1 **A chromosome-level genome assembly of *Zasmidium syzygii* isolated from**

2 **banana leaves**

3

4 Anouk C. van Westerhoven[1,2*], Rahim Mehrabi[1,3*], Reza Talebi[3], Maikel B.F

5 Steentjes[1], Benny Corcolon[4], Pablo A. Chong[5], Gert H.J. Kema[1,#,*], Michael F.

6 Seidl[2,#,*]

7

8 [1]Laboratory of Phytopathology, Wageningen University & Research, Wageningen,

9 6700 AA, the Netherlands

10 [2]Theoretical Biology & Bioinformatics, Department of Biology, Utrecht University,

11 Utrecht, 3584 CS, the Netherlands

12 [3]Keygene N.V., P.O. Box 216, Wageningen, 6700 AE, Netherlands

13 [4]Research, Information, Compliance Department, Tadeco Inc., Panabo, Davao del

14 Norte 8105, Philippines

15 [5]ESPOL Polytechnic University, Escuela Superior Politécnica del Litoral, Centro de

16 Investigaciones Biotecnológicas del Ecuador, Laboratorio de Biología Molecular, Vía

17 Perimetral 5, Guayaquil, 090112, Ecuador

18

19 *Authors contributed equally

20 [#]Corresponding authors: Gert H.J. Kema and Michael F. Seidl

© The Author(s) 2023. Published by Oxford University Press on behalf of The Genetics Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

Keywords: *Zasmidium syzygii,* genome assembly, banana leaf disease, foliar blight, plant disease

Abstract

Accurate taxonomic classification of samples from infected host material is essential for disease diagnostics and genome analyses. Despite the importance, diagnosis of fungal pathogens causing banana leaf diseases remains challenging. Foliar diseases of bananas are mainly caused by three *Pseudocercospora* species, of which the most predominant causal agent is *P. fijiensis*. Here, we sequenced and assembled four fungal isolates obtained from necrotic banana leaves in Bohol (Philippines) and obtained a high-quality genome assembly for one of these isolates. The samples were initially identified as *P. fijiensis* using PCR diagnostics, however, the assembly size was consistently 30 Mb smaller than expected. Based on the ITS sequences, we identified the samples as *Zasmidium syzygii* (98.7% identity). The high-quality *Zasmidium syzygii* assembly is 42.5 Mb in size, comprising 16 contigs, of which 11 are most likely complete chromosomes. The genome contains 98.6% of the expected single-copy BUSCO genes and contains 14,789 genes and 10.3% repeats. The three short-read assemblies are less continuous but have similar genome sizes (40.4 – 42.4 Mb) and contain between 96.5% and 98.4% BUSCO genes. All four isolates have identical ITS sequences and are distinct from *Zasmidium* isolates that were previously sampled from banana leaves. We thus report the first continuous genome assembly of a member of the *Zasmidium* genus, forming an essential resource for further analysis to enhance our understanding of the diversity of pathogenic fungal isolates as well as fungal diversity.

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

Introduction

Genome sequencing is important for the diagnostics and monitoring of diseases and is an important step to understand the biology of pathogens and diseases. To enable sample identification and downstream genome analyses, accurate taxonomic classification and prevention of contamination in genome assemblies are essential (Francois et al., 2020; Lu & Salzberg, 2018; Rachtman et al., 2020). However, publicly available genome assemblies are occasionally reported to contain contaminants, which can lead to incorrect species classification (Cornet & Baurain, 2022; Kusch et al., 2023; Steinegger & Salzberg, 2020). Obtaining a clean genome assembly is especially challenging for pathogens that live in close association with their host; samples from these pathogens are often contaminated with host material or other organisms that proliferate in proximity to the host such as endophytic fungi (Kusch et al., 2023; Zaccaron & Stergiopoulos, 2021).

Banana is an important food crop providing food security in tropical and subtropical regions worldwide. Foliar blights are a major constraint to banana production and are mainly caused by a complex of three *Pseudocercospora* species with *P. fijiensis* as a major constituent that causes black leaf streak disease or black Sigatoka. Control of this disease is responsible for up to 25% of the total costs of banana production (Drenth & Kema, 2021). The two other species also cause foliar blights, but are currently less prevalent, *P. musae* causes yellow Sigatoka and *P. eumusae* causes eumusae leaf spot (Chang et al., 2016). However, besides these three *Pseudocercospora* species, other fungal species can appear in association with symptomatic banana foliage. For example, a recent study that analysed fungal isolates associated with banana foliar diseases revealed the presence of over 30 other fungal species, primarily belonging to the *Mycosphaerellacea* family (Crous et

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

al., 2021). Interestingly, before this study revealed the identity of these isolates, most of these fungal species were considered to belong to the *Pseudocercospora* genus, highlighting that accurate diagnosis of the causal agents of leaf symptoms remains challenging. Yet, accurate identification and classification of the obtained isolates is important, particularly for developing disease diagnostics and enhancing the effectiveness of disease management strategies (Kusch et al., 2023; Lu & Salzberg, 2018).

Here we sampled four fungal isolates from banana foliage showing blight symptoms in Bohol, Philippines, and identified the fungal isolates by disease symptoms, morphology, and diagnostic PCR assays (Arzanlou et al., 2008). Interestingly, initial observations classified the isolates as *P. fijiensis*, but further genome analyses revealed that these isolates rather represent *Zasmidium syzygii*. Until now, only three fragmented genome assemblies of *Zasmidium* species are publicly available (Haridas et al., 2020; Xu et al., 2017). Here, we assembled the first chromosome-level genome assembly of a representative of the *Zasmidium* genus. The availability of this genome assembly will help to improve molecular diagnostics for pathogenic *Pseudocercospora* spp. on banana foliage. Moreover, the genome adds to the diversity of available fungal genomes, providing a resource for future genomic studies.

