DMA Genome-F 8

Draft genome of of Cercospora zeina, Fusarium pininemorale, Hawksworthiomyces lignivorus, H. decipiens and Ophiostoma ips

Brenda D. Wingfield1, Dave K. Berger2, Emma T. Steenkamp3, Hye-Jin Lim1, Tuan A. Duong1, Burton H. Bluhm1, Z. Wilhelm de Beer1, Lieschen De Vos2, G. Fourie2, Kershney Naidoo2, Nicky Olivier3,6, Yao-Cheng Lin4,6, Yves Van de Peer4,5, Fourie Joubert6, Bridget G. Crampton3, Velushka Swart3, Nicole Soal3, Catherine Tatham3, Magriet A. van der Nest1, Nicolaas A. van der Merwe1, Stephanie van Wyk2, P. Markus Wilken1, and Michael J. Wingfield1

1Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0002, South Africa; corresponding author e-mail: Brenda.wingfield@up.ac.za
2Department of Microbiology and Plant Pathology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0002, South Africa
3Department of Plant and Soil Sciences, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0002, South Africa
4VIB Department of Plant Systems Biology, Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium
5Department of Genetics, Genomics Research Institute, University of Pretoria, Pretoria, South Africa
6Centre for Bioinformatics and Computational Biology, Department of Biochemistry, University of Pretoria, Pretoria, South Africa
7Department of Plant Pathology, University of Arkansas, Fayetteville, USA

Abstract: The genomes of Cercospora zeina, Fusarium pininemorale, Hawksworthiomyces lignivorus, H. decipiens, and Ophiostoma ips are presented in this genome announcement. Three of these genomes are from plant pathogens and otherwise economically important fungal species. Fusarium pininemorale and H. decipiens are not known to cause significant disease but are closely related to species of economic importance. The genome sizes range from 25.99 Mb in the case of O. ips to 4.82 Mb for H. lignivorus. These genomes include the first reports of a genome from the genus Hawksworthiomyces. The availability of these genome data will allow the resolution of longstanding questions regarding the taxonomy of these species. In addition these genome sequences through comparative studies with closely related organisms will increase our understanding of how these species or close relatives cause disease.

Key words: Gray Leaf Spot maize insect vectored fungi Fusarium fujikuroi species complex Ophiostomatales wood decay Ceratocystidaceae blue stain

Article info: Submitted: 27 October 2017; Accepted: 9 November 2017; Published: 21 November 2017.

IMA Genome-F 8A

Draft genome of an African isolate of the maize fungal pathogen, Cercospora zeina

Maize (Zea mays) is one of the most important staple food crops, especially in Sub-Saharan Africa. In the last three decades, Gray Leaf Spot (GLS) has become a widespread foliar disease of maize globally (Latterell & Rossi 1983, Ward et al. 1999, Meisel et al. 2009, Berger et al. 2014). Since GLS leads to substantial yield losses, it poses a threat to food security, especially in Africa (Meisel et al. 2009). Two causal agents of GLS have been identified, with Cercospora zeae-maydis the predominant pathogen in the USA (Wang et al. 1998), and to date, only Cercospora zeina reported from Africa (Crous et al. 2006, Meisel et al. 2009). These species belong to the class Dothidiomycetes, that contains several important plant pathogens (Ohm et al. 2012). However, little is known about pathogenicity mechanisms of C. zeina.


© 2017 International Mycological Association

You are free to share - to copy, distribute and transmit the work, under the following conditions:

Attribution: You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).

Non-commercial: You may not use this work for commercial purposes.

No derivative works: You may not alter, transform, or build upon this work.

For any reuse or distribution, you must make clear to others the license terms of this work, which can be found at http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode. Any of the above conditions can be waived if you get permission from the copyright holder. Nothing in this license impairs or restricts the author’s moral rights.
NUCLEOTIDE SEQUENCE ACCESSION NUMBER

The draft genome has been deposited at DDBJ/ENA/GenBank and is available under the accession number MVDW00000000; Biosample SAMN06067857; Bioproject PRJNA355276. This paper describes the first version of the genome. RNA sequencing data have been deposited in the NCBI Gene Expression Omnibus (Accession number GSE90705).

MATERIALS AND METHODS

Conidia of Cercospora zeina were grown on V8 agar (20 % (v/v) Campbells V8 juice, 2 % (w/v) Bacterial Agar, 0.349 % (w/v) CaCO₃) at ambient room temperature in constant darkness to promote conidiation. Conidia were collected and 1x10⁵ conidia/ml cultured in Potato Dextrose Broth (PDB) at 25 °C with gentle shaking. Prior to cultures reaching the melanized stage DNA was isolated as described previously (Ma et al. 2010).

