best fitting the sequence alignment, using jModelTest and the Akaike Information Criterion (Darriba et al. 2012; Guindon and Gascuel 2003), and incorporated in the maximum likelihood analyses. A maximum likelihood phylogenetic tree was constructed using PHYML v. 3.0 (Guindon et al. 2010), applying 1000 bootstrap replicates. The tree was rooted with sequences of *A. hinnulea*.

### Results and discussion

The 4 PacBio read length ranged between 289 and 15627 bases. The 2 × 151 bp Illumina HiSeq paired-end libraries yielded a total of 14,927,540,182 reads, amounting to 98,857,882 nucleotides. The PacBio and Illumina reads

were assembled into 840 contigs with a total assembly size of 54.95 Mbp. All contigs were longer than 1000 bp, with the largest contig being 1,463,441 bp. The N50 and N75 values were 128,967 bp and 45,059 bp, respectively. The L50 and L75 values were 85 and 270 contigs, respectively. The GC content of the assembled genome was 46.53%. Genome completeness was estimated to be 98%, corresponding to 96.8% complete and single-copy BUSCOs, 1.2% complete and duplicated BUSCOs, 0.1% fragmented BUSCOs, and 1.9% missing BUSCOs (n = 3870). AUGUSTUS predicted 13,600 protein coding genes.

The genome statistics of the sequenced *Armillaria* strain correlated with that reported for the genomes of other species in the genus (Table 1). The assembly size fell within the range of 53.00–73.74 Mbp, though the number of predicted protein coding genes (13,600 genes) was somewhat lower than the 14,473–26,261 genes reported in the assembled genomes of other species of *Armillaria*. The GC contents of the genomes of other *Armillaria* spp. (47.4–49.1%) are similar to the GC content of 46.53% reported here.

The sequenced genome of *Armillaria* sp. strain CMW4456 grouped with other strains of the African *Armillaria* Clade B from Cameroon, Zambia and Zimbabwe (Coetzee et al. 2005), confirming its identity (Fig. 1). The *tef1- $\alpha$*  sequence from the genome was identical to the *tef1- $\alpha$*  sequence of CMW4456 on GenBank (accession number DQ435617.1). This genome sequence will serve as a useful resource for investigating the biology, chemistry, and pathogenicity of *Armillaria* species from Africa in comparison to those from other continents.

**Authors:** Deborah L. Narh Mensah<sup>1</sup>, Brenda D. Wingfield<sup>1</sup>, Mkhululi Maphosa<sup>1</sup>, Tuan A. Duong<sup>1</sup>, and Martin P. A. Coetzee<sup>1,2</sup>.

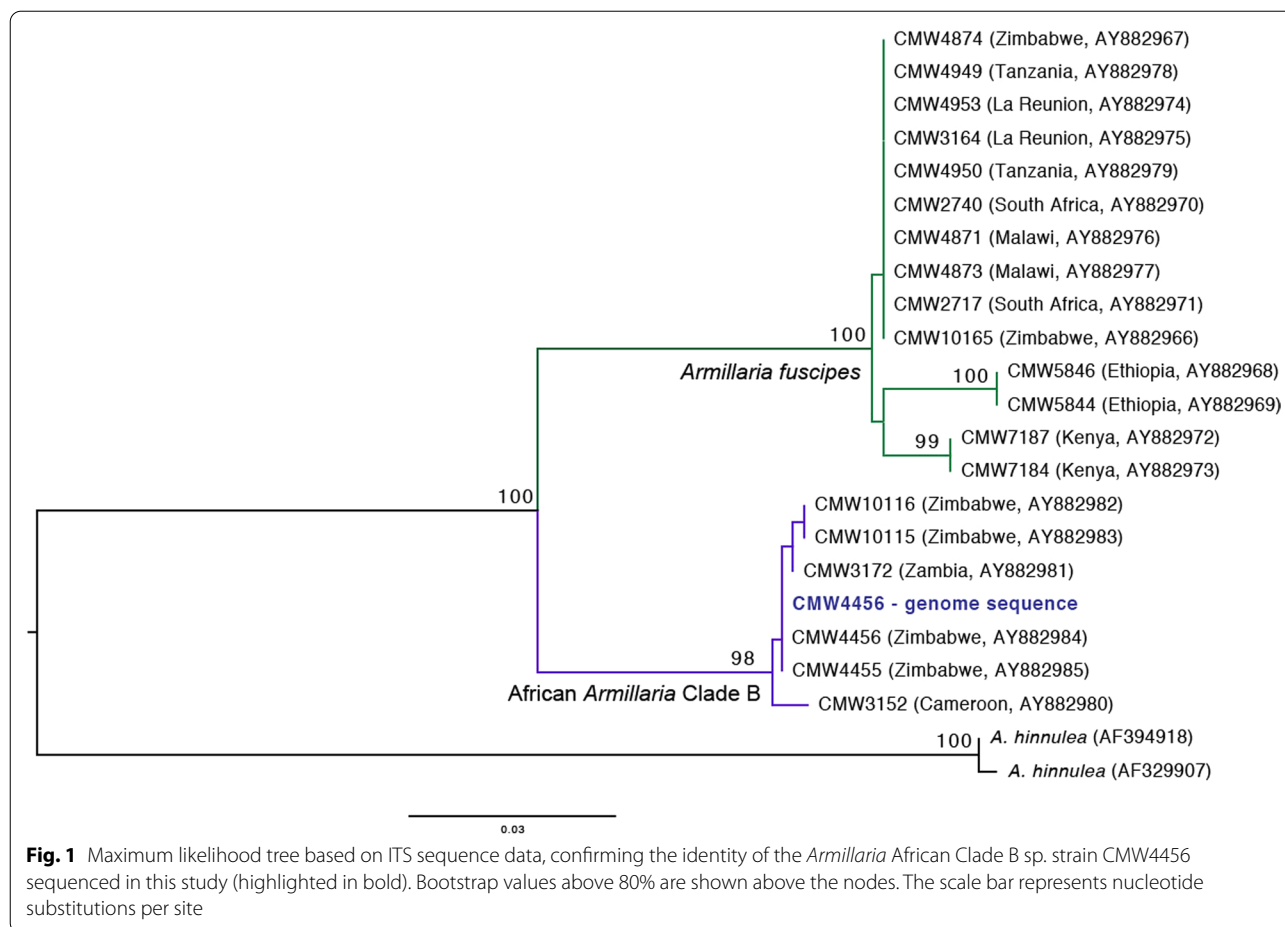
\*Contact: martin.coetzee@fabi.up.ac.za.

## IMA GENOME-F 17B

### Draft genome sequence of *Ceratocystis colombiana*

#### Introduction

The ascomycete fungal genus *Ceratocystis* consists of over 40 species, many of which are important pathogens of forestry and agricultural tree crops worldwide (de Beer et al. 2014). While many species of *Ceratocystis* only cause problems in their regional distribution, some managed to spread across continents and pose considerable threats to the health of forestry and plantations worldwide (Engelbrecht et al. 2004; Liu et al. 2021). The most notable pathogens of the genus include *C. fimbriata* that causes black root of sweet potatoes (Halsted 1890), *C. platani* that causes cankers and wilt disease of plane trees (Tsopeles et al. 2017), *C. eucalypticola* that causes vascular wilt in *Eucalyptus* (Roux et al. 2020), and *C. manginecans* that causes vascular wilt disease in Acacia and



Mango trees (Tarigan et al. 2011; Al Adawi et al. 2013). It also includes the recently described *C. huliohia* and *C. lukuohia* associated with the rapid death of native `ōhi`a lehua (*Metrosideros polymorpha*) in Hawai`i (Barnes et al. 2018).

The genomic resources for studying *Ceratocystis* species have been increasing in recent years thanks to the increasing affordability in sequencing fungal genomes with next generation sequencing technologies. There are a number of genome sequences available for species of *Ceratocystis* from various hosts including *C. fimbriata* from sweet potato (Wilken et al. 2013), *C. manginecans* from *Acacia mangium* (van der Nest et al. 2014b), *C. eucalypticola* from *Eucalyptus* (Wingfield et al. 2015a, b), *C. albifundus* from *Acacia mearnsii* (van der Nest et al. 2014a, 2019), *C. harringtonii* from poplar (Wingfield et al. 2016a, b), *C. smalleyi* from hickory (Wingfield et al. 2018), and *C. cacaofunesta* from cacao (Molano et al. 2018). In this study, we report the genome sequence of an isolate of *C. colombiana* from *Coffea arabica* in Colombia. This species is known

from coffee, citrus, and *Schizolobium parahyba* and but so far is only known from Colombia (Van Wyk et al. 2010). Pathogenicity assays showed that that *C. colombiana* can cause disease on *Coffea arabica* and hence it could pose a serious threat to the coffee industry in the region and the availability of genome sequence data will help in understanding its biology and pathogenicity.

#### Sequenced strain

**Colombia:** Valle del Cauca, isol. ex *Coffea arabica*, 2000, *M. Marin* (CMW5751; CBS 121792; PREM 59434—dried culture).

#### Nucleotide sequence accession number

The *Ceratocystis colombiana* genomic sequence data has been deposited at DDBJ/EMBL/GenBank under the accession JAOSLS000000000. The version described in this paper is version JAOSLS010000000.



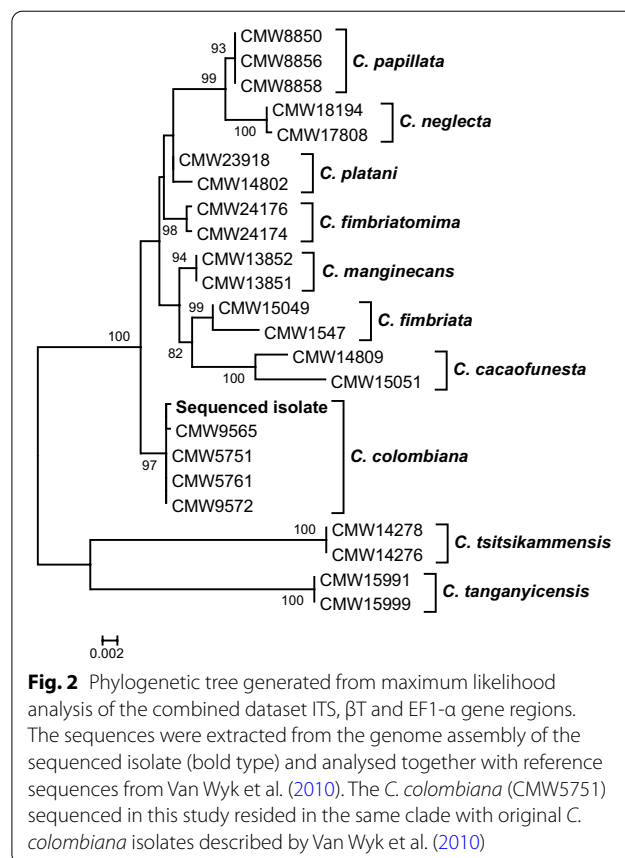
### Materials and methods

A single spore culture of *C. colombiana* CMW5751 was grown on malt extract broth (2% malt extract; 0.5% yeast extract) for 3 d at room temperature, after which the mycelium was harvested and freeze-dried. DNA was extracted from freeze-dried mycelium using the method described in Duong et al. (2013). Whole genome sequencing was by Macrogen (Seoul, South Korea) where a paired-end library was constructed using the TruSeq PCR free protocol and sequenced on the HiSeq 2500 platform to obtain 251 bp paired-end reads. The illumina data were trimmed using Trimmomatic v.0.38.1 (Bolger et al. 2014) and the genome was assembled using SPAdes v.3.14.0 (Bankevich et al. 2012). The resulting assembled scaffolds were filtered based on k-mer coverage ( $\geq 20\%$  of the medium coverage) and size ( $\geq 500$  bp). Assembly completeness was assessed using BUSCO v.4.1.4 (Simão et al. 2015) using the *sordariomycetes\_odb10* dataset. The number of protein coding gene models was estimated using AUGUSTUS v.3.2.3 (Keller et al. 2011) using the pre-defined species model for *Fusarium graminearum*. To validate the identity of the sequenced isolate, ITS,  $\beta$ T and EF1- $\alpha$  gene regions were extracted from the assembly and a maximum likelihood phylogeny was constructed using the reference sequences of Van Wyk et al. (2010).

### Results and discussion

A total of ~2. million paired-end reads were obtained from Illumina paired-end sequencing, of which ~1.96 million pairs remained after trimming. *De novo* genome assembly with SPAdes followed by filtering steps resulted in a final assembly with 973 scaffolds and a N50 of 80.73 Kb. The genome assembly of *C. colombiana* is 31.19 Mb with a mean GC of 47.93%. The genome size of this species is in the same range as for other species of *Ceratocystis* sequenced to date; 27.31 Mb for *C. smalleyi* (Wingfield et al. 2018) and 32.15 Mb for *C. manginecans* (Fourie et al. 2020). Phylogenetic analysis using the extracted gene regions (ITS,  $\beta$ T and EF1- $\alpha$ ) confirmed that the sequenced isolate was *C. colombiana* (Fig. 2), residing in the same clade with isolates originally described by Van Wyk et al. (2010). BUSCO analysis of the genome assembly using the *sordariomycetes\_odb10* dataset yielded a complete score of 94.6% and Augustus predicted 7358 protein coding genes encoded by the genome. The genome of *C. colombiana* from this study, together with those available for the genus and the larger *Ceratocystidaceae* will facilitate comparative genomics studies to understand evolutionary, pathogenicity and host adaptation of these economically important group of pathogens.

**Authors:** Tuan A. Duong\* and Brenda D. Wingfield.  
\*Contact: Tuan.duong@fabu.up.ac.za.



### IMA GENOME-F 17C

#### Long-read genome assembly of the maize grey leaf spot pathogen *Cercospora zeina* gives insight into its genomic architecture

##### Introduction

Grey leaf spot (GLS) is a foliar disease of maize that is widespread in sub-Saharan Africa (Meisel et al. 2009; Nsibo et al. 2021). The causal agents of GLS are two closely related species of *Mycosphaerellales* (*Dothideomycetes*), *Cercospora zeina* and *Cercospora zeaemaydis* (Crous et al. 2006). *Cercospora zeina* has a wide geographical distribution including the USA, Brazil, and China (Wang et al. 1998; Liu and Xu 2013; Neves et al. 2015). It is the causal agent of GLS in Africa, as DNA analysis confirmed that only *C. zeina* was present amongst 964 isolates from Kenya, Uganda, Zambia, Zimbabwe, and South Africa (Nsibo et al. 2021). The causal agent of GLS in Africa was named as *C. zeaemaydis* prior to splitting the GLS causal agent into two species, and the use of a reliable DNA diagnostic to distinguish it from *C. zeina* (Ward et al. 1999; Crous et al. 2006; Swart et al. 2017).

