




Genome and Transcriptome Sequencing Analysis of *Fusarium commune* Provides Insights into the Pathogenic Mechanisms of the Lotus Rhizome Rot

 Weigang Kuang,^a Lianhu Zhang,^a Lifang Ye,^a Jian Ma,^a Xugen Shi,^a Yachun Lin,^a Xiaotang Sun,^a Ruqiang Cui^a

^aCollege of Agronomy, Jiangxi Agricultural University, Nanchang, Jiangxi, China

ABSTRACT *Fusarium* wilt, a vascular wilt caused by *F. commune*, has been a serious problem for the lotus. Although some *F. commune* isolate genomes have been sequenced, little is known about the genomic information of the strain that causes *Fusarium* wilt of aquatic plants. In this study, the genome of *F. commune* FCN23 isolated from lotuses in China was sequenced using Illumina and PacBio sequencing platforms. The FCN23 genome consisted of 53 scaffolds with a combined size of 46,211,149 bp. According to the reference genome, *F. oxysporum* f. sp. *lycopersici* 4287 isolated from tomato, it was finally assembled into 14 putative chromosomes, including 10 core and 4 lineage-specific chromosomes. The genome contains about 3.45% repeats and encodes 14,698 putative protein-coding genes. Among these, 1,038 and 296 proteins were potentially secreted proteins and candidate effector proteins, respectively. Comparative genomic analysis showed that the CAZyme-coding genes and secondary metabolite biosynthesis genes of FCN23 were similar to those of other Ascomycetes. Additionally, the transcriptome of FCN23 during infection of lotus was analyzed and 7,013 differentially expressed genes were identified. Eight putative effectors that were upregulated in the infection stage were cloned. Among them, F23a002499 exhibited strong hypersensitive response after transiently expressed in *Nicotiana benthamiana* leaves. Our results provide a valuable genetic basis for understanding the molecular mechanism of the interaction between *F. commune* and aquatic plants.

IMPORTANCE *Fusarium commune* is an important soilborne pathogen with a wide range of hosts and can cause *Fusarium* wilt of land plants. However, there are few studies on *Fusarium* wilt of aquatic plants. Lotus rhizome rot mainly caused by *F. commune* is a devastating disease that causes extensive yield and quality losses in China. Here, we obtained high-quality genomic information of the FCN23 using Illumina NovaSeq and the third-generation sequencing technology PacBio Sequel II. Compared to the reference genome *F. oxysporum* f. sp. *lycopersici* strain 4287, it contains 11 core and 3 lineage-specific chromosomes. Many differentially expressed genes associated with pathogenicity were identified by RNA sequencing. The genome and transcriptome sequences of FCN23 will provide important genomic information and insights into the infection mechanisms of *F. commune* on aquatic plants.

KEYWORDS *Fusarium commune*, lotus, comparative genomics, transcriptome, plant-microbe interactions, effector

Lotus (*Nelumbo nucifera* Gaertn.), belonging to Nelumbonaceae, is a perennial aquatic plant. It is an ornamental and edible plant in Asian countries, such as China, Japan, and India (1). It is a well-known and economically important plant, which cultivated for more than 2,000 years (2). The rhizome is used as a vegetable, and part of the leaves, flowers and fruits can be made into food and medicinal materials (3). Although the lotus is an

Editor Luxin Wang, University of California Davis

Copyright © 2022 Kuang et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Xiaotang Sun, sunxiaotang80@foxmail.com, or Ruqiang Cui, cuiuruqiang@jxau.edu.cn.

The authors declare no conflict of interest.

Received 17 January 2022

Accepted 4 June 2022

Published 5 July 2022

aquatic plant, about 25 species of plant parasitic fungi have been recorded to have successfully colonized the rhizome and leaves of the lotus (4, 5). Other pathogenic microorganisms, such as *Hirschmanniella imamuri* and *H. diversa*, are known to cause serious damage to Indian lotus and are considered a major nematode threat to lotus production in Tokushima Prefecture, Japan (6).

Fusarium wilt is a common vascular fungal disease in plants, which can result in significant yield losses. *Fusarium commune* is one of the species associated with wilt and root rot. It is a recently characterized species closely related to the sister group *F. oxysporum* species complex (7). The *F. commune* in the soil initially invades the roots asymptotically, and then colonizes the vascular tissues, causing wilting, necrosis, and yellowing of the above-ground parts. Cultivating a single crop for a long period on the same plot easily leads to the accumulation of pathogens and aggravates the occurrence of *Fusarium* wilt (8). Rhizome rot was first reported to be caused by *F. oxysporum* Schl. f. sp. *nelumbicola* (Nis. & Wat.) Booth in 1953 (9, 10). Recent studies show that the main pathogen was *F. commune* and has been a serious problem for lotuses in China (9, 11–13). The lotus will be damaged by *Fusarium* wilt throughout its growth period, causing the underground stems to rot and the ground to wither. Due to the consecutive monoculture of lotus plants in most parts of China, the disease has become more and more serious. Symptoms of wilt usually begin to appear in early June in southern China. When the disease is serious, the value of lotus products can be reduced by more than 60%, or even no harvest.

To date, the studies on *Fusarium* wilt have focused mainly on terrestrial plants, with a lack of investigation in aquatic plants. The mechanism of replant disease was still unknown in aquatic crops. Therefore, it is of great significance to investigate the interaction mechanism between the lotus and *F. commune* for the comprehensive prevention and control of the disease. However, the understanding of the mechanism is very limited as there are no relevant genetic data available for this pathogen. With the advantages of next-generation sequencing technology, the number of sequenced *Fusarium* genomes is increasing. The availability of these sequences has made it possible to research *F. commune* at the genomic level. Genome sequencing of plant pathogens and comparison are effective methods to reveal the pathogenic mechanism and evolution of plant pathogens (14–16).

In this study, we sequenced the genome of *F. commune* used the PacBio and Illumina high-throughput sequencing technology, which is isolated from the aquatic lotus plant. Whole-genome and comparative genome analyses revealed that the genome encodes a diverse range of genes related to virulence, including carbohydrate-active enzymes, secreted proteins, effector proteins and genes involved in secondary metabolism. Our research has laid a solid foundation for exploring the pathogenic process of *F. commune* infecting aquatic plants.

