



The Genome of Banana Leaf Blight Pathogen *Fusarium sacchari* str. FS66 Harbors Widespread Gene Transfer From *Fusarium oxysporum*

Yiping Cui¹, Bo Wu^{2*}, Aitian Peng^{1*}, Xiaobing Song¹ and Xia Chen¹

¹ Guangdong Provincial Key Laboratory of High Technology for Plant Protection, Plant Protection Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou, China, ² School of Computing, Clemson University, Clemson, SC, United States

OPEN ACCESS

Edited by:

Hyang Bum Lee,
Chonnam National University,
South Korea

Reviewed by:

Wei-Hua Tang,
Shanghai Institutes for Biological
Sciences (CAS), China
David J. Studholme,
University of Exeter, United Kingdom

*Correspondence:

Bo Wu
aragornwubo@163.com
Aitian Peng
pengait@163.com

Specialty section:

This article was submitted to
Plant Pathogen Interactions,
a section of the journal
Frontiers in Plant Science

Received: 16 November 2020

Accepted: 12 January 2021

Published: 04 February 2021

Citation:

Cui Y, Wu B, Peng A, Song X and
Chen X (2021) The Genome
of Banana Leaf Blight Pathogen
Fusarium sacchari str. FS66 Harbors
Widespread Gene Transfer From
Fusarium oxysporum.
Front. Plant Sci. 12:629859.
doi: 10.3389/fpls.2021.629859

Fusarium species have been identified as pathogens causing many different plant diseases, and here we report an emerging banana leaf blight (BLB) caused by *F. sacchari* (*Fs*) discovered in Guangdong, China. From the symptomatic tissues collected in the field, a fungal isolate was obtained, which induced similar symptoms on healthy banana seedlings after inoculation. Koch's postulates were fulfilled after the re-isolation of the pathogen. Phylogenetic analysis on two gene segments and the whole genome sequence identified the pathogen belonging to *Fs* and named as *Fs* str. FS66. A 45.74 Mb genome of FS66 was acquired through *de novo* assembly using long-read sequencing data, and its contig N50 (1.97 Mb) is more than 10-fold larger than the previously available genome in the species. Based on transcriptome sequencing and *ab initio* gene annotation, a total of 14,486 protein-encoding genes and 418 non-coding RNAs were predicted. A total of 48 metabolite biosynthetic gene clusters including the fusaric acid biosynthesis gene cluster were predicted *in silico* in the FS66 genome. Comparison between FS66 and other 11 *Fusarium* genomes identified tens to hundreds of genes specifically gained and lost in FS66, including some previously correlated with *Fusarium* pathogenicity. The FS66 genome also harbors widespread gene transfer on the core chromosomes putatively from *F. oxysporum* species complex (FOSC), including 30 involved in *Fusarium* pathogenicity/virulence. This study not only reports the BLB caused by *Fs*, but also provides important information and clues for further understanding of the genome evolution among pathogenic *Fusarium* species.

Keywords: banana leaf blight, *Fusarium sacchari*, *de novo* assembly, comparative genomics, gene transfer

INTRODUCTION

Bananas (*Musa* spp.) are one of the most important crops in the world that has been widely planted as staple food or fruit in many tropical and subtropical regions. China is one of the largest banana production countries, and most banana cultivars planted in China belong to the Cavendish subgroup of the AAA banana cultivar group (*Musa acuminata*). In recent years, the production

of banana in China has been affected by several diseases, especially the catastrophic Panama disease (or *Fusarium* wilt) caused by *Fusarium oxysporum* f. sp. *cubense* (*Foc*) tropical race 4 (Ploetz, 2006; Dong et al., 2013), and new diseases have been keeping emerging (Fan et al., 2016; Du et al., 2017; Zhou et al., 2017). Immediate report of new diseases and understanding of their characteristics are essential for controlling their damage to the banana industry.

The genus *Fusarium* includes more than 20 monophyletic species complexes and several monotypic lineages involving many destructive plant pathogens (O'Donnell et al., 2015). In this study, the pathogen *Fusarium sacchari* (*Fs*) of the reported banana leaf blight (BLB) disease belongs to the *Fusarium fujikuroi* species complex (FFSC) which includes approximately 50 species (O'Donnell et al., 2015). FFSC is phylogenetically close to *Fusarium oxysporum* species complex (FOSC) which includes *Foc* (Maryani et al., 2019). *Fs* could cause growth disease on sugarcane (Bao et al., 2020; Yao et al., 2020) and oil palm (Suwandi et al., 2018), and postharvest fruit rot on lady finger banana (Riolo et al., 2020). Several other studies suggested that different *Fs* strains could harbor distinct pathogenicity toward bananas. An *Fs* isolate was isolated from banana fruit in 2013, but its pathogenicity was untested (Zeng et al., 2013). In another study, *Fs* and 10 other species were isolated from rot banana fruit, and only *Fs* isolates could not induce fruit rot after inoculation on healthy fruit (Abd Murad et al., 2017). Two FFSC isolates, *F. verticillioides* str. M108 and *Fs* str. M7, were isolated from pseudo-stem and leaves of *Musa* spp. ABB plants with Panama disease in Southern Mexico, respectively, but no evidence robustly supported they could cause symptoms through root inoculation (Maldonado-Bonilla et al., 2019). Overall, to our knowledge, *Fs* has not been related to any banana growth disease yet. Except for *Foc* and *Fs*, at least seven other *Fusarium* species have been linked to distinct banana diseases (Jones, 1997; Du et al., 2017). The large number of pathogenic *Fusarium* species and frequently identified new pathogenic strains toward bananas suggest that many *Fusarium* species have the potential to develop pathogenicity toward bananas.

Gene transfer through horizontal gene transfer (HGT) or introgression might have contributed to the development of new pathogenicity in the *Fusarium* genus. In recent years, more and more knowledge has been achieved on the impact of HGT on eukaryotic evolution, especially for unicellular organisms (Fitzpatrick, 2012), which refers to the non-sexual transmission of genetic material between organisms in nature (Keeling and Palmer, 2008). HGT could confer nutrition-related fitness (Milner et al., 2019) and modify the pathogenicity and host range (Alexander et al., 2016; Zhang et al., 2019) of fungi. In *Fusarium*, HGT of genetic materials has been observed in multiple species from divergent organisms (Ma et al., 2013; Sieber et al., 2014; Gao et al., 2019; Kim et al., 2020) and between close relatives within the genus (Liu et al., 2019; Bredeweg and Baker, 2020). Because no sexual cycle has been observed in FOSC, the homologous recombination events were all considered as HGT events in the group (Liu et al., 2019), but the possibility of introgression through rare inter-population or inter-species hybridization followed by multiple backcrossings could not be

completely excluded. In FOSC, HGT of effector genes on the lineage-specific accessory or chromosomes has been related to pathogenicity gain and host range alteration (Ma et al., 2010; Laurence et al., 2015; van Dam et al., 2017; Czislowski et al., 2018). Recently, HGT has also been observed on multiple genes located on the core chromosomes within FOSC and between FOSC and FFSC (Liu et al., 2019; Bredeweg and Baker, 2020), suggesting that HGT in the genus could be more widespread than previously expected. HGT and introgression occur through different mechanisms and both cause gene transfer (Schmickl et al., 2017), and HGT could occur among more divergent organisms. When happening between phylogenetically close populations or species, the outcome of HGT and introgression could converge and be indistinguishable. Introgression was even considered as a path of HGT in some literature (Choudhuri, 2014; Bredeweg and Baker, 2020), which is controversial since it conflicts with the definition of HGT. In FFSC, teleomorphs have been observed in only about one-fifth of the species in the group including *Fs* (Leslie et al., 2005; Montoya-Martínez et al., 2019), suggesting that introgression could also have played a role in gene transfer from other species in *Fs* genomes.

In this study, we reported the BLB disease caused by *Fs* str. FS66 for the first time. The pathogen of the disease was confirmed by fulfilling Koch's postulates. We acquired a high-quality genome assembly and the whole-genome gene annotation for *Fs*. Through comparative genomics and phylogenetic analysis, we also revealed the lineage-specific gene gain and loss and gene transfer at the whole-genome scale in the FS66 genome.

MATERIALS AND METHODS

Plant Materials and Pathogen Isolation

Banana leaves showing leaf blight symptoms were collected from three different 'Baxi' banana plants in a banana orchard (113°58' E, 23°34' N) located in Longmen, Guangdong, in July 2016. Pieces (0.5 × 0.5 cm) were cut from the margin of the infected lesions, surface sterilized with 70% ethanol for 30 s, followed by 0.3% NaClO solution for 1–2 min, and then rinsed three times in sterile water. Three to five pieces from each original plant were placed on potato dextrose agar (PDA) medium and incubated in darkness at 25°C. Single spore cultures were obtained from each growing colony and subjected to morphological characterization.

Inoculation of Candidate Pathogen on Banana Seedlings

Only one type of fungal colony was obtained from the last step. One isolate from each plant was subjected to a pathogenicity test. Two different inoculation methods were used to inoculate the isolates on 1-month 'Baxi' banana seedling leaves. In the first method, we dipped the leaves of three replicate seedlings in conidial suspensions (1.0×10^6 CFU/ml) of each isolate for 10 s, and another three seedlings were dipped in distilled water as a control. In the second method, five puncture wounds were made on both sides of the leaves of banana seedlings using sterilized needles. On three replicate seedlings for each isolate, one side of each wounded leaf was inoculated with three conidial

suspensions (1.0×10^6 CFU/ml) drops (5 μ l) of the isolate while the other half were treated with distilled water as a control. The plants from different treatment groups were grown in separate growth chambers at 25°C with a 16 h photoperiod each day. After leaf blight symptoms were observed on the inoculated leaves, the fungus was re-isolated using the same method described in the previous section. The isolates from different plants were later confirmed to be identical and named FS66, which was also inoculated onto five pseudo-stems of 1-month old 'Baxi' banana seedling using both the methods applied on the leaves. The same number of negative controls were inoculated with water. To test if the isolate could invade through banana roots, we applied root inoculation on five 1-month old 'Baxi' banana seedlings as described in the study of Araújo et al. (2017). A pathogenic *Foc* tropical race 4 isolate was inoculated as the positive control, and distilled sterile water was used as the negative control on an equal amount of seedlings. The plants were kept in the greenhouse, and the symptom development was recorded over 2 months.

Extraction of DNA and RNA

DNA mixture of the host plant and the pathogen was extracted from symptomatic banana leaf tissues was carried out using the modified CTAB method (Gawel and Jarret, 1991). Total genomic DNA was extracted from three single conidium cultures from different plants grown on PDA using the AxyPrep Multisource Genomic Miniprep DNA kit (Axygen, New York, NY, United States) following the instruction. RNA for transcriptome sequencing was extracted from one single conidium fungal culture using the AxyPrep Multisource Genomic Miniprep RNA kit (Axygen, New York, NY, United States). The quality of the DNA and RNA were checked using both Nanodrop 2000c Spectrophotometers (Thermo Fisher Scientific, Waltham, MA, United States) and agarose gel electrophoresis.

Detection of Putatively Existing *Foc* in Symptomatic Tissues

We applied the *Foc* detection method described by Li et al. (2012) on both DNA extracted from the symptomatic leaf tissues and that from the fungal culture, except that primers in the multiplex PCR for amplifying the race 1 and race 4 were run in separate tubes. The *Foc* race 1 specific primers are W1805F: 5'-GTTGAGTCTCGATAAACAGCAAT-3', and W1805R: 5'-GACGAGGGGAGATATGGTC-3'. The *Foc* race 4 specific primers are W2987F: 5'-GCCGATGTCTTCGTCAGGTA-3' and W2987R: 5'-CTGAGACTCGTGCTGCATGA-3'. The amplicons were subjected to agarose gel electrophoresis, and PCR amplicons using DNA of our previously obtained *Foc* race 1 and race 4 isolates FOC1 and FOC4 as templates were added as the positive control.