For RNA isolations C. zeina was cultured under seven different in vitro growth conditions. For solid media, conidia were transferred from V8 agar onto sterile cellophane sheets overlain onto the particular media by subculturing. For liquid media, conidia were washed from the V8 agar using the relevant media and a sterile L-spreader, and transferred to a flask containing the growth media. Media used included Complete media (1 % (w/v) glucose, 0.1 % (w/v) yeast extract, 0.1 % (w/v) casein hydrolysate, 0.1 % (w/v) Ca(NO₃)₂.4H₂O, 1 % (v/v) mineral solution [2 % (w/v) KH₂PO₄, 2.5 % (w/v) MgSO₄.7H₂O, 1.5 % (w/v) NaCl]; 0.2x PDA (0.3 % (w/v) Potato Dextrose Agar (PDA), 1.2 % (w/v) Bacterial Agar), 0.2x PDA supplemented with 10mM NH₄H₂PO₄; Cornmeal agar (1.7 % (w/v)); PDA pH8 (3.9 % PDA, pH adjusted with 0.29 % (w/v) Na₂CO₃ and 0.76 % NaHCO₃); PDB pH3 (2.4 % (w/v) PDB, pH adjusted with 1.67 % (w/v) citric acid and 0.58 % (w/v) Na₂HPO₄); and Yeast Extract Peptone Dextrose (YPD, 0.05 % (w/v) peptone, 0.05 % (w/v) yeast extract, 0.5 % (w/v) glucose, 1.8 % (w/v) NaCl). With the exception of V8 agar, the cultures were kept in constant light at 25 °C for 7 d.

RNA was isolated using the Qiagen QIAzol Lysis Reagent and on-column DNase treatment and RNA purification was performed with the Qiagen RNaseq Mini kit, all according to the manufacturer’s specifications. Illumina HiSeq2000 100 bp read-length sequencing was performed on three DNA libraries (paired-end (PE), 3 kb and 8 kb mate-pair). Sequence reads were quality filtered and trimmed using Trimmomatic v. 0.30 (Bolger et al. 2014) using the default parameters. Due to nucleotide content inconsistencies in the PE and 3 kb libraries the 8 kb mate-pair library reads were assembled as single-end reads using Velvet (Zerbino & Birney 2008) with optimal k-mer size of 57 nt determined with VelvetOptimizer. Contigs were scaffolded using SSPACE v. 2.0 (Boetzer et al. 2011) and scaffold gaps filled with Gapfiller v. 1.11 (Boetzer &

Table 1. Species and GenBank accessions for the Translation Elongation Factor 1-alpha (TEF 1-alpha) and Internal Transcribed Spacer (ITS) sequences used for Maximum Parsimony inference.