GLS caused by *C. zeina* results in great economic losses. Severe blighting of leaves in GLS susceptible

maize genotypes reduces the supply of photosynthate during grain filling, with yield losses of 40–67% reported in South African field trials (Ward et al. 1999). A Zambian strain of *C. zeina* (CMW25467) has been used as an experimental model for studying infection of the maize host and the molecular biology of the pathogen (Meisel et al. 2009; Korsman et al. 2012; Swart et al. 2017; Meyer et al. 2017). A draft genome sequence of this strain was previously generated using short read data (Wingfield et al. 2017). However, this assembly was highly fragmented which limits its use in studying genomic architecture and evolutionary processes. To improve the genome assembly for the same strain (CMW25467) we employed long-read sequencing technology, since this approach can generate contiguous assemblies that span entire regions of repetitive DNA. The improved genome assembly will help to gain insight into the genomic architecture, proximity of coding genes to mobile elements, and serve as a foundation for syntenic comparisons with related phytopathogens.

#### Sequenced strain

**Zambia:** *Central region (Mkushi):* isol. ex *Zea mays* (maize), March 2007, F.J. Kloppers & B. Meisel (CMW25467, MUCL 51677, CBS142763, PREM 61898—dried culture).

#### Nucleotide sequence accession number

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number MVDW00000000. The version described in this paper is version MVDW02000000. Biosample SAMN06067857; Bioproject PRJNA355276.

#### Materials and methods

*Cercospora zeina* strain CMW25467 (Meisel et al. 2009) was recovered from a glycerol stock stored at  $-80^{\circ}\text{C}$  and maintained on V8 media. Greenhouse-grown maize B73 plants were inoculated with the strain. Once GLS lesions formed, single spore isolation was made to obtain a culture used for DNA extraction and sequencing. The gDNA was extracted using the CTAB protocol (Allen et al. 2006). The quality and quantity of the extracted DNA (3.2  $\mu\text{g}$ ) were assessed using 1% agarose gel electrophoresis and a Qubit 4 fluorometer.

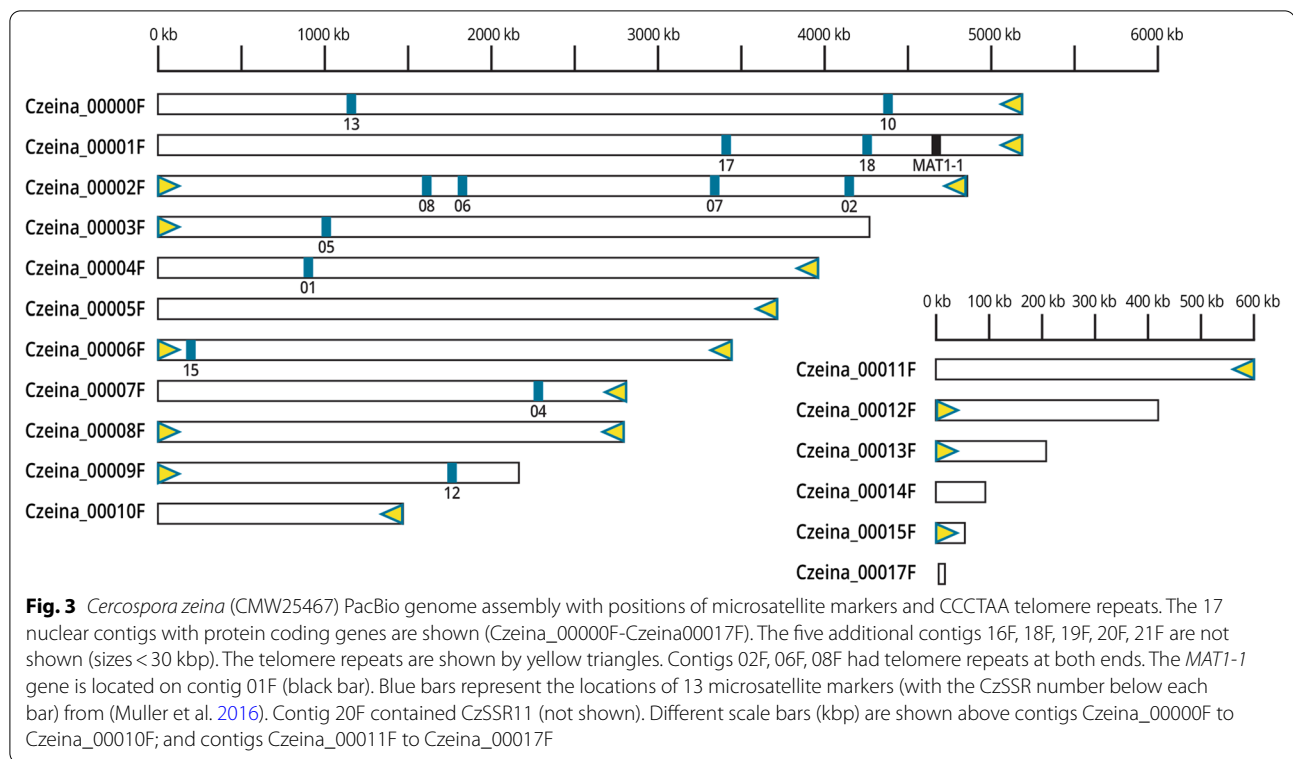
PacBio sequencing was performed using one cell of a Single Molecule Real Time (SMRT) sequencer. The PacBio raw reads were assembled using the Hierarchical Genome Assembly Process (HGAP v4) included in SMRTLink v6 (Pacific Biosciences, CA) with default parameters. A polished assembly was generated using Arrow, which uses a statistical approach to generate a consensus sequence from the PacBio reads, as

implemented in SMRTLink. The depth of coverage obtained by realignment of reads on the draft assembly was used to filter the assembly: we discarded contigs with a depth deviating by more than 1.5X from the average coverage across all contigs weighted by the contig length as in Plissonneau et al. (2016). To further improve the PacBio-based assembly, we used previously generated paired-end Illumina reads from the same *C. zeina* strain (Wingfield et al. 2017). We aligned the Illumina reads to the PacBio assembly using the Burrows-Wheeler Alignment (BWA) tool (Li and Durbin 2009), followed by one round of polishing with Pilon (Walker et al. 2014) with default settings. Statistics related to the assembly quality such as the N50 were measured with QUAST v4.4 (Gurevich et al. 2013). We addressed to which extent the PacBio genome assembly comprised complete chromosomes. To identify chromosome ends we used the programme Bowtie to identify the telomeric repeat CCC TAA in the final genome assembly (Langmead et al. 2009). This repeat motif has been identified in the telomeric regions of diverse organisms including plant pathogenic fungi (Fulnečková et al. 2013; King et al. 2015). We considered loci with at least ten times the repeat length (including one potential incorrect repeat) as putative sub-telomeric regions.

To annotate the genome, RNA-Seq reads from several *in planta* and *in vitro* sources (GSE99005, GSE94442, GSE90705, (Swart et al. 2017; Meyer et al. 2017)) were mapped onto the genome assembly using HISAT2 v2.1.0 (Kim et al. 2015) with the following parameters: --pen-noncansplice 18, --mp 6,0, --no-softclip, --max-intronlen 10,000, -t-reorder. The BAM file generated from HISAT was filtered to retain only concordant pair alignments and then used in the annotation of the genome using BRAKER v2.0.6 (Hoff et al. 2016) with the—fungus parameter. Genome assembly and annotation completeness were assessed with BUSCO v3.2.0 using the Ascomycota dataset (Simão et al. 2015).

OcculterCut v1.1 (Testa et al. 2016) was used to investigate whether the *C. zeina* genome was compartmentalized with respect to GC content. This method scans along the genome and detects adjacent segments of at least 1 kbp that have statistically significant GC content differences and are on either side of a position where the Jensen-Shannon divergence is maximized (Testa et al. 2016).

The species identity of the sequenced strain was verified by extraction of the translation elongation factor 1-alpha (TEF) and ITS sequences from the PacBio genome assembly, and phylogenetic comparison with related fungal species. Phylogenetic analysis was conducted using RAxML by applying the GTR + F0 + G10m model (Kozlov et al. 2019). The datasets for *TEF* and *ITS*



nucleotide sequences were the same as those used to verify the species identity in the previous version of the *C. zeina* genome (Wingfield et al. 2017), with the exception that we used a more closely related outgroup (*Pseudocercospora oxalidis*) in this study.

### Results and discussion

Genome assembly using HGAP v4 resulted in a 41 Mbp assembly consisting of 22 contigs of which 17 contigs had a mean coverage of 260X. Analysis focused on the 17 contigs since they corresponded to nuclear genome regions with protein-coding genes (Fig. 3; contig names shortened from Czeina\_xxxxF to xxF in the text). The five contigs not included in the set of 17 had either a low (< 260X; contigs 16F, 18F, and 20F) or a very high mean coverage (> 1000X, contigs 19F and 21F). The assembled genome had an N50 of 4 Mbp and an L50 of 5. The longest contig was 5,172,692 bp (contig 00F) and the shortest contig was 29,391 bp (contig 17F). The telomeric repeated motif CCCTAA was present at both ends of three contigs (02F, 06F, 08F), indicating that these may represent complete chromosomes (Fig. 3). The telomeric repeated motif was also found at one end of twelve other contigs (Fig. 3).

The average GC content of the *C. zeina* genome was found to be 48%. We next investigated the variation in distribution of GC content along the genome assembly using the software OcculterCut (Testa et al. 2016).

This analysis showed that approximately a third of the genome is made up of AT-rich segments (less than 41% GC content).

A total of 11,570 gene models were predicted from this new genome assembly, which is 1377 more than the 10,193 gene models predicted from a previously published Illumina-based genome assembly and annotation (Wingfield et al. 2017). Long-read sequencing with PacBio therefore revealed additional gene models that would not have been identified in a more fragmented assembly. The number of protein-coding genes predicted in the newly assembled *C. zeina* genome was similar to that of other Dothideomycetes determined by long-read sequencing such as 10,528 to 12,386 in *Zymoseptoria* species (Feurtey et al. 2020), 11,257 in *Ascochyta rabiei* (Shah et al. 2020), and 14,186 in *Pyrenochaeta lycopersici* (Dal Molin et al. 2018). BUSCO analysis indicated that the assembly had a completeness score of 97.2% with 0.1% duplicated, 1.7% fragmented and 1.1% missing genes. The PacBio assembly reported here recovered five additional BUSCO genes that were not identified in the previous version of the assembly (Wingfield et al. 2017).

The five contigs that we had previously filtered out due to high or low coverage were also annotated to determine their gene content. BLASTN searches of the contigs 16F and 18F against the nr/nt database (NCBI) revealed sequences with similarity to annotated transposable



element sequences of *C. zea-maydis*. Contig 20F did not have any significant BLASTN match other than to a *C. zeina* microsatellite sequence (CzSSR11) (Fig. 3) (Muller et al. 2016). Contig 19F had sequence similarity to ribosomal RNA sequences including the 18S ribosomal RNA gene of various *Mycosphaerellales* including *Cercospora sojina*, *Zymoseptora tritici*, and *Cladosporium fulvum*. This suggested that contig 19F contains the ribosomal RNA cistron of *C. zeina*. The BLASTN search of contig 21F resulted in similarity to mitochondrial genome sequences of fungi, indicating that this may represent part of the *C. zeina* mitochondrial genome. None of the *in planta* nor *in vitro* RNA-Seq reads mapped to contigs 16F, 18F or 20F (data not shown). Finally, we also conducted an *ab initio* gene prediction of the sequences of the three contigs using AUGUSTUS, which did not identify additional coding sequences. We conclude that the additional contigs are largely made up of repetitive sequences. However, further repeat element annotations are needed to verify this, and to identify the type of elements that they putatively encode.

The availability of the improved *C. zeina* assembly allowed us to locate the positions of microsatellites developed for population genetics studies (Muller et al. 2016) to determine whether linkage disequilibrium could affect their use in future. Previously, the sequences of the *C. zeina* CMW25467 alleles for 14 micro satellite markers had been determined by Sanger sequencing (NCBI accessions KP015832-42, KP015844, KP015846-47) (Muller et al. 2016). We located these on the *C. zeina* PacBio assembly using BLASTN, which showed that most of them were on different contigs or spaced more than 1 Mbp apart, except for CzSSR06 and CzSSR08 which were 200 kbp apart (Fig. 3). That most of these markers are not closely linked corroborates their usefulness where they have been used for population genetics studies of *C. zeina* isolates from East and southern Africa, and for future studies (Muller et al. 2016; Nsibo et al. 2019, 2021). The genome assembly from long-read sequencing determined in this study (MVDW02) was confirmed to be derived from *C. zeina*. This was shown by extraction of the sequences of elongation factor 1-alpha (*TEF1*) and *ITS* from the assembled genome sequence, followed by phylogenetic analysis with a dataset of these genes from related fungi. The phylogenetic tree showed that the source of the long-read assembly clustered with the same *C. zeina* strain (CMW25467) and other *C. zeina* strains, and that it was distinct from *C. zea-maydis* and other *Cercospora* species, with strong bootstrap support (Fig. 4).

**Authors:** E.N.K. Kabwe, T. Welgemoed, T.A. Duong, L. Potgieter, A. Feurtey, D.L. Nsibo, E.H. Stukenbrock, and D.K. Berger\*

\*Contact: dave.berger@fabi.up.ac.za.

## IMA GENOME-F 17D

### Draft genome sequence of *Elsinoë necatrix*: the causal agent of an emerging new and serious *Eucalyptus* foliar disease

#### Introduction

*Elsinoë* (*Elsinoaceae*) was introduced by Raciborski (1900) to accommodate a fungus that causes scab-like lesions on plant tissue. Members of these necrotrophic fungal pathogens are globally distributed and infecting approximately 70 hosts including important forestry, agricultural and horticultural crops as well as ornamental plants (Fan et al. 2017; Marin-Felix et al. 2019). Important diseases include the citrus scab pathogens *Elsinoë fawcettii* and *E. australis* (Chung 2011), the causal agent of grapevine spot anthracnose *E. ampelina* (Li et al. 2021), and the recently described *E. necatrix* that causes a devastating scab and shoot malformation disease on plantation-grown *Eucalyptus* (Pham et al. 2021).