Identification of the Pathogen by Molecular Phylogeny

To determine the taxon of the pathogen, we sequenced a segment of the internal transcribed spacer (ITS) and partial coding sequence of the second-largest subunit of RNA polymerase II

(*RPB2*) of three isolates from three different plants. The primers used for ITS amplification were ITS1 and ITS4 (White et al., 1990), and *RPB2*-7cR and *RPB2*-5F (Reeb et al., 2004) were applied in amplification of the *RPB2* segment. Each amplification reaction included 25 μ l of 2 \times EasyTaq PCR SuperMix (TransGen Biotech, China), 2 μ l (10 mM) of each primer, and 17 μ l of double-distilled water in a final volume of 50 μ l. The amplification program was set as follows: an initial denaturation step at 94°C for 5 min, followed by 36 cycles consisting of 30 s at 94°C, 1 min at 54°C, and 5 min at 72°C, and a final elongation step of 10 min at 72°C. PCR reactions were performed on a BIO-RAD T100™ PCR machine (Bio-Rad Laboratories, Hercules, CA, United States). Sanger sequencing of the amplicons was carried out by Sangon Biotech (Shanghai, China).

We searched the sequences of both the segments against the NCBI nt (non-redundant nucleotide) database using BLASTN algorithm (Boratyn et al., 2013), and sequences of a few isolates from *Fs* and several close *Fusarium* species were downloaded for phylogenetic analysis. Multiple sequence alignment was carried out using Clustal X v2.1 (Larkin et al., 2007), and phylogenetic trees were constructed using both Neighbor-joining (NJ) and Maximum likelihood (ML) methods implemented with 1000 rounds of bootstrap tests in MEGA X v10.1 (Kumar et al., 2018).

Optimal Temperature and pH for the Pathogen Growth on PDA

We carried out three replicates for each culture condition in this section. To find the optimal temperature for the growth of FS66, we inoculated it on PDA (pH = 7.0) plates and kept them in incubators set at 5, 10, 20, 25, 30, 35, and 40°C under darkness. To study its tolerance to different pH values, we inoculated it on PDA mediums adjusted to pH 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0 with NaOH and HCl solutions, which were incubated in darkness at 25°C. The diameters of the colonies were measured and recorded for a week after inoculation.

Whole-Genome Sequencing and *de novo* Assembly

The whole genome of FS66 was sequenced using both PacBio sequencing and next-generation sequencing (NGS) by NovoGene (Beijing, China). PacBio sequencing libraries were constructed following the 20 kb SMRTbell™ (Pacific Biosciences, CA, United States) protocol and subjected to sequencing on the PacBio Sequel System (Pacific biosciences, CA, United States). We acquired 1,682,281 PacBio sequencing reads totaling 13.39 Gb, with an average length of 7,959 bp and an N50 of 10,432 bp. For NGS, a library with on average 350 bp insertion size was constructed, and 2 \times 150 bp pair-end sequencing was implemented on an Illumina HiSeq 4000 instrument (Illumina, CA, United States). We acquired a total of 2.18 Gb clean reads by NGS.

De novo assembly using Pacbio reads was carried out with Miniasm v0.3 (Li, 2016). Polishing of the assembly was carried out using Racon v1.4.3 (Vaser et al., 2017) and Pilon v1.23 (Walker et al., 2014) with Pacbio reads and NGS sequencing reads respectively. The mitochondrial genome was manually corrected

by integrating contigs and removing redundant regions caused by the circular topology at the ends. The assembly parameters, including N50, L50, et al., were evaluated using QUAST v5.0.2 (Gurevich et al., 2013).

Transcriptome Sequencing and Assembly

The RNA extracted from the fungal culture was subjected to RNA-seq by NovoGene (Beijing, China). mRNAs were enriched with oligo(dT) before the 350 bp insertion-size pair-end RNA-seq library was constructed. High-throughput sequencing was carried out on an Illumina HiSeq 4000 instrument and a total of 7.35 Gb pair-end reads (2×150 bp) were acquired. The RNA-seq data were mapped to the FS66 assembly by HISAT2 v2.2.1 (Kim et al., 2019) and assembled using both StringTie v2.1.4 (Pertea et al., 2015) and Class2 (Song et al., 2016). Then we picked up the high-quality assembled transcripts with intact coding regions from transcriptome sequencing using Mikado v1.5 (Venturini et al., 2018).

Gene Structure Annotation

Mitochondrial and nuclear genomes were annotated separately for FS66. We predicted the repetitive elements in the genome with the Extensive *de novo* TE Annotator (EDTA) pipeline (Ou et al., 2019) and soft masked the assembly before *ab initio* gene structure annotation. Nuclear protein-encoding genes were predicted using both *ab initio* prediction and evidence-based inference methods. First, the high-quality transcripts assembled in the last step were mapped to FS66 via minimap2 v2.17 (Li, 2018) in long mRNA read mode. Based on the transcript assembly and *Fusarium* protein sequences downloaded from NCBI database, Augustus v3.2.3 (Stanke et al., 2006) was trained in *ab initio* gene prediction for FS66 following the protocol of Hoff and Stanke (2019). Homology based annotation was carried out by GeMoMa v1.7.1 (Keilwagen et al., 2019) using *F. fujikuroi* genome (GCA_900079805.1) as reference. We merged the annotation files from RNA-seq data assembly, homology-based annotation, and *ab initio* prediction, which were further subjected to redundancy removal using GffRead v0.12.3 and GffCompare v0.12.1 (Pertea and Pertea, 2020). Coding regions were inferred from all transcripts using TransDecoder v5.5.0¹. The completeness of the gene structure annotation was assessed with BUSCO v4.0.5 (Simão et al., 2015). The function of the whole-genome genes was annotated by searching against the InterPro database (Hunter et al., 2009) and KEGG database (Kanehisa et al., 2017) through BLAST. GO annotation of all protein-encoding genes was carried out using BLAST2GO v5.2.5 (Götz et al., 2008).

We applied both RNAmmer v1.2 (Lagesen et al., 2007) and Barrnap v0.9² in predicting 5S, 5.8S, 18S, and 28S rRNAs in the genome. The tRNAs were predicted using tRNAscan-SE v2.06 (Chan and Lowe, 2019). Other non-coding RNAs including small nuclear RNA (snRNA), small nucleolar RNAs (snoRNAs), spliceosome RNAs, et al. were annotated via homology analysis

on both sequence and secondary structure by Rfam 14.3 (Kalvari et al., 2018) and Infernal v1.1.2 (Nawrocki and Eddy, 2013). Genes encoding proteins and non-coding RNAs in the mitochondria genome were predicted using MITOS2 web server³ (accessed on 10/03/2020) with the reference set as 'RefSeq 63 Fungi' and genetic code as '4 mold' (Bernt et al., 2013).

All *F. fujikuroi* mating-type proteins (Martin et al., 2011) including MAT 1-1 (G3G2C1, G3G2C2, G3G2C3, O93924, O93924, O93925, Q9C461, Q9URK7) and MAT 1-2 (G3G2C4, G3G2C5, O93926, O94158) types were downloaded from the UniProt database. A local blast database was created using all the annotated proteins in FS66 genome, and all the MAT proteins were used as query sequences to search for their orthologs in FS66.

All pathogenicity or virulence-related *Fusarium* genes in the Pathogen Host Interactions (PHI) database (Urban et al., 2017) were searched against the acquired orthogroups using MMseqs2 v12.113e3 (Steinegger and Söding, 2017), and the best hits were considered as the orthologous groups to the query genes.

Identification of Gene Orthogroups From Multiple *Fusarium* Genomes and Species Tree Construction

The genome sequences of several *Fusarium* species (Ma et al., 2010; Wiemann et al., 2013; Niehaus et al., 2016, 2017; King et al., 2018; Asai et al., 2019; Warmington et al., 2019; Wang et al., 2020) close to FS66 were downloaded from GenBank assembly database (**Supplementary Table 1**). We assessed the completeness of their gene structure annotations with BUSCO v4.0.5. We required the genomes to have no less than 98% tested genes classified as complete to include them in the further analysis except for the *Fs* str. NRRL 66326 (GCA_013759005.1), which was an isolate derived from the mating of two isolates collected from sugarcane in Taiwan and the only available *Fs* genome in the public database. The genome of *F. venenatum* str. A3/5 (GCF_900007375.1) was selected as an outgroup. No gene structure annotation was available for *Fs* str. NRRL 66326, so we predicted the protein-encoding genes in the genome using GeMoMa v1.7.1 with FS66 as reference. Whole-genome alignment between FS66 and NRRL 66326 or FO4287 (GCA_000149955.2) and genetic variation identification were carried out using Mummer v4.0.0.beta5 (Marçais et al., 2018).

Whole-genome protein sequences were obtained for the 12 analyzed genomes either through downloading from NCBI GenBank database or output by GffRead v0.12.3. For genes annotated with more than one alternatively spliced transcripts, only proteins encoded by the main transcripts were used. To identify the orthologous and paralogous (duplicated) relationship among the genes, the OrthoFinder v2.4.0 pipeline (Emms and Kelly, 2019) was then applied to all the proteins. The pipeline includes the following step: (1) the proteins were clustered into orthogroups based on normalized genetic distance. An orthogroup has been defined as homologous genes from extant species which were descended from a single gene in the last

¹<https://github.com/TransDecoder/TransDecoder>

²<https://github.com/tseemann/barrnap>

³<http://mitos2.bioinf.uni-leipzig.de/index.py>

common ancestor of the species; (2) multiple alignments and maximum likelihood tree construction were carried out for each orthogroup; (3) A species tree was constructed by STAG v1.0.0 (Emms and Kelly, 2018) based on all the individual gene trees. The supporting rate of each bipartition was calculated as the proportion of gene trees with the bipartition, which has been supposed to be a more stringent measure than standard bootstrap support (Emms and Kelly, 2018); (4) Gene duplication events were inferred based on the species tree and gene trees.

Gene Gain and Loss

We obtained the number of gene copies in each orthogroup for each genome from the results of the OrthoFinder pipeline. A *t*-test was carried out to check if there was a significant ($p < 0.01$) difference in the gene copy number in each orthogroup between FSCC and FOSC genomes. Gene gain and loss in FS66 compared with other FSCC genomes were analyzed using the following approach. In each orthogroup, the mean value (m) and the standard deviation (s) of gene number in the six FSCC genomes (not including FS66 and NRRL 66326) were calculated. If the number of genes of FS66 in an orthogroup is smaller than $(m - 2 \times s)$ or larger than $(m + 2 \times s)$, gene loss or gene gain was inferred in FS66, respectively.

Gene Cluster Prediction and Comparison

Putative gene clusters involved in secondary metabolites biosynthesis were predicted using antiSmash v5.1.2 (Blin et al., 2019) locally, and the borders of the gene clusters were inferred using the implemented cassis algorithm. The gene clusters from the 12 genomes were merged and formatted into a MultiGeneBlast database, and each predicted gene cluster was searched against the database using MultiGeneBlast v1.1.14 as done by Hoogendoorn et al. (2018).