<table>
<thead>
<tr>
<th>Species</th>
<th>TEF 1-alpha</th>
<th>ITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cercospora alchemillicola strain CPC 5259</td>
<td>JX143279.1</td>
<td>JX143525.1</td>
</tr>
<tr>
<td>Cercospora althaeina strain CBS 126.26</td>
<td>JX143282.1</td>
<td>JX143528.1</td>
</tr>
<tr>
<td>Cercospora apii strain CBS 116504</td>
<td>AY840487.1</td>
<td>AY840520.1</td>
</tr>
<tr>
<td>Cercospora beticola strain CBS 116502</td>
<td>AY840496.1</td>
<td>AY840529.1</td>
</tr>
<tr>
<td>Cercospora chinensis strain CBS 132612</td>
<td>JX143334.1</td>
<td>JX143578.1</td>
</tr>
<tr>
<td>Cercospora mercuralis strain CBS 549.71</td>
<td>JX143386.1</td>
<td>JX143627.1</td>
</tr>
<tr>
<td>Cercospora olivascens strain CBS 253.67</td>
<td>JX143391.1</td>
<td>JX143632.1</td>
</tr>
<tr>
<td>Cercospora ricinella strain CBS 132605</td>
<td>JX143405.1</td>
<td>JX143646.1</td>
</tr>
<tr>
<td>Cercospora sp. F JZG-2013 strain CPC 12062</td>
<td>DQ185083.1</td>
<td>DQ185071.1</td>
</tr>
<tr>
<td>Cercospora vignigena strain CBS 132611</td>
<td>JX143493.1</td>
<td>JX143734.1</td>
</tr>
<tr>
<td>Cercospora violae strain CBS 251.67</td>
<td>JX143496.1</td>
<td>JX143737.1</td>
</tr>
<tr>
<td>Cercospora zeae-maydis strain CBS 117755</td>
<td>DQ185084.1</td>
<td>DQ185072.1</td>
</tr>
<tr>
<td>Cercospora zeae-maydis strain CBS 117756</td>
<td>DQ185085.1</td>
<td>DQ185073.1</td>
</tr>
<tr>
<td>Cercospora zeae-maydis strain CBS 117757</td>
<td>DQ185086.1</td>
<td>DQ185074.1</td>
</tr>
<tr>
<td>Cercospora zeae-maydis strain CBS 117758</td>
<td>DQ185087.1</td>
<td>DQ185075.1</td>
</tr>
<tr>
<td>Cercospora zeae-maydis strain CBS 117759</td>
<td>DQ185088.1</td>
<td>DQ185076.1</td>
</tr>
<tr>
<td>Cercospora zeina strain CMW 25445</td>
<td>EU569217.1</td>
<td>EU569225.1</td>
</tr>
<tr>
<td>Cercospora zeina strain CMW 25459</td>
<td>EU569215.1</td>
<td>EU569226.1</td>
</tr>
<tr>
<td>Cercospora zeina strain CMW 25462</td>
<td>EU569210.1</td>
<td>EU569224.1</td>
</tr>
<tr>
<td>Cercospora zeina strain CMW 25467</td>
<td>EU569218.1</td>
<td>EU569227.1</td>
</tr>
<tr>
<td>Cercospora zeina strain CMW 25467 genome</td>
<td>MVDW00000000</td>
<td></td>
</tr>
<tr>
<td>Mycosphaerella thailandica strain CBS 116367</td>
<td>AY840476.1</td>
<td>KF901776.1</td>
</tr>
<tr>
<td>Mycosphaerella thailandica strain CPC 10548</td>
<td>AY840477.2</td>
<td>AY752157.1</td>
</tr>
</tbody>
</table>
Pirovano 2012) using all the sequenced libraries and allowing a 25 % error in library insert sizes. Completeness of the assembly was evaluated using BUSCO (Simao et al. 2015).

Transcriptome sequencing was performed at BGI (HK) using HiSeq2000 100bp PE sequencing. RNAseq reads were mapped to the C. zeina reference genome using TopHat2 (Kim et al. 2013). Manual gene finding of 150 C. zeina genes using Genomeview (Abeel et al. 2012) was guided by BLAST (Altschul et al. 1990) alignments of C. zeae-maydis (JGI) genes to the C. zeina reference genome together with C. zeina transcriptome mapping information for gene and splice site identification. Automated gene prediction was performed using the MAKER (Cantarel et al. 2008) pipeline incorporating the SNAP (Korf 2004) and AUGUSTUS (Stanke et al. 2006) ab initio gene predictors. The SNAP training data were the 150 manually curated genes, while the AUGUSTUS training data comprised of high-confidence predicted genes from MAKER using only SNAP predictions (AED score <0.2).

To verify the species identity of the sequenced strain, the Translation Elongation Factor 1-alpha and ITS sequences for selected Cercospora species (Table 1) were concatenated and aligned with ClustalW (Thompson et al. 1994). The relevant sequences were extracted from the genome assembly using blastn from the C. zeina strain CMW 25467 data, and are listed under the genome assembly accession. The analysis workflow was similar to Meisel et al. (2009). The species relationships were inferred using the Maximum Parsimony method in MEGA7 (Kumar et al. 2016), with confidence at nodes gained using bootstrap analysis (Felsenstein 1985). Branches corresponding to partitions reproduced in less than 50 bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). The MP tree was obtained using the Tree-Bisection-Regrafting algorithm (Nei & Kumar 2000) with search level 1 in which the initial trees were obtained by the random addition of sequences (10 replicates). All positions containing gaps and missing data were eliminated. Two strains of Mycosphaerella thailandica (CBS 116367 and CPC 10548) were used to root the cladogram (Fig. 1).