Methods & Materials

Fungal isolation and sequencing

Banana leaves with symptoms of foliar disease were collected from a field in Bohol, Philippines. Leaf samples were cut into 1 cm^2 segments and used to discharge

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1 ascospores as described previously (Chong et al., 2019). Four single ascospore
2 isolates (P121, P122, P123, and P124) were collected and grown on potato dextrose
3 agar (PDA) plates supplemented with streptomycin (100 µg/ml) for three weeks at 25
4 °C. To obtain sufficient fungal biomass for DNA isolation, a piece of mycelium (2
5 cm^2) was blended for 20 sec. at 6,000 rpm in an Ultra Turrax Tube Drive
6 homogenizer (IKA, Staufen, Germany) in 15 ml water using a sterile DT-20 tube
7 (IKA, Staufen, Germany). The fragmented mycelium was transferred into a flask
8 containing 100 ml of PDB amended with streptomycin (100 µg/ml) and kept at 25 °C
9 on a rotary shaker with 150 rpm for about two weeks. The fungal mycelium was
10 filtered through miracloth and subsequently washed with sterile water. Fungal
11 mycelium was freeze-dried overnight and used for high-molecular-weight (HMW)
12 DNA isolation based on the CTAB method (Murray & Thompson, 1980). After adding
13 isopropanol, HMW DNA was collected from the extraction buffer using a sterile
14 needle. AMPure XP purification kit (Becman Cluter Life Sciences, USA) was used to
15 clean up the DNA. DNA quality and quantity were checked by gel electrophoresis,
16 Nanodrop micro-volume spectrophotometers, and Qubit Fluorometric Quantitation
17 (Thermo Fisher Scientific, USA). The HMW DNA of isolate P124 was sequenced
18 using PromethION Oxford Nanopore (ONT) sequencing technologies. Additionally,
19 all isolates were sequenced using the Illumina HiSeq platform, both sequencing
20 platforms were located at Keygene B.V. (Wageningen, the Netherlands).
21
22 *Fungal species diagnosis by PCR*
23 To identify the isolated fungal species, the DNA from each isolate was extracted and
24 subjected to *P. fijiensis*-specific diagnostic PCR according to a previously
25 established protocol (Arzanlou et al., 2008). The list of primers used in PCRs is

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

shown in Table S1. Actin (820 bp) was used as a positive control to ensure successful amplification and to assess the quality of DNA.

Pathogenicity assay

The pathogenicity of isolate P124 was tested on two banana genotypes; Cavendish cv. Grand Naine (AAA) and a diploid (AA) genotype 'Pisang Berlin'. The inoculum for isolate P124 as well as for the *P. fijiensis* reference isolate P78 were prepared similarly. A piece of 1 cm^2 mycelium from three-week-old colonies grown on PDA was collected in an Eppendorf tube containing 3-4 metal beads (3 mm diameter) and was blended for 20 seconds at 3,000 rpm in a Bead Beater Homogenizer. About one ml of sterile water was added to each tube and fragmented mycelium was spread on PDA plates amended with 100 μg/ml streptomycin. The plates were kept at 25°C for 3-4 weeks and then a piece of mycelium (10 cm^2) was blended for 40 sec. at 6,000 rpm in an Ultra Turrax Tube Drive homogenizer (IKA, Staufen, Germany) in 15 ml of distilled water using a sterile DT-20 tube (IKA, Staufen, Germany). The suspension was passed through miracloth to remove non-fragmented mycelium. The collected mycelial fragments were further diluted and adjusted to $5x10^5$ fragments ml^{-1} and supplemented with 0.15% Tween 20. This suspension was used to inoculate two-month-old banana plants on both sides of the leaves. Each treatment was repeated tree times, as a control, water was used for mock inoculation. Inoculated plants were kept for 48 h at 90% relative humidity at 25ºC in the dark in a growth cabinet and subsequently for eight weeks in a greenhouse with >85% RH, and with a day length of 12 hours.

De novo *genome assembly*

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1 To construct a complete and continuous genome assembly, we assembled the ONT
2 long-read sequences of isolate P124. The raw Fast5 files were basecalled with
3 Guppy (v. 5.1.13) and filtered for a minimum quality score of 10. The filtered reads
4 were assembled *de novo* using Canu (v. 2.2, Koren et al., 2017) with the parameters
5 '-OvlMerThreshold=300' and '-corMaxEvidenceErate=0.15' to improve the assembly
6 of repetitive regions. The initial genome assembly was polished with Racon (v. 1.5.0
7 Vaser et al., 2017) using the raw ONT long reads. After polishing, missing telomeres
8 were reconstructed, when possible, using teloclip (v. 0.0.3,
9 https://github.com/Adamtaranto/teloclip) and a final polishing step was performed
10 based on shrot-read Illumina data using Pilon (v. 1.24, Walker et al., 2014). The
11 presence of telomeres (TTAGGG) and coverage of the chromosomes in the final
12 polished assembly were determined using tapestry (v. 1.0.0, Davey et al., 2020). To
13 estimate the completeness of the genome assembly, the presence of conserved
14 single-copy genes was determined using BUSCO with the Capnodiales database
15 (Simão et al., 2015). The assembly continuity and potential contaminations were
16 accessed using BlobTools (v. 4.1.5, Challis et al., 2020). To estimate the size of the
17 assembled genome, all Illumina sequences were analyzed using GenomeScope (v.
18 2.0, Vurture et al., 2017) with the 21-mer profile determined by JellyFish (v 2.3.0,
19 Marçais & Kingsford, 2011). Isolate P121 showed high k-mer coverage outliers,
20 therefore the maximal k-mer coverage was set to 350x, similar to the maximum k-
21 mer coverage observed in the other genome. Repeats were predicted using
22 RepeatModeler (v. 2.0.3, Flynn et al., 2020) and annotated using RepeatMasker (v.
23 4.1.2, Smit et al., 2015). Finally, protein-coding genes were annotated using
24 Funannotate (v. 1.8.9, Palmer & Stajich, 2019) with the BUSCO lineage
25 pezizomycotina utilizing predicted *Zasmidium cellare* proteins as evidence

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1 (GCF_010093935.1, Haridas et al., 2020). Additionally, we assembled the Illumina
2 sequenced isolates P121-P123 with Spades (v. 3.13, Bankevich et al., 2012). To
3 assess the quality and completeness of these assemblies, we ran QUAST (v. 5.2,
4 Gurevich et al., 2013) and BUSCO (v. 5.3.2, Simão et al., 2015) with the
5 Capnodiales odb10 database.