RESULTS

Pathogen identification. Symptoms of *Fusarium* wilt disease on lotuses in the field and morphological characteristics of FCN23 strain are shown in Fig. 1. The colonies were characterized by an abundant white cottony mycelium, which became light to dark purple with age. Macroconidia had 3–5 septa, slightly curved at the apex, and measured 32.62 to 48.81×2.70 to $4.93 \mu\text{m}$ ($n = 50$). Microconidia had 0–1 septa, oval or ellipsoid, and were 5.33 to 8.38×2.05 to $3.64 \mu\text{m}$ ($n = 50$). Chlamydospores were spherical to oval. Phylogenetic trees based on the sequences of the intergenic spacer region (IGS) and translation elongation factor 1-alpha (*EF-1 α*) confirmed the isolate as *F. commune* (Fig. S1). In addition, a 295-bp size fragment was produced from FCN23 gDNA using specific PCR primers (efFc100F/efFc385R) for detecting *F. commune*, and no amplicon was detected using specific PCR primers (FOF1/FOR1) for *F. oxysporum* (Fig. S2). The typical symptoms appeared on all inoculated plants after 5 days of inoculation at 25°C, while the controls showed no symptoms (Fig. S3). The same pathogen was reisolated from the lesions, thereby fulfilling Koch's postulates. These characteristics were consistent to those

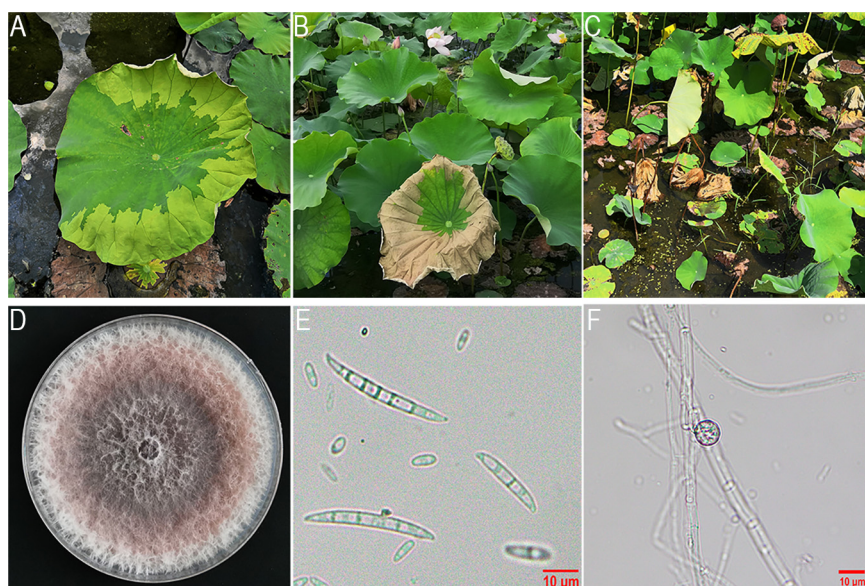


FIG 1 Symptoms of *Fusarium* wilt disease on lotus plants and morphological characteristics of strain FCN23. (A-C) Symptoms of *Fusarium* wilt disease on lotus plants in the early, middle, and late stages. (D) Mycelium colony on PDA medium at 25°C after 7 days of incubation. (E) Morphological characteristics of macroconidia and microconidium. (F) Morphological characteristics of chlamydospores.

described for *F. commune*. Reisolations from infected plants confirmed that the reisolated pathogens possessed identical morphological characteristics to those of the original pathogens.

Genome sequencing, assembly and annotation. The genome of FCN23 was assembled from the data generated by the Illumina and PacBio sequencing platforms. After quality control of raw data, 9,570.6 Mb NovaSeq-data and 467,707,267 reads with an average length of 6,525 bp PacBio-data were generated. K-mer analysis confirmed the high quality of the library (Fig. S4). As showed in Table 1, the assembled FCN23 genome consisted of 53 scaffolds with an N50 length of 2,194,425 bp and a combined size of 46,211,149 bp. The average GC ratio and repeats rate were 47.55% and 3.45%, respectively. A total of 14,698 protein-coding genes were predicted with a total length of 21,370,401 bp. The average gene length was 1,454 bp, and the gene density was ~326 genes per one Mb of the genome. These predicted genes account for 46.25% of the genome. The genomic features for FCN23 are shown in Table 1.

About 14,472 (98.46%), 3,976 (27.05%), 7,990 (54.36%), 4,781(32.53%), and 7,843 (53.36%) of the predicted genes presented homologies with known functions in the NCBI Non-Redundant (NR), Gene Ontology (GO), eggNOG, Kyoto Encyclopedia of Genes and Genomes (KEGG), and Swiss-Prot databases, respectively (Table S1 and S2). In addition, 303 tRNA and 88 rRNA were predicted in the assembly genome (Table S3).

TABLE 1 General features of the *F. commune* FCN23 genome

Features	FCN23
Size (bp)	46,211,149
Coverage (fold)	200x
%G+C content	47.55
Repeat rate (%)	3.45
N50 length (bp)	2,194,425
No. of scaffolds	53
Protein-coding genes	14,698
Gene total length (bp)	21,370,401
Avg gene length (bp)	1,454
Gene length/Genome (%)	46.25
Gene density (no. genes per Mb)	326

TABLE 2 General features of repeat element types in the FCN23 genome

Type	Total length (bp)	No. elements	Percentage in genome (%)
Long terminal repeat (LTR)	431,295	1,820	0.9333
DNA transposons	800,188	1,545	1.7316
Non-LTR retrotransposon (LINEs)	124,881	637	0.2702
Short interspersed repeated (SINE)	1,214	18	0.0026
Rolling circle (RC)	5,380	67	0.0116
Unknown	34,823	263	0.0754
Total interspersed repeated	1,386,185	4,350	2.9997
Minisatellite DNA	121,086	2,722	0.2620
Microsatellite DNA	19,577	476	0.0424
Total tandem repeats	210,429	3,688	0.4554

The repeat elements identified in FCN23 constituted the genome, including 2.9997% as interspersed repeats and 0.4554% as tandem repeats. The most abundant repetitive element was the DNA transposons (1.73%), followed by long terminal repeat (LTR) (0.93%), non-LTR retrotransposon LINEs (long interspersed nuclear elements) (0.27%), rolling circle (RC) (0.01%) (Table 2). Among the tandem repeats, a total 121,086 bp minisatellite was identified, accounting for 0.2620% of the genome. In addition, a 19,577 bp (0.0424%) microsatellite was identified. Compared with the reference genomes Fo47 (17) and Fo4287 (18), FCN23 has a smaller genome size and repeat rate, and fewer genes.

Comparative genomic and orthologous clusters analysis. Genomic sequences of FCN23 was aligned with reference genome sequences of the well characterized strain Fo4287 (GenBank accession number [AAXH00000000.1](#)). Based on the reference genome, putative core chromosomes 1, 2, 4, 5, 7, 8, 9, 10, 11, 12, and 13 were assembled. In addition, putative lineage-specific (LS) chromosomes 6, 14, and 15 were identified in the FCN23 genome. Lineage-specific genes present in FCN23 was distinct from those in Fo4287. The gene distribution characteristics of core and lineage-specific chromosomes in the FCN23 and Fo4287 genomes is presented in Fig. 2. We further verified the results of the putative chromosome assembly by PCR amplification. The results were consistent with the assembly results (Table S4 and Fig. S5).