RESULTS AND DISCUSSION

Assembly of the sequenced reads yielded a draft genome 37 Mb in size in 10 027 scaffolds >200 bp with N50 of 161 kb. The average scaffold size was 4 059 bp with the largest contig of 938 kb. BUSCO evaluation using the Ascomycota dataset yielded a completeness report of C: 95.4 %: (95.4 % Complete and single-copy BUSCOs, D: 0.0 % duplicated BUSCOs, 2.1 % fragmented BUSCOs, M: 2.5 % missing BUSCOs, total 1315 genes evaluated). Branches corresponding to partitions reproduced in less than 50 bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). The MP tree was obtained using the Tree-Bisection-Regrafting algorithm (Nei & Kumar 2000) with search level 1 in which the initial trees were obtained by the random addition of sequences (10 replicates). All positions containing gaps and missing data were eliminated. Two strains of Mycosphaerella thailandica (CBS 116367 and CPC 10548) were used to root the cladogram (Fig. 1).
genes are similar to other genomes in the order Capnodiales (Stanke et al. 2006).

Authors: N. Olivier, Y.-C. Lin, Y. Van de Peer, F. Joubert, B.G. Crampton, V. Swart, B. Bluhm, and D.K. Berger*  
*Contact: Dave.Berger@fabi.up.ac.za

IMA Genome-F 8B

Draft genome sequence of Fusarium pininemorale

INTRODUCTION

The Fusarium fujikuroi species complex (FFSC) represents an assemblage of diverse fungal species. The so-called “American” clade of the FFSC contains known emerging pathogens of many cultivated crops and trees including pine (F. circinatum; Hepting & Roth 1946), maize (F. temperatum; Desjardins et al. 2000), and mango (F. parvisorum; Liew et al. 2016). Fusarium pininemorale, a recently recognized member of this clade, was isolated from diseased pine trees in plantations of Colombia (Herron et al. 2015). Even though this species shares numerous morphological and biological traits with other FFSC pine pathogens, F. pininemorale itself was not found to cause significant disease symptoms on pine as is the case for F. circinatum (Herron et al. 2015).

Overall, little is known regarding the biology and genetics of this species, even less so for the genetic and determinants of host range in the broader “American” clade. The exact genetic basis and genomic processes underlying pathogenicity also remains elusive. The aim of this study was to determine the full genome sequence for F. pininemorale, which will allow further studies to investigate genomic aspects of not only pathogenicity but also of its biology and evolution.

SEQUENCED STRAIN

Colombia: Angela Maria (Santa Rosa), Risaralda 75°360.2100’ W 4°490.1800’ N, Isolated from Pinus tecunumanii, 2012, (CBS 137240 = CMW 25243; FCC 5383 – cultures; Herron et al. 2015).

NUCLEOTIDE ACCESSION NUMBER

The Fusarium pininemorale whole genomic sequence data has been deposited at DDBJ/EMBL/GenBank under the accession NFZR00000000. The version described in this paper is version NFZR00000000.

METHODS

The Fusarium pininemorale isolate was grown on ½ PDA for 7 d and DNA was extracted as described previously (Möller et al. 1992). A pair-end library (350 bp average insert size) and a mate-pair library (5 kb average insert size) were prepared and sequenced using Illumina HiSeq platforms at Macrogen (Seoul, Korea). All sequences had an average read length of 58 bp. Poor quality reads, terminal nucleotides, as well as duplicate reads were removed using CLC Genomics Workbench v. 8.0.1 (CLCBio, Aarhus). De novo genome assembly was performed using ABYSS v. 1.3.7 (Simpson et al. 2009), further scaffolding was performed using SSPACE v. 2.0 (Boetzer et al. 2011) and gapped genomic regions were closed using GapFiller v. 1.11 (Boetzer & Pirovano 2012). Genome completeness was assessed using the software BUSCO (Benchmarking Universal Single-Copy Orthologs) v2 with the Sordariomycete dataset (Simão et al. 2015). Chromosome-sized scaffolds were compared using LASTZ alignments (Harris 2007) against genomes of F. circinatum (Wingfield et al. 2012) and F. temperatum (Wingfield et al. 2015a). Lastly, gene prediction was performed using AUGUSTUS 2.5.5 (Hoff & Stanke 2013) based on gene models for F. graminearum (http://bioinf.uni-greifswald.de/ augustus), together with mRNA sequence data from the F. circinatum genome (Wingfield et al. 2012).

RESULTS AND DISCUSSION

The draft nuclear genome of Fusarium pininemorale had an estimated genome size of 47 778 776 bp. The N50 value was 1 358 616 bp and the GC content was 46.0 %. The assembly consisted of 200 contigs which ranged between 200 bp to 6 144 005 bp in size. The average scaffold size was determined to be 310 252 bp in length of which 13 contigs were larger than 1 000 000 bp. The assembled genome was predicted to encode for 14 640 open reading frames (ORFs) with an average gene length of 1472 bp, and an overall gene density of 306 ORFs/MB was observed. BUSCO suggested that the assembly was 99 % complete (i.e., complete BUSCOs = 99 %; complete and single-copy BUSCOs = 98.2 %; complete and duplicated BUSCOs = 0.8 %; fragmented BUSCOs = 0.6 %; missing BUSCOs = 0.4 %; number of BUSCOs searched = 3725). The taxonomic identity of the genome was confirmed using a phylogenetic analysis using authenticated sequences (Fig. 2).