6

7 *Identification of internal transcribed spaces sequence*

8 The internal transcribed spacer (ITS) sequences from the four assemblies (P121-
9 P124) were retrieved using ITSx (v. 1.1.3, Bengtsson-Palme et al., 2013) and blastn
10 was used to detect similar ITS sequences in the blast database (accessed 19th
11 November 2022). Additionally, a maximum-likelihood tree was constructed to
12 compare the four ITS sequences to 227 publicly available ITS sequences obtained
13 from NCBI (accessed 13th January 2023). Sequences were aligned using mafft (v.7.
14 453, Katoh et al., 2002) and a maximum-likelihood phylogenetic tree was
15 constructed using RAxML, with 500 bootstrap replicates (-m GTRCAT -p 1234 -b
16 100 -N 500) (v. 8.2.12, Stamatakis, 2014). The resulting phylogeny was visualized
17 using iTOL (Letunic & Bork, 2021).

18

19

20 *Genetic diversity*

21 To assess the genetic diversity among isolates, short-read sequencing data of P121-
22 P123 were aligned to the P124 long-read genome assembly using BWA-mem (v.
23 0.7.17, Li, 2013). GATK4 was used to call variants (Van der Auwera et al., 2013),

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

and these variants were subsequently filtered using GATK Variant Filtration based on the GATK best practices (Van der Auwera et al., 2013). The filtering process involved excluding variants from reads with low mapping quality, variants predominantly located at the edge of reads, and variants exhibiting a bias towards reverse/forward strands.

Results and Discussion

Fungal isolates obtained from symptomatic banana foliage in Bohol

Foliar blights of banana is mainly caused by a complex of *Pseudocercospora* species (Chang et al., 2016; Drenth & Kema, 2021). However, additional fungal species predominantly belonging to the *Mycosphaerella* genus have also been associated with symptomatic banana leaves (Crous et al., 2021). To further analyze the fungal pathogens that cause disease on banana leaves, four samples were obtained from banana plants with necrotic lesions in Bohol, Philippines (Chong et al., 2021). For all four isolates (P121 – P124) we were able to amplify a PCR product using *P. fijiensis*-specific PCR primers (Arzanlou et al., 2008), indicating that the isolates can be identified as *P. fijiensis* (**Figure 1a**). To corroborate the identity of the isolates, we tested the pathogenicity of one representative (isolate P124) on banana cultivars Cavendish (cv. Grand Naine AAA) and Pisang Berlin (AA) and compared the pathogenicity to the reference *P. fijiensis* isolate P78 originating from Tanzania. No leave spot symptoms were observed on Cavendish upon inoculation with P124, in contrast to the necrotic lesions caused by P78 (**Figure 1b**). However, both isolates caused necrotic lesions on Pisang Berlin four weeks after inoculation (**Figure 1c**). The necrotic lesions caused by isolate P124 were less severe than the disease

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1 symptoms caused by *P. fijiensis* isolate P78. Although we observe a difference in
2 pathogenicity between the two isolates differ in pathogenicity, both isolates but
3 cause necrosis on banana foliage.
4
5 **Chromosome-level genome assembly of P124**

6 To generate a high-quality genome assembly for the isolates from Bohol, we
7 randomly selected strain P124 for sequencing using Oxford Nanopore Technology.
8 This yielded 11.9 Gb of reads with an average size of ~9.1 kb and a read N50 of 12
9 kb, corresponding to a ~150x genome coverage based on the estimated genome
10 size of 74 Mb (Arango et al., 2016). The *de novo* genome assembly resulted in 16
11 nuclear contigs and one contig representing the mitochondrial genome. The
12 assembly has an N50 of 3.4 Mb and a total genome size of 42.5 Mb. Eleven of the
13 16 contigs had telomeric sequences (TTAGGG) at both ends, and thus the assembly
14 is highly contiguous and mostly represents complete chromosomes (Figure 2). To
15 assess genome completeness, we queried the genome assembly for the presence of
16 single-copy BUSCO genes and identified 98.6% of the single-copy BUSCO genes
17 that are expected in the fungal order Capnodiales, indicating that the genome
18 assembly of P124 covers the conserved gene space (Figure 2a). We predicted a
19 total of 14,789 protein-coding genes in the genome and identified that 10.3% of the
20 genome consists of repetitive elements (Figure 2b). *De novo* genome assembly of
21 isolates P121 – P123, sequenced with short-read sequencing technology only,
22 resulted in genome assemblies with similar sizes (40.4 – 42.5 Mb). Although the
23 assemblies of P121 - P123 are more fragmented compared to the assembly of the
24 nanopore-sequenced isolate P124, they approximately contain an equally high
25 number of expected BUSCO genes (Table 1).

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1

2 The *P. fijiensis* reference genome is 73.6 Mb (Arango et al., 2016), and remarkably,
3 the genome assemblies of the here sequenced fungal isolates are approximately 30
4 Mb smaller than expected. To verify the assembly size, we estimated genome sizes
5 using k-mer profiles, which resulted in an estimated genome size ranging from 38.9
6 to 41.9 Mb, similar to the size of the *de novo* assembled genomes (Table 1).
7 Furthermore, we mapped the sequencing reads of isolates P121 – P124 to the
8 chromosome-level genome assembly of P124, mapping on average 95% of the
9 reads with limited genetic diversity between isolates (on average 15 SNPs per kb;
10 Figure 2b**).** An increased read mapping coverage is observed in most telomeric
11 regions (Figure 2c), while the coverage drops in contig 1, contig 6, and contig 16,
12 suggesting that these regions possibly contain assembly artefacts. Contig 16 is only
13 0.2 Mb in size and only contains a telomeric repeats at one end. The low coverage
14 region together with read mappings that support a possible link to either contig 3 or
15 14, which also lack one of the telomeric repeats, suggests that this contig might be
16 associated to one of these two contigs. Apart from these regions, mapping of the
17 short reads of P124 to the P124 chromosome-level genome assembly revealed an
18 overall constant read mapping coverage over the chromosomes, indicating that the
19 P124 genome assembly does not contain extensively collapsed repetitive regions
20 that could account for the smaller genome size compared to *P. fijiensis*. This
21 validates that, although the genome assembly of strain P124 is smaller than
22 expected, it is highly complete and continuous, suggesting that the length variation is
23 likely not due to assembly errors. To capture possible contaminations in the genome
24 assemblies, we queried all contigs in the BLAST database to determine their identity.
25 All contigs showed a high similarity to members of the *Mycosphaerellacea* family;

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1 most contigs were similar to *Zasmidium* species (13 contigs, 35 Mb) and only one
2 contig (3.3 Mb) showed similarity to *Pseudocercospora* species (**Figure S1**). Whole-
3 genome alignments of our assembly to the *P. fijiensis* reference genome assembly
4 revealed that none of the contigs display significant similarity (**Figure S2**). Therefore,
5 we considered that the difference in genome size between isolates P121 – P124 and
6 the *P. fijiensis* reference isolate P78 is likely caused by the isolation of a different
7 fungal species associated with necrotic symptoms on banana foliage (Crous et al.,
8 2021), most likely by a member of genus *Zasmidium*.