The comparative analysis of FCN23 and Fo4287 showed that the two strains shared 12,325 orthologous clusters. In addition, 102 clusters containing 224 proteins were found to be unique to the FCN23 (Fig. 3A and Table S5). GO analysis among the group of 102 unique clusters showed an enrichment of genes associated with a nitrogen

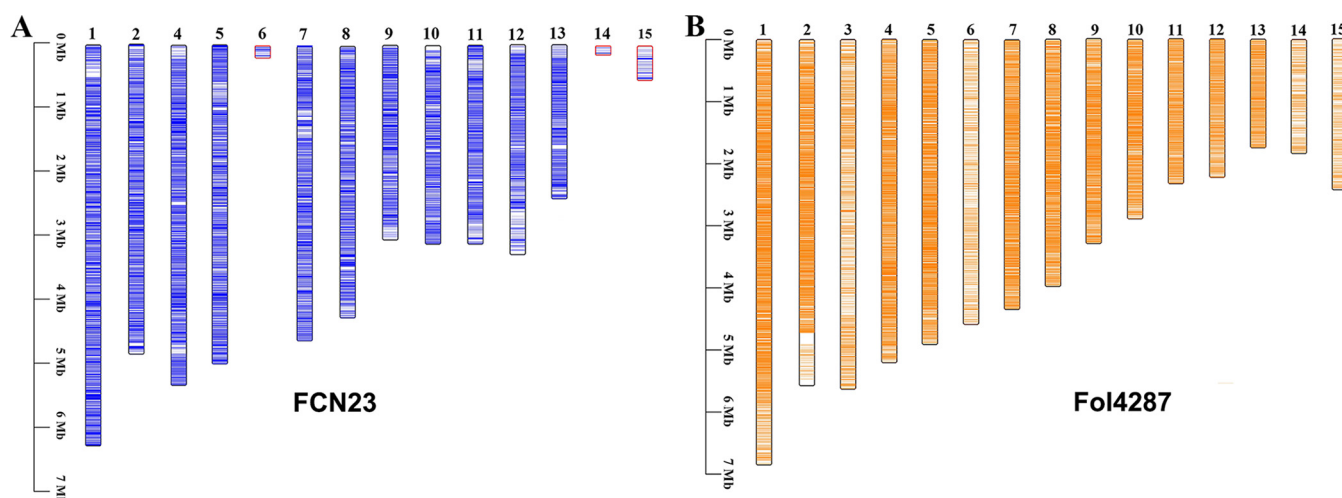


FIG 2 Comparison of gene distribution characteristics of core and lineage-specific chromosomes in FCN23 and Fo4287 genomes. (A) FCN23, *Fusarium commune* FCN23. (B) Fo4287, *F. oxysporum* f. sp. *lycopersici* 4287. Core chromosomes: 1, 2, 4, 5, 7, 8, 9, 10, 11, 12, and 13. Lineage-specific chromosomes: 3, 6, 14, and 15. The color represents the location of the gene.

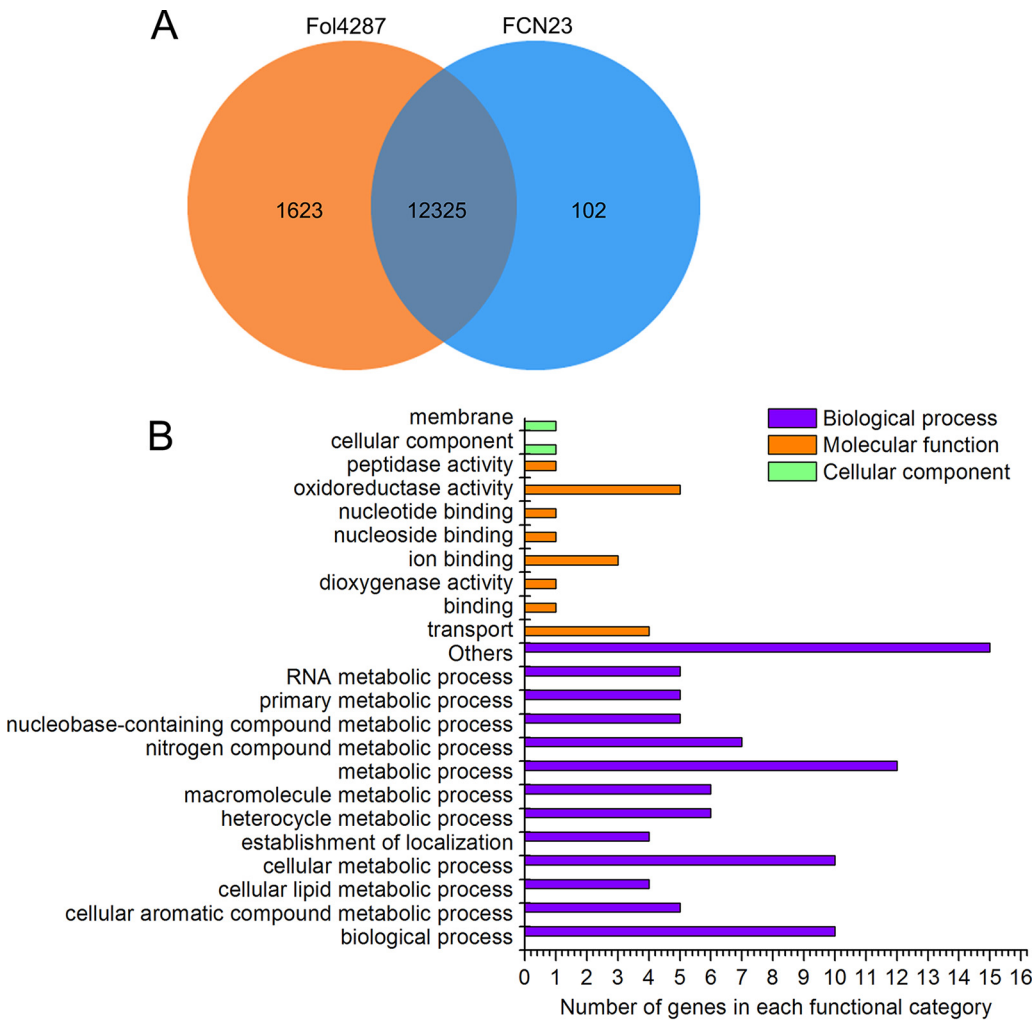


FIG 3 Comparison of orthologous genes, FCN23 and the reference *F. oxysporum* f. sp. *lycopersici* genome. (A) Venn diagram of gene clusters, FCN23, *F. commune*; Fol4287, *F. oxysporum* f. sp. *lycopersici* 4287. (B) GO enrichment analysis of annotated genes exists in FCN23 genome, but not in Fol4287.

compound metabolic process, cellular metabolic process, and heterocycle metabolic process (Fig. 3B).

Similarly, we compared four genomes, including *F. graminearum*, *F. poae*, *F. verticillioides* and FCN23 using OrthoVenn. In total, 56,936 genes were grouped into 13,626 clusters. Among these clusters, 5,264 orthologous clusters (containing at least two species) and 8,362 single copy gene clusters were identified. A total of 8,814 core gene clusters were identified from the four species, as shown in Fig. S6. FCN23 contains 108 specific gene clusters, including 237 genes, of which 101 genes were enriched in the biological process and 15 genes were enriched in molecular function. These unique genes can be explored as a potential target to affect the metabolism and pathogenicity of FCN23 (Table S6). The genomic information of strains used for comparison in this study is shown in Table S7.