The GC content and number of identified ORFs is comparable to that of other FFSC genomes (Ma et al. 2010, Wingfield et al. 2012, Wiemann et al. 2013, Wingfield et al. 2015a, 2015b, Niehaus et al. 2016). However, the genome assembly of F. pininemorale is notably larger than those of other species in the “American” clade for which genome sequences are available. In comparison, the genome sizes of F. circinatum and F. temperatum is 4 534 176 bp and 2 319 967 bp smaller than that of F. pininemorale, respectively. Nevertheless, sequence comparisons based on chromosome-sized scaffolds suggests that F. pininemorale harbours 12 chromosomes similar to that of species in the FFSC. The putative 12th chromosome of F. pininemorale is 968 722 bp size and spans two scaffolds in the current assembly. Also, the reciprocal chromosome translocation (involving parts of chromosomes 8 and 11) confirmed previously in F. circinatum and F. temperatum (De
Vos et al. 2014, Wingfield et al. 2015a) is also present in F. pininemorale. Clearly, the addition of the F. pininemorale genome assembly would facilitate detailed studies of genome evolution, host adaption and pathogenicity in the “American” clade as well as the broader FFSC.

Fig. 2. Maximum likelihood (ML) tree based on partial gene sequences of β-tubulin and translation elongation factor 1-α. Sequence alignments were assembled with MAFFT version 7 (Katoh & Standley 2013). The program jModelTest v 2.1.7 (Guindon & Gascuel 2003, Darriba et al. 2012) was used to determine the best-fit substitution model (GTR+I+G substitution model) with gamma correction (Tavaré 1986). The ML phylogenetic analysis was performed using PhyML v 3.1 (Guindon & Gascuel 2010). Values at branch nodes are the bootstrapping confidence values with those ≥ 85 % shown. Indicated in bold are sequences from the Fusarium genomes of F. pininemorale and those of F. nygamai, F. temperatum and F. circinatum strain GL1327 which have been previously published in IMA fungus (Wingfield et al. 2012, 2015a, 2015b).

Draft genome sequence of Hawksworthiomyces lignivorus

INTRODUCTION

Hawksworthiomyces lignivorus (Ascomycota: Ophiostomatales) was first described from decaying Eucalyptus utility poles in South Africa (De Meyer et al. 2008). It was initially described in the genus Sporothrix which, under the dual nomenclature system, accommodated mycelial asexual morphs of Ophiostomatales (De Beer & Wingfield 2013,
De Hoog 1993). However, DNA sequences of multiple genes and more inclusive phylogenies showed that that this species was distinct from all other species residing in Sporothrix as redefined by De Beer et al. (2016a), and from other genera in Ophiostomatales (De Beer & Wingfield 2013). De Beer et al. (2016b) therefore described the new genus Hawksworthiomyces, to accommodate H. lignivorus together with four other newly described species, H. crousii, H. hibbettii, H. sequentia, and H. taylorii.

Species of Hawksworthiomyces differ from all other ophiostomatoid fungi in their biology and ecology. Most species residing in the numerous genera of Ophiostomatales are associates of bark or ambrosia beetles that inhabit the cambium and sapwood of trees (De Beer & Wingfield 2013). Many species of Sporothrix are exceptions to this norm and are found in soil and Protea infructescences (De Beer et al. 2016a, Roets et al. 2013). All species of Hawksworthiomyces have either been isolated from decaying wood or environments associated with degrading plant materials (De Beer et al. 2016b). An inoculation study conducted by De Beer et al. (2006) also suggested that H. lignivorus is capable of degrading lignocellulose components of wood. This is different to other species in Ophiostomatales that specifically degrade resinous compounds in wood (Farrell et al. 1993) and that do not display any marked lignocellulose degradation (Seifert 1993).

In this study, we generated the draft genome sequence for H. lignivorus, which will provide basal data required to explore its unique biology and ecology. As the type species of Hawksworthiomyces, the genome sequence of H. lignivorus will be useful for future phylogenetic studies aimed at a better understanding the evolutionary history of this genus and other genera in Ophiostomatales.