*Eucalyptus* scab and shoot malformation was first reported in North Sumatra, Indonesia, in 2014 (Pham et al. 2021). The disease is characterized by black necrotic spots that first emerge on young leaves and petioles, which become scab-like as the lesions age. Infected trees respond to infection by producing shoots with small leaves that commonly appear feathered. Severely affected *Eucalyptus* clones usually die after a number of successive infection cycles, generally over a period of two to three years. The disease has become prevalent across a large area of planted *Eucalyptus* in the region, resulting in significant damage (Pham et al. 2021).

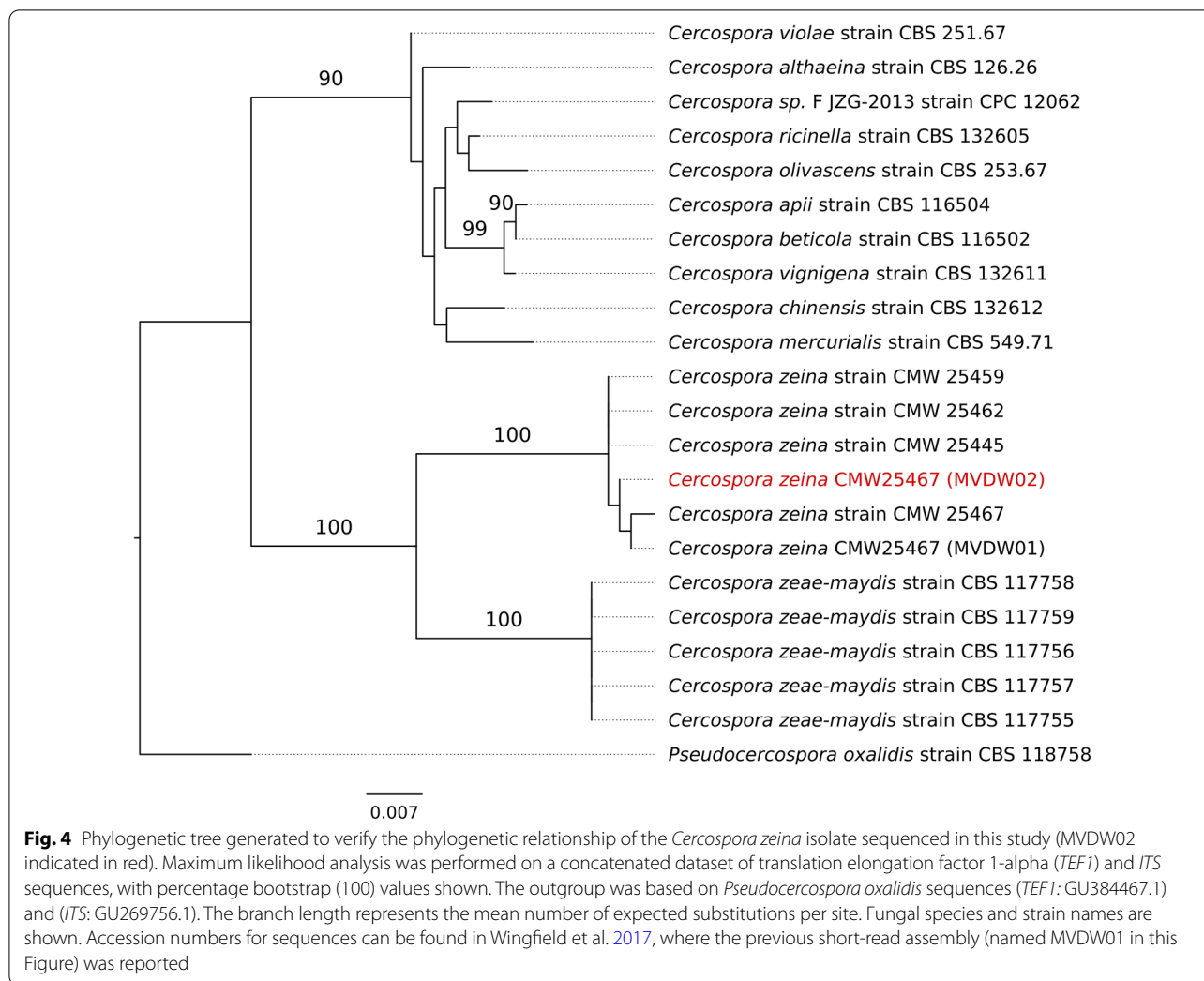
Genome sequences are currently available for six economically important species of *Elsinoë*. These include *E. ampelina* (Haridas et al. 2020; Li et al. 2020), *E. arachidis* (Jiao et al. 2021; Su et al. 2022), *E. australis* (Shanmugam et al. 2020, Zhao et al. 2020), *E. batatas* (Zhang et al. 2022), *E. fawcettii* (Shanmugam et al. 2020; Jeffress et al. 2020) and *E. murrayae* (NCBI; <https://www.ncbi.nlm.nih.gov>). *Elsinoë necatrix* is amongst the more destructive of these pathogens and consequently a threat to commercial forestry in Asia and globally. The availability of its genome sequence will contribute to comparative genomics studies aimed at further understanding the biology of this important but relatively unknown group of pathogens.

#### Sequenced strain

**Indonesia:** North Sumatra: symptomatic leaf of *Eucalyptus* sp., 2020, N.Q. Pham (ex-holotype culture CMW56134 = CBS 147439; PREM 63209—holotype).

#### Sequence accession numbers

The genome sequence of *Elsinoë necatrix* (CMW56134) has been deposited in DDBJ/EMBL/GenBank databases



under the accession number JANZYH000000000. The version described in this paper is JANZYH010000000.

#### Materials and methods

Genomic DNA was extracted from freeze-dried 5 day-old mycelium grown in malt yeast broth (2% malt extract, 0.5% yeast extract; Biolab, Midrand, South Africa) following the method described by Duong et al. (2013). Nanopore sequencing was conducted using the MinION sequencing device. The sequencing library was prepared using the Genomic DNA by Ligation (SQK-LSK109) protocol. The library was loaded on a MinION flow cell (R10.3) and sequencing was run for 48 h. Base calling was conducted using ONT Guppy base calling software v. 4.0.14 (<https://community.nanoporetech.com>). Porechop v. 0.2.1 (<https://github.com/rrwick/Porechop>) was used to remove adapters from the Nanopore reads. Illumina sequencing was carried out by Macrogen I

(Seoul, South Korea), where the paired-end library was constructed and sequenced on NovaSeq 6000 Sequencing System to obtain 151 bp paired-end reads. The quality of the data obtained was assessed using the software FastQC v. 0.11.5 (Afgan et al. 2016). Trimmomatic v. 0.38 (Bolger et al. 2014) was used to remove poor quality data and the remaining Illumina adapters.

The genome was assembled with Nanopore data using Flye v. 2.7 (Kolmogorov et al. 2019) followed by polishing with raw nanopore reads using Racon v. 1.4.13 (Vaser et al. 2017) and Medaka v 1.0.3 (<https://github.com/nanoporetech/medaka>). To further improve the accuracy of the Nanopore assembly, two rounds of polishing with Illumina data were carried out using Pilon (Walker et al. 2014), where trimmed Illumina reads were aligned to the long-read contigs to generate a bam file used as input to polish the assembly. Protein coding gene models were annotated using the fungal version of GeneMark-ES

(Ter-Hovhannisyan et al. 2008). The assembled genome completeness was evaluated using BUSCO v. 5.1.2 by using the Dothideomycetes dataset (Manni et al. 2021).

To validate the identity of the isolate, the internal transcribed spacer (ITS) region, the nuclear large subunit (LSU), part of the DNA-directed RNA polymerase II second largest subunit (*RPB2*) and the partial translation elongation factor 1- $\alpha$  gene (*TEF1*) regions were extracted from the assembly and analyzed together with references sequences of *E. necatrix* and other *Elsinoë* species obtained from GenBank. The sequence data set was aligned using the online version of MAFFT v. 7 (<http://mafft.cbrc.jp/alignment/server/>), (Katoh and Standley 2013). Phylogenetic analysis using maximum likelihood (ML) was performed with RaxML v. 8.2.4 on the CIPRES Science Gateway v. 3.3 (Stamatakis 2014) with GTR substitution model and 1,000 rapid bootstraps.

### Results and discussion

Nanopore sequencing generated 5.66 Gb data with read N50 value of 2.07 kb. Illumina sequencing generated 22.8 million paired-end 151 bp reads. The final assembly of *E. necatrix* (isolate CMW56134) consisted of 69 contigs, with the N50 of 0.73 Mb and L50 of 10. Phylogenetic analysis of four regions (ITS, LSU, *RPB2*, *TEF1*) confirmed the taxonomic identity of the isolate as *E. necatrix* (Fig. 5). The assembled genome of *E. necatrix* was approximately 24.07 Mb with a GC content of 51.59%. The BUSCO completeness of the genome was estimated to be 93.42%: of the 3786 Dothideomycetes BUSCO groups searched, 53 BUSCO orthologs were reported to be fragmented, and 196 BUSCO groups were missing. GeneMark-ES predicted 9180 protein coding gene models in the assembled genome. The genome size and gene

number for *E. necatrix* is relatively similar to that of other *Elsinoë* spp. (Table 1).

The draft genome sequence of *E. necatrix* generated here will facilitate future research regarding the biology and pathogenicity of this fungus. In particular, the genome sequence will be useful for developing molecular markers for population genetic studies to determine its origin and pathways of movement. This will have implications for the management of the disease and contribute towards a better understanding of the growing disease threats (Wingfield et al. 2013, 2015a, b) to *Eucalyptus* plantation forestry globally.

**Authors:** Pham NQ, Duong TA\*, Wingfield BD, Barnes I, Durán A, and Wingfield MJ.

\*Contact: Tuan.Duong@fabi.up.ac.za.

## IMA GENOME-F 17E

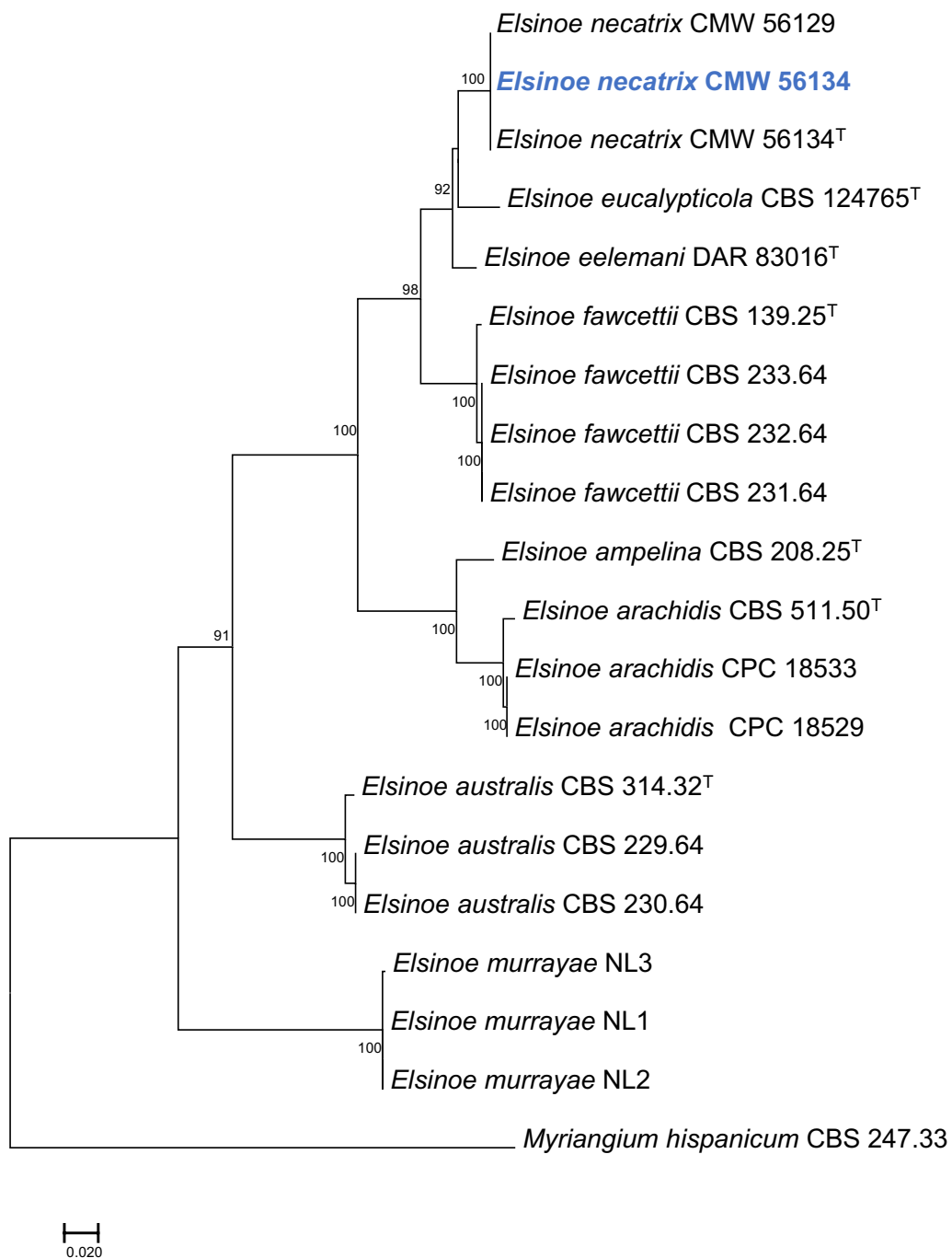
### Short-read genome assemblies and annotations of four *Pyrenophora teres* isolates collected from barley grass