Inference of Genomic Regions Transferred From Other Species

The FS66 assembly was split into 1 kbp sliding windows with 500 bp overlap across the genome using Bedtools v2.29.2 (Quinlan, 2014). The genomic segments in these windows were aligned using BLASTN to the previously compared 11 *Fusarium* genomes and another three outgroup *Fusarium* genomes, *F. tricinctum* (GCA_900382705.2) (Ponts et al., 2018), *F. graminearum* (GCA_900044135.1) (King et al., 2015), and *F. avenaceum* (GCA_000769295.1) (Lysøe et al., 2014). These genomes except for FS66 were classified into FFSC (seven genomes), FOSC (three genomes), and the outgroup (four genomes). When the best BLASTN hit segment in a genome is longer than 500 bp and shares >50% nucleotide similarity with the query FS66 segment, it was identified as an ortholog of the query. Each query segment from FS66 and all its orthologs found in the 14 genomes were output in an individual fasta file. When at least one ortholog of a segment was available in each of the FFSC, FOSC, phylogenetic incongruence test was carried out to check if there was gene transfer from other groups. Because the FFSC species were phylogenetically close and incomplete lineage sorting could also result in many incongruent gene trees

(Maddison and Knowles, 2006), only gene transfer from FOSC was inferred by this test. To reduce the impact of gene transfer in the other genomes on our analysis, the test was carried out on all the possible ortholog combinations among FS66 and the three groups. The test includes the following steps: (1) one ortholog was randomly selected from each of the three groups and output together with the corresponding FS66 segment into a FASTA file, and all the available combinations were output in separate files for each FS66 window; (2) all the FASTA files were subjected to multiple alignments with MUSCLE v3.8.1551 (Edgar, 2004) and ML tree search using RAXML v8.2.12 (Stamatakis, 2014), and the best tree in 100 iterations of ML tree search was acquired for each FASTA file; (3) the DendroPy v4.4.0 (Sukumaran and Holder, 2010) python phylogenetic package was applied to check if the best ML tree inferred from each FASTA file supported local gene transfer from FOSC in FS66; (4) FS66 genome windows with more than 50% FASTA files supporting gene transfer were considered as genomic regions putatively transferred from FOSC. All the circular graphs in this study were drawn using Circos v0.69-5 (Krzywinski et al., 2009).

RESULTS

Banana Leaf Blight Disease Observed in Guangdong, China

In July 2016, leaf blight was observed on about 20% 'Baxi' (*Musa* spp. AAA Cavendish) plants in a 2 ha banana plantation located in Longmen, Guangdong. Yellow to brown lesions expanded from the leaf blade edge to the inner area were observed on the symptomatic tree (Figure 1A). The symptoms were more often observed on the young leaves than older leaves. A few trees (~5% of the infected trees) developed very severe symptoms and stopped growing months after the symptom first appeared. The disease was continuously observed in the following 3 years, and it was more severe and widespread between early May and late September when the amount of rainfall was relatively high within a year in the region. We dissected the roots and pseudo-stems of five diseased trees with the most severe symptoms, and no discoloration of vessels was observed, showing it was not a wilt disease.

Isolation of a Fungal Pathogen of the Disease

The same type of reddish-brown colonies with abundant white wooly mycelia was obtained ~2 days after the inoculation of the symptomatic tissues from three different trees on PDA (Figures 1B,C). The mycelia were straight, septate, and only a few branches were observed. There were two different types of conidia (Figure 1D): the large conidia were falcate, septal, hyaline, and $11.88 \sim 63.5 \mu\text{m} \times 2.5 \sim 6.2 \mu\text{m}$ in size; the small conidia were elliptical, hyaline, and $5 \sim 10.5 \mu\text{m} \times 2.0 \sim 4.8 \mu\text{m}$ in size. We also observed some ascospores, which were finger-shaped, containing at least one septum, and $11.5 \sim 95 \mu\text{m} \times 3.5 \sim 7.5 \mu\text{m}$ in size (Figure 1D). As shown in Supplementary Figure 1, the optimal temperature and pH of the

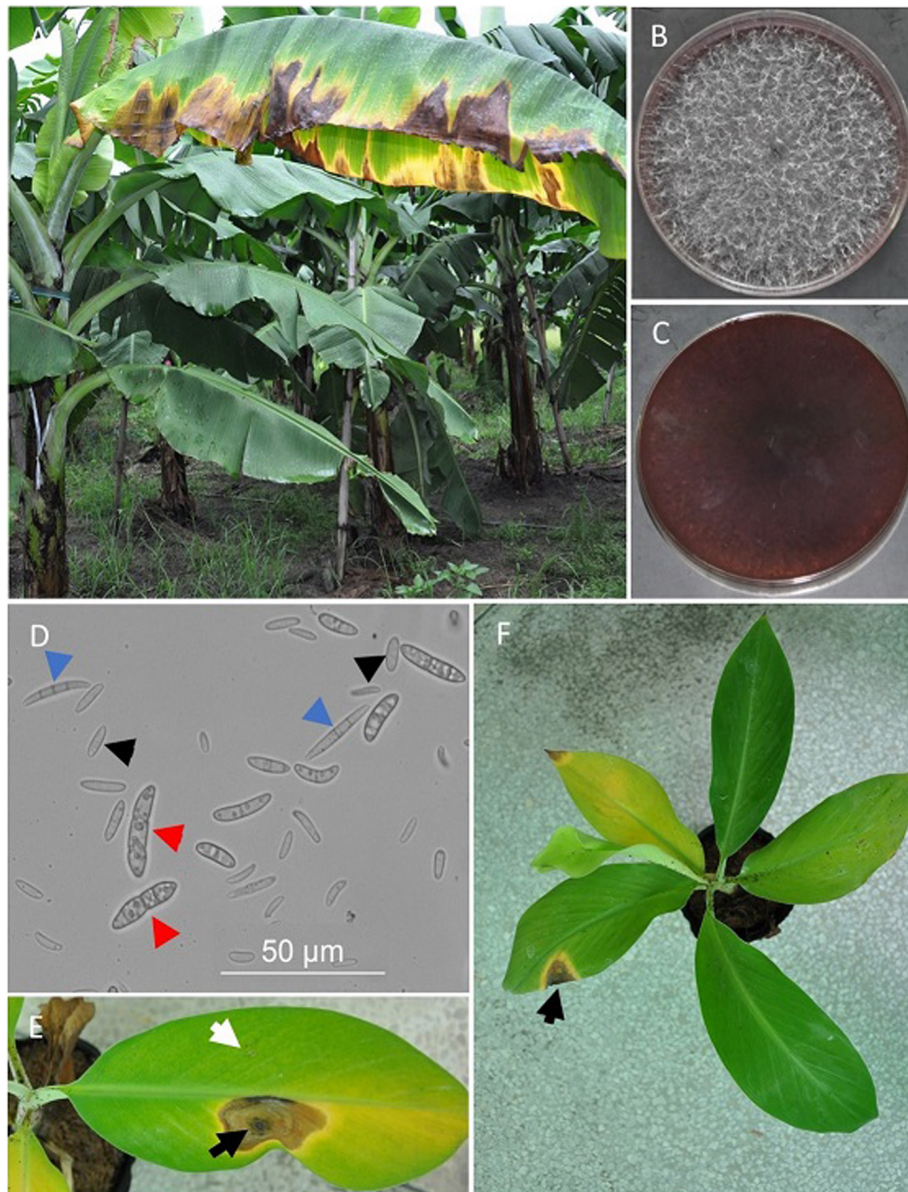
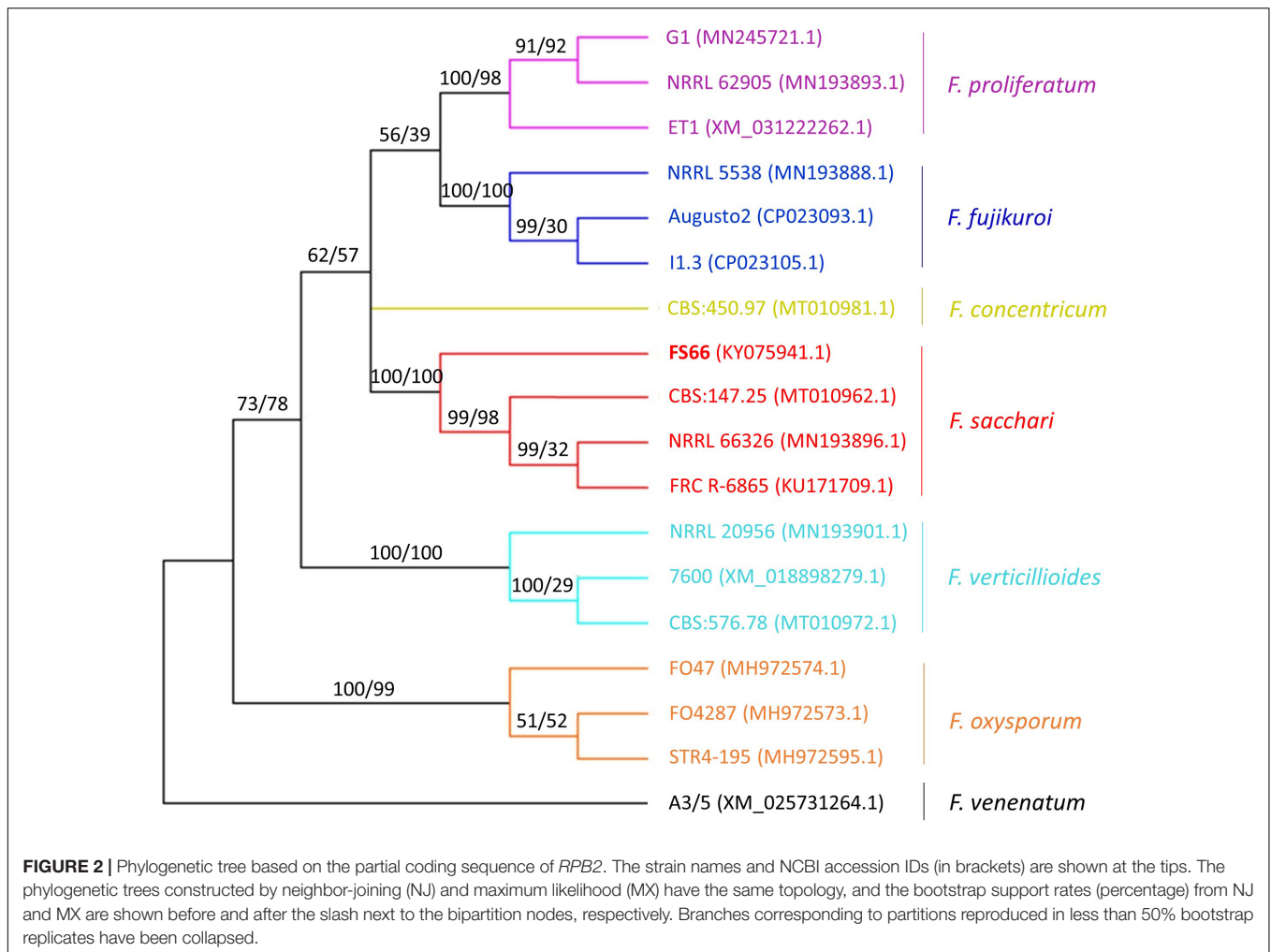


FIGURE 1 | Banana leaf blight disease and pathogen isolation. **(A)** Symptoms of the BLB disease in the field. **(B,C)** The appearances of the fungal colony at 7 days after single-spore inoculation on PDA plates from the front **(B)** and the back **(C)**. **(D)** Morphology of the fungal conidia and spores under the optical microscope. Blue, black, and red arrows denote large conidia, small conidia, and ascospores, respectively. **(E)** Symptoms on leaves inoculated with both the candidate pathogen and sterile water after the puncture. The white and black arrows pointed to the area punctured with needles and inoculated with water and FS66, respectively. **(F)** Symptoms on a banana seedling treated with leaf non-wound inoculation. Pictures **(E,F)** were taken at 30 days after inoculation.

isolates were identified to be around 25°C and 7.0, respectively (**Supplementary Figure 1**).