RESULTS AND DISCUSSION

Sequencing of the Hawksworthiomyces lignivorus DNA libraries yielded 14 203733 paired-end reads with average read length of around 100 bp. Trimming recovered 12 704 558 pair-end reads and 1 362 623 single reads. De-novo genome assembly with SPAdes resulted in an assembly of 43.81 Mb in size, distributed in 280 scaffolds larger than 500 bp. The number of scaffolds was further reduced to 214 after scaffolding with SSPACE and filling gaps with GapFiller. The current genome assembly of H. lignivorus has a total sequence length of 43 822 585 bases, with an N50 value of 383 563 and an average GC content of 51.27 %. Hawksworthiomyces lignivorus had the largest genome size when compared to all species of Ophiostomatales for which whole genome data are available; the smallest genome reported was that of Ceratocystiopsis minuta (21.3 Mb) (Wingfield et al. 2016a), and the second largest to H. lignivorus was that of of Sporothrix pallida (37.8 Mb) (D’Alessandro et al. 2016).

The assembly had a BUSCO completeness score of 95.7 %. Out of the 3725 BUSCO groups searched, 3556 were complete single-copy BUSCOs, nine were complete duplicated BUSCOs, 64 were fragmented BUSCOs, and 96 were missing BUSCOs. AUGUSTUS predicted a total of 11 216 protein-coding genes encoded by H. lignivorus genome. The taxonomic identity of the genome was confirmed using a phylogenetic analysis using authenticated sequences (Fig. 3).

Authors: H.-J. Lim, B.D. Wingfield, M.J. Wingfield, Z.W. De Beer, and T.A. Duong*

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

IMA Genome-F 8D

Draft genome assembly for Huntiella decipiens

INTRODUCTION

The family Ceratocystidaceae as defined by De Beer et al. (2014) includes economically important plant pathogens, as
**Fig. 3.** Identity verification of Ophiostomataceae isolates sequenced in this study and in all previous IMA Genome Announcements (IMA Genome-F: 3–7; Van der Nest et al. 2014a, Wingfield et al. 2015a, 2016a, 2015b, 2016b). Gene regions (LSU, βT) used for verification were extracted from assembled genomes. Other reference isolates and their corresponding sequences were obtained from published papers (De Beer et al. 2016, Linnakoski et al. 2012, Yin et al. 2015, Zipfel et al. 2006). The phylogeny was constructed using RAxML with the GTRGAMMA substitution model. Bootstrap values greater than 75 are indicated at the nodes.

well as agents of blue stain in timber, many of which result in substantial economic losses (Roux et al. 2000, Baker et al. 2003, Barnes et al. 2003, Van Wyk et al. 2007, Heath et al. 2009). De Beer et al. (2014), Fig. 4, revised this fungal family based on morphological, phylogenetic and ecological evidence and it now includes numerous clearly circumscribed genera such as Ceratocystis, Endoconidiophora, and Huntiella.

Huntiella species previously formed part of the Ceratocystis moniliformis complex (De Beer et al. 2014). Species of Huntiella are generally weak pathogens or saprobes, although some cause sapstain which reduces the value of timber (Van Wyk et al. 2006, Kiritsis et al. 2013). These fungi are associated with insects, particularly sap-feeding beetles (Nitidulidae) that are thought to primarily aid in their spread and distribution (Kiritsis 2004). An unusual example of a Huntiella species is H. bhutanensis that lives in association with the bark beetle Ips schmutzenhoferi (Scolytidae) (Van Wyk et al. 2004, Kiritsis et al. 2013). In general, however, little is known regarding the biology or ecology of Huntiella species. For example, H. decipiens, that forms the basis of the present study, is known only from one region in the Limpopo Province of South Africa, where it was isolated from wounds on plantation-grown Eucalyptus...
species and from a Staphilinid (Staphylinidae) beetle found on freshly cut E. saligna stumps (Nkuekam et al. 2012).

The aim of this study was to produce a good quality draft genome assembly for H. decipiens. The genomes of several members of Ceratocystidaceae are already available in the public domain (Belbahri 2015, Van der Nest et al. 2014a, 2014b, Wilken et al. 2013, Wingfield et al. 2015a, 2015b, 2016) and the genome sequence for H. decipiens will provide valuable opportunities for comparative genomic studies on this important group of fungi.