#### Introduction

*Pyrenophora teres* (syn. *Drechslera teres*) is the causative fungus of net blotch disease in barley (*Hordeum vulgare*). Yield loss due to net blotch in susceptible barley cultivars can range from 10 to 70% (Jayasena et al. 2007; Wallwork et al. 2016) making *P. teres* one of the most important fungal pathogens of the barley industry world-wide. *Pyrenophora teres* exists as two forms; *Pyrenophora teres* f. *teres* (Pitt) and *Pyrenophora teres* f. *maculata* (Ptm), causing net-form net blotch and spot-form net blotch in barley, respectively (Liu et al. 2011). Leaf symptoms due to *Pyrenophora teres* f. *teres* appear as dark-brown net-like transverse and longitudinal necrotic striations, while symptoms due to *P. teres* f. *maculata* develop as

**Table 1** Genome assembly features of *Elsinoe* spp.

Species	Strain	Accession number	Genome size (Mb)	Gene number	References
<i>Elsinoe ampelina</i>	YL-1	SMYM000000000	28.30	8057	Li et al. (2020)
	CECT 201119	JAAEIW000000000	28.27	10,207	Haridas et al. (2020)
<i>Elsinoe arachidis</i>	LNFT-H01	JAAPAX000000000	33.18	9174	Jiao et al. (2021)
	LY-HS-1	GWHBFXO00000000	32.44	9435	Su et al. (2022)
<i>Elsinoe australis</i>	Ea-1	SWCS000000000	23.79	9002	Shanmugam et al. (2020)
	NL1	NHZQ000000000	23.34	9223	Zhao et al. (2020)
<i>Elsinoe batatas</i>	CRI-CJ2	JAESVG000000000	26.49	9521	Zhang et al. (2022)
<i>Elsinoe fawcettii</i>	SM16-1	VAAB000000000	26.65	10,340	Shanmugam et al. (2020)
	DAR-70024	SWCR000000000	26.32	9930	Shanmugam et al. (2020)
	BRIP 53147a	SDJM000000000	26.01	10,080	Jeffress et al. (2020)
<i>Elsinoe murrayae</i>	CQ-2017a	NKHZ000000000	20.72	8281	NCBI
<i>Elsinoe necatrix</i>	CMW56134	JANZYH000000000	24.07	9180	This study



**Fig. 5** Maximum Likelihood tree based on ITS, LSU, *RPB2* and *TEF1* sequences for *Elsinoë* spp. Bootstrap values  $\geq 90\%$  for ML analyses are indicated at the nodes. Isolates representing ex-type material are marked with "T". *Myriangium hispanicum* (CBS 247.33) represents the outgroup. The genome sequenced in this study, of which the sequences were extracted from the assembly, is indicated in blue

dark-brown circular to elliptic lesions on susceptible barley cultivars (Smedegård-Petersen 1971).

In addition to barley, *P. teres* can be found on weeds such as barley grass (*Hordeum leporinum*, subspecies of *H. murinum*) and other Gramineae crops such as wheat

(*Triticum aestivum*) and oat (*Avena sativa*) (Shipton 1966; McLean et al. 2009; Khan and Boyd 1969). Even though *P. teres* has been reported to infect both cultivated and weed-like barley, host specificity of the pathogen is controversial (Linde and Smith 2019). Some studies



reported that *P. teres* can infect host plant species belonging to the *Hordeum* family without being specific to any species in the family (Bakke 1912; Braverman 1960; Kenneth 1962), while other studies reported strict host specificity of *P. teres* to its host species (Khan 1973; Linde and Smith 2019). Either way, ancillary hosts like barley grass growing alongside cultivated barley can act as a source of inoculum for *P. teres* (MacNish 1964; Shipton 1966) and may play an important role in the evolution of this pathogen (Linde et al. 2016). Whole genome sequence data for *P. teres* isolates associated with weed-like barley hosts will contribute to better the understanding of the molecular mechanisms underlying its host association.

### Sequenced strains

**Australia:** *Victoria:* Curyo, isolated from *Hordeum leporinum*, 2014, J. Fanning [Ptm14015] (BRIP 71574); Joel South, isolated from *Hordeum leporinum*, 2012, M. Mclean [Ptt12013] (BRIP 71573). *Queensland:* Yelarbon, isolated from *Hordeum leporinum*, 2010, R. Fowler HRS10128 (BRIP 71572). *Western Australia:* Mt Barker, isolated from *Hordeum leporinum*, 1995, Rob Loughman [SNB172] (BRIP 74832).

### Nucleotide sequence accession number

This whole-genome shotgun project was deposited in the NCBI GenBank database under accession numbers JAMGBO000000000 (*Ptm14015*), JAMGBN000000000 (HRS10128), JAMGBM000000000 (*Ptt12013*) and JAMGBL000000000 (SNB172) [BioProject: PRJNA838266 and BioSamples: SAMN28416484 (*Ptm14015*), SAMN28416485 (HRS10128), SAMN28416486 (*Ptt12013*) and SAMN28416487 (SNB172)]. This paper describes the first versions of the four genomes.

### Materials and methods

Genomic DNA of four barley grass isolates, *Ptm14015*, HRS10128, *Ptt12013* and SNB172, was extracted from 14 to 20 day old fungal mycelium grown on half strength potato dextrose agar (PDA) medium [20 g/litre; Biolab Merck Darmstadt, Germany] using a Wizard<sup>®</sup> Genomic DNA Purification kit (Promega, Sydney, Australia) as per the manufacturer's protocol. The integrity of the DNA was assessed under ultraviolet light (Fusion FX, VILBER, Marne-la-Vallée, France) after agarose gel electrophoresis. DNA quantity and quality was measured with a NanoDrop<sup>™</sup>TC 2000/2000c and a NanoPhotometer P300 spectrophotometer (IMPLEN, Munich, Germany). The shotgun DNA libraries of *Ptm14015*, HRS10128 and *Ptt12013* were constructed by Australian Genome

Research Facility (AGRE, Melbourne) with 125 bp pair-end reads using TruSeq Nano library preparation kit. The shotgun DNA libraries of SNB172 was constructed with Nextera DNA XT library preparation kit by Macrogen (Seoul, South Korea) as per the Illumina short read protocol and sequenced with 150 bp pair-end reads on Illumina HiSeq 2000. The sequence quality of the pair-end reads of four genomes were examined by FastQC v0.11.8 (Andrews 2010). The adapter sequences of pair-end reads were trimmed by Trimmomatic v0.39 (Bolger et al. 2014). Sequences shorter than 40 bp and sequences with average phred score lower than 33 bp were also removed. The sequence quality of trimmed sequences was examined using FastQC. Trimmed sequences were used to perform *de novo* whole genome assembly with SPAdes v3.15.2 (Bankevich et al. 2012) by adjusting k-mer size from 20 to 40 bp. The quality and the completeness of all four genome assemblies were assessed using QUAST v5.0.2 (Gurevich et al. 2013) and BUSCO v4.1.2 (Simão et al. 2015) with the fungi\_odb10 database.

Repeat elements present in the genome assemblies were detected by RepeatModeler v1.0.11 (Smit and Hubley 2008), using Repbase v20.4 library (Bao et al. 2015) and masked using RepeatMasker v4.0.9 (Nishimura 2000). Gene prediction was conducted using BRAKER2 v2.1.6 (Hoff et al. 2019) genome annotation pipeline, using the protein sequences of *P. teres* f. *teres* 0–1 (Ellwood et al. 2010) as protein evidences to train AUGUSTUS (Stanke et al. 2006).

The verification of the sequenced isolates, *Ptm14015*, HRS10128, *Ptt12013* and SNB172 as *Pyrenophora teres* was carried out by multi-locus phylogenetic analysis (Fig. 6) of four loci; internal transcribed spacers and the intervening 5.8S sequence of the nrDNA (ITS), partial large subunit of the nrDNA (LSU), partial glyceraldehyde-3-phosphate dehydrogenase gene (*gapdh*), and partial DNA-directed RNA polymerase II second largest subunit gene (*rpb2*). The sequences of the four loci of reference *Pyrenophora* strains were obtained from Marin-Felix et al. (2019) and Duong et al. (2021). The sequences of the four loci of the isolates in the current study were extracted from their respective genome assemblies. Multiple sequence alignment of each loci was conducted in MAFFT v7.450 (Katoh and Standley 2013). Four multiple alignments from the four loci were then concatenated in Geneious Prime v2021.2.2 (Kearse et al. 2012). A maximum likelihood phylogram of *Pyrenophora* species was derived from the concatenated alignment by RAxML v.8 (Stamatakis 2014) using the GTR substitution model with gamma-distribution rate variation for individual partitions and 1000 bootstraps. *Pyrenophora poae* was used as the outgroup for the phylogram (Duong et al. 2021).



**Table 1** Genome assembly statistics of the four *Pyrenophora teres* isolates from barley grass

	<i>Ptm14015</i>	<i>HRS10128</i>	<i>Ptt12013</i>	<i>SNB172</i>
<i>Assembly</i>				
Total assembly size (Mbp)	42.42	42.99	43.03	41.31
Number of contigs ( $\geq 1000$ bp)	1295	956	1225	3384
N <sub>50</sub> (kb)	111.57	331.43	109.55	348.92
Largest contigs (kb)	596.68	1508.41	712.94	323.58
GC-content (%)	46.71	46.57	46.59	47.92
<i>BUSCO analysis</i>				
Completeness	94.5	96.3	96.1	95.5
Complete and single-copy BUSCOs	3564	3630	3628	3604
Complete and duplicated BUSCOs	14	14	10	11
Fragmented BUSCOs	28	26	27	40
Missing BUSCOs	180	116	121	131
<i>Repeat annotation</i>				
DNA repeat elements (%)	1.16	1.24	1.52	2.66
LINEs (%)	0.28	3.20	0.31	1.16
LTR (%)	18.98	17.83	20.70	15.87
Unclassified (%)	6.26	4.95	5.28	6.88
Simple repeats (%)	0.73	0.66	0.73	0.66
Total gene annotation	11,069	11,038	11,068	11,314

### Results and discussion

Illumina paired end (125 and 150 bp) sequencing of *Ptm14015*, *HRS10128*, *Ptt12013* and *SNB172* resulted in around 50 million reads each with  $\sim 100\times$  coverage of the whole genome. The final assemblies of *Ptm14015*, *HRS10128*, *Ptt12013* and *SNB172* genomes included 1295, 956, 1225 and 3384 contigs/scaffolds ( $\geq 1000$ bp) with N<sub>50</sub> values of 111.57, 331.43, 109.55 and 348.92 kb, respectively. The CG contents of *Ptm14015*, *HRS10128*, *Ptt12013* and *SNB172* genomes were 46.71, 46.57, 46.59 and 47.92% and the largest contig size of each assembly was 596.68, 1508.41, 712.94 and 323.58 kb, respectively. The BUSCO completeness of the four assemblies ranged from 94.5 to 96.3% (Table 1). Out of 3786 total BUSCO genes searched, 3564, 3630, 3628 and 3604 BUSCO genes were found from the *Ptm14015*, *HRS10128*, *Ptt12013* and *SNB172* genomes, respectively (Table 1). The high completeness of these four genomes confirmed the high quality of the assemblies.

The compositions of the DNA transposons of the four genomes ranges from 1.16 to 2.66% of the total genome. The long interspersed nuclear element (LINE) composition of the genomes ranged from 0.28 to 1.16% and the long terminal repeat (LTR) retrotransposon ranged from 15.87 to 20.70% of the whole genome. The compositions of DNA transposons, LINEs and LTRs in the current study were similar to those previously reported for *P. teres* (Duong et al. 2021; Syme et al. 2018; Wyatt et al. 2020). The total number of protein-coding genes

generated using BRAKER2 ranged from 11,038 to 11,314 among the four *P. teres* isolates (Table 1), which was greater than the 10,051 protein-coding genes reported for the *Ptt* and *Ptm* hybrid genome (Duong et al. 2021). These four genomes are the first published genomes for *P. teres* collected from barley grass (*Hordeum leporinum*) and will be highly valuable for future comparative studies of *Pyrenophora* species.

**Authors:** Buddhika Amarasinghe Dahanayaka, Barsha Poudel, Niloofar Vaghefi, Kiruba Shankari Arun-Chinnappa, and Anke Martin.

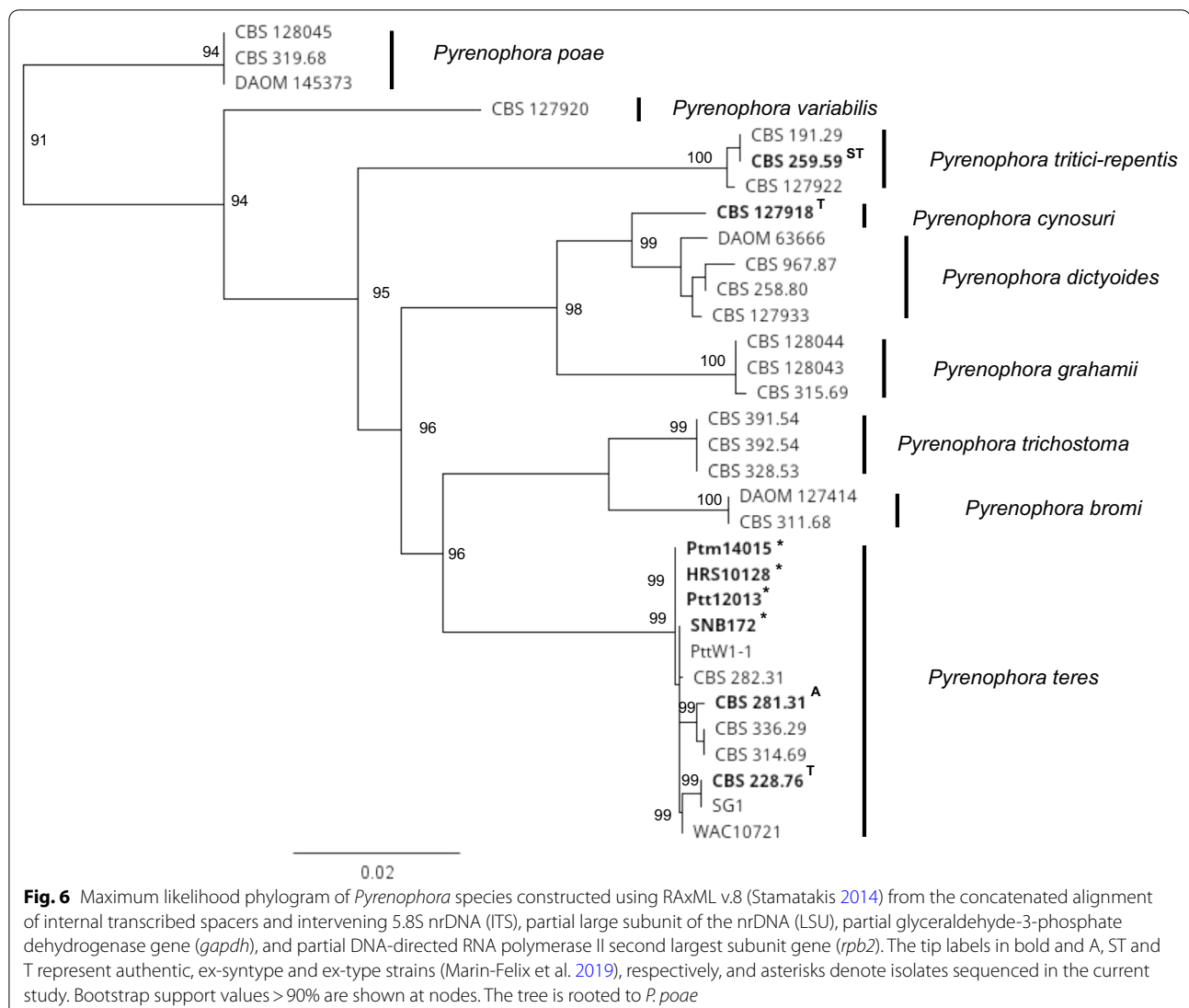
\*Contact: Anke.Martin@usq.edu.au.