Three isolates from different trees were inoculated on banana leaves and pseudo-stems through both non-wound and wound inoculations. Both types of tissues with puncture wounds started developing blight symptoms as early as 4 days after inoculation with the candidate pathogen, and the symptoms on leaves and pseudo-stems at 30 days and 45 days after inoculation were shown in **Figure 1E** and **Supplementary Figure 2**, respectively. No symptom was observed on any of the negative controls

inoculated with water. When non-wound inoculation of the isolates was applied to the leaves, only two of the nine seedlings developed leaf blight symptoms (**Figure 1F**). None of the five pseudo-stems subjected to non-wound inoculation developed any symptom. Five banana seedlings were subjected to root inoculation of the candidate pathogen, and no symptom was observed on any of them, while all positive controls inoculated with a pathogenic *Foc* tropical race 4 isolate developed symptoms of the Panama disease as early as 15 days after inoculation.



Identification of the Taxonomy of the Pathogen

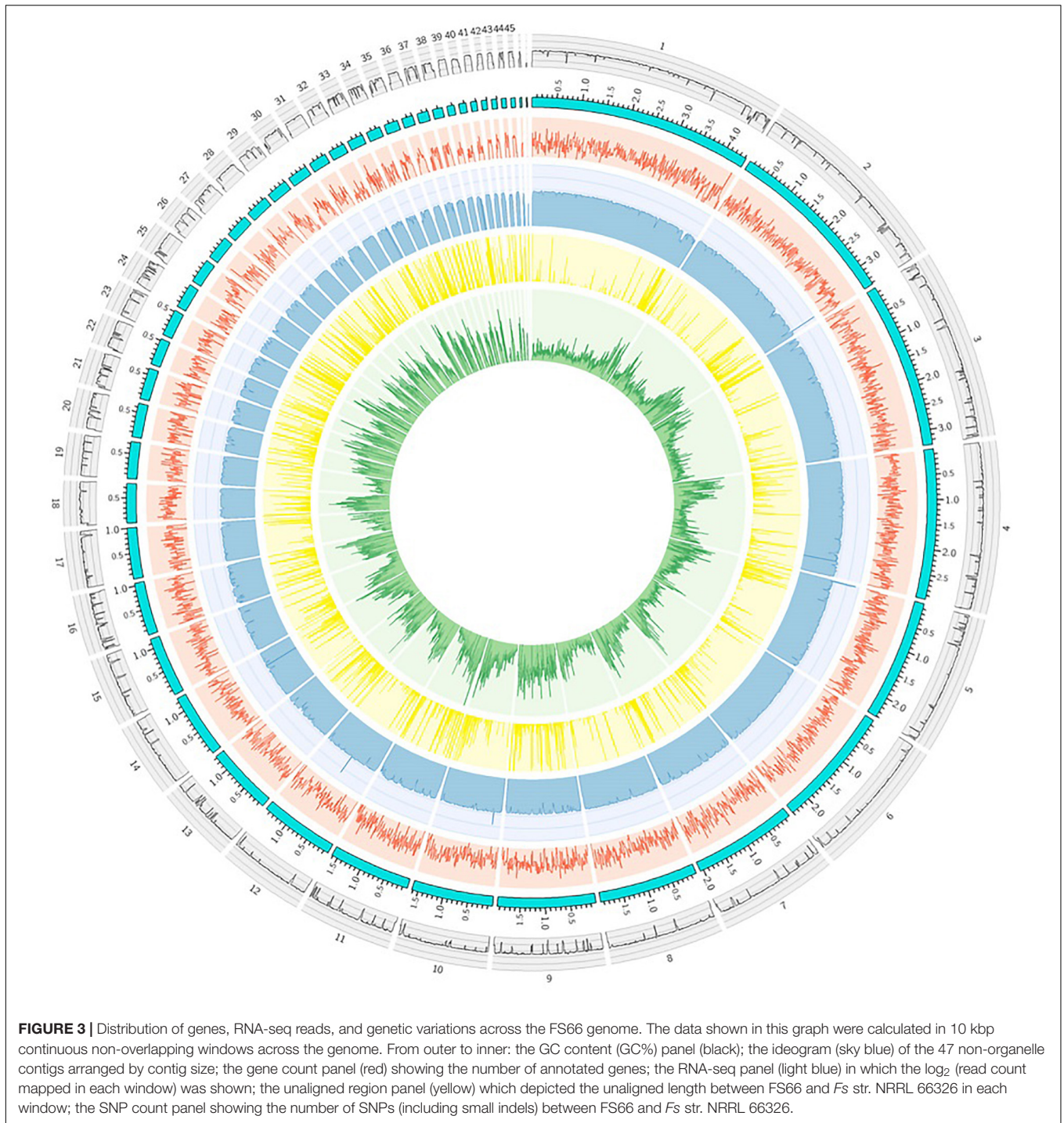
Considering Panama disease had been detected in adjacent orchards, we suspected there could be *Foc* involved in the disease at the beginning and carried out PCR test using both *Foc* race 1 and race 4 specific primers. With either the DNA extracted from symptomatic plant tissues or the fungal isolates as templates, no

Foc target amplicon was obtained using either *Foc* race 1 or race 4 specific primers. Instead, a >1000 bp amplicon was amplified by the race 1 primers, and no (symptomatic plant tissues DNA as templates) or only a very weak amplicon band (fungal culture DNA as templates) shorter than the target amplicon was observed for race 4 primers (**Supplementary Figure 3**).

We sequenced the internal transcribed spacer (ITS, KY075940.1) and segment the second-largest subunit of RNA polymerase II (*RPB2*, KY075941.1) segments of the three isolates from different trees. The results showed that the two segments shared 100% nucleotide similarity among them. Considering the three isolates had nearly the same colony phenotypes and pathogenicity, they were supposed to be the same fungal strain which we named FS66. Through searching in the NCBI nt database, we found that the ITS and *RPB2* segments of FS66 shared the highest nucleotide similarity (both 100%) with *Fs* isolates. Phylogenetic analysis based on ITS showed that FS66 should belong to FFSC, and *RPB2* supported that FS66 was clustered with the *Fs* strains (**Figure 2**). Based on the two segments and the whole-genome assembly in the following section, we concluded that FS66 belongs to *Fs*.

TABLE 1 | Statistics on two *Fusarium sacchari* genome assemblies.

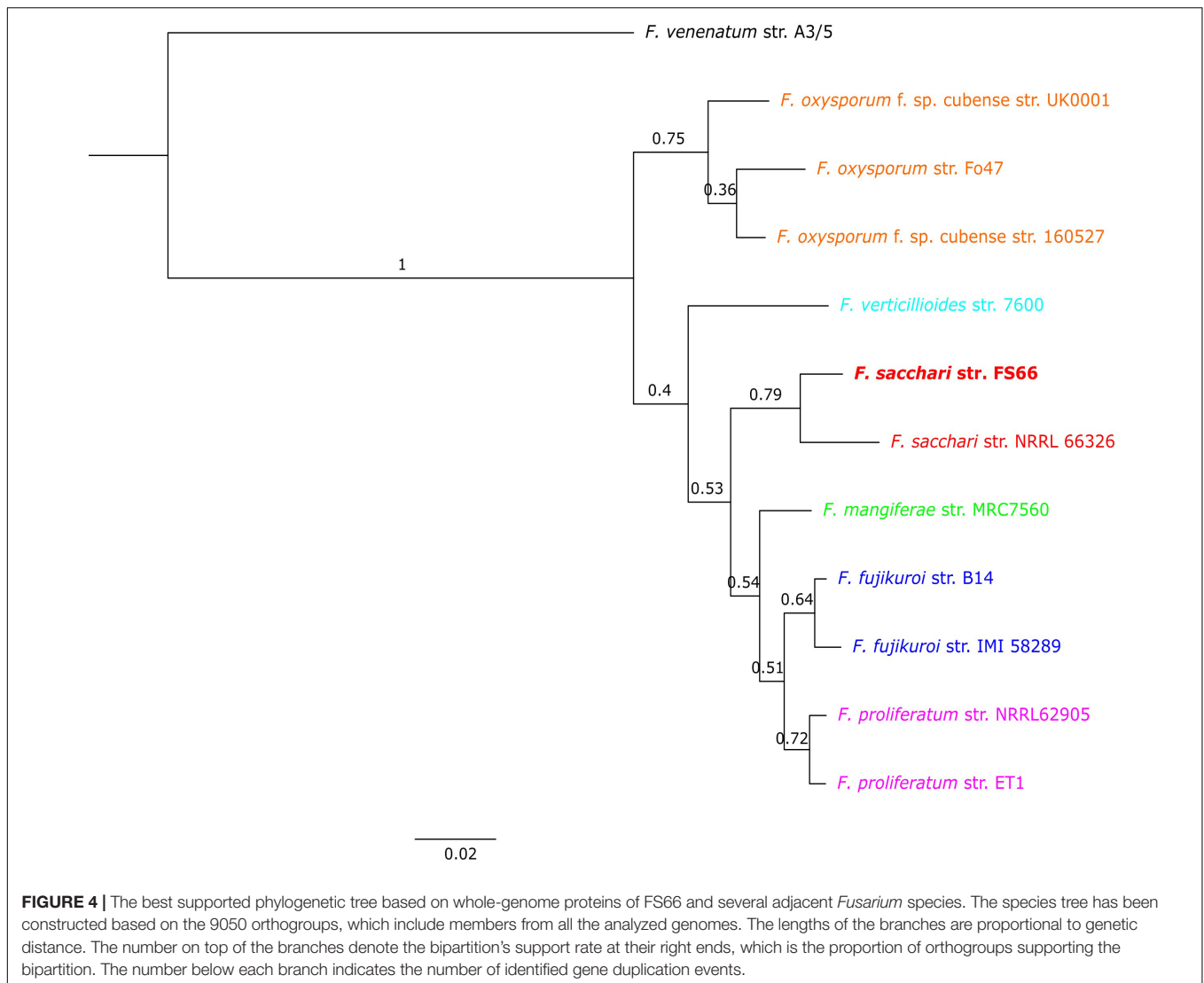
<i>Fs</i> strain name	FS66	NRRL 66326
GenBank ID	JADANP000000000	GCA_013759005.1
Number of contigs	48	479
Largest contig (bp)	4,434,236	997,752
Total length (bp)	45,739,938	42,740,848
N50 (bp)	1,965,716	187,816
N75 (bp)	1,048,338	96,836
L50	8	63
L75	16	142
GC (%)	46.3	49.03



De novo Assembly and Gene Structure Annotation of FS66 Genome

We sequenced *Fs* strain FS66 using both PacBio long-read sequencing ($\sim 290\times$) and NGS ($\geq 40\times$), based on which *de novo* assembly was carried out. The acquired genome consisted of 47 nuclear contigs and a complete 59,755 bp circular mitochondrial genome, totaling 45.74 Mb with an average GC content of 46.30% (Table 1 and Figure 3). The FS66 assembly has high

completeness, and both high mapping rates of whole-genome sequencing reads (97.96%) and RNA-seq reads (93.61%) were achieved. The FS66 assembly is about 2.99 Mb larger than *Fs* str. NRRL 66326 (GCA_013759005.1) (Table 1), and its contig N50 size (1,965,716 bp) is more than 10-fold larger than that of NRRL 66326 (N50 = 187,816 bp). The largest contig of our FS66 assembly is 4,434,236 bp long, and a minimum of 8 and 16 contigs could cover >50% and >75% of the genome,