Carbohydrate-active enzymes. The structure and composition of plant cell walls are highly complex and diverse, mainly composed of cellulose, hemicellulose and pectin. Pathogenic fungi can decompose and utilize plant cell wall polysaccharides by carbohydrate-active enzymes. In total, 662 CAZyme-coding gene homologs were predicted by the domain-based annotation dbCAN in combination with CAT in the genome of FCN23, which comprised 325 glycoside hydrolases (GHs), 26 polysaccharide lyases (PLs), 54 carbohydrate esterases (CEs), 126 auxiliary activities (AAs), 107 glycosyl transferases (GTs),

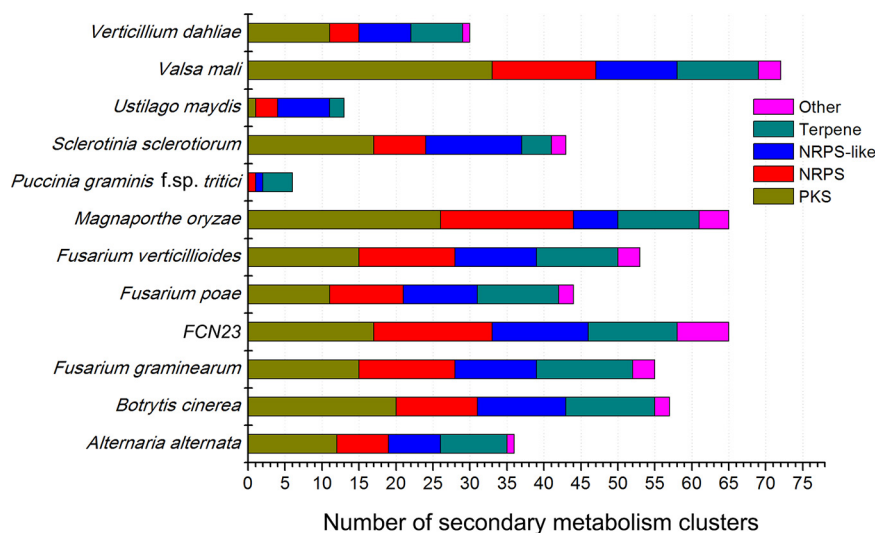
TABLE 3 Statistical summary of carbohydrate active enzymes in FCN23 and other 11 fungal genomes

CAZymes	AAs	CBMs	CEs	GHs	GTs	PLs	Total	Pathogen lifestyle
<i>Ustilago maydis</i>	88	6	49	265	113	5	526	Biotroph
<i>Puccinia graminis</i> f. sp. <i>tritici</i>	28	1	37	152	80	3	301	Biotroph
FCN23	126	24	54	325	107	26	662	Hemibiotroph
<i>Fusarium graminearum</i>	100	17	42	246	97	22	524	Hemibiotroph
<i>Fusarium poae</i>	100	17	44	238	106	19	524	Hemibiotroph
<i>Fusarium verticillioides</i>	112	14	49	290	97	23	585	Hemibiotroph
<i>Magnaporthe oryzae</i>	114	14	53	251	94	6	532	Hemibiotroph
<i>Alternaria alternata</i>	164	10	58	280	94	25	631	Hemibiotroph
<i>Botrytis cinerea</i>	109	20	41	274	107	10	561	Necrotroph
<i>Valsa mali</i>	104	10	32	266	88	13	513	Necrotroph
<i>Verticillium dahliae</i>	99	18	47	247	85	31	527	Necrotroph
<i>Sclerotinia sclerotiorum</i>	78	16	35	222	84	5	440	Necrotroph

and 24 carbohydrate-binding modules (CBMs) (Table 3). Compared with other fungi, FCN23 has a larger potential glycoside hydrolase group, which may be an important virulence factor for its infection of the lotus (Table S8).

Prediction of secondary metabolite biosynthetic gene clusters. Fungi can produce many secondary metabolites, which have a variety of functions and great pharmacological potential (19). Related genes encoding enzymes that control the biosynthetic pathway of secondary metabolites are usually clustered and continuous in the fungal genome. In addition, the gene cluster for the synthesis of secondary metabolites usually also includes genes involved in the completion of the product metabolism pathway, transmembrane, and activation of related specific transcription factors. In the present study, a total of 65 gene clusters were distinguished in the strain FCN23 genome using antiSMASH software, including 1 tRNA-dependent cyclodipeptide synthases (CDPS), 2 betalactones, and 4 indoless. In addition, 29 nonribosomal peptide synthetases (NRPS), 17 polyketide synthases (PKS), and 12 terpene gene clusters were revealed, accounting for 89.23% of the total number of predicted gene clusters. Similar results were also observed in other ascomycetes and were more abundant than basidiomycetes. The details of the gene clusters in FCN23 and other fungi are displayed in Fig. 4.

Comparative transcriptome analysis of host-pathogen interactions. To analyze pathogenicity-associated genes of FCN23, RNA of infected lotus tissue and of pure mycelium were sequenced by paired-end sequencing on an Illumina NovaSeq 6000 platform.

**FIG 4** Comparison of secondary metabolism gene clusters, FCN23 and other fungi. PKS, polyketide synthase; NRPS, nonribosomal peptide synthetase; NRPS-like, NRPS-like fragment.