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

The assembled genome sequence reveals that isolate P124 belongs to the fungal genus *Zasmidium*

To determine the identity of isolate P124, we retrieved the ITS sequence from the genome assembly and searched for related species in the non-redundant BLAST database on NCBI using the P124 ITS sequence as a query. We obtained a highly similar match (98.7% nucleotide identity) to *Zasmidium syzygii* (NR_111826.1), supporting that the assembled isolate belongs to the *Zasmidium* genus. In line with this finding, previous studies have found *Zasmidium* isolates on symptomatic banana leaves (Arzanlou et al., 2008; Crous et al., 2021), and other *Zasmidium* species are reported to cause leaf spot diseases on other plant species such as citrus (Aguilera-Cogley et al., 2017; An et al., 2021; Han et al., 2015). The *Zasmidium* genome assembly we report here is less fragmented and 4 Mb larger than the previous genome assembly of a *Zasmidium* species; *Z. cellare*. The species, known as the 'wine cellar fungus' because it thrives in walls and ceilings of wine cellars (Tribe et al., 2006), was sequenced and its genome was assembled into >267 scaffolds with a genome size of ~38 Mb (Haridas et al., 2020), illustrating that our genome offers a more continuous genome representation of a member of the *Zasmidium* genus.

To validate the identity and to determine the diversity of the *Zasmidium* isolates, we compared the ITS sequences of isolates P121-P124 with 43 other *Zasmidium* ITS sequences from NCBI (20 December 2022) as well as ITS sequences of 120 *Mycosphaerella* strains obtained from banana leaves (Crous et al., 2021). A maximum-likelihood phylogeny based on the aligned ITS sequences confirmed that isolates P121 – P124 belong to the *Zasmidium* genus and shows that these four isolates encode identical ITS sequences (Figure 3). Notably, the set of ITS sequences also contains sequences from five other *Zasmidium* isolates that had

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1 been previously sampled from banana leaves from different geographic locations
2 (Martinique, Tonga, and Gabon) (Crous et al., 2021). Interestingly, these do not
3 cluster with isolates P121-P124, suggesting that pathogenicity towards banana is a
4 polyphyletic trait within the *Zasmidium* genus (Figure 3). Although *Zasmidium*
5 species have been linked to foliar blights in various hosts (Aguilera-Cogley et al.,
6 2017; Han et al., 2015; Osorio et al., 2021), the pathogenicity and global spread of
7 *Zasmidium* species has not been studied in depth. Based on our data, we conclude
8 that Z. *syzygii* occurs on banana leaves with necrotic symptoms and can be the
9 cause of mild necrotic lesions on the foliage. However, the abundance and role of Z.
10 *syzygii* as a banana pathogen remains unknown, which requires further research to
11 understand its prevalence, significance, and potential impact on banana cultivation.
12
13 Accurate diagnostics of pathogens is essential to detect the emergence and
14 trace the dispersal of diseases, which is pivotal for effective disease management.
15 However, our data reveal that *P. fijiensis* and *Z. syzygii* are indistinguishable with the
16 current PCR diagnostic (Arzanlou et al., 2008). To compare the similarity of the PCR
17 primers between *P. fijiensis* and *Z. syzygii*, we *in silico* detected the amplicon of the
18 supposedly *Pseudocercospora*-specific primers in the *Zasmidium syzygii* P124
19 genome assembly and in the *P. fijiensis* reference genome assembly (Cirad86;
20 Arango et al., 2016). Both isolates possess the primer sequence used to distinguish
21 *P. fijiensis* (Table S1) from *P. musae* and *P. eumusae* and produce a similar sized
22 amplicon of 480 bp in *Z. syzygii* P124 and 478 bp in *P. fijiensis* Cirad86, which
23 explains the positive result for *Z. syzygii* in our PCR assay (Figure 1). The amplicons
24 share 88% sequence identity and *Zasmidium* or *Pseudocercospora* isolates can
25 therefore be distinguished only upon amplicon sequencing. Thus, novel primer pairs

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

need to be developed to enable easy and accurate diagnosis of fungal species present in foliar blights of banana.

Conclusion

Here, we report the first chromosome-scale genome assembly of a *Zasmidium* species, this adds a high-quality genome sequence to the thus far limited genetic resources available for this genus. Our data show that *Z. syzygii* occurs on banana foliage in Bohol and can cause leaf necrosis, comparable to the foliar blight symptoms observed for *P. fijiensis*. The availability of the genome assembly will facilitate further research into the association of *Zasmidium, Pseudocercospora*, and possibly other fungal species related to foliar blights of banana. Moreover, it will serve as a valuable resource for developing novel molecular diagnostics, enabling the accurate identification and characterization of these fungal species.

Data Availability Statement

All sequencing data are deposited in the NCBI database under accession number PRJNA931964. Genome assembly, genome annotation, and vcf files are available on figshare: https://doi.org/10.25387/g3.24117504. Scripts used for the genome assembly and analysis are available on github https://github.com/Anouk-vw/Zasmidium.