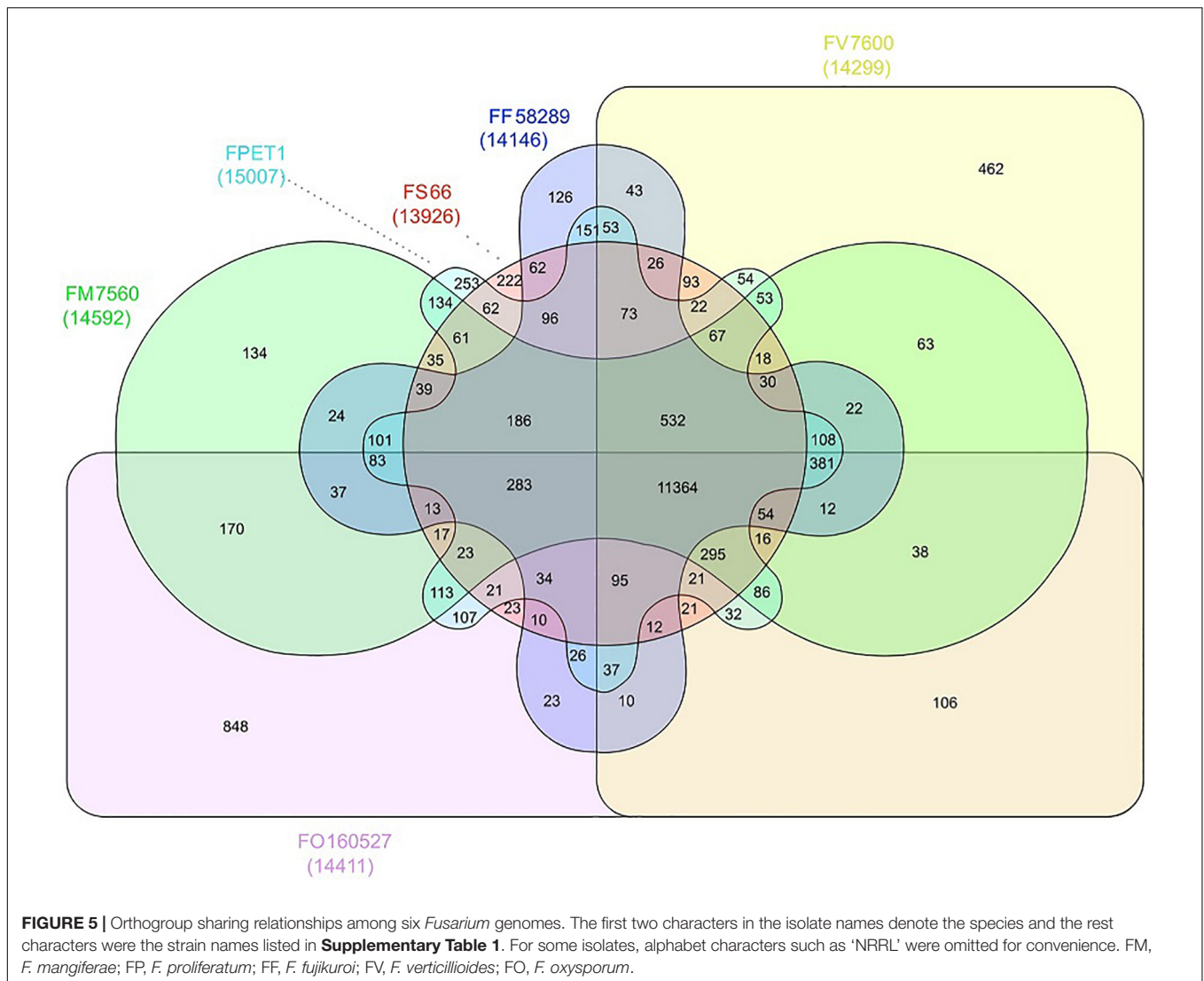
respectively. The whole-genome alignment showed an overall nucleotide similarity of 96.74% between FS66 and NRRL 66326. Only 85.53% of the entire FS66 genome and 91.42% of the NRRL 66326 genome could be aligned with each other, and the single nucleotide variations (SNV) between the two genomes were distributed heterogeneously across the genome (Figure 3). The FS66 assembly was also aligned to the genome *F. oxysporum* f. sp. *lycopersici* 4287 (FO4287), and the four lineage-specific accessory chromosomes in FO4287 (Ma et al., 2010) were almost completely absent (99.7%) in FS66.

We carried out gene structure annotation for FS66 based on transcriptome sequencing and *ab initio* gene prediction. A total of 14,486 protein-encoding genes were predicted in the genome (Figure 3). Assessment by BUSCO showed that 99.5% of tested Hypocreales core genes (4,494) were successfully annotated as complete BUSCOs in our annotation, which is as high as the best genome assemblies in several other *Fusarium* species. Besides protein-encoding genes, 418 non-coding RNAs, including 309 tRNAs, 67 rRNAs, and 42 other non-coding RNAs were predicted

in the genome. The mitochondrial genome was annotated separately, in which 29 protein-encoding genes, 3 rRNAs, and 27 tRNAs were predicted. Gene functional annotation including Gene Ontology (GO), InterPro annotation, and KEGG pathway annotation were obtained for all the nuclear and mitochondrial protein-encoding genes.

Gene Gain and Loss in FS66 Compared With Close *Fusarium* Species

Orthologous gene groups (orthogroup) were inferred among FS66 and 11 other *Fusarium* genomes, including seven from FFSC, three from FO SC, and one from *F. venenatum* as outgroup (Supplementary Table 1). A total of 18,151 orthogroups were detected (Supplementary Table 2), 14,126 (77.8%) of which included genes from no less than six of the genomes (Supplementary Figure 4). There were 9050 orthogroups with all species present and 7725 of these consisted entirely of single-copy genes. A species tree for FS66 and six other *Fusarium* species was

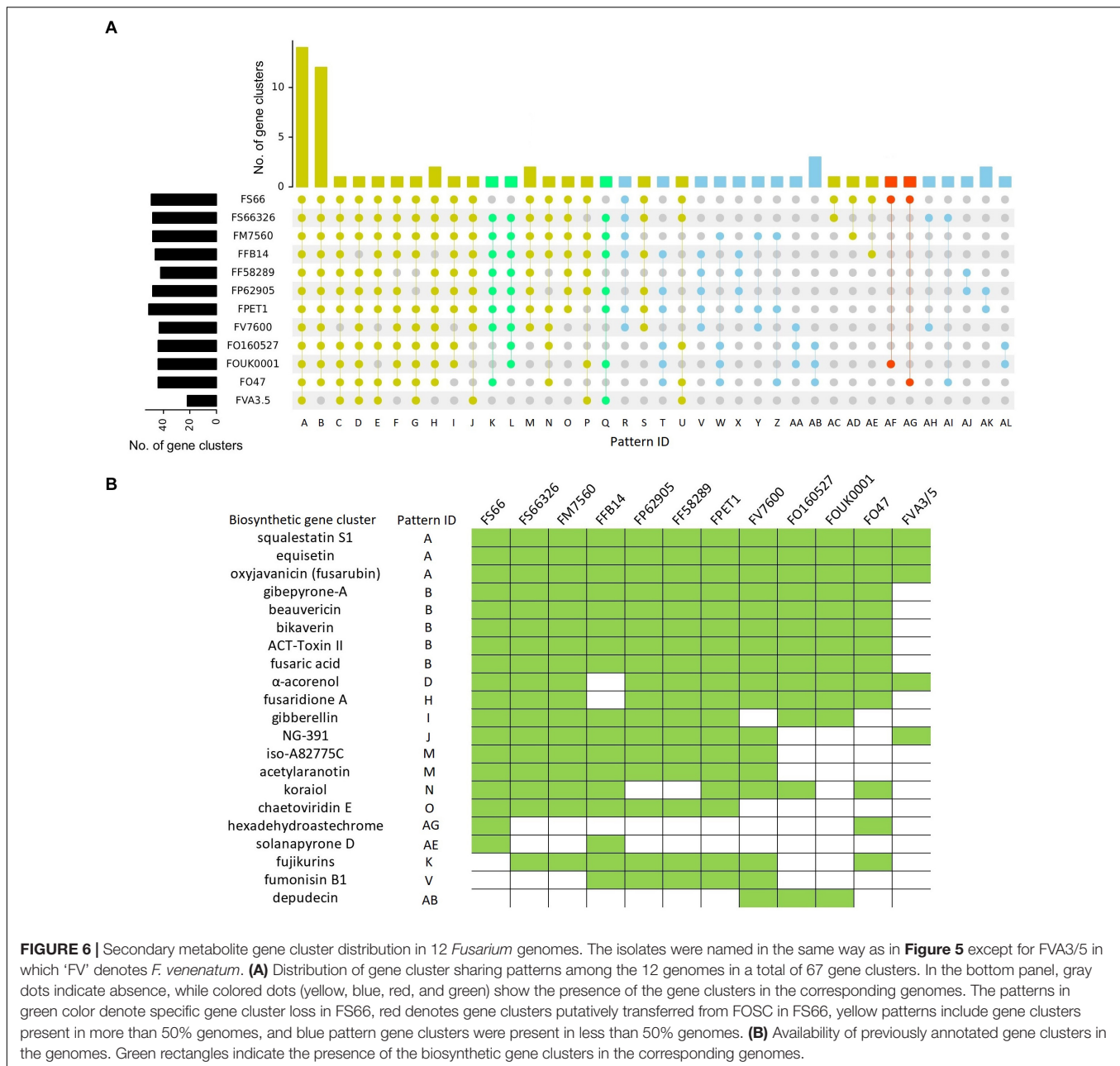


obtained based on the orthogroups, which confirmed that FS66 is the closest to *Fs* str. NRRL 66326 and belongs to FFSC (**Figure 4**).

FS66 has at least one gene in 13,866 orthogroups, and it generally shares more (on average 823.3) orthogroups with FFSC genomes than with FO SC genomes (**Figure 5** and **Supplementary Table 3**). In 347 and 246 orthogroups, FFSC genomes have significantly ($p < 0.01$) fewer and more gene copies than the FO SC genomes, respectively. In comparison to other FFSC genomes, FS66 had gene copy gain in 539 orthogroups and gene copy loss in 778 orthogroups. Through phylogenetic analysis on all individual genes a total of 40 lineage-specific gene duplications were detected in FS66 (**Figure 4**). Four of the orthogroups with gene gain and nine with gene loss in FS66 have been correlated with phytopathogenicity in other *Fusarium* species in the Pathogen Host Interactions (PHI) database (**Supplementary Table 4**). The lost genes included three genes involved in the biosynthesis of fumonisins, FUM6, 8, and 21. No ortholog of mating-type MAT1-1 protein-encoding gene was found in FS66, and two MAT 1-2 type

protein genes, *MAT 1-2-1* and *MAT 1-2-9*, were identified in the genome, indicating that FS66 belongs to the MAT1-2 mating type.

All secondary metabolite biosynthesis gene clusters were predicted in the 12 *Fusarium* genomes. As a result, a total of 67 putative secondary metabolite biosynthesis gene clusters were detected, of which 48 were detected in the FS66 genome (**Figure 6A** and **Supplementary Table 5**). Among the 67 gene clusters, 21 are homologous to previously annotated gene clusters in the Minimum Information about a Biosynthetic Gene cluster (MIBiG) database (**Figure 6B**), including 18 existing in the FS66 genome. Three gene clusters present in *Fs* str. NRRL 66326 and at least one other FFSC genome were absent in FS66, including the annotated fujikurins biosynthesis gene clusters (BGC0001305). As shown in **Figures 6A,B**, at least three gene clusters were putatively transferred from FO SC or more divergent species to *Fs* (black arrow) or FS66 (red arrows), including the previously annotated hexahydroastechrome biosynthetic gene cluster (BGC0000372).