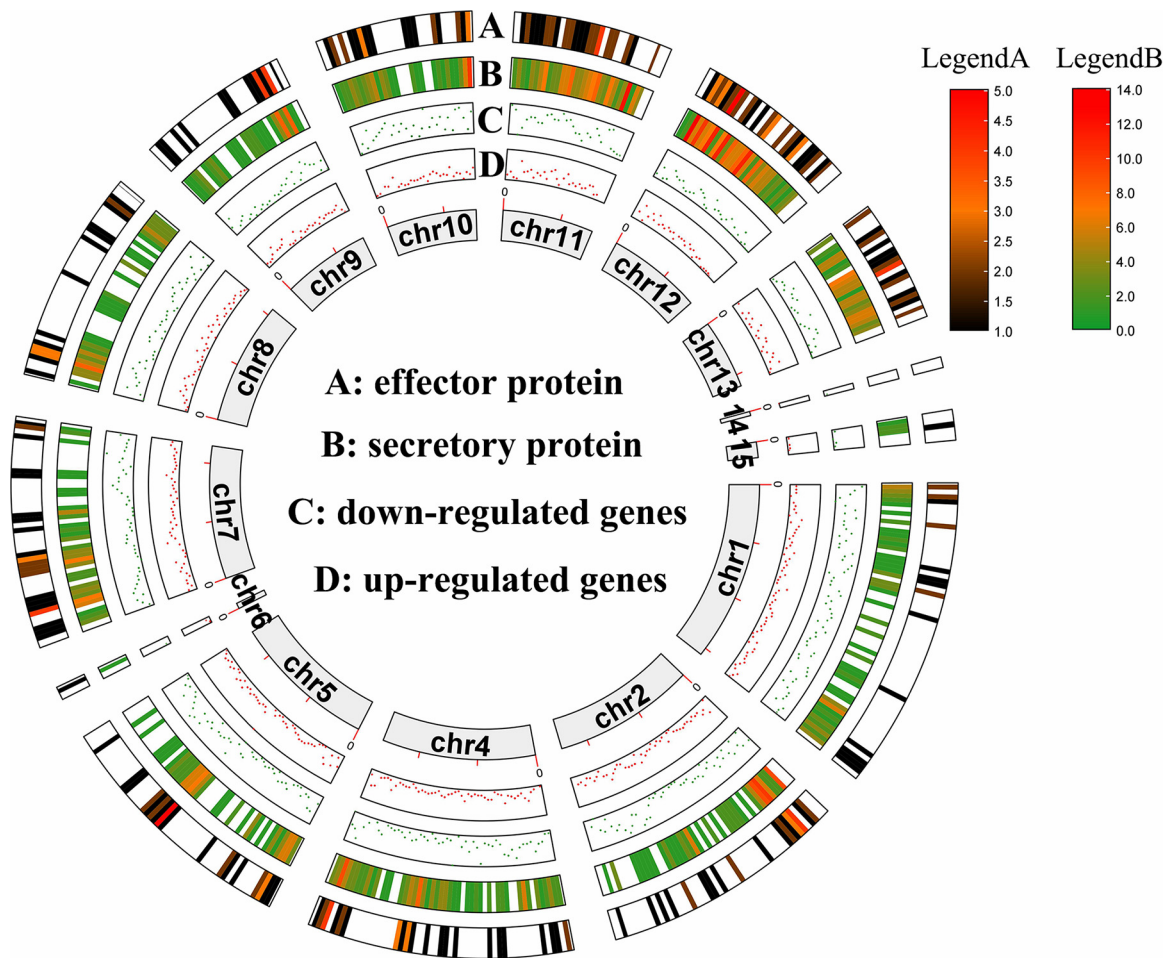


FIG 5 Global view of differentially-expressed genes, secreted proteins, and putative effector proteins in the FCN23 genome. (A) Chromosomal localization of putative effector proteins in the FCN23 genome. Color regions represent the location and number of putative effector proteins. (B) Chromosomal localization of putative secreted proteins. Color regions represent the location and number of putative secreted proteins. (C) The heat map of downregulated genes in FCN23 during infection of ‘Taikong lotus 36’. Each dot represents a gene. (D) The heat map of upregulated genes.

After trimming low-quality sequences, the cleaned RNA-seq reads were mapped to the reference genome of FCN23. The clean reads Q20 and Q30 value was higher than 93%. About 47.43 M and 43.70 M cleaned reads were obtained from the mycelium growth on medium and infected tissue, respectively. Approximately 97.37% and 38.43% of the clean reads were mapped to the FCN23 genome (Table S9). We performed correlation analysis between expression values of different samples. The Pearson’s correlation coefficient of the three biological replicates in the experimental group (LotD) exceeded 0.91, and in the control group (FCN) exceeded 0.98; the Pearson’s correlation coefficient between the experimental group and the control group was 0.40. The results showed higher similarities among samples with the same treatment (Fig. S7).

A total of 7,013 differentially expressed genes (DEGs) were identified, with 3,000 genes exhibiting a significant increase in abundance of transcripts during infection (Fig. S8). Through the analysis of DEGs locations, it is shown that the DEGs were mainly located on the core chromosome. Genomic mapping analysis showed that the DEGs were mainly located on the 11 core chromosomes (Fig. 5).

GO assignments were used for the functional classification of the upregulated genes. The three most basic ‘molecular function’ categories are transporter activity, transmembrane transporter activity, and coenzyme binding. The top three ‘biological processes’ are the carbohydrate metabolic process, small molecule metabolic process,

and organonitrogen compound biosynthetic process; and the top three “cell components” categories are cytoplasm, protein-containing complex, and cytoplasmic part.

To investigate biological pathways that are active during the fungal infection, the transcriptome of FCN23 was fit to the reference canonical pathways in KEGG. The 1,168 upregulated genes were mapped to 93 pathways. Among these pathways, ‘Biosynthesis of amino acids’ contained the highest percentage of genes (58 genes), followed by ‘Biosynthesis of cofactors’ and ‘Carbon metabolism’ (Fig. 6). These results suggest that FCN23 was active in secondary metabolite biosynthesis and catalytic activity during infection. Carbon metabolism plays an important role in cellulose degradation, and amino acid metabolism promotes humus synthesis.

Prediction and functional analysis of secretory effector proteins. In the present study, we first used SignalP to screen secretory proteins and identify 1,471 proteins with a secretion signal. Next, 1,218 proteins without transmembrane domains were identified in these proteins using TMHMM. Finally, 1,038 proteins contained glycoposphatidylinositol anchor motifs were predicted by PredGPI, accounting for 7.06% of FCN23 genome (Table S10). The region with the most amino acid length distribution was concentrated between 70 and 500 amino acid sequences, accounting for 69.56% of the putative secretory proteins.

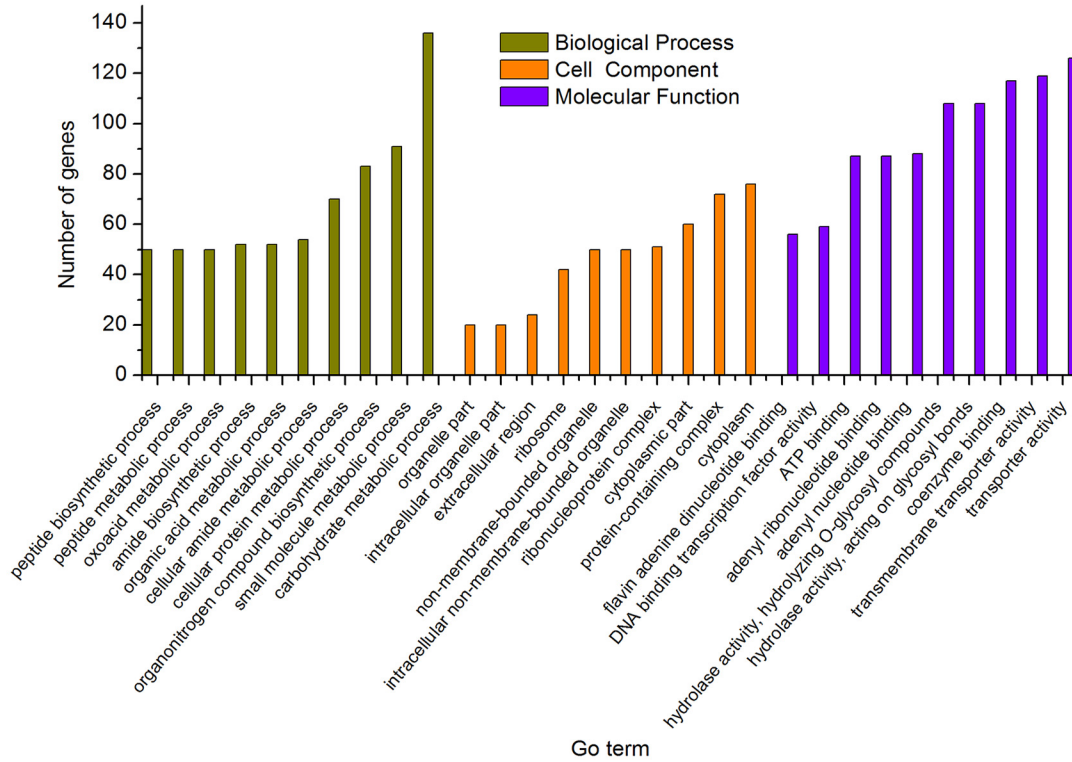
We further analyzed these predicted secretory proteins and screened the possible effector proteins. Finally, 296 candidates of small-secreted effector proteins were predicted using EffectorP 3.0, accounting for 2.0% of the total proteins and 28.5% of the total secreted proteins (Table S11). Most of them were identified as a hypothetical protein or uncharacterized protein. Chromosomal localization of putative secreted and effector proteins are shown in Fig. 5.