**Fig. 4.** Identity verification of Ceratocystidaceae isolates sequenced in this study, and in all previous IMA Genome Announcements (IMA Genome-F: 1–7; Van der Nest et al. 2014a, 2014b, Wilken et al. 2013, Wingfield et al. 2015a, 2016a, 2015b, 2016b). Gene regions (60S, LSU, MCM7) used for verification were extracted from assembled genomes. Other reference isolates and their corresponding sequences were obtained from De Beer et al. (2014). The phylogeny was constructed using RAxML with the GTRGAMMA substitution model. Bootstrap values greater than 75 are indicated at the nodes.
**RESULTS AND DISCUSSION**

The estimated size of the assembled *Huntiella decipiens* draft genome was 26.7 Mb, with 638 scaffolds larger than 500 bases. AUGUSTUS analysis predicted 7254 ORFs, which corresponds to an average gene density of 271.7 ORFs/Mb. The assembly contained 1 403 Complete Single-Copy BUSCOs, 90 Complete Duplicated BUSCOs, 11 Fragmented BUSCOs and 24 missing BUSCOs. The taxonomic identity of the genome was confirmed using a phylogenetic analysis using authenticated sequences (Fig. 4).

Relative to other species in *Ceratocystidaceae*, *H. decipiens* has a similar size genome than those of *H. omanensis* (31.5 Mb, 8395 ORFs), *Ceratocystis manginecans* (31.7 Mb, 7494 ORFs), *C. fimbriata* (29.4 Mb, 7266 ORFs), *E. lariicola* (33.3 Mb, 6 897 ORFs), and *D. virescens* (33.7 Mb, 6953 ORFs) (Van der Nest *et al.* 2014a, 2014b, Wilken *et al.* 2013, Wingfield *et al.* 2015, 2016). Whether these differences in genome size are linked to the different lifestyles of these fungi requires further research, which will be the subject of future studies.

**SEQUENCED STRAIN**


**NUCLEOTIDE SEQUENCE ACCESSION NUMBER**

The *Huntiella decipiens* isolate CMW 30855 Whole Genome Shotgun project has been deposited in GenBank under accession no. NETU00000000.

**MATERIALS AND METHODS**

Genomic DNA was isolated from *H. decipiens* isolate CMW 30855 (Barnes *et al.* 2001) and sequenced using the Genomics Analyzer IIx, Illumina platform from the UC Davis Genome Centre (University of California, Davis, A). For this purpose, paired-end libraries of 350- p and 600 bp insert sizes were prepared and sequenced following the Illumina protocol. The CLC Genomics Workbench v. 8.0.1 (CLCBio, Aarhus) was employed to quality trim reads and de-novo assemble a draft genome sequence using the default parameters. Thereafter, the assemblies were scaffolded using SSPACE v. 2.0 (Boetzer *et al.* 2011). GapFiller v. 2.2.1 (Boetzer & Pirovano 2012) was used to fill the gaps created during the scaffolding. The number of putative open reading frames (ORFs) was predicted with the web-based de novo gene prediction software AUGUSTUS using the *Fusarium graminearum* gene models (Stanke *et al.* 2008). The "create detailed mapping report" command of the CLC Genomics Workbench was used to produce statistics for the draft sequence. The Benchmarking Universal Single-Copy Orthologs (BUSCO v. 1.22) tool was used to assess the genome completeness (Simão *et al.* 2015) using the fungal data set.

**IMAGENOME-F 8E**

**Draft genome sequence of *Ophiostoma ips***

**INTRODUCTION**

*Ophiostoma ips* is an ascomycete fungus of the *O. ips* species complex (*Ophiostomatales*, *Ascomycota*) (De Beer & Wingfield 2013). The fungus was first described as the causal agent of blue stain on *Pinus* lumber in the USA (Rumbold 1931). *Ophiostoma ips* is commonly associated with conifer-infesting bark beetles in the genera *ips*, *Orthotomicus*, and *Hylurgus*, which are native to the Northern Hemisphere. Outside of its native range, *O. ips* has been reported in various countries of the Southern Hemisphere including Australia, Chile, New Zealand, and South Africa, where it has been accidentally introduced with its bark beetle vectors (Wingfield & Marasas 1980, Zhou *et al.* 2001, 2004). A population genetic study using microsatellite markers (Zhou *et al.* 2007) revealed a high level of admixture between *O. ips* populations, suggesting frequent movement of this species between countries and continents. Although there have been some studies suggesting weak pathogenicity on conifers (Lieutier *et al.* 1989, Zhou *et al.* 2002), like most other *Ophiostoma* spp. and their relatives (Six & Wingfield 2011), *O. ips* is generally considered not to be a primary pathogen and its relevance is usually as a consequence of the reduced value of the wood due to sap stain (Seifert 1993).