Acknowledgments

We thank the Research Support Group (RSG) at KeyGene B.V. (Wageningen, the Netherlands) for performing the DNA isolation, library preparation and genome

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1 sequencing. We thank Xiaoqian Shi-Kunne (Wageningen University & Research, the
2 Netherlands) for the data management and bioinformatics support.
3
4 *Funder Information*
5 ACW and GHJK were supported by the Bill and Melinda Gates Foundation, grant
6 number AG – 4425. Banana research at Wageningen University has been supported
7 by the Dutch Dioraphte Foundation, grant number 20 04 04 02.
8
9 *Conflict of Interest*
10 The authors report no conflict of interests.
11

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1 **References**

2 Aguilera-Cogley, V. A., Berbegal, M., Català, S., Brentu, F. C., Armengol, J., &
3 Vicent, A. (2017). Characterization of *Mycosphaerellaceae* species associated
4 with citrus greasy spot in Panama and Spain. *PLOS ONE*, *12*(12), e0189585.
5 https://doi.org/10.1371/journal.pone.0189585
6 An, Y.-Y., Zeng, X.-Y., Geng, K., Hyde, K. D., & Wang, Y. (2021). One new species
7 and one new record of Zasmidium in China. *Biodiversity Data Journal*, *9*,
8 e59001. https://doi.org/10.3897/BDJ.9.e59001
9 Arango, R. E. isaza, Diaz-Trujillo, C., Dhillon, B., Aerts, A., Carlier, J., Crane, C. F.,
10 Jong, T. V. de, Vries, I. de, Dietrich, R., Farmer, A. D., Fereira, C. F., Garcia,
11 S., Guzman, M., Hamelin, R. C., Lindquist, E. A., Mehrabi, R., Quiros, O.,
12 Schmutz, J., Shapiro, H., … Kema, G. H. J. (2016). Combating a global threat
13 to a clonal crop: Banana Black Sigatoka pathogen *Pseudocercospora fijiensis*
14 (synonym *Mycosphaerella fijiensis*) genomes reveal clues for disease control.
15 *PLOS Genetics*, *12*(8), e1005876.
16 https://doi.org/10.1371/journal.pgen.1005876
17 Arzanlou, M., Groenewald, J. Z., Fullerton, R. A., Abeln, E. C. A., Carlier, J., Zapater,
18 M.-F., Buddenhagen, I. W., Viljoen, A., & Crous, P. W. (2008). Multiple gene
19 genealogies and phenotypic characters differentiate several novel species of
20 *Mycosphaerella* and related anamorphs on banana. *Persoonia : Molecular*
21 *Phylogeny and Evolution of Fungi*, *20*, 19–37.
22 https://doi.org/10.3767/003158508X302212
23 Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S.,
24 Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V.,
25 Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A.

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

(2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., De Wit, P., Sánchez-García, M., Ebersberger, I., de Sousa, F., Amend, A., Jumpponen, A., Unterseher, M., Kristiansson, E., Abarenkov, K., Bertrand, Y. J. K., Sanli, K., Eriksson, K. M., Vik, U., … Nilsson, R. H. (2013). Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in Ecology and Evolution*, *4*(10), 914–919. https://doi.org/10.1111/2041-210X.12073

Challis, R., Richards, E., Rajan, J., Cochrane, G., & Blaxter, M. (2020). BlobToolKit – interactive quality assessment of genome assemblies. *G3 Genes|Genomes|Genetics*, *10*(4), 1361–1374. https://doi.org/10.1534/g3.119.400908

Chang, T.-C., Salvucci, A., Crous, P. W., & Stergiopoulos, I. (2016). Comparative genomics of the Sigatoka disease complex on banana suggests a link between parallel evolutionary changes in *Pseudocercospora fijiensis* and *Pseudocercospora eumusae* and increased virulence on the banana host. *PLOS Genetics*, *12*(8), e1005904. https://doi.org/10.1371/journal.pgen.1005904

Chong, P., Essoh, J. N., Arango Isaza, R. E., Keizer, P., Stergiopoulos, I., Seidl, M. F., Guzman, M., Sandoval, J., Verweij, P. E., Scalliet, G., Sierotzski, H., de Lapeyre de Bellaire, L., Crous, P. W., Carlier, J., Cros, S., Meijer, H. J. G., Peralta, E. L., & Kema, G. H. J. (2021). A world-wide analysis of reduced

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1 sensitivity to DMI fungicides in the banana pathogen *Pseudocercospora*
2 *fijiensis*. *Pest Management Science*, *77*(7), 3273–3288.
3 https://doi.org/10.1002/ps.6372
4 Chong, P., Vichou, A.-E., Schouten, H. J., Meijer, H. J. G., Arango Isaza, R. E., &
5 Kema, G. H. J. (2019). Pfcyp51 exclusively determines reduced sensitivity to
6 14α-demethylase inhibitor fungicides in the banana black Sigatoka pathogen
7 *Pseudocercospora fijiensis*. *PLOS ONE*, *14*(10), e0223858.
8 https://doi.org/10.1371/journal.pone.0223858
9 Cornet, L., & Baurain, D. (2022). Contamination detection in genomic data: More is
10 not enough. *Genome Biology*, *23*(1), 60. https://doi.org/10.1186/s13059-022-
11 02619-9
12 Crous, P. W., Carlier, J., Roussel, V., & Groenewald, J. Z. (2021).
13 Pseudocercospora and allied genera associated with leaf spots of banana
14 (Musa spp.). *Fungal Systematics and Evolution*, *7*, 1–19.
15 https://doi.org/10.3114/fuse.2021.07.01
16 Davey, J. W., Davis, S. J., Mottram, J. C., & Ashton, P. D. (2020). *Tapestry: Validate*
17 *and edit small eukaryotic genome assemblies with long reads* (p.
18 2020.04.24.059402). bioRxiv. https://doi.org/10.1101/2020.04.24.059402
19 Drenth, A., & Kema, G. (2021). The vulnerability of bananas to globally emerging
20 disease threats. *Phytopathology®*, *111*(12), 2146–2161.
21 https://doi.org/10.1094/phyto-07-20-0311-rvw
22 Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., & Smit,
23 A. F. (2020). RepeatModeler2 for automated genomic discovery of
24 transposable element families. *Proceedings of the National Academy of*
25 *Sciences*, *117*(17), 9451–9457. https://doi.org/10.1073/pnas.1921046117