Fungal effectors play an important role in the interaction between pathogenic fungi and hosts, which directly affect the invasion, expansion, and disease occurrence of pathogenic fungi (20, 21). In addition, the function of most fungal effector proteins is unclear. In order to obtain information about the function of putative effectors, we selected eight candidate effector proteins (CEP) that were upregulated in the infection stage for transient expression in *N. benthamiana*. Primers used for functional analysis of putative effectors were listed in Table S12. F23a002499 shows a significant ability to trigger a hypersensitive response in plants, but all other seven candidate effectors do not, suggesting that F23a002499 is involved in the infection process of the pathogen (Fig. 7A). Quantitative RT-PCR analysis was applied to analyze the dynamic expression patterns of the F23a002499 gene at different time. The results showed that the expression of F23a002499 gene were highest at the early infection stage (especially at 12 h postinoculation) (Fig. 7B).

DISCUSSION

Fusarium commune is a typical soilborne disease, which can cause *Fusarium* wilt of many plants. The lotus is an important aquatic crop of considerable agricultural, ornamental, religious, and medical importance in Asia. The rhizome rot of lotus plants caused by *F. commune* is one of the most common and destructive diseases of lotuses. However, the study of pathogenic mechanism is very limited. With the popularization of high-throughput sequencing technology, the genomes of more and more species have been sequenced, including many important plant pathogenic fungi. Genome sequencing has laid an important foundation for accelerating the study of pathogenic biology and pathogenesis. In this study, the genome of *F. commune* isolated from aquatic plant was assembled using short-read sequencing technologies combined with the PacBio long reads. The size of the genome was approximately 46,211,149 bp, which is smaller than the reference genome sequence isolate from other land crops (17, 18). The FCN23 genome contains 11 core and 3 LS chromosomes, that were distinct from Fo4287 *F. oxysporum* f. sp. *Lycopersici* and Fo47. The complete LS chromosome 3 was missing in FCN23 genome. Comparative genomic studies show that the genome of *F. Oxysporum* is composed of 10–11 core chromosomes and varying

A



B

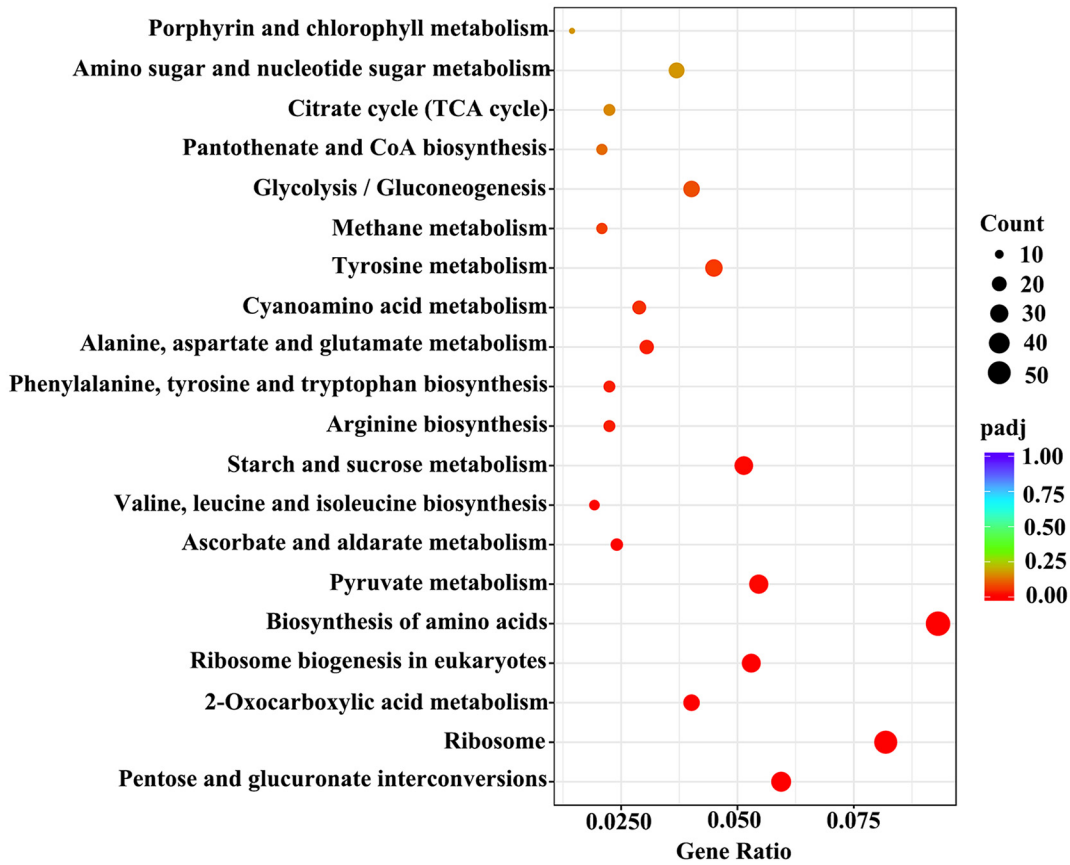


FIG 6 Enrichment analysis of upregulated differentially expressed genes. (A) Histogram of GO classifications. (B) KEGG enrichment scatter diagram.