### IMA GENOME-F 17F

#### Draft genome sequences for two isolates of the plant pathogen *Sclerotinia minor*

##### Introduction

*Sclerotinia minor* is a fungal pathogen that infects numerous plant hosts, including several crop species of economic importance (Melzer et al. 1997). It is closely related to the broad host-range pathogen *S. sclerotiorum* (Holst-Jensen et al. 1998), and these two species are often studied together due to their overlapping host ranges and similarities in both life cycles and mating systems (Subbarao 1998; Wu et al. 2008; Chitrampalam et al. 2013; Chitrampalam and Pryor 2015). Genome data for *S. minor* would be a useful resource for future research endeavours that could ultimately assist with managing this agricultural pathogen.



Limited genomic data is already available for *S. minor*. A high-quality reference genome sequence was produced for a single isolate isolated from lettuce in the Hubei province of China (Yang et al. 2022). This reference genome is 39 279 639 bp in length and consists of 24 scaffolds, with a GC content of 41.91%, which is similar to that of *S. sclerotiorum* (Amselem et al. 2011). Additionally, an unassembled genomic dataset generated by Illumina HiSeq 2500 sequencing for a single isolate of *S. minor*, isolate SsChi (Curtin University 2019), is available on the SRA database of the NCBI platform (Curtin University 2019). This isolate was obtained from common chicory (*Cichorium intybus*) in 1960, and the availability of a draft assembly for this isolate would be a useful addition to the current genomic resources for this pathogen.

In this genome announcement, the sequencing and assembly of a draft genome for a single isolate of *S. minor*

is presented. This draft assembly is presented together with that of the isolate SsChi which is currently unassembled, and both these genomes were submitted to a public database. These genomes contribute to the available genomic resources for *S. minor* and other plant pathogens in the genus *Sclerotinia*.

#### Sequenced strains

**Italy:** isolated from *Lactuca sativa*, *G. Goidánich* (CBS 339.39, NBRC 6767, PREM 63312-dried culture).

#### Nucleotide sequence accession number

The Whole Genome Shotgun Project has been deposited at DDBJ/ENA/GenBank under the Accession Number JAMLG000000000. The version described in this paper is JAMLG010000000.

### Materials and methods

*Sclerotinia minor* isolate CBS339.39 was obtained from the Westerdijk Fungal Biodiversity Institute and maintained on 2% PDA-ST media [20 g/L potato dextrose agar (Biolab, Merk, South Africa) supplemented with 150 mg/L streptomycin and 100 mg/L thiamine (Sigma, Steinheim, Germany)] at 25 °C for the duration of the study. For DNA isolation, the isolate was grown in a glass bottle containing 15 ml of 2% PDB medium (20 g/L potato dextrose broth) on an electronic shaker for 1–3 days at 25 °C. Mycelia were harvested by centrifugation in 50 mL centrifuge tubes at 4 °C for 10 min at 5000 rpm. The mycelia was freeze-dried before being subjected to genomic DNA extraction using a previously published method (Murray and Thompson 1980).

Genomic DNA was submitted to Macrogen (South Korea) for whole-genome sequencing. A TruSeq Nano library preparation was used for library construction with a 350 bp insert size. Sequencing was carried out on the NovaSeq platform, with target read lengths of 151 bp. The raw reads were imported as individual forward and reverse libraries into the Galaxy online platform (<https://usegalaxy.org>) (Afgan et al. 2018), and were subjected to a FastQC (version: 0.11.8) analysis to assess read quality. SPAdes (version: 3.12.0) was used for a *de novo* assembly of the draft genome sequence (Bankevich et al. 2012) using single-cell mode, k-mer options of 21, 33 and 55, and activating the careful correction option to minimize the number of short indels and mismatches. QUAST (version: 5.0.2) was used to assess the quality of the genome assembly and to determine the general statistics of the genomes (including genome size, GC content, N50, L50, number of contigs, largest contig size, and the average number of mismatches per 100 kbp) (Mikheenko et al. 2018). The BUSCO pipeline (version: 5.2.2) implemented in Galaxy was used to perform a quantitative assessment of the genome assembly completeness based on the fungi\_odb10 dataset (Simão et al. 2015; Manni et al. 2021).

For a comparison, a second *S. minor* genome was also assembled. A single pair-end library (accession number SRX5407461) was obtained from the sequence read archive (SRA) database on NCBI (<https://www.ncbi.nlm.nih.gov/sra>) (Cochrane et al. 2011). This sequencing library was deposited by Curtin University in 2019 using the Illumina HiSeq 2500 sequencing platform (Curtin University 2019), and was derived from isolate SsChi that was isolated by F. Mujica from common chicory (*Cichorium intybus*) in 1960. The 33,971,728 reads of 125 bp average length were used in a stand-alone SPAdes assembly (version: 3.14) using “careful correction” and automated k-mer selection. The resulting assembly was again

**Table 1** The main metrics of the three available genome sequences of *Sclerotinia minor*

	CBS 339.39 <sup>a</sup>	SsChi <sup>b</sup>	LC41 <sup>c</sup>
<i>General genome statistics</i>			
Total genome length (bp)	37,906,295	36,885,191	39,279,639
Number of contigs (≥ 500 bp)	1908	2091	151 contigs / 24 scaffolds <sup>d</sup>
Total length (in contigs ≥ 500 bp)	35,755,541	35,701,198	39,279,639
Coverage	142x	115x	134–645x <sup>e</sup>
Largest contig size (bp)	202,369	150,738	4,022,415
GC (%)	41.82	41.80	41.91
N50	39,276	34,541	443,861
L50	276	319	7
Number of N's per 100 kbp	32.27	36.89	60.86
<i>BUSCO completeness statistics</i>			
Overall completeness %	99.5	99.5	95.1
Total BUSCO terms	758	758	758
Single copy terms	754	754	718
Fragmented terms	1	1	3
Missing terms	3	3	34

<sup>a</sup> Isolate sequenced and assembled in this study

<sup>b</sup> Genome assembled from publicly available genomic data

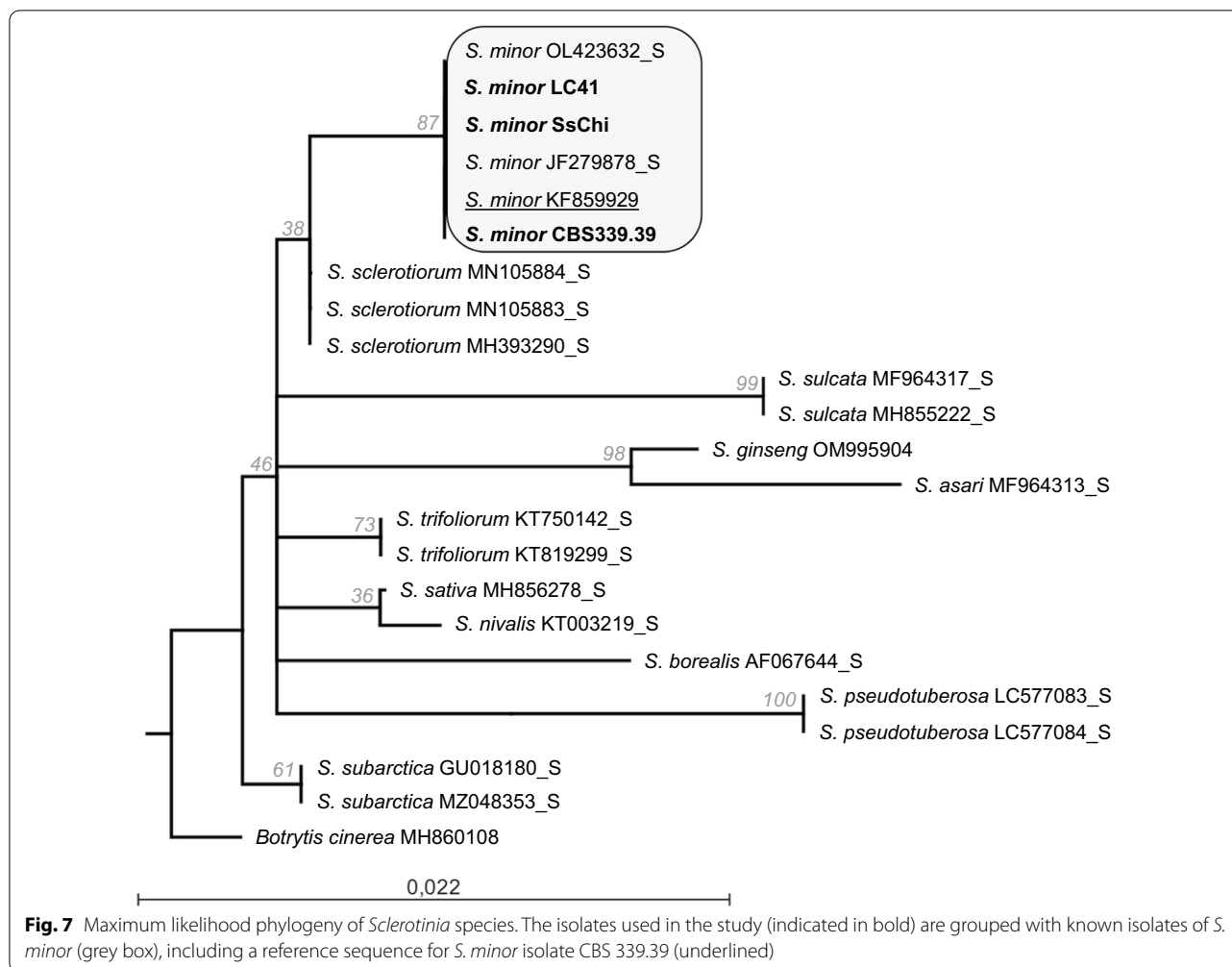
<sup>c</sup> Previously published genome sequence (Yang et al. 2022)

<sup>d</sup> Assembled contigs were scaffolded using a published genome of *S. sclerotiorum* (Derbyshire et al. 2017)

<sup>e</sup> Coverage was reported for PacBio and MGISEQ-2000 sequencing, respectively

submitted to QUAST and BUSCO analyses to determine genome statistics and completeness, respectively.

To confirm the identity of the two *S. minor* draft genomes presented in this announcement, the partial small subunit ribosomal RNA gene, complete internal transcribed spacer 1 and 5.8S ribosomal RNA gene, and partial internal transcribed spacer 2 (ITS region) were extracted from the two assembled genomes (CBS 339.39 and SsChi) as well as from the LC41 reference genome (Yang et al. 2022). The sequences were added to a custom dataset of publicly available ITS sequences representing known *Sclerotinia* species, as well as a single sequence of *Botrytis cinerea* as outgroup. The dataset was aligned using MUSCLE (Edgar 2004) and curated with Gblocks using the Phylogeny.fr server (Castresana 2000; Dereeper et al. 2008). The curated dataset was used to construct a maximum likelihood tree in CLC Main Workbench (version 22.0.1, Qiagen). Additionally, an *in-silico* PCR amplification targeting the laccase 2 gene (*Lcc2*) was conducted in the CLC Main Workbench program using the “Find Binding Sites and Create Fragments” command. The PCR amplification made use of the laccase-specific primers developed as a diagnostic for *S. minor* (Abd-Elmagid



et al. 2013), and was used to confirm the identity of isolate LC41 (Yang et al. 2022).

### Results and discussion

Sequencing of the genomic DNA for *S. minor* strain CBS 339.39 on the Novaseq platform produced 35,898,334 reads with an average size of 151 bp. FastQC analysis confirmed that all reads were of high quality and were subsequently assembled into a draft genome of 37.9 Mb consisting of 16,992 contigs, 1908 of which were 500 bp or larger (Table 1). In comparison, the sequencing dataset of isolate SsChi obtained from the SRA archive was assembled into a genome of 36.8 Mb, consisting of 12,079 contigs of which 2091 were larger than 500 bp. This sequence was also submitted to the DDBJ/ENA/GenBank database under the Accession Number JAMLFZ000000000, with the specific version described here as JAMLFZ010000000. Despite the discrepancy in estimated genome size, both isolates had an identical GC

content of 41.8% and a BUSCO completeness score of 99.5% (Table 1).

Although the ITS region does not provide sufficient phylogenetic signal to distinguish *Sclerotinia* species (Holst-Jensen et al. 1998), the maximum likelihood tree produced from this region did produce a single clade containing all included isolates of *S. minor*, confirming the identity of the sequenced isolates (Fig. 7). The alignment of the ITS region for all the included *S. minor* isolates also showed 100% sequence identity (data not shown). The *in silico* PCR amplification predicted a 257 bp amplicon from both genome sequences using the *S. minor* diagnostic PCR primers, in line with the predicted amplicon size of 264 bp (Abd-Elmagid et al. 2013).