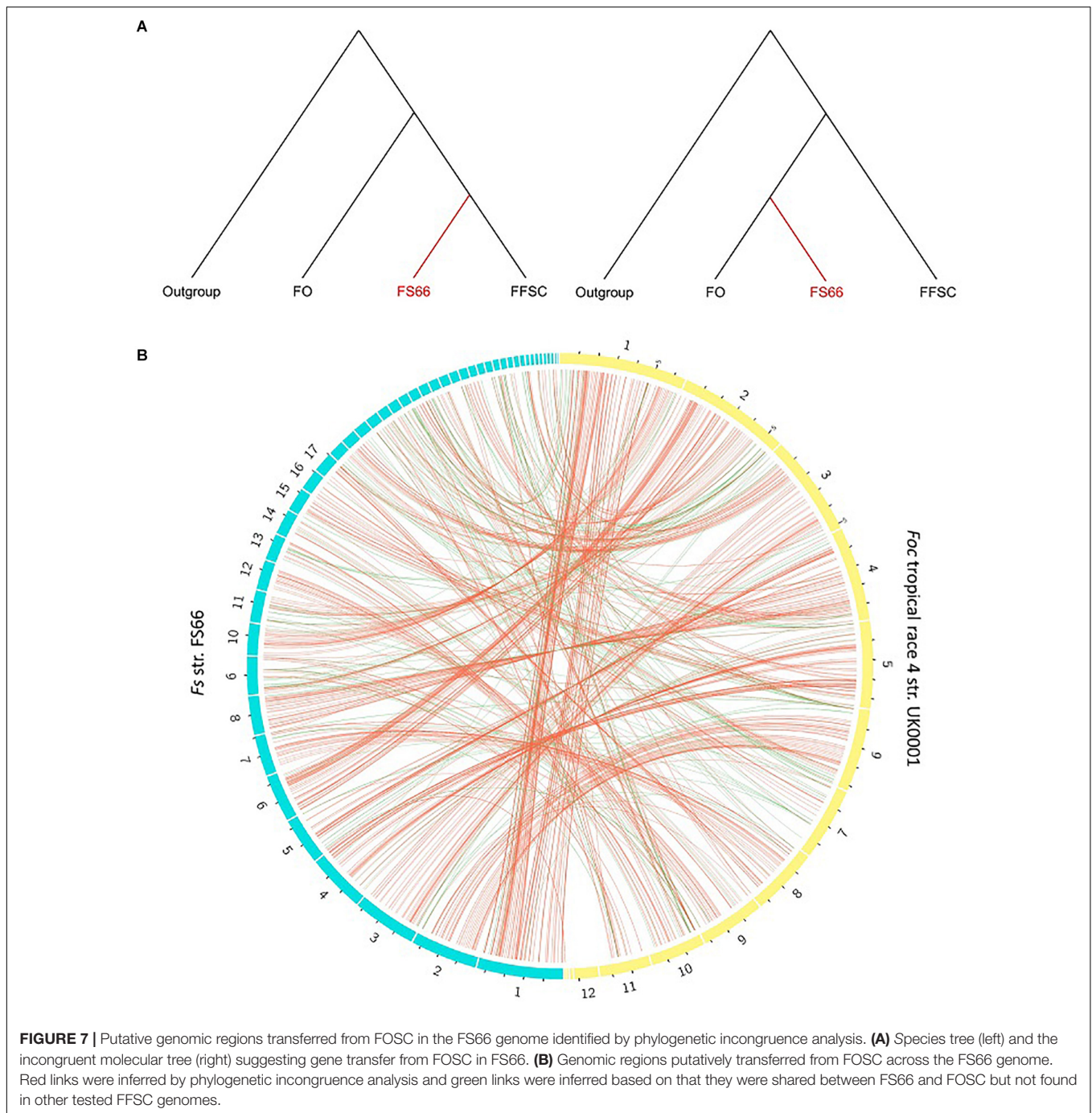


FS66 Genomic Regions Putatively Transferred From Other Species

Genomic regions putatively transferred from FOSC or other non-FFSC species in the FS66 were investigated. The whole-genome of FS66 was divided into 1 kbp windows and searched against the above 11 *Fusarium* genomes and 3 additional outgroup *Fusarium* genomes. Homologous regions of 92.3% (84,341/91,383) of the segments were detected in at least one genome. FS66 shared 65,399 to 78,302 homologous segments with the FFSC genomes, 64,732 to 64,996 segments with the FOSC genomes, and only 11,522 to 15,505 segments with the outgroup genomes. A total of 56 homologous segments (38 regions) involving 11 genes

(three fully covered and eight partially covered) were only present in the outgroup.

On 351 segments (221 regions), the homologs were only detected in the FOSC group but not in any of the other FFSC genomes or the four outgroup genomes, suggesting they could be derived from FOSC or its close relatives. In **Figure 7B**, 207 of the 351 segments have been shown as green links. On 21,827 FS66 genomic segments on which homologs were identified, genomic segments in FS66 putatively introduced from FOSC were further inferred based on phylogenetic incongruence between the species tree and the molecular trees (**Figure 7A**). Incongruence was detected on additional 555 segments (497 discontinuous regions), of which 546 have been shown as red links in **Figure 7B**. All



these 960 segments which were putatively introduced from FOSC accounted for 814,500 bp in the genome, covering 27 genes entirely and 529 genes partially. Orthologs of 30 of the overlapped genes have been associated with pathogenicity and/or virulence in at least one *Fusarium* species (**Supplementary Table 6**), including two genes *FoSl2* and *CHS2* reported to be related to the virulence of *F. oxysporum*. Considering only 23.9% (21827/91383) of genomic segments were subjected to the phylogenetic analysis due to lack of enough outgroup segments, there still could be many un-detected transferred regions in FS66.

DISCUSSION

In this study, we reported the BLB disease caused by *Fs* str. FS66 in the world for the first time, which is another disease on banana caused by a *Fusarium* pathogen. Previously, *Fs* strains were isolated from banana fruit and leaves in Asia, but they either had no (Abd Murad et al., 2017) or untested (Zeng et al., 2013) pathogenicity, suggesting the FS66 strain could have evolved the pathogenicity toward banana only recently. We not only completed the *de novo* assembly of the FS66 genome,

which is more superior on contig N50 than the previously available genome of *Fs* str. NRRL 66326 from sugarcane, but also acquired gene structure annotation for the species for the first time. Through comparative genomic analysis among FS66 and closely related *Fusarium* genomes, lineage-specific characteristics were identified in FS66 including gene gain and loss and gene transfer from other species. The discovery and analysis of this study should have provided important information and clues for further understanding of the molecular mechanism underlying the development of phytopathogenicity in *Fusarium* species.

The BLB disease caused by *Fs* str. FS66, which mostly caused only growth retard of plants in the field, is much less destructive than the deadly Panama disease caused by *Foc* (Dita et al., 2018). *Fs* has been reported to be transmitted by air currents and easily dispersed by rain splash (Nordahlwate et al., 2008), and it has also been reported to cause leaf diseases on *Rhynchosyilis gigantea* (Dekham and Kanchanawatee, 2020) and sugarcane (Yao et al., 2020). It has also been reported to cause wilt on sugarcane (Viswanathan et al., 2017; Bao et al., 2020), indicating it has the potential to become a soil-borne disease. A most notorious soil-borne banana pathogen is *Foc*, which first invades the roots and then climbs up the pseudo-stems through the vessels of the plant's vascular system (Czislowski et al., 2018). Thus in the Panama disease, discolorations of root and pseudo-stem vascular vessels are generally observed, and the leaf wilt symptoms usually first appear on older leaves. No root or pseudo-stem lesion has been observed on banana plants with BLB disease in the field but only on the leaves, and no symptom was observed on plants subjected to root inoculation, showing that FS66 is not a soil-borne pathogen like *Foc*. Another banana disease caused by an FFSC species (*F. verticillioides*) is the pseudo-stem heart rot disease, which also could not invade through the roots and requires inoculation with a wound in the pseudo-stem above the corm for the symptoms to develop (Shalaby et al., 2006). The low successful rate of non-wound inoculation and high successful rate of wound inoculation of FS66 suggests that wounds facilitate or are even necessary in the BLB disease.

Widespread gene transfer signals have been detected in the genome of FS66, suggesting the scale of gene transfer in *Fusarium* genomes could have been underestimated previously. Gene transfer studies in *Fusarium* have mainly been focused on specific genomic regions or a limited number of genes. A review on *Fusarium* pathogenomics suggested that the transfer of an entire plasmid or non-core chromosome was the more common type of HGT in *Fusarium* (Ma et al., 2013). Recent studies showed that gene transfer not only occurred on lineage-specific supernumerary chromosomes (Ma et al., 2010; Rocha et al., 2016; van Dam and Rep, 2017; van Dam et al., 2017; Yang et al., 2020) but also on genes located on the core chromosomes (Gardiner et al., 2012; Sieber et al., 2014; Vlaardingerbroek et al., 2016; Hoogendoorn et al., 2018; Liu et al., 2019; Tralamazza et al., 2019) or mitochondrial genomic regions (Brankovics et al., 2020). Most of these studies focused on effectors such as several *SIX* (Secreted In Xylem) genes (van Dam and Rep, 2017; Czislowski et al., 2018) and secondary metabolite gene clusters (Liu et al., 2019; Kim et al., 2020), which are related

to the pathogenicity or virulence of *Fusarium* isolates (Lanubile et al., 2016; An et al., 2019). In this study, no homolog of the *Fo* accessory chromosomes was found in FS66, but hundreds of genes were inferred to be putatively transferred from FOOSC or other species on the core chromosomes, not only including genes related to pathogenicity or virulence in other *Fusarium* pathogens but also many genes involved in multiple functional categories. The mechanism of gene transfer among *Fusarium* isolates remains largely unknown. Considering *Fusarium* harbors species both with known sexual cycle and those without (Ma et al., 2013), both sexual (introgression) and asexual (HGT) processes should have contributed to the observed gene transfer events.

In conclusion, we reported the BLB disease caused by *Fs* str. FS66 for the first time and obtained both a high-quality assembly and gene structure annotation for the genome. The genome of FS66 shares relatively high nucleotide similarity (96.74%) with *Fs* str. NRRL 66326 from sugarcane, a larger proportion (8.58% of NRRL 66326 and 14.47% of FS66) of the genomes could not be aligned, which could probably be explained by lineage-specific deletion and gene transfer events. Hundreds of genes across the FS66 genome putatively transferred from FOOSC and other outgroup species were detected in the FS66 genome, including a few involved in pathogenicity/virulence. These results and analyses should be valuable for further understanding of the genomic evolution underlying new pathogenicity development in *Fusarium*.

DATA AVAILABILITY STATEMENT

The genome sequence (doi: 10.6084/m9.figshare.13524434) and gene structure annotations (doi: 10.6084/m9.figshare.13524443) have been deposited to both NCBI (BioProject PRJNA667564) and the FigShare repository. The DDBJ/ENA/GenBank accession ID of the FS66 genome assembly is JADANP000000000, and the version described in this paper is version JADANP010000000. Raw sequencing data of genome and transcriptome have been deposited under the same NCBI BioProject. Whole-genome mRNA sequences (doi: 10.6084/m9.figshare.13524452), coding sequences (doi: 10.6084/m9.figshare.13524458), protein sequences (doi: 10.6084/m9.figshare.13524449), gene function annotation (doi: 10.6084/m9.figshare.13524446), repeat element annotation (doi: 10.6084/m9.figshare.13524497), non-coding RNA annotation (doi: 10.6084/m9.figshare.13524464), and mitochondria genome annotation (doi: 10.6084/m9.figshare.13524467) files are accessible from FigShare.

AUTHOR CONTRIBUTIONS

YC and BW conducted data analysis and manuscript writing. YC, AP, and XS carried out field investigation. YC carried out all the experiments. All authors participated in the revision of the manuscript. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research was funded by Guangdong Academy of Agricultural Sciences Dean's Fund, grant number 201708.