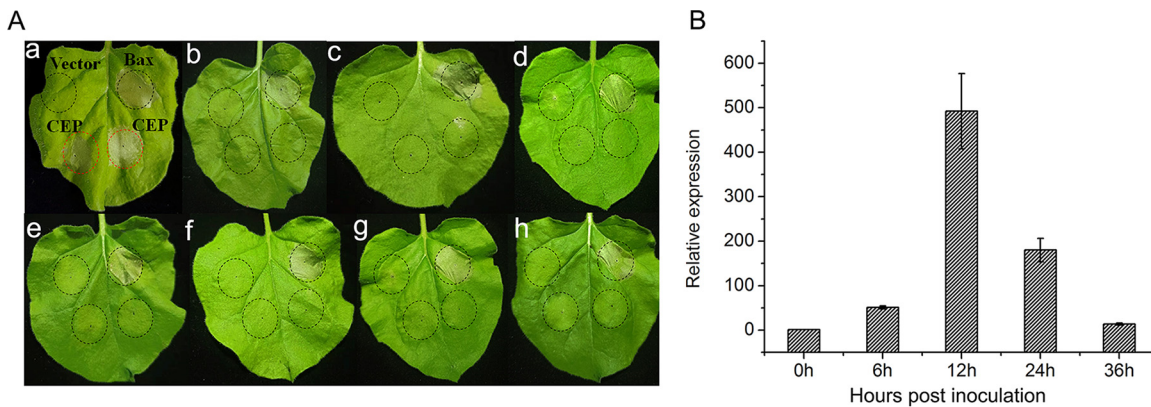


FIG 7 Functional analysis of putative FCN23 effectors. (A) Transient expression of the putative effector in *N. benthamiana* leaves via agro-infiltration. Leaves of *N. benthamiana* inoculated with *Agrobacteria* carrying putative effectors were assessed at 3 days postinoculation (positive control, Bax; negative control, pBinGFP); (a) F23a002499 induced hypersensitive reaction, (b-h) failed to induce hypersensitive reaction. (B) The dynamic expression patterns of the F23a002499 gene at 6, 12, 24, and 36 h postinoculation.

numbers of LS chromosomes (22). LS chromosomes are related to host-specific pathogenicity in phytopathogenic *F. oxysporum* (18). The differences of LS chromosomes 3, 6, 14, and 15 between the FCN23 and Fol4287 genomes may be caused by their host characteristics. These specific genomic differences can be explored as potential targets for studying the pathogenic mechanism of the FCN23.

The plant cell wall is the first barrier encountered by pathogens when they infect plants. The complexity of the plant cell wall also affected the diversity of CAZymes produced by pathogens. During the interaction between plant pathogenic fungi and their host, *F. oxysporum* will secrete a large number of CAZymes to overcome the plant cell wall, including pectinase, cellulase, xylanase, hemicellulose, and ligninase, etc. (23). CAZymes are conducive to the invasion, colonization and expansion of pathogenic fungi. There are 662 putative CAZyme-coding genes predicted in the FCN23 genome. The number is higher than that of the CAZymes in biotroph and necrotroph pathogens. The CAZyme profiles of FCN23 are close to that of soilborne pathogenic fungus, such as *F. verticillioides* and *V. dahliae*. Each of these CAZymes may play a role in different stages of the infection and expansion process. Although these enzymes may be important for the virulence of pathogens, they are a double-edged sword. When they are secreted into plants, they may cause plant defense reactions, such as callose deposition and programmed cell death (24, 25). Previous studies have shown that some CAZymes in *F. oxysporum* and other soilborne pathogens not only have the role of virulence factors, but also have the function of pathogen-associated molecular patterns (PAMPs), which can induce plant immune response (26–28). Most of the CAZymes identified in FCN23 belong to the GHs family with glycoside hydrolase domain, which indicates that there may be many protein hydrolases related to pathogen infection or plant immune induction in the GHs.

To further analyze the genes associated with pathogenicity, the transcriptome of FCN23 during infection of lotuses was analyzed and 7,013 DEGs were identified. GO and KEGG enrichment analysis of the upregulated genes suggests prevalence of genes associated with biosynthesis of amino acids, cofactors, and carbon metabolism during infection. Carbon metabolism plays an important role in plant cell wall degradation, and amino acid metabolism is essential for pathogen growth. Studies have shown that several genes involved in the biosynthesis of amino acids, cofactors and carbon metabolism are necessary for complete pathogenicity. During the interaction between pathogenic fungi and host plants, effectors proteins secreted by fungi are an important weapon for successful infection of plants (29, 30). We predicted 296 hypothetical effector proteins in this study. F23a002499 encoded 131 amino acids, belongs to the fungal type RNase, and was found to trigger a significant hypersensitive response by transient

expression in *N. benthamiana*. Bioinformatic analysis revealed that F23a002499 contains four cysteine residues and typical signal peptide sequences at the N-terminal. Secreted ribonucleases play a variety of roles in many different aspects of host pathogen interactions in plant systems. Recent studies have shown that plant pathogens such as *Zymoseptoria tritici* (31), *Blumeria graminis* f. sp. *hordei* (32, 33), and *F. graminearum* (34, 35) secrete several RNase-like effectors, which contribute to the pathogenic process of pathogens.

F. commune is an important soilborne plant pathogen, but there is limited research on how it causes aquatic plant *Fusarium* wilt. In this study, the high-quality assembled genome of *F. commune* isolated from aquatic lotuses in China was obtained using Illumina and PacBio sequencing platforms. In addition, we analyzed DEGs associated with pathogenesis by transcriptome sequencing. The genome and transcriptome sequences of FCN23 will provide important genomic information and insights into how pathogens interfere with host immunity for the success of infection. Further functional analysis of pathogenic candidate genes will improve our understanding of the interaction between *F. commune* and aquatic plants.

MATERIALS AND METHODS

Fungal strain and genome sequencing. The strain of FCN23 was isolated from diseased rhizomes of lotus plants in Jiangxi province of China (26°50'14"N, 116°19'32"E), and purified by single-spore subcultures on potato dextrose agar (PDA) for morphological identification. For molecular identification, the intergenic spacer region (IGS) and part of the translation elongation factor 1-alpha (*EF-1 α*) were amplified and sequenced with primer pairs CNS1/CNL12 (36) and EF1/EF2-21 (37). The resulting sequences were deposited in GenBank (accession numbers [ON642544](#) and [ON642545](#)). A phylogenetic analysis was performed using MEGA7 software with 1,000 bootstrap replicates. The FCN23 was further identified with specific primers (eFfc100F and eFfc385R) for *F. commune* and specific primers (FOF1 and FOR1) for *F. oxysporum* (38, 39). In the pathogenicity test, healthy lotus plants (cv. Taikong lotus 36) at the three-leaf stage were inoculated by dipping the roots into a conidial suspension of 1×10^6 conidia/mL for 3 h. The wounds of roots were inoculated with sterile distilled water as negative controls. The experiment was performed independently three times.

The strain was cultured on PDA medium at 25°C, and fresh mycelia were harvested and ground in liquid nitrogen. Total DNA of FCN23 was isolated from the mycelia using the EZNA Fungal DNA Kit (Omega Bio-Tek, USA) following the manufacturer's instructions. The genome sequencing was performed using the Illumina NovaSeq and the PacBio Sequel II sequencing platform. For Illumina sequencing, the library was constructed with the initial amount of 1 μ g DNA, and the DNA was interrupted to 300~500 bp with a Covaris M220 Focused Ultrasonicator (Covaris Inc, Woburn, MA, USA). The sequencing library was constructed according to the TruSeq DNA Sample Prep Kit (Illumina, San Diego, CA, USA) method. The Illumina NovaSeq platform was used for whole-genome sequencing. For PacBio sequencing, genomic DNA was sheared to 15–20 kb using G-tubes (Covaris Inc, USA) and converted into the proprietary SMRT Bell library format using the RS DNA Template Preparation Kit (Pacific Biosciences, Menlo Park, CA). The library was sequenced on the PacBio Sequel II platform.