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1 Francois, C. M., Durand, F., Figuet, E., & Galtier, N. (2020). Prevalence and
2 implications of contamination in public genomic resources: A case study of 43
3 reference arthropod assemblies. *G3 Genes|Genomes|Genetics*, *10*(2), 721–
4 730. https://doi.org/10.1534/g3.119.400758
5 Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality
6 assessment tool for genome assemblies. *Bioinformatics (Oxford, England)*,
7 *29*(8), 1072–1075. https://doi.org/10.1093/bioinformatics/btt086
8 Han, K. S., Park, M. J., Choi, I. Y., Park, J. H., & Shin, H. D. (2015). First report of
9 leaf spot caused by *zasmidium lythri* on *lythrum salicaria* in korea. *Plant*
10 *Disease*, *99*(5), 728–728. https://doi.org/10.1094/PDIS-10-14-1039-PDN
11 Haridas, S., Albert, R., Binder, M., Bloem, J., LaButti, K., Salamov, A.,
12 Andreopoulos, B., Baker, S. E., Barry, K., Bills, G., Bluhm, B. H., Cannon, C.,
13 Castanera, R., Culley, D. E., Daum, C., Ezra, D., González, J. B., Henrissat,
14 B., Kuo, A., … Grigoriev, I. V. (2020). 101 Dothideomycetes genomes: A test
15 case for predicting lifestyles and emergence of pathogens. *Studies in*
16 *Mycology*, *96*, 141–153. https://doi.org/10.1016/j.simyco.2020.01.003
17 Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). Mafft: A novel method for rapid
18 multiple sequence alignment based on Fast Fourier transform. *Nucleic Acids*
19 *Research*, *30*(14), 3059–3066. https://doi.org/10.1093/nar/gkf436
20 Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M.
21 (2017). Canu: Scalable and accurate long-read assembly via adaptive k-mer
22 weighting and repeat separation. *Genome Research*, *27*(5), 722–736.
23 https://doi.org/10.1101/gr.215087.116

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1 Kusch, S., Vaghefi, N., & Kiss, L. (2023). The good, the bad, and the misleading:
2 How to improve the quality of 'genome announcements'? *Molecular Plant-*
3 *Microbe Interactions®*. https://doi.org/10.1094/MPMI-01-23-0009-LE
4 Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: An online tool for
5 phylogenetic tree display and annotation. *Nucleic Acids Research*, *49*(W1),
6 W293–W296. https://doi.org/10.1093/nar/gkab301
7 Li, H. (2013). *Aligning sequence reads, clone sequences and assembly contigs with*
8 *BWA-mem*. https://doi.org/10.6084/m9.figshare.963153.v1
9 Lu, J., & Salzberg, S. L. (2018). Removing contaminants from databases of draft
10 genomes. *PLOS Computational Biology*, *14*(6), e1006277.
11 https://doi.org/10.1371/journal.pcbi.1006277
12 Marçais, G., & Kingsford, C. (2011). A fast, lock-free approach for efficient parallel
13 counting of occurrences of k-mers. *Bioinformatics*, *27*(6), 764–770.
14 https://doi.org/10.1093/bioinformatics/btr011
15 Murray, M. G., & Thompson, W. F. (1980). Rapid isolation of high molecular weight
16 plant DNA. *Nucleic Acids Research*, *8*(19), 4321–4325.
17 https://doi.org/10.1093/nar/8.19.4321
18 Osorio, J. A., Linnakoski, R., Wingfield, M. J., de Beer, Z. W., & Roux, J. (2021). A
19 new species in the *Mycosphaerellaceae from Cecidomyiidae leaf galls* on
20 *Avicennia marina* in South Africa. *Antonie van Leeuwenhoek*, *114*(5), 515–
21 526. https://doi.org/10.1007/s10482-021-01537-3
22 Palmer, J., & Stajich, J. (2019). *nextgenusfs/funannotate: Funannotate v1.5.3*
23 [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.2604804
24 Rachtman, E., Balaban, M., Bafna, V., & Mirarab, S. (2020). The impact of
25 contaminants on the accuracy of genome skimming and the effectiveness of

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

exclusion read filters. *Molecular Ecology Resources*, *20*(3), 649–661. https://doi.org/10.1111/1755-0998.13135

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Smit, A. F. A., Hubley, R., & Green, P. (2015). *RepeatMasker Open-4.0*. http://www.repeatmasker.org

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Steinegger, M., & Salzberg, S. L. (2020). Terminating contamination: Large-scale search identifies more than 2,000,000 contaminated entries in GenBank. *Genome Biology*, *21*(1), 115. https://doi.org/10.1186/s13059-020-02023-1

Tribe, H. T., Thines, E., & Weber, R. W. S. (2006). Moulds that should be better known: The wine cellar mould, *Racodium cellare*. *Mycologist*, *20*(4), 171–175. https://doi.org/10.1016/j.mycol.2006.09.016

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics*, *43*(1), 11.10.1-11.10.33. https://doi.org/10.1002/0471250953.bi1110s43

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

Vaser, R., Sović, I., Nagarajan, N., & Šikić, M. (2017). Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Research*, *27*(5), 737–746. https://doi.org/10.1101/gr.214270.116

Vurture, G. W., Sedlazeck, F. J., Nattestad, M., Underwood, C. J., Fang, H., Gurtowski, J., & Schatz, M. C. (2017). GenomeScope: Fast reference-free genome profiling from short reads. *Bioinformatics*, *33*(14), 2202–2204. https://doi.org/10.1093/bioinformatics/btx153

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., & Earl, A. M. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLOS ONE*, *9*(11), e112963. https://doi.org/10.1371/journal.pone.0112963

Xu, C., Zhang, R., Sun, G., & Gleason, M. L. (2017). Comparative genome analysis reveals adaptation to the ectophytic lifestyle of sooty blotch and flyspeck fungi. *Genome Biology and Evolution*, *9*(11), 3137–3151. https://doi.org/10.1093/gbe/evx229

Zaccaron, A. Z., & Stergiopoulos, I. (2021). Characterization of the mitochondrial genomes of three powdery mildew pathogens reveals remarkable variation in size and nucleotide composition. *Microbial Genomics*, *7*(12), 000720. https://doi.org/10.1099/mgen.0.000720

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1 **Figures and tables**

2 **Table 1. The genome assembly statistics of four sequenced isolates obtained**