REFERENCES

- Abd Murad, N. B., Mohamed Nor, N. M. I., Shohaimi, S., and MohdZainudin, N. A. I. (2017). Genetic diversity and pathogenicity of *Fusarium* species associated with fruit rot disease in banana across Peninsular Malaysia. *J. Appl. Microbiol.* 123, 1533–1546. doi: 10.1111/jam.13582
- Alexander, W. G., Wisecaver, J. H., Rokas, A., and Hittinger, C. T. (2016). Horizontally acquired genes in early-diverging pathogenic fungi enable the use of host nucleosides and nucleotides. *PNAS* 113, 4116–4121. doi: 10.1073/pnas.1517242113
- An, B., Hou, X., Guo, Y., Zhao, S., Luo, H., He, C., et al. (2019). The effector SIX8 is required for virulence of *Fusarium oxysporum* sp. *cubense* tropical race 4 to Cavendish banana. *Fungal Biol.* 123, 423–430. doi: 10.1016/j.funbio.2019.03.001
- Araújo, N. A. F., Pasqual, M., Pio, L. A. S., Alves, E., de Matos Moura, N., and Costa, S. D. S. (2017). Identification and aggressiveness of four isolates of *Fusarium oxysporum* sp. *cubense* from Latundan banana in Brazil. *J. Phytopathol.* 165, 257–264. doi: 10.1111/jph.12557
- Asai, S., Ayukawa, Y., Gan, P., Masuda, S., Komatsu, K., Shirasu, K., et al. (2019). High-quality draft genome sequence of *Fusarium oxysporum* f. sp. *cubense* strain 160527, a causal agent of Panama disease. *Microbiol. Resour. Announc.* 8:e00654–19.
- Bao, Y., Xu, Y., Wang, S., Yao, Z., Rao, G. P., and Zhang, M. (2020). First Report of *Fusarium sacchari* that causes sugarcane wilt disease in China. *Plant Dis.* 104:2289. doi: 10.1094/PDIS-02-20-0229-PDN
- Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., et al. (2013).) MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* 69, 313–319. doi: 10.1016/j.ympev.2012.08.023
- Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., et al. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, W81–W87. doi: 10.1093/nar/gkz310
- Boratyn, G. M., Camacho, C., Cooper, P. S., Coulouris, G., Fong, A., Ma, N., et al. (2013). BLAST: a more efficient report with usability improvements. *Nucleic Acids Res.* 41, W29–W33. doi: 10.1093/nar/gkt282
- Brankovics, B., van Diepeningen, A. D., Hoog, G. S., van der Lee, T. A. J., and Waalwijk, C. (2020). Detecting introgression between members of the *Fusarium fujikuroi* and *F. oxysporum* species complexes by comparative mitogenomics. *Front. Microbiol.* 11:1092. doi: 10.3389/fmicb.2020.01092
- Bredeweg, E. L., and Baker, S. E. (2020). “Horizontal gene transfer in fungi,” in *Grand Challenges in Fungal Biotechnology*, ed. H. Nevalainen (Cham: Springer), 317–332.
- Chan, P. P., and Lowe, T. M. (2019). tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods Mol. Biol.* 1962, 1–14. doi: 10.1007/978-1-4939-9173-0_1
- Choudhuri, S. (2014). “Genes, genomes, molecular evolution, databases, and analytical tools,” in *Bioinformatics for Beginners*, ed. M. Kotewicz (Amsterdam: Elsevier/AP), 27–53.
- Czislowski, E., Fraser-Smith, S., Zander, M., O'Neill, W. T., Meldrum, R. A., Tran-Nguyen, L. T., et al. (2018). Investigation of the diversity of effector genes in the banana pathogen, *Fusarium oxysporum* f. sp. *cubense*, reveals evidence of horizontal gene transfer. *Mol. Plant Pathol.* 19, 1155–1171. doi: 10.1111/mpp.12594
- Dekham, K., and Kanchanawatee, K. (2020). The first report of *Fusarium sacchari* causing yellow leaf spot disease on *Rhynchosstylis gigantea* orchids in Thailand. *Am. J. Agric. Biol. Sci.* 15, 68–74. doi: 10.3844/ajabssp.2020.68.74
- Dita, M., Barquero, M., Heck, D., Mizubuti, E. S. G., and Staver, C. P. (2018). *Fusarium* wilt of banana: current knowledge on epidemiology and research needs toward sustainable disease management. *Front. Plant Sci.* 9:1468. doi: 10.3389/fpls.2018.01468
- Dong, T., Chen, X. J., Fan, C., Jin, Y., and Yi, G. J. (2013). Analysis on the development of the banana industry in China. *Guangdong Agric. Sci.* 11, 220–223.
- Du, Y. X., Chen, F. R., Shi, N. N., and Ruan, H. C. (2017). First report of *Fusarium chlamydosporum* causing banana crown rot in Fujian province, China. *Plant Dis.* 101:1048. doi: 10.1094/PDIS-11-16-1674-PDN
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Emms, D. M., and Kelly, S. (2018). STAG: species tree inference from all genes. *BioRxiv* doi: 10.1101/267914
- Emms, D. M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20:238. doi: 10.1186/s13059-019-1832-y
- Fan, H. C., Zeng, L., Yang, P. W., Guo, Z. X., and Bai, T. T. (2016). First report of banana soft rot caused by *Klebsiella variicola* in China. *Plant Dis.* 100:517. doi: 10.1094/PDIS-05-15-0586-PDN
- Fitzpatrick, D. A. (2012). Horizontal gene transfer in fungi. *FEMS Microbiol. Lett.* 329, 1–8. doi: 10.1111/j.1574-6968.2011.02465.x
- Gao, S., Gold, S. E., Wisecaver, J. H., Zhang, Y., Guo, L., Ma, L. J., et al. (2019). Genome-wide analysis of *Fusarium verticillioides* reveals inter-kingdom contribution of horizontal gene transfer to the expansion of metabolism. *Fungal Genet. Biol.* 128, 60–73. doi: 10.1016/j.fgb.2019.04.002
- Gardiner, D. M., McDonald, M. C., Covarelli, L., Solomon, P. S., Rusu, A. G., Marshall, M., et al. (2012). Comparative pathogenomics reveals horizontally acquired novel virulence genes in fungi infecting cereal hosts. *PLoS Pathog.* 8:e1002952. doi: 10.1371/journal.ppat.1002952
- Gawel, N. J., and Jarret, R. L. (1991). A modified CTAB DNA extraction procedure for *Musa* and *Ipomoea*. *Plant Mol. Biol. Rep.* 9, 262–266. doi: 10.1007/BF02672076
- Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., et al. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36, 3420–3435. doi: 10.1093/nar/gkn176
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086
- Hoff, K. J., and Stanke, M. (2019). Predicting genes in single genomes with AUGUSTUS. *Curr. Protoc. Bioinform.* 65:e57. doi: 10.1002/cpbi.57
- Hoogendoorn, K., Barra, L., Waalwijk, C., Dickschat, J. S., van der Lee, T. A. J., and Medema, M. H. (2018). Evolution and diversity of biosynthetic gene clusters in *Fusarium*. *Front. Microbiol.* 9:1158. doi: 10.3389/fmicb.2018.01158
- Hunter, S., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., et al. (2009). InterPro: the integrative protein signature database. *Nucleic Acids Res.* 37, D211–D215. doi: 10.1093/nar/gkn785
- Jones, D. R. (1997). *Diseases of Banana and Plantain (Musa spp.)*. Available online at: <https://www.apsnet.org/edcenter/resources/commonnames/Pages/BananaandPlantain.aspx> (accessed October 03, 2020).
- Kalvari, I., Argasinska, J., Quinones-Olvera, N., Nawrocki, E. P., Rivas, E., Eddy, S. R., et al. (2018). Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res.* 46, D335–D342. doi: 10.1093/nar/gkx1038
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–D361. doi: 10.1093/nar/gkw1092
- Keeling, P. J., and Palmer, J. D. (2008). Horizontal gene transfer in eukaryotic evolution. *Nat. Rev. Genet.* 9, 605–618. doi: 10.1038/nrg2386
- Keilwagen, J., Hartung, F., and Grau, J. (2019). “GeMoMa: homology-based gene prediction utilizing intron position conservation and RNA-seq Data,” in *Gene Prediction: Methods and Protocols*, ed. M. Kollmar (New York, NY: Humana), 161–177.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.629859/full#supplementary-material>

- Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4
- Kim, H. S., Lohmar, J. M., Busman, M., Brown, D. W., Naumann, T. A., Divon, H. H., et al. (2020). Identification and distribution of gene clusters required for synthesis of sphingolipid metabolism inhibitors in diverse species of the filamentous fungus *Fusarium*. *BMC Genomics* 21:510. doi: 10.1186/s12864-020-06896-1
- King, R., Brown, N. A., Urban, M., and Hammond-Kosack, K. E. (2018). Inter-genome comparison of the Quorn fungus *Fusarium venenatum* and the closely related plant infecting pathogen *Fusarium graminearum*. *BMC Genomics* 19:269. doi: 10.1186/s12864-018-4612-2
- King, R., Urban, M., Hammond-Kosack, M. C., Hassani-Pak, K., and Hammond-Kosack, K. E. (2015). The completed genome sequence of the pathogenic ascomycete fungus *Fusarium graminearum*. *BMC Genomics* 16:544. doi: 10.1186/s12864-015-1756-1
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109
- Kumar, S., Stecher, G., Li, M., Nkayaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096
- Lagesen, K., Hallin, P., Rødland, E. A., Staerfeldt, H.-H., Rognes, T., and Ussery, D. W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35, 3100–3108. doi: 10.1093/nar/gkm160
- Lanubile, A., Ellis, M. L., Marocco, A., and Munkvold, G. P. (2016). Association of effector Six6 with vascular wilt symptoms caused by *Fusarium oxysporum* on soybean. *Phytopathology* 106, 1404–1412. doi: 10.1094/PHYTO-03-16-0118-R
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404
- Laurence, M. H., Summerell, B. A., and Liew, E. C. Y. (2015). *Fusarium oxysporum* f. sp. *canariensis*: evidence for horizontal gene transfer of putative pathogenicity genes. *Plant Pathol.* 64, 1068–1075. doi: 10.1111/ppa.12350
- Leslie, J. F., Summerell, B. A., Bullock, S., and Doe, F. J. (2005). Description of *Gibberella sacchari* and neotypification of its anamorph *Fusarium sacchari*. *Mycologia* 97, 718–724. doi: 10.1080/15572536.2006.11832801
- Li, H. (2016). Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* 32, 2103–2110. doi: 10.1093/bioinformatics/btw152
- Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 10.1093/bioinformatics/bty191
- Li, M. H., Yu, X. T., Wang, H. F., Zhou, J. N., and Jiang, Z. D. (2012). Rapid detection and identification of *Fusarium oxysporum* f. sp. *cubense* race 1 and race 4. *Sci. Agric. Sin.* 45, 3971–3979.
- Liu, S., Wu, B., Lv, S., Shen, Z., Li, R., Yi, G., et al. (2019). Genetic diversity in FUB genes of *Fusarium oxysporum* f. sp. *cubense* suggests horizontal gene transfer. *Front. Plant Sci.* 10:1069. doi: 10.3389/fpls.2019.01069
- Lysøe, E., Harris, L. J., Walkowiak, S., Subramaniam, R., Divon, H. H., Riiser, E. S., et al. (2014). The genome of the generalist plant pathogen *Fusarium avenaceum* is enriched with genes involved in redox, signaling and secondary metabolism. *PLoS One* 9:e112703.
- Ma, L. J., Geiser, D. M., Proctor, R. H., Rooney, A. P., O'Donnell, K., Trail, F., et al. (2013). *Fusarium* pathogenomics. *Annu. Rev. Microbiol.* 67, 399–416. doi: 10.1146/annurev-micro-092412-155650
- Ma, L. J., Van Der Does, H. C., Borkovich, K. A., Coleman, J. J., Daboussi, M. J., Di Pietro, A., et al. (2010). Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* 464, 367–373. doi: 10.1038/nature08850
- Maddison, W. P., and Knowles, L. L. (2006). Inferring phylogeny despite incomplete lineage sorting. *Syst. Biol.* 55, 21–30. doi: 10.1080/10635150500354928
- Maldonado-Bonilla, L. D., Calderón-Oropeza, M. A., Villarruel-Ordaz, J. L., and Sánchez-Espinosa, A. C. (2019). Identification of novel potential causal agents of *Fusarium* wilt of *Musa* sp. AAB in southern Mexico. *Plant Pathol. Microbiol.* 10:479. doi: 10.24105/2157-7471.10.479
- Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., and Zimin, A. (2018). MUMmer4: a fast and versatile genome alignment system. *PLoS Comput. Biol.* 14:e1005944. doi: 10.1371/journal.pcbi.1005944
- Martin, S. H., Wingfield, B. D., Wingfield, M. J., and Steenkamp, E. T. (2011). Structure and evolution of the *Fusarium* mating type locus: new insights from the *Gibberella fujikuroi* complex. *Fungal Genet. Biol.* 48, 731–740.
- Maryani, N., Lombard, L., Poerba, Y. S., Subandiyah, S., Crous, P. W., and Kema, G. H. J. (2019). Phylogeny and genetic diversity of the banana *Fusarium* wilt pathogen *Fusarium oxysporum* f. sp. *cubense* in the Indonesian centre of origin. *Stud. Mycol.* 92, 155–194. doi: 10.1016/j.smyco.2018.06.003
- Milner, D. S., Attah, V., Cook, E., Maguire, F., Savory, F. R., Morrison, M., et al. (2019). Environment-dependent fitness gains can be driven by horizontal gene transfer of transporter-encoding genes. *PNAS* 116, 5613–5622. doi: 10.1073/pnas.1815994116
- Montoya-Martínez, A. C., Rodríguez-Alvarado, G., Fernández-Pavía, S. P., Proctor, R. H., Kim, H.-S., and O'Donnell, K. (2019). Design and validation of a robust multiplex polymerase chain reaction assay for MAT idiomorph within the *Fusarium fujikuroi* species complex. *Mycologia* 111, 772–781. doi: 10.1080/00275514.2019.1649956
- Nawrocki, E. P., and Eddy, S. R. (2013). Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29, 2933–2935. doi: 10.1093/bioinformatics/btt509
- Niehaus, E. M., Kim, H. K., Münsterkötter, M., Janevska, S., Arndt, B., Kalinina, S. A., et al. (2017). Comparative genomics of geographically distant *Fusarium fujikuroi* isolates revealed two distinct pathotypes correlating with secondary metabolite profiles. *PLoS Pathog.* 13:e1006670. doi: 10.1371/journal.ppat.1006670
- Niehaus, E. M., Münsterkötter, M., Proctor, R. H., Brown, D. W., Sharon, A., Idan, Y., et al. (2016). Comparative “omics” of the *Fusarium fujikuroi* species complex highlights differences in genetic potential and metabolite synthesis. *Genome Biol. Evol.* 8, 3574–3599.
- Nordahlwate, M. S., Izzati, M. Z., Azmi, A. R., and Salleh, B. (2008). Distribution, morphological characterization and pathogenicity of *Fusarium sacchari* associated with pokkahboeng disease of sugarcane in Peninsular Malaysia. *Pertanika J. Trop. Agric. Sci.* 31, 279–286.
- O'Donnell, K., Ward, T. J., Robert, V. A., Crous, P. W., Geiser, D. M., and Kang, S. (2015). DNA sequence-based identification of *Fusarium*: current status and future directions. *Phytoparasitica* 43, 583–595. doi: 10.1007/s12600-015-0484-z
- Ou, S., Su, W., Liao, Y., Chougule, K., Agda, J. R., Helling, A. J., et al. (2019). Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* 20:275.
- Pertea, G., and Pertea, M. (2020). GFF Utilities: GffRead and GffCompare. *F1000Res* 9:304. doi: 10.12688/f1000research.23297.2
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122
- Ploetz, R. C. (2006). *Fusarium* wilt of banana is caused by several pathogens referred to as *Fusarium oxysporum* f. sp. *cubense*. *Phytopathology* 96, 653–656. doi: 10.1094/PHYTO-96-0653
- Ponts, N., Richard-Forget, F., Zhang, H., Barroso, G., and Zhao, C. (2018). Genome sequence of the emerging mycotoxin-producing filamentous fungus *Fusarium tricinctum* strain INRA104. *Genome Announc.* 6, e509–e518.
- Quinlan, A. R. (2014). BEDTools: the swiss-army tool for genome feature analysis. *Curr. Protoc. Bioinform.* 47, 11.12.1–11.12.34. doi: 10.1002/0471250953.bi1112s47
- Reeb, V., Lutzoni, F., and Roux, C. (2004). Contribution of *RPB2* to multilocus phylogenetic studies of the euascomycetes (Pezizomycotina, Fungi) with special emphasis on the lichen-forming Acarosporaceae and evolution of polyspory. *Mol. Phylogenet. Evol.* 32, 1036–1060. doi: 10.1016/j.ympev.2004.04.012
- Riolo, M., Aloï, F., Faedda, R., Cacciola, S. O., and Pane, A. (2020). First report of postharvest fruit rot caused by *Fusarium sacchari* on lady finger banana in Italy. *Plant Dis.* 104:2290. doi: 10.1094/PDIS-01-20-0143-PDN
- Rocha, L. O., Laurence, M. H., Ludowici, V. A., Puno, V. I., Lim, C. C., Tesoriero, L. A., et al. (2016). Putative effector genes detected in *Fusarium oxysporum* from natural ecosystems of Australia. *Plant Pathol.* 65, 914–929. doi: 10.1111/ppa.12472
- Schmickl, R., Marburger, S., Bray, S., and Yant, L. (2017). Hybrids and horizontal transfer: introgression allows adaptive allele discovery. *J. Exp. Bot.* 68, 5453–5470. doi: 10.1093/jxb/erx297

- Shalaby, M. S., Kamhawy, M. A., and Ammar, M. I. (2006). Pathological studies on heart rot disease of banana in Egypt. *Egypt. J. Phytopathol.* 34, 17–27.
- Sieber, C. M. K., Lee, W., Wong, P., Münsterkötter, M., Mewes, H.-W., Schmeitzl, C., et al. (2014). The *Fusarium graminearum* genome reveals more secondary metabolite gene clusters and hints of horizontal gene transfer. *PLoS One* 9:e110311. doi: 10.1371/journal.pone.0110311
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351
- Song, L., Sabunciyani, S., and Florea, L. (2016). CLASS2: accurate and efficient splice variant annotation from RNA-seq reads. *Nucleic Acids Res.* 44:e98. doi: 10.1093/nar/gkw158
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., and Morgenstern, B. (2006). AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res.* 34, W435–W439. doi: 10.1093/nar/gkl200
- Steinberger, M., and Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* 35, 1026–1028. doi: 10.1038/nbt.3988
- Sukumaran, J., and Holder, M. T. (2010). DendroPy : a Python library for phylogenetic computing. *Bioinformatics* 26, 1569–1571. doi: 10.1093/bioinformatics/btq228
- Suwandi, S., Akino, S., and Kondo, N. (2018). Enhanced virulence of *Fusarium* species associated with spear rot of oil palm following recovery from osmotic stress. *Mycology* 9, 20–28. doi: 10.1080/21501203.2017.1336497
- Tralamazza, S. M., Rocha, L. O., Oggenfuss, U., Corrêa, B., and Croll, D. (2019). Complex evolutionary origins of specialized metabolite gene cluster diversity among the plant pathogenic fungi of the *Fusarium graminearum* species complex. *Genome Biol. Evol.* 11, 3106–3122. doi: 10.1093/gbe/evz225
- Urban, M., Cuzick, A., Rutherford, K., Irvine, A., Pedro, H., Pant, R., et al. (2017). PHI-base: a new interface and further additions for the multi-species pathogen-host interactions database. *Nucleic Acids Res.* 45, D604–D610. doi: 10.1093/nar/gkw1089
- van Dam, P., Fokkens, L., Ayukawa, Y., van der Gragt, M., Ter Horst, A., Brankovics, B., et al. (2017). A mobile pathogenicity chromosome in *Fusarium oxysporum* for infection of multiple cucurbit species. *Sci. Rep.* 7:9042. doi: 10.1038/s41598-017-07995-y
- van Dam, P., and Rep, M. (2017). The distribution of miniature impala elements and six genes in the *Fusarium* Genus is suggestive of horizontal gene transfer. *J. Mol. Evol.* 85, 14–25. doi: 10.1007/s00239-017-9801-0
- Vaser, R., Sović, I., Nagarajan, N., and Šikić, M. (2017). Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* 27, 737–746. doi: 10.1101/gr.214270.116
- Venturini, L., Caim, S., Kaithakottil, G. G., Mapleson, D. L., and Swarbreck, D. (2018). Leveraging multiple transcriptome assembly methods for improved gene structure annotation. *Gigascience* 7:giy093. doi: 10.1093/gigascience/gy093
- Viswanathan, R., Balaji, C. G., Selvakumar, R., Malathi, P., Sundar, A. R., Prasanth, C. N., et al. (2017). Epidemiology of *Fusarium* diseases in sugarcane: a new discovery of same *Fusarium sacchari* causing two distinct diseases, wilt and pokkahnoeng. *Sugar. Tech.* 19, 638–646. doi: 10.1007/s12355-017-0552-4
- Vlaardingerbroek, I., Beerens, B., Rose, L., Fokkens, L., Cornelissen, B. J., and Rep, M. (2016). Exchange of core chromosomes and horizontal transfer of lineage-specific chromosomes in *Fusarium oxysporum*. *Environ. Microbiol.* 18, 3702–3713. doi: 10.1111/1462-2920.13281
- Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. doi: 10.1371/journal.pone.0112963
- Wang, B., Yu, H., Jia, Y., Dong, Q., Steinberg, C., Alabouvette, C., et al. (2020). Chromosome-scale genome assembly of *Fusarium oxysporum* strain Fo47, a fungal endophyte and biocontrol agent. *Mol. Plant Microbe Interact.* 33, 1108–1111.
- Warmington, R. J., Kay, W., Jeffries, A., O'Neill, P., Farbos, A., Moore, K., et al. (2019). High-quality draft genome sequence of the causal agent of the current Panama disease epidemic. *Microbiol. Resour. Announc.* 8, e904–e919.
- White, T. J., Bruns, T., Lee, S. J. W. T., and Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protoc. Guide Methods Appl.* 18, 315–322.
- Wiemann, P., Sieber, C. M., Von Barga, K. W., Studt, L., Niehaus, E. M., Espino, J. J., et al. (2013). Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. *PLoS Pathog.* 9:e1003475.
- Yang, H., Yu, H., and Ma, L. J. (2020). Accessory chromosomes in *Fusarium oxysporum*. *Phytopathology* 110, 1488–1496. doi: 10.1094/PHYTO-03-20-0069-IA
- Yao, Z., Zou, C., Peng, N., Zhu, Y., Bao, Y., Zhou, Q., et al. (2020). Virome identification and characterization of *Fusarium sacchari* and *F. andiyazi*: causative agents of PokkahBoeng disease in sugarcane. *Front. Microbiol.* 11:240. doi: 10.3389/fmicb.2020.00240
- Zeng, L. S., Zhao, Z. H., Lü, S., Xi, Z. J., Li, M. H., Xi, P. G., et al. (2013). The *Fusarium* species isolated from banana and their phylogenetic relationships. *Mycosystema* 32, 617–632.
- Zhang, Q., Chen, X., Xu, C., Zhao, H., Zhang, X., Zeng, G., et al. (2019). Horizontal gene transfer allowed the emergence of broad host range entomopathogens. *PNAS* 116, 7982–7989. doi: 10.1073/pnas.1816430116
- Zhou, Y., Huang, J. S., Yang, L. Y., Wang, G. F., and Li, J. Q. (2017). First report of banana anthracnose caused by *Colletotrichum scovillei* in China. *Plant Dis.* 101:381. doi: 10.1094/PDIS-08-16-1135-PDN

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Cui, Wu, Peng, Song and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.