We generated the genome sequence for *O. ips*, which adds to a growing number of genomes available from species of *Ophiostomatales*. Together, these genomes will serve as a valuable resource, enabling future comparative genomics studies seeking to gain insight into the biology, ecology and evolution of species in *Ophiostomatales*.

**SEQUENCED STRAIN**

NUCLEOTIDE SEQUENCE ACCESSION NUMBER

The genomic sequence of *Ophiostoma ips* (CMW 19371 = CBS 138721) has been deposited at DDBJ/EMBL/GenBank under the accession NTMB00000000. The version described in this paper is version NTMB01000000.

METHODS

An isolate of *Ophiostoma ips* (CMW 19371 = CBS 138721) was obtained from the culture collection (CMW) of the Forestry and Agricultural Biotechnology Institute, University of Pretoria. DNA was extracted from a single conidium culture following a method described previously (Duong et al. 2013). Two pair-end libraries (350 bp and 550 bp medium insert sizes) were prepared and sequenced using the Illumina HiSeq 2500 platform. The program Trimmomatic v. 0.36 (Bolger et al. 2014) was used for quality and adapters trimming of pair-end reads. The genome was assembled from trimmed reads using Spades v. 3.10 (Bankevich et al. 2012) and was further placed into scaffolds using SSPACE Standard v. 3.0 (Boetzer et al. 2011). Gaps were filled or reduced with GapFiller v. 1.10 (Boetzer & Pirovano 2012). Several runs of genome assembly were conducted using different parameters. Assemblies obtained from these runs were subjected to quality and completeness assessment using the program BUSCO v. 2.0 (Simão et al. 2015) using the dataset for *Sordariomycetes*. The best assembly based on the best BUSCO statistics in term of completeness, was selected and presented in this study. The program AUGUSTUS v. 3.2.2 (Stanke et al. 2006) was used to estimate the number of protein coding genes encoded by the genome using the species model for *Neurospora crassa*.

RESULTS AND DISCUSSION

The genome of *Ophiostoma ips* was assembled into 351 scaffolds. The genome is 25.99 Mb and the assembly has an N50 of 140.6 Kb. The average coverage across the whole genome was 80 times. The assembled *O. ips* genome has an average GC content of 56 %. The genome of *O. ips* has smaller genome size and higher GC genome content when compared to that of other *Ophiostoma* species such as *O. ulmi* (31.5 Mb, GC = 50.02 %; Khoshraftar et al. 2013), *O. novo-ulmi* (31.8 Mb, GC = 50.01 %; Forgetta et al. 2013) and *O. piceae* (32.8 Mb, GC = 52.8 %; Haridas et al. 2013). The taxonomic identity of the genome was confirmed using a phylogenetic analysis using authenticated sequences (Fig. 3). Assessment of the assembly using BUSCO with the *Sordariomycetes* dataset resulted in a completeness of 96.8 % (C:3608 [S:3603, D:3], F:42, M:77, n:3725), indicating that the assembly should cover most of the organism’s gene space. AUGUSTUS prediction using species model for *Neurospora crassa* resulted in 7607 protein coding genes, which is slightly lower than that of *O. ulmi* (8639 genes; Khoshraftar et al. 2013), *O. novo-ulmi* (8640 genes; Comeau et al. 2015), and *O. piceae* (8884 genes; Haridas et al. 2013).

The draft genome sequence from *O. ips* presented in this study will be a valuable addition to a number of genomes already available for species in *Ophiostomatales*. These will enable further studies to better understand this interesting group of fungi.

Authors: T.A. Duong*, M.J. Wingfield, Z.W. De Beer, and B.D. Wingfield
*Contact: Tuan.Duong@fabi.up.ac.za

ACKNOWLEDGEMENTS

Genome sequencing of *Cercospora zeina* was done by the Phillip SanMiguel, Purdue Genomics Core Facility (USA), and was funded by National Research Foundation (NRF) CPRR, South Africa. The genome sequencing assembly and analyses of the Fusarium pininemorale, Haws worthio myces ignivorus, Huntiella decipiens, and *Ophiostoma ips* genomes were co-funded by the Tree Protection Co-operative Programme (TPCP) and the DST-NRF Centre of Excellence in Tree Health Biotechnology (CTHB) at the Forestry and Agricultural Biotechnology Institute (FABI), and the DST-NRF South African Research Chairs Initiative (SARCHI) in Fungal Genomics.

REFERENCES


*CONTACT: Tuan.Duong@fabi.up.ac.za*