3 **from symptomatic banana foliage in Bohol, Philippines.**

	P124*	**P121****	**P122****	**P123****
Genome size (Mb) based on k-mer estimation (k=21)	40.6	39.8	39.9	41.9
Genome assembly size (Mb)	42.5	40.4	40.6	42.2
Number of contigs	16	1,225	1,563	796
Largest contig (Mb)	6.97	0.44	0.31	0.84
GC%	51.70	52.55	52.37	52.01
N50 (Mb)	3.40	0.98	0.60	1.9
N75 (Mb)	2.08	0.49	0.30	1.05
L50	2	134	201	67
BUSCO completeness ($n = 3,578$)	98.6%	97.8%	96.5%	98.4%
BUSCO fragmented genes	0.5%	0.9%	1.5%	0.5%
BUSCO missing genes	1.0%	1.3%	2.0%	1.1%

4 *P124 is assembled using long reads from Oxford Nanopore Technology (ONT).

5 **P121-P123 are assembled using Illumina short reads.

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

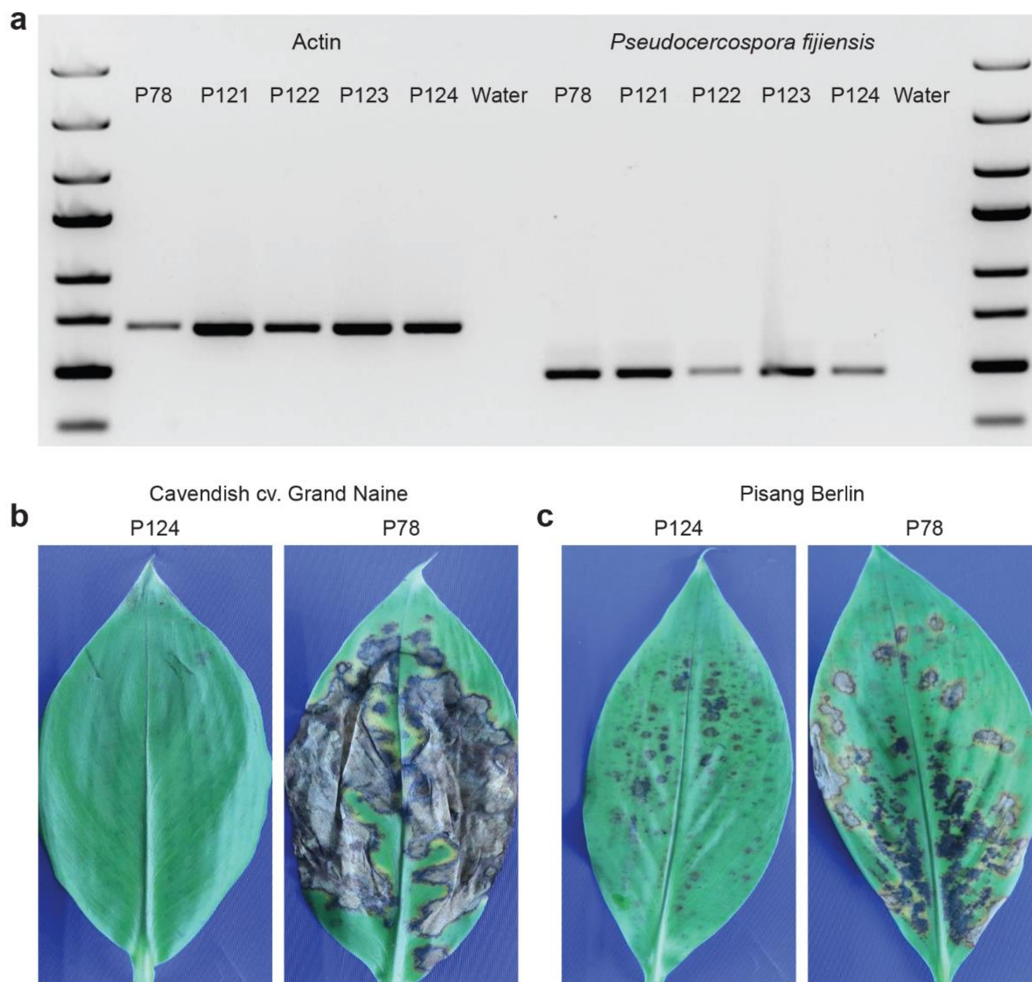
Figure 1. PCR diagnostics and pathogenicity assay of isolate P124 in comparison to the *Pseudocercospora fijiensis* reference strain P78 a) All isolates show amplification with actin primers (left panel). *Pseudocercospora*-specific primers amplified PCR products from *P. fijiensis* strain P78 (positive control) as well as from isolates P121, P122, P123, and P124 (right panel), suggesting that the isolates can be identified as *P. fijiensis*. b) Isolate P124 does not cause necrotic lesions on Cavendish banana, in contrast to necrotic symptoms caused by P78. c) Both P124 and P78 cause necrotic leaf symptoms on Pisang Berlin. Disease symptoms were scored eight weeks after inoculation.

Figure 2. Chromosome-level genome assembly of *Zasmidium syzygii* isolate P124. a) Genome assembly statistics for the *de novo* assembly of *Z. syzygii* isolate P124 based on Oxford Nanopore Technology (ONT). The genome assembly has 17 contigs (16 nuclear contigs of which at least 11 are complete chromosomes, and the mitochondrial genome) with a total genome size of 42.5 Mb and contains 98.7% complete single-copy BUSCO genes. **b**) A circular representation of the contigs in P124. Dashed lines indicate missing telomeres for chromosome 3, 11, 13, 15, and 16. Genes (14,786) and repeats (10.3%) are distributed evenly over the chromosomes. Single nucleotide polymorphisms are found over all contigs, with two SNP dense regions on contig 8 and 9. **c**) Short-read coverage of P124 mapped to the assembled contigs of P124 shows minimal regions with exceptionally high or low coverage, suggesting that the assembly does not contain large repetitive regions that may have been collapsed during the assembly process. Mean coverage (green line) and Median coverage (red line) are indicated in the figure.