A previously published reference genome currently serves as the reference sequence for *S. minor* (Yang et al. 2022). This genome was from a hybrid assembly of long-read and short-read sequences that overall produces genome sequences of higher contiguity. The contigs

were further scaffolded using the previously published genome sequence of the closely-related species *S. sclerotiorum* (Derbyshire et al. 2017) as reference. When compared to the two genome sequences presented here, the genome size of the LC41 reference strain is 1.3 Mb larger. This variation might have been caused by the choice of sequencing platform or assembly algorithm, or could reflect errors in the assembly and scaffolding process (Nagarajan and Pop 2013). Another interesting possibility is that these variations might reflect biological differences between the strains. The three isolates vary both in geographic origin (Italy, Chile and China) and host species (lettuce and chicory). Although some work has been done to unravel the genetic determinants of host specificity and population diversity in *S. sclerotiorum* (Aldrich-Wolfé et al. 2015; Derbyshire et al. 2022), this has not been echoed in *S. minor*. The availability of three genome sequences provides an exciting opportunity for future studies investigating the genomic basis of host specificity and pathogenicity for this important agricultural pathogen.

**Authors:** Chanel Thomas, Sikelela Buthelezi, Brenda D. Wingfield, and P. Markus Wilken.

\*Contact: Markus.Wilken@fabu.up.ac.za.

## IMA GENOME-F 17G

### Draft genome sequence of *Rosellinia necatrix* from avocado in South Africa

#### Introduction

*Rosellinia necatrix* is an ascomycete pathogen of plants in tropical, subtropical and temperate regions (Sivanesan and Holliday 1972; Petrini 1992). This fungus causes white root rot on many important crops, including apple, avocado, mango, and pear. *Rosellinia necatrix* invades the roots and crown which leads to the collapse of conducting vessels and eventually results in wilt and sudden death of infected plants (Pliego et al. 2009). Due to its resistance to chemical treatments, control and eradication of *R. necatrix* have become notoriously difficult (Pasini et al. 2016).

In South Africa, *R. necatrix* has been detected on apple and pear since the 1970s (Van der Merwe and Matthee 1974). In 2018, it was first reported to cause tree decline in avocado in the country (van den Berg et al. 2018). Since then, it has been spreading rapidly and is now present in all three major avocado-producing provinces of South Africa (Hartley et al. 2022). Nothing is known regarding the sources of *R. necatrix* introduction into South Africa nor its genetic diversity in the country. In this study, we report the genome sequence of a *R. necatrix* strain isolated from avocado in South Africa. This genomic resource, together with genome sequences of the species from other hosts

and locations (Shimizu et al. 2018); <https://www.ncbi.nlm.nih.gov/bioproject/727191>), will facilitate comparative genomic studies and the development of genetic markers to study this important plant pathogen.

#### Sequenced strain

**South Africa:** Limpopo, isol. *Persea americana*, 2016, J. Engelbrecht (CMW50482, PREM 63335-dried culture).

#### Nucleotide sequence accession number

The Whole Genome Shotgun Project has been deposited at DDBJ/ENA/GenBank under the BioProject number PRJNA884201.

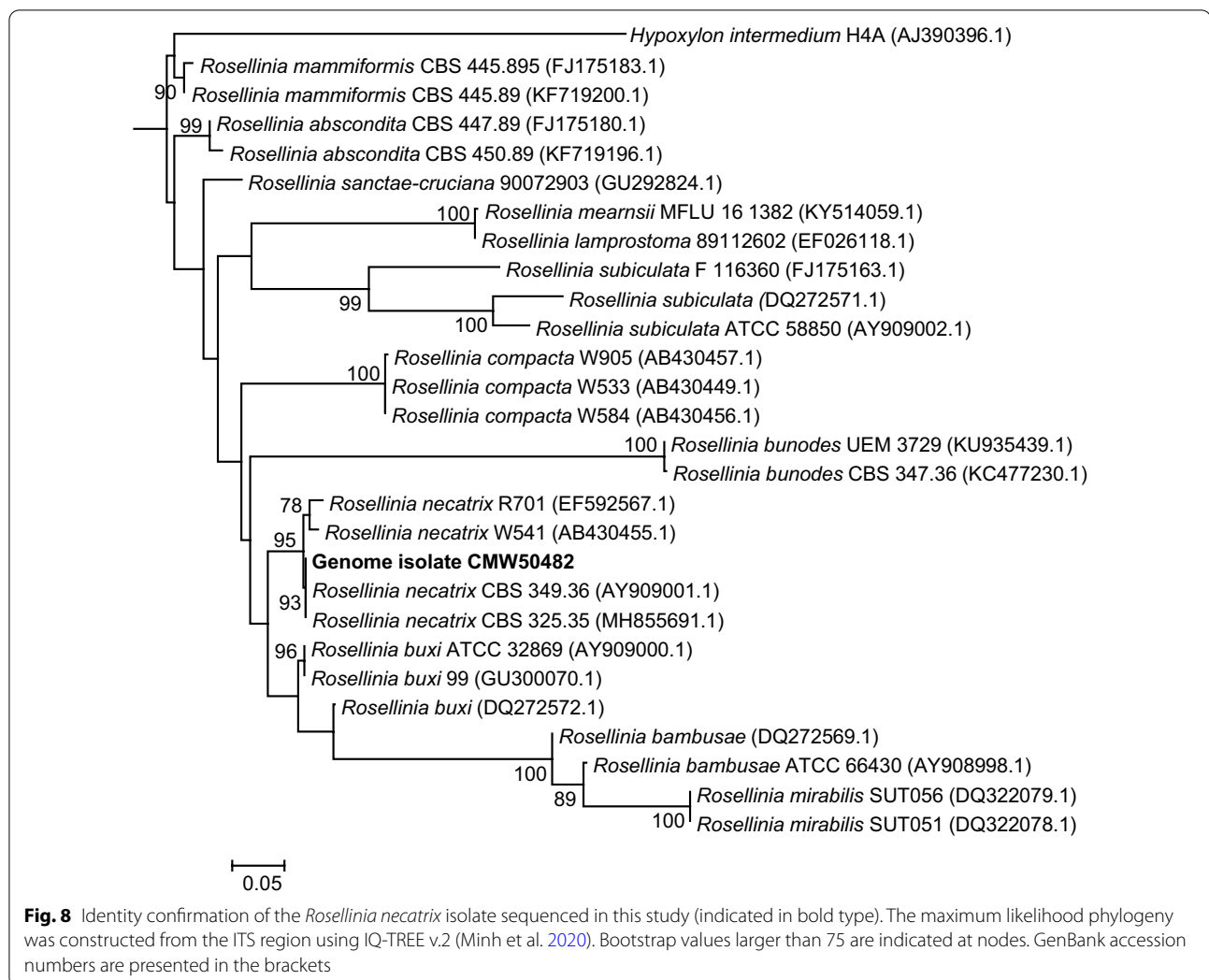
#### Materials and methods

Genomic DNA was extracted from a single hyphal culture of *R. necatrix* CMW50482 using the salt-based DNA extraction protocol of Aljanabi and Martinez (1997) with modifications (Duong et al. 2013). DNA was sent to Macrogen (Seoul, Korea) where a TruSeq PCR free library was constructed and sequenced on the HiSeq 2500 platform to generate paired-end reads of 251 bp. The sequence data were quality filtered and assembled in CLC Genomics Workbench v.22.0.1 (QIAGEN, Aarhus, Denmark). Quality filter parameters included a minimum quality limit of 0.02, removal of ambiguous nucleotides, removal of homopolymers at the 5' and 3' end when 9 out of 10 in a window are homopolymers and final reads had to be a minimum of 100 nucleotides. A *de novo* genome assembly was performed with default parameters, but with a bubble size of 100 and a word size of 40. Reads were mapped back to the contigs during assembly, which also updates the assembled contigs based on mapping information. The Genome Finishing Module v.22 was used to extend the contigs, using a minimum coverage cut-off of 10 reads and a maximum unaligned ends coverage cut-off of 30%. Contigs with a minimum length of 1 kb were selected to be included in the final assembly. Genome completeness was assessed based on the set of 1706 conserved Ascomycete single copy orthologs using BUSCO v.5.3.2 (Manni et al. 2021). AUGUSTUS v. 3.2.3 (Stanke et al. 2006) was used to annotate protein coding genes using the species model for *Fusarium graminearum*. To validate the identity of the isolate that was used for sequencing, the ITS gene region was extracted from the assembly and phylogenetic analysis was conducted with a reference dataset of *R. necatrix* and closely related species obtained from Hartley et al. (2022).

#### Results and discussion

A total of 25.6 million paired-end reads were generated, of which 24.6 million paired reads with an average read





length of 223 bp remained after trimming and quality filtering. The genome assembly resulted in 1362 contigs that were larger than 1000 bp with a maximum contig length of 560 kb and an N50 value of 124 kb, and thus the genome was sequenced with 65× coverage. The assembled genome size was 48.79 Mb, with a GC content of 46.25%. Phylogenetic analysis using the ITS gene region extracted from the assembly confirmed the taxonomic identity of the sequenced isolate as *R. necatrix* (Fig. 8). The assembly size of *R. necatrix* in this study is larger than that of a previous sequenced isolate from pear in Japan (44 Mb: Shimizu et al. 2018), but in the same range with those from roses (48.7–49.2 Mb: <https://www.ncbi.nlm.nih.gov/bioproject/727191>), indicating the great variability in genomic composition of the species. BUSCO analysis indicated that the assembled genome had a 96.2% completeness score (0.1% duplicated and 3.8% missing). AUGUSTUS predicted a total of 10 714 protein-coding genes from the genome assembly. The genomic resource presented here represents

the first *R. necatrix* genome from an isolate associated with avocado (*Persea americana*) and will contribute to a better understanding of the genetic mechanisms underlying pathogenicity and virulence of the species. This genome resource will also be useful for the development of molecular markers for population studies of this important species.

**Authors:** Juanita Engelbrecht, Arista Fourie, Jesse Hartley, Gerda Fourie, Noelani van den Berg.  
\*Contact: noelani.vandenberg@up.ac.za.

#### Funding

*Cercospora zeina* genome sequencing was carried out at the Max Planck Genome-centre Cologne (MP-GC), Germany, and was funded by the National Research Foundation, South Africa, grants 85782, 98977 and 118503. The authors of the *Rosellinia necatrix* genome announcement acknowledge funding from the Hans Merensky Chair in Avocado Research. The authors of the *Ceratocystis colombiana* and *Sclerotinia minor* genome announcements acknowledge funding received from the Department of Science and Technology (DSI)/National Research Foundation (NRF) Centre of Excellence in Plant Health Biotechnology (CPHB), South Africa, and the DSI-NRF SARCH chair in Fungal Genomics.

**Availability of data and materials**

The datasets generated during the current study are available in the NCBI repository, <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA355276/>.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

All authors read and approved the final manuscript.

**Competing interests**

Brenda D. Wingfield is a Senior Editor for *IMA Fungus*. The authors declare that they have no other competing interests.

**Adherence to national and international regulations**

Not applicable.

**Author details**

<sup>1</sup>Department of Biochemistry, Genetics and Microbiology, Faculty of Natural and Agricultural Sciences, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa. <sup>2</sup>Department of Plant and Soil Sciences, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0028, South Africa. <sup>3</sup>Centre for Crop Health, University of Southern Queensland, Toowoomba, QLD 4350, Australia. <sup>4</sup>PerkinElmer Pty Ltd., Level 2, Building 5, Brandon Business Park, 530-540, Springvale Road, Glen Waverley, VIC 3150, Australia. <sup>5</sup>Plant Health Program, Research and Development, Asia Pacific Resources International Holdings Ltd. (APRIL), Pangkalan Kerinci, Riau 28300, Indonesia. <sup>6</sup>Environmental Genomics, Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany. <sup>7</sup>Environmental Genomics, Christian-Albrechts University of Kiel, 24118 Kiel, Germany. <sup>8</sup>Department of Biochemistry, Genetics and Microbiology, Centre for Bioinformatics and Computational Biology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa. <sup>9</sup>School of Agriculture and Food, University of Melbourne, Parkville, VIC 3010, Australia. <sup>10</sup>CSIR, Food Research Institute, Accra, Ghana.

Accepted: 30 September 2022

Published online: 21 November 2022

**References**

- Abd-Elmagid A, Garrido PA, Hunger R, Lyles JL, Mansfield MA et al. (2013) Discriminatory simplex and multiplex PCR for four species of the genus *Sclerotinia*. *J Microbiol Methods* 92:293–300
- Afgan E, Baker D, Van den Beek M, Blankenberg D, Bouvier D et al. (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res* 44:W3–W10
- Afgan E, Baker D, Batut B, van den Beek M, Bouvier D et al. (2018) The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 46:W537–W544
- Akulova VS, Sharov VV, Aksyonova AI, Putintseva YA, Oreshkova, et al. (2020) *De novo* sequencing, assembly and functional annotation of *Armillaria borealis* genome. *BMC Genomics* 21:534
- Al Adawi AO, Barnes I, Khan IA, Al Subhi AM, Al Jahwari AA, Deadman ML, Wingfield BD, Wingfield MJ (2013) *Ceratocystis manginecans* associated with a serious wilt disease of two native legume trees in Oman and Pakistan. *Australas Plant Pathol* 42:179–193
- Aldrich-Wolfe L, Travers S, Nelson BD (2015) Genetic variation of *Sclerotinia sclerotiorum* from multiple crops in the North Central United States. *PLoS ONE* 10:e0139188
- Aljanabi SM, Martinez I (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res* 25:4692–4693
- Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF (2006) A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat Protoc* 1:2320–2325
- Amselem J, Cuomo CA, Van Kan JAL, Viaud M, Benito EP et al. (2011) Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *PLoS Genet* 7:e1002230
- Anderson JB, Bruhn JN, Kasimer D, Wang H, Rodrigue N et al. (2018) Clonal evolution and genome stability in a 2500-year-old fungal individual. *Proc R Soc B Biol Sci* 285:20182233
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Accessed 17 May 2018
- Bakke AL (1912) The late blight of barley (*Helminthosporium teres* Sacc.). In: Proceedings of the Iowa academy of science, vol 1. pp 93–102
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M et al. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477
- Bao W, Kojima KK, Kohany O (2015) Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob DNA* 6:1–6
- Barnes I, Fourie A, Wingfield MJ, Harrington TC, McNew DL et al. (2018) New *Ceratocystis* species associated with rapid death of *Metrosideros polymorpha* in Hawai'i. *Persoonia* 40:154–181
- Baumgartner K, Rizzo DM (2001) Ecology of *Armillaria* spp. in mixed-hardwood forests of California. *Plant Dis* 85:947–951
- Baumgartner K, Rizzo DM (2002) Spread of *Armillaria* root disease in a California vineyard. *Am J Enol Vitic* 53:197–203
- Baumgartner K, Coetzee MPA, Hoffmeister D (2011) Secrets of the subterranean pathosystem of *Armillaria*. *Mol Plant Pathol* 12:515–534
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120
- Braverman S (1960) The *Helminthosporium-gramineum* complex and related species on cereals and forage grasses. *Phytopathology* 50:688–691
- Caballero JRI, Lalande BM, Hanna JW, Klopfenstein NB, Kim MS et al. (2022) Genomic comparisons of two *Armillaria* species with different ecological behaviors and their associated soil microbial communities. *Microb Ecol*. <https://doi.org/10.1007/s00248-022-01989-8>
- Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17:540–552
- Chitrampalam P, Pryor BM (2015) Characterization of mating type (MAT) alleles differentiated by a natural inversion in *Sclerotinia minor*. *Plant Pathol* 64:911–920
- Chitrampalam P, Inderbitzin P, Maruthachalam K, Wu B-M, Subbarao KV (2013) The *Sclerotinia sclerotiorum* mating type locus (MAT) contains a 3.6-kb region that is inverted in every meiotic generation. *PLoS ONE* 8:e56895
- Chung KR (2011) *Elsinoë fawcettii* and *Elsinoë australis*: the fungal pathogens causing citrus scab. *Mol Plant Pathol* 12:123–135
- Cochrane G, Karsch-Mizrachi I, Nakamura Y, on behalf of the International Nucleotide Sequence Database C (2011) The international nucleotide sequence database collaboration. *Nucleic Acids Res* 39:D15–D18
- Coetzee MPA, Wingfield BD, Bloomer P, Wingfield MJ (2005) Phylogenetic analyses of DNA sequences reveal species partitions amongst isolates of *Armillaria* from Africa. *Mycol Res* 109:1223–1234
- Coetzee MPA, Bloomer P, Wingfield MJ, Wingfield BD (2011) Paleogene radiation of a plant pathogenic mushroom. *PLoS ONE* 6(12):e28545
- Collins C, Keane TM, Turner DJ, O'Keefe G, Fitzpatrick DA et al. (2013) Genomic and proteomic dissection of the ubiquitous plant pathogen, *Armillaria mellea*: toward a new infection model system. *J Proteome Res* 12:2552–2570
- Collins C, Hurley R, Almutlaqah N, O'Keefe G, Keane TM et al. (2017) Proteomic characterization of *Armillaria mellea* reveals oxidative stress response mechanisms and altered secondary metabolism profiles. *Microorganisms* 5:60
- Crous PW, Groenewald JZ, Groenewald M et al. (2006) Species of *Cercospora* associated with grey leaf spot of maize. *Stud Mycol* 55:189–197
- Curtin University (2019) Genome sequencing of *Sclerotinia minor* <https://www.ncbi.nlm.nih.gov/sra/SRX5407461> [accn]. Accessed 27 Apr 2022
- Dal Molin A, Minio A, Delledonne GF, M, Infantino A, et al. (2018) The genome assembly of the fungal pathogen *Pyrenochaeta lycopersici* from single-molecule real-time sequencing sheds new light on its biological complexity. *PLoS ONE* 13:e0200217
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9:772–772

- de Beer ZW, Duong TA, Barnes I, Wingfield BD, Wingfield MJ (2014) Redefining *Ceratocystis* and allied genera. *Stud Mycol* 79:187–219
- Derbyshire M, Denton-Giles M, Hegedus D, Seifbarghy S, Rollins J et al. (2017) The complete genome sequence of the phytopathogenic fungus *Sclerotinia sclerotiorum* reveals insights into the genome architecture of broad host range pathogens. *Genome Biol Evol* 9:593–618
- Derbyshire MC, Newman TE, Khenry Y, Owolabi Taiwo A (2022) The evolutionary and molecular features of the broad-host-range plant pathogen *Sclerotinia sclerotiorum*. *Molecular plant pathology*. Early View. <https://doi.org/10.1111/mp.13221>
- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 36:W465–W469
- Duong TA, De Beer ZW, Wingfield BD, Wingfield MJ (2013) Characterization of the mating-type genes in *Leptographium procerum* and *Leptographium profanum*. *Fungal Biol* 117:411–421
- Duong TA, Aylward J, Ametrano CG, Poudel B, Santana QC et al. (2021) Draft genome assembly of *Fusarium pilosicola*, *Meredithiella fracta*, *Niebla homalea*, *Pyrenophora teres* hybrid WAC10721, and *Teratosphaeria viscida*. *IMA Fungus* 12:1–17
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
- Ellwood SR, Liu Z, Syme RA, Lai Z, Hane JK et al. (2010) A first genome assembly of the barley fungal pathogen *Pyrenophora teres* f. *teres*. *Genome Biol* 11:R109. <https://doi.org/10.1186/gb-2010-11-11-r109>
- Engelbrecht CJB, Harrington TC, Steimel J, Capretti P (2004) Genetic variation in eastern North American and putatively introduced populations of *Ceratocystis fimbriata* f. *platanii*. *Mol Ecol* 13:2995–3005
- Fan XL, Barreto RW, Groenewald JZ, Bezerra JDP, Pereira OL et al. (2017) Phylogeny and taxonomy of the scab and spot anthracnose fungus *Elsinoë* (Myriangiales, Dothideomycetes). *Stud Mycol* 87:1–41
- Feurtey A, Lorrain C, Croll D, Eschenbrenner C, Freitag M et al. (2020) Genome compartmentalization predates species divergence in the plant pathogen genus *Zymoseptoria*. *BMC Genom* 21(1):588
- Fourie A, de Jonge R, van der Nest MA, Duong TA, Wingfield MJ (2020) Genome comparisons suggest an association between *Ceratocystis* host adaptations and effector clusters in unique transposable element families. *Fungal Genet Biol* 143:103433
- Fulnečková J, Ševčíková T, Fajkus J, Lukešová A, Lukeš M et al. (2013) A broad phylogenetic survey unveils the diversity and evolution of telomeres in eukaryotes. *Genome Biol Evol* 5:468–483
- The Galaxy Community (2022) The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Res*: gkac247. <https://doi.org/10.1093/nar/gkac247>. Online ahead of print
- Gregory SC, Rishbeth J (1991) Pathogenicity and virulence. In: Shaw CG III, Kile GA (eds) *Agriculture handbook (USA)*, vol 691. USDA Forest Service, Washington, District of Columbia, pp 76–87
- Guillaumin J-J, Mohammed C, Anselmi N, Courtecuisse R, Gregory SC (1993) Geographical distribution and ecology of the *Armillaria* species in western Europe. *Eur J for Pathol* 23:321–341
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696–704
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59:307–321
- Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013) QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075
- Halsted BD (1890) Some fungous diseases of the sweet potato. *Bull N J Agric Exp Stn* 76: 3–32
- Haridas S, Albert R, Binder M, Bloem J, LaButti K et al. (2020) 101 Dothideomycetes genomes: a test case for predicting lifestyles and emergence of pathogens. *Stud Mycol* 95:5–169
- Hartley J, Engelbrecht J, van den Berg N (2022) Detection and prevalence of *Rosellinia necatrix* in South African avocado orchards. *Eur J Plant Pathol*. <https://doi.org/10.1007/s10658-022-02532-8>
- Heinzelmann R, Dutech C, Tsykun T, Labbé F, Soularue J-P (2019) Latest advances and future perspectives in *Armillaria* research. *Can J Plant Path* 41:1–23
- Heinzelmann R, Rigling D, Sipos G, Münsterkötter M, Croll D (2020) Chromosomal assembly and analyses of genome-wide recombination rates in the forest pathogenic fungus *Armillaria ostoyae*. *Heredity* 124:699–713
- Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M (2016) BRAKER1: unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics* 32:767–769
- Hoff KJ, Lomsadze A, Borodovsky M, Stanke M (2019) Whole-Genome Annotation with BRAKER. In: Kollmar M (ed) *Gene prediction. Methods in Molecular Biology*, vol 1962. Humana Press, New York, pp 65–95
- Holst-Jensen A, Vaage M, Schumacher T (1998) An approximation to the phylogeny of *Sclerotinia* and related genera. *Nord J Bot* 18:705–719
- Jayasena K, Van Burgel A, Tanaka K, Majewski J, Loughman R (2007) Yield reduction in barley in relation to spot-type net blotch. *Australas Plant Pathol* 36:429–433
- Jeffress S, Arun-Chinnappa K, Stodart B, Vaghefi N, Tan YP et al. (2020) Genome mining of the citrus pathogen *Elsinoë fawcettii*; prediction and prioritisation of candidate effectors, cell wall degrading enzymes and secondary metabolite gene clusters. *PLoS ONE* 15:e0227396
- Jiao W, Xu M, Zhou R, Fu Y, Li Z et al. (2021) Genomic analysis of *Elsinoë arachidis* reveals its potential pathogenic mechanism and the biosynthesis pathway of elsinochrome toxin. *PLoS ONE* 16:e0261487
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version7: improvements in performance and usability. *Mol Biol Evol* 30:772–780
- Katoh K, Rozewicki J, Yamada KD (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 20(4):1160–1166
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M et al. (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649
- Keller O, Kollmar M, Stanke M, Waack S (2011) A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics* 27:757–763
- Kenneth R (1962) On the taxonomy, morphology and geographic origins of *Pyrenophora teres* Drechsler and allied species. *Bull Res Council Isr* 11:55–82
- Khan T (1973) Host specialization by Western Australian isolates causing net blotch symptoms on *Hordeum*. *Trans Br Mycol Soc* 61:215–220
- Khan T, Boyd W (1969) Physiologic specialization in *Drechslera teres*. *Aust J Biol Sci* 22:1229–1236
- Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 12:357–360
- King R, Urban M, Hammond-Kosack MCU, Hassani-Pak K, Hammond-Kosack KE (2015) The completed genome sequence of the pathogenic ascomycete fungus *Fusarium graminearum*. *BMC Genomics* 16:544
- Koch RA, Wilson AW, Séné O, Henkel TW, Aime MC (2017) Resolved phylogeny and biogeography of the root pathogen *Armillaria* and its gasteroid relative *Guyanagaster*. *BMC Evol Biol* 17:33
- Kolesnikova AI, Putintseva YA, Simonov EP, Biriukov VV, Oreshkova NV et al. (2019) Mobile genetic elements explain size variation in the mitochondrial genomes of four closely-related *Armillaria* species. *BMC Genomics* 20:351
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA (2019) Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546
- Korsman J, Meisel B, Kloppers FJ, Crampton BG, Berger DK (2012) Quantitative phenotyping of grey leaf spot disease in maize using real-time PCR. *Eur J Plant Pathol* 133(2):461–471
- Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A (2019) RAXML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35:4453–4455
- Labbé F, Marçais B, Dupouey J-L, Bélouard T, Capdevielle X (2015) Pre-existing forests as sources of pathogens? The emergence of *Armillaria ostoyae* in a recently planted pine forest. *For Ecol Manage* 357:248–258
- Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25
- Legrand P, Ghahari S, Guillaumin J-J (1996) Occurrence of genets of *Armillaria* spp. in four mountain forests in Central France: the colonization strategy of *Armillaria ostoyae*. *New Phytol* 133:321–332
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754–1760