Figure 3. The four isolates sampled from Bohol (P121-P124) belong to the genus *Zasmidium*. Maximum-likelihood phylogenetic tree is constructed using the internally transcribed spacer (ITS) sequences of 43 *Zasmidium* species from NCBI and 120 *Mycosphaerella* isolates from banana leaves (Crous et al., 2021). The ITS sequence extracted from the chromosome-level genome assembly (P124) as well as from the other isolates (P121-123) associated with an isolate classified as *Z. syzgii* are not related to *Zasmidium* strains previously isolated from banana leaves. *Zasmidium* isolates associated with infected banana leaves (yellow labels) are genetically diverse and are distributed across various branches of the phylogenetic tree.

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

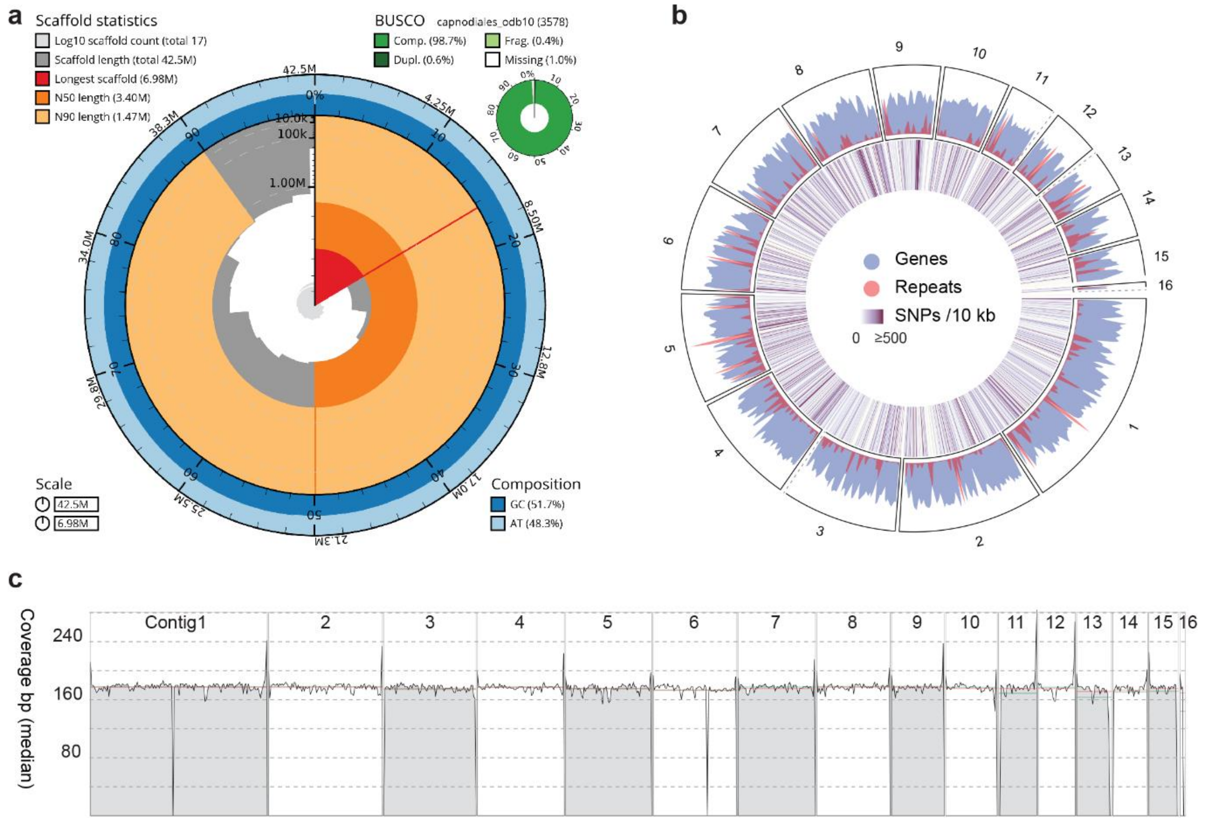
1



2 *Figure 1*
3 *139x131 mm (x DPI)*

4

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023



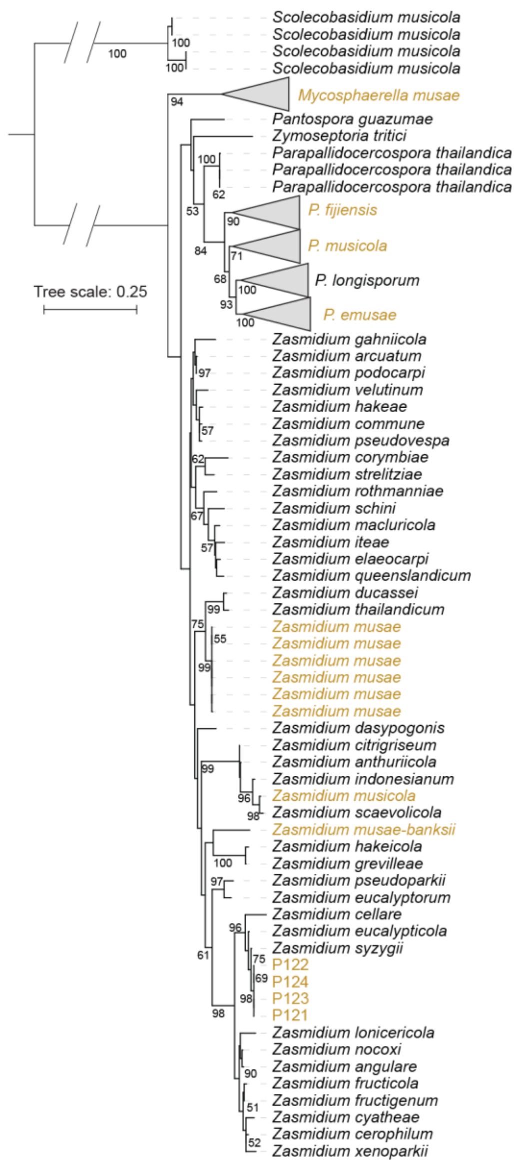
1

2 *Figure 2*
3 *187x128 mm (x DPI)*

4

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023

1



2 *Figure 3*

3 *67x152 mm (x DPI)*

Downloaded from https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkad262/7424864 by Wageningen UR Library user on 08 December 2023