- Li Z, Fan Y, Chang P, Gao L, Wang X (2020) Genome sequence resource for *Elsinoë ampelina*, the causal organism of grapevine anthracnose. *Mol Plant Microbe Interact* 33:576–579
- Li Z, Dos Santos RF, Gao L, Chang P, Wang X (2021) Current status and future prospects of grapevine anthracnose caused by *Elsinoë ampelina*: an important disease in humid grape-growing regions. *Mol Plant Pathol* 22:899–910
- Linde CC, Smith LM (2019) Host specialisation and disparate evolution of *Pyrenophora teres f. teres* on barley and barley grass. *BMC Evol Biol* 19:139–153
- Linde CC, Smith LM, Peakall R (2016) Weeds, as ancillary hosts, pose disproportionate risk for virulent pathogen transfer to crops. *BMC Evol Biol* 16:1–12
- Linnakoski R, Sutela S, Coetzee MPA, Duong TA, Pavlov IN (2021) *Armillaria* root rot fungi host single-stranded RNA viruses. *Sci Rep* 11:7336
- Liu K-J, Xu X-D (2013) First report of gray leaf spot of maize caused by *Cercospora zeina* in China. *Plant Dis* 97:1656–1656
- Liu Z, Ellwood SR, Oliver RP, Friesen TL (2011) *Pyrenophora teres*: profile of an increasingly damaging barley pathogen. *Mol Plant Pathol* 12:1–19
- Liu F, Duong TA, Barnes I, Wingfield MJ, Chen S (2021) Population diversity and genetic structure reveal patterns of host association and anthropogenic impact for the globally important fungal tree pathogen *Ceratocystis manginecans*. *J Fungi* 7:759
- MacNish G (1964) Supplementary list of disease recorded on various hosts in Western Australia. *J Dep Agric West Aust Ser* 45(12):991–995
- Manni M, Berkeley MR, Seppely 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
- Marin-Felix Y, Hernández-Restrepo M, Iturrieta-González I, García D, Gené J et al. (2019) Genera of phytopathogenic fungi: GOPHY 3. *Stud Mycol* 94:1–124
- McLean MS, Howlett BJ, Hollaway GJ (2009) Epidemiology and control of spot form of net blotch (*Pyrenophora teres f. maculata*) of barley: a review. *Crop Pasture Sciences* 60:303–315
- Meisel B, Korsman J, Kloppers FJ, Berger DK (2009) *Cercospora zeina* is the causal agent of grey leaf spot disease of maize in southern Africa. *Eur J Plant Pathol* 124:577–583
- Melzer MS, Smith EA, Boland GJ (1997) Index of plant hosts of *Sclerotinia minor*. *Can J Plant Path* 19:272–280
- Meyer J, Berger DK, Christensen SA, Murray SL (2017) RNA-Seq analysis of resistant and susceptible sub-tropical maize lines reveals a role for kauralexins in resistance to grey leaf spot disease, caused by *Cercospora zeina*. *BMC Plant Biol* 17:197
- Mikheenko A, Pribelski A, Saveliev V, Antipov D, Gurevich A (2018) Versatile genome assembly evaluation with QUASt-LG. *Bioinformatics (Oxford, England)* 34:i142–i150
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD et al. (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 37:1530–1534
- Misiek M, Hoffmeister D (2012) Sesquiterpene aryl ester natural products in North American *Armillaria* species. *Mycol Prog* 11:7–15
- Misiek M, Braesel J, Hoffmeister D (2011) Characterisation of the ArmA adenylation domain implies a more diverse secondary metabolism in the genus *Armillaria*. *Fungal Biol* 115:775–781
- Molano EPL, Cabrera OG, Jose J, do Nascimento LC, Carazzolle MF et al. (2018) *Ceratocystis cacaofunesta* genome analysis reveals a large expansion of extracellular phosphatidylinositol-specific phospholipase-C genes (PI-PLC). *BMC Genomics* 19:58
- Morrison DJ (1991) Identity of *Armillaria* isolates used in studies of rhizomorph behaviour and pathogenicity. *Mycol Res* 95:1437–1438
- Muller MF, Barnes I, Kunene NT, Crampton BG, Bluhm BH et al. (2016) *Cercospora zeina* from maize in South Africa exhibits high genetic diversity and lack of regional population differentiation. *Phytopathology* 106:1194–1205
- Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* 8:4321–4326
- Nagarajan N, Pop M (2013) Sequence assembly demystified. *Nat Rev Genet* 14:157–167
- Neves DL, Silva CN, Pereira CB, Campos HD, Tessmann DJ (2015) *Cercospora zeina* is the main species causing gray leaf spot in southern and central Brazilian maize regions. *Trop Plant Pathology* 40:368–374
- Nishimura D (2000) RepeatMasker. *Biotech Softw Internet Rep* 1:36–39
- Nsibo DL, Barnes I, Kunene NT, Berger DK (2019) Influence of farming practices on the population genetics of the maize pathogen *Cercospora zeina* in South Africa. *Fungal Genetics Biol* 125:36–44
- Nsibo DL, Barnes I, Omondi DO, Dida MM, Berger DK (2021) Population genetic structure and migration patterns of the maize pathogenic fungus, *Cercospora zeina* in East and Southern Africa. *Fungal Genetics Biol* 149(103527):1–14
- Pasini L, Prodrutti D, Pastorelli S, Pertot I (2016) Genetic diversity and bio-control of *Rosellinia necatrix* infecting apple in northern Italy. *Plant Dis* 100:444–452
- Petrini LE (1992) *Rosellinia* species of the temperate zones. *Sydowia* 44:169–281
- Pham NQ, Marincowitz S, Solís M, Duong TA, Wingfield BD et al. (2021) Eucalyptus scab and shoot malformation: a new and serious foliar disease of *Eucalyptus* caused by *Elsinoë necatrix* sp. nov. *Plant Pathol* 70:1230–1242
- Pliego C, Kanematsu S, Ruano-Rosa D, de Vicente A, López-Herrera C et al. (2009) GFP sheds light on the infection process of avocado roots by *Rosellinia necatrix*. *Fungal Genet Biol* 46:137–145
- Plissonneau C, Stürchler A, Croll D (2016) The evolution of avocado regions in genomes of a fungal pathogen of wheat. *Mbio* 7(5):e01231-16
- Prospero S, Holdenrieder O, Rigling D (2004) Comparison of the virulence of *Armillaria cepistipes* and *Armillaria ostoyae* on four Norway spruce provenances. *For Pathol* 34:1–14
- Raciborski M (1900) Parasitische Algen und Pilze Java's I. *Botanischen Institut zu Buitenzorg, Batavia*
- Roux J, Wingfield MJ, Fourie A, Noeth K, Barnes I (2020) *Ceratocystis wilt* on Eucalyptus: first record from South Africa. *Southern Forests* 82:24–31
- Shah RM, Williams AH, Hane JK, Lawrence JA, Farfan-Caceres LM et al. (2020) Reference Genome Assembly for Australian *Ascochyta rabiei* isolate ArME14. *G3 Genes|Genomes|Genetics* 10(7):2131–2140
- Shanmugam G, Jeon J, Hyun JW (2020) Draft genome sequences of *Elsinoë fawcettii* and *Elsinoë australis* causing scab diseases on citrus. *Mol Plant Microbe Interact* 33:135–137
- Shimizu T, Kanematsu S, Yaegashi H (2018) Draft genome sequence and transcriptional analysis of *Rosellinia necatrix* infected with a virulent mycovirus. *Phytopathology* 108:1206–1211
- Shipton WA (1966) Net blotch of barley. *J Dep Agric West Aust Ser* 47(3):132–133
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212
- Sipos G, Prasanna AN, Walter MC, O'Connor E, Bálint B et al. (2017) Genome expansion and lineage-specific genetic innovations in the forest pathogenic fungi *Armillaria*. *Nat Ecol Evol* 1:1931–1941
- Sivanesan A, Holliday P (1972) *Rosellinia necatrix*. In: Kew, Surrey, UK: Commonwealth Mycological Institute, CMI Descriptions of pathogenic fungi and bacteria Set 36. No 352
- Smedegård-Petersen V (1971) *Pyrenophora teres f. maculata f. nov.* and *Pyrenophora teres f. tereson* in Denmark. In: 'Yearbook of the Royal Veterinary and Agricultural University (Copenhagen)', pp 124–144
- Smit AF, Hubley R (2008) RepeatModeler Open-1.0.
- Sonnenbichler J, Guillaumin J-J, Peipp H, Schwarz D (1997) Secondary metabolites from dual cultures of genetically different *Armillaria* isolates. *Eur J For Pathol* 27:241–249
- Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313
- Stanke M, Waack S (2003) Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* 19(suppl\_2):ii215–ii225
- Stanke M, Schöffmann O, Morgenstern B, Waack S (2006) Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinform* 7:1–11
- Stanke M, Diekhans M, Baertsch R, Haussler D (2008) Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* 24:637–644

- Su J, Liu J, Hu Y, Wang Y, Jia Y et al. (2022) High-quality genome sequence resource of *Elsinoë arachidis* strain LY-HS-1, causing scab disease of peanut. *Plant Dis* 106:1506–1509
- Subbarao KV (1998) Progress toward integrated management of lettuce drop. *Plant Dis* 82:1068–1078
- Sun X, Zhang T, Zhao Y, Zhu H, Cai E (2020) Protoilludane sesquiterpenoid aromatic esters from *Armillaria mellea* improve depressive-like behavior induced by chronic unpredictable mild stress in mice. *J Funct Foods* 66:103799
- Swart V, Crampton BG, Ridenour JB, Bluhm BH, Olivier NA et al. (2017) Complementation of *CTB7* in the Maize Pathogen *Cercospora zeina* overcomes the lack of *in vitro* cercosporin production. *Mol Plant Microbe Interact* 30:710–724
- Syme R, Martin A, Wyatt N, Lawrence J, Muria-Gonzalez M (2018) Transposable element genomic fissioning in *Pyrenophora teres* is associated with genome expansion and dynamics of host-pathogen genetic interactions. *Front Genet* 9:130–145
- Tarigan M, Roux J, Van Wyk M, Tjahjono B, Wingfield MJ (2011) A new wilt and die-back disease of *Acacia mangium* associated with *Ceratocystis manginecans* and *C. acaciivora* sp. nov. in Indonesia. *S Afr J Bot* 77:292–304
- Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M (2008) Gene prediction in novel fungal genomes using an *earch* algorithm with unsupervised training. *Genome Res* 18:1979–1990
- Testa AC, Oliver RP, Hane JK (2016) OcculterCut: a comprehensive survey of AT-rich regions in fungal genomes. *Genome Biol Evol* 8:2044–2064
- Tsopelas P, Santini A, Wingfield MJ, Wilhelm de Beer Z (2017) Canker stain: a lethal disease destroying iconic plane trees. *Plant Dis* 101:645–658
- Tsykun T, Rigling D, Nikolaychuk V, Prospero S (2012) Diversity and ecology of *Armillaria* species in virgin forests in the Ukrainian Carpathians. *Mycol Prog* 11:403–414
- van den Berg N, Hartley J, Engelbrecht J, Mufamadi Z, van Rooyen Z et al. (2018) First report of white root rot caused by *Rosellinia necatrix* on *Persea americana* in South Africa. *Plant Dis* 102:1850
- Van der Merwe JJH, Matthee FN (1974) *Rosellinia*-root rot of apple and pear trees in South Africa. *Phytophylactica* 6:199
- van der Nest MA, Beirn LA, Crouch JA, Demers JE, de Beer ZW et al. (2014a) IMA genome-F 3: draft genomes of *Amanita jacksonii*, *Ceratocystis albifundus*, *Fusarium circinatum*, *Huntia omanensis*, *Leptographium procerum*, *Rutstroemia sydowiana*, and *Sclerotinia echinophila*. *IMA Fungus* 5:473–486
- van der Nest MA, Bihon W, De Vos L, Naidoo K, Roodt D et al. (2014b) IMA genome-F 2: draft genome sequences of *Diplodia sapinea*, *Ceratocystis manginecans*, and *Ceratocystis moniliformis*. *IMA Fungus* 5:135–140
- van der Nest MA, Steenkamp ET, Roodt D, Soal NC, Palmer M et al. (2019) Genomic analysis of the aggressive tree pathogen *Ceratocystis albifundus*. *Fungal Biol* 123:351–363
- Van Wyk M, Wingfield BD, Marin M, Wingfield MJ (2010) New *Ceratocystis* species infecting coffee, cacao, citrus and native trees in Colombia. *Fungal Diversity* 40:103–117
- Vaser R, Sovic I, Nagarajan N, Sikic M (2017) Fast and accurate *de novo* genome assembly from long uncorrected reads. *Genome Res* 27:737–746
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A et al (2014) Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* 9(11):e112963
- Wallwork H, Butt M, Capio E (2016) Pathogen diversity and screening for minor gene resistance to *Pyrenophora teres f. teres* in barley and its use for plant breeding. *Australas Plant Pathol* 45:527–531
- Wang J, Levy M, Dunkle LD (1998) Sibling species of *Cercospora* associated with gray leaf spot of maize. *Phytopathology* 88:1269–1275
- Ward JMJ, Stromberg EL, Nowell DC, Nutter FW (1999) Gray leaf spot: a disease of global importance in maize production. *Plant Dis* 83:884–895
- Wilken PM, Steenkamp ET, Wingfield MJ, de Beer ZW, Wingfield BD (2013) IMA genome-F 1: draft nuclear genome sequence for the plant pathogen, *Ceratocystis fimbriata*. *IMA Fungus* 4:357–358
- Wingfield MJ, Roux J, Slippers B, Hurlley BP, Garnas J et al. (2013) Established and new technologies reduce increasing pest and pathogen threats to Eucalypt plantations. *For Ecol Manag* 301:35–42
- Wingfield BD, Barnes I, Wilhelm de Beer Z, De Vos L, Duong TA et al. (2015a) Draft genome sequences of *Ceratocystis eucalypticola*, *Chrysosporthe cubensis*, *C. deuterocubensis*, *Davidsoniella virescens*, *Fusarium temperatum*, *Graphilbum fragrans*, *Penicillium nordicum*, and *Thielaviopsis musarum*. *IMA Fungus* 6:493–506
- Wingfield MJ, Brockerhoff EG, Wingfield BD, Slippers B (2015b) Planted forest health: the need for a global strategy. *Science* 349:832–836
- Wingfield BD, Ambler J, Coetzee M, De Beer ZW, Duong TA et al. (2016a) Draft genome sequences of *Armillaria fuscipes*, *Ceratocystis minuta*, *Ceratocystis adiposa*, *Endoconidiophora laricicola*, *E. polonica* and *Penicillium frei* DAOMC 242723. *IMA Fungus* 7:217–227
- Wingfield BD, Duong TA, Hammerbacher A, van der Nest MA, Wilson A et al. (2016b) Draft genome sequences for *Ceratocystis fagacearum*, *C. harringtonii*, *Grossmannia penicillata*, and *Huntia bhutanensis*. *IMA Fungus* 7:317–323
- Wingfield BD, Berger DK, Steenkamp ET, Lim HJ, Duong TA et al. (2017) Draft genome of *Cercospora zeina*, *Fusarium pininemorale*, *Hawksworthiomyces lignivorus*, *Huntia decipiens* and *Ophiostoma ips*. *IMA Fungus* 8:385–396
- Wingfield BD, Bills GF, Dong Y, Huang W, Nel WJ (2018) Draft genome sequence of *Annulohyphoxylon stygium*, *Aspergillus mulundensis*, *Berkeleyomyces basicola* (syn. *Thielaviopsis basicola*), *Ceratocystis smalleyi*, two *Cercospora beticola* strains, *Coleophoma cylindrospora*, *Fusarium fracticaudum*, *Phialophora cf. hyalina*, and *Morchella septimelata*. *IMA Fungus* 9:199–223
- Wu BM, Subbarao KV, Liu YB (2008) Comparative survival of sclerotia of *Sclerotinia minor* and *S. sclerotiorum*. *Phytopathology* 98:659–665
- Wyatt NA, Richards JK, Brueggeman RS, Friesen TL (2020) A comparative genomic analysis of the barley pathogen *Pyrenophora teres f. teres* identifies subtelomeric regions as drivers of virulence. *Mol Plant Microbe Interact* 33:173–188
- Yang D, Luo T, Wei J, Cao C, Li G et al. (2022) High-quality genome resource of the phytopathogenic fungus *Sclerotinia minor* LC41, the causal agent of Sclerotinia blight on lettuce in China. *Plant Dis* 106:1042–1044
- Zhan M, Tian M, Wang W, Li G, Lu X et al. (2020) Draft genomic sequence of *Armillaria gallica* 012m: insights into its symbiotic relationship with *Gastrodia elata*. *Braz J Microbiol* 51:1539–1552
- Zhang X, Zou H, Yang Y, Fang B, Huang L (2022) Genome resource for *Elsinoë batatas*, the causal agent of stem and foliage scab disease of sweet potato. *Phytopathology* 112:973–975
- Zhao L, Xiao H, Ma X, Cheng Q (2020) *Elsinoë australis* causing spot anthracnose on poplar in China. *Plant Dis* 104:2202–2209

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