Genome assembly. In order to make the subsequent assembly more accurate, low-quality bases and sequencing adapter sequences were trimmed and filtered from the raw Illumina reads using Trimmomatic v0.39 (40). The following operations were performed: (i) reads with a certain proportion of low quality (less than Q20) bases were removed, (ii) reads with the proportion of N bases up to 10% were removed, (iii) the adapter and the small fragments less than 75 bp in length after quality trimming were discarded. The following operations were performed in order to filter the raw data of the PacBio Sequel II platform: (i) polymerase reads with length <200 bp were filtered out, (ii) polymerase reads with quality score <0.80 were filtered out, (iii) adapter sequences were filtered out, then subreads were produced, (iv) subreads with length <200 bp were filtered out for further analysis. The clean data was then analyzed by k-mer (k-mer = 21) to estimate the genomic size, heterozygosity, and repetition rate (41).

Canu v2.1.1 (42) and MaSuRCA v3.4.2 were tested to assemble the FCN23 genome. Canu used PacBio long-reads to build contigs, while MaSuRCA combined both Illumina reads and PacBio long-reads. Racon v1.3.3 and Pilon v1.23 were used for further error corrections.

Non-encoded RNA and repetitive elements analysis. RNAMmer v1.2 (43) and tRNAscan-SE v2.0 (44) were used to predict the rRNA and tRNA contained in the genome. Infernal v1.1 (45) was used to make other kinds of ncRNA predictions based on the Rfam database and to classify statistics. The repeats were identified by RepeatMasker software (<https://github.com/rmhubble/RepeatMasker>). Tandem repeats finder (TRF) searches for tandem repeats in DNA sequences. RepeatMasker searches for scattered repetitive sequences by comparing sequences with a known repetitive database. The series repeat sequence was simulated by using the TRF software to copy the indel frequency by percentage and the adjacency mode, and the series repeat sequence was identified using statistical standards.

Protein-coding genes analysis. We identified the protein coding genes in the FCN23 genome by using a combination of *de novo* prediction, homology-based prediction, and transcriptome prediction. For *de novo* prediction, gene prediction was performed with Augustus v3.2.3 using *F. graminearum* as

the training set (46). For homology-based prediction, the protein sequences of homologous species were aligned against the matching proteins using Genewise v2.4.1 (47). In addition, RNA-seq data generated in this study were mapped to the FCN23 genome using TopHat v2.1.1, and transcriptome-based gene structures were obtained by Trinity v2.11.0. Finally, all gene evidence was integrated using EvidenceModeler (<http://evidencemodeler.github.io/>) (48). The gene set integrity was then evaluated with BUSCO software (49). The protein sequences of the predicted genes were compared with NCBI Non-Redundant (NR), Kyoto Encyclopedia of Genes and Genomes (KEGG), Swiss-Prot protein databases, and Gene Ontology (GO) databases by BLASTp (BLAST + 2.7.1 with E-value cut-off $<1e-05$).

Comparative genomics and PCR verification. The protein sequences of the FCN23 genome were compared with other fungi genomes downloaded from the Ensembl Fungi and NCBI databases. OrthoVenn (<https://orthovenn2.bioinfotoolkits.net/>) was used to analyze the orthologs, co-orthologs, and in-paralog pairs, as well as gene clusters function annotation (50). The tools were executed with the E-value cutoff of $1e-5$ for all-to-all protein similarity comparisons, and an inflation value of 1.5 for the generation of orthologous clusters using the Markov Cluster Algorithm. TBtools v1.098684 was used to analyze the gene distribution characteristics of core and lineage-specific chromosomes (51). We further verified the results of the putative chromosome assembly by PCR amplification. Twelve genes were randomly selected for PCR amplification using the FCN23 genomic DNA as template. The primer sets were listed in Table S4.

CAZymes and secondary metabolism gene clusters. The identification and annotation of carbohydrate active enzymes were performed on the dbCAN2 meta server (E-Value $<1e-15$, coverage >0.35) (52). Gene clusters related to secondary metabolite biosynthesis were identified using the antibiotics and secondary metabolite analysis shell (antiSMASH v6.0) online tool (<https://fungismash.secondarymetabolites.org>) (53). The parameter settings were maintained as the default parameter values.

Secretome and putative effectors. The signal peptides of FCN23 amino acid sequences were screened using SignalP 5.0 (54). The transmembrane domain of proteins with N-terminal signal peptides were analyzed by TMHMM 2.0. TargetP 2.0 was used to predict the subcellular localization of proteins, and proteins with signal peptides located in organelles were excluded (55). PredGPI (<http://gpcr.biocomp.unibo.it/predgpi/>) was used to predict glycosyl phosphatidyl inositol anchor signals and retain non-GPI-anchored proteins.

In order to obtain more reliable effector proteins, we further screened these secreted proteins. The EffectorP 3.0 (<http://effectorp.csiro.au/>) was used to distinguish effector proteins and secreted proteins in the FCN23 genome (56, 57). Fungal effectors have the characteristics of signal peptide, low molecular weight and rich cysteine residues (58, 59). Amino acid length (≤ 500) and the number of cysteine residues (≥ 3) were used as screening conditions to further predict and analyze the candidate effector proteins.

Inoculation treatment and RNA isolation. Cultivar 'Taikong lotus 36' was used as the plant host. The FCN23 strain was incubated in a 250 mL-flask with 150 mL of PDB (potato dextrose broth) on a rotary shaker at 25°C at 180 rpm for 3 d, and then the spores were harvested. The spore suspension with a concentration of 10^6 was prepared for inoculation. The young lotus whips, about 10 cm in length, were disinfected with 75% alcohol and rinsed with sterile water three times. For inoculation treatment, the lotus whip was punctured with a needle, and a $10\text{-}\mu\text{L}$ spore suspension was added to the surface of the wound. The lotus whip with sterile water was added as a control. Each treatment was repeated three times.

For transcriptome analysis of FCN23 during infection, the samples of pure culture and inoculation for 96 h were collected. RNA isolation and reverse transcription for cDNA synthesis were carried out using TRIzol reagent (Invitrogen, USA) and the PrimeScript RT reagent kit with gDNA Eraser (TaKaRa, Japan) according to the manufacturer's direction.

Transcriptome sequencing and analyses. A total of $1.5\ \mu\text{g}$ RNA from each sample was used as input material for RNA library preparation using the NEBNext UltraTM RNA Library Prep Kit for Illumina (NEB, USA) following manufacturer's instructions. The library was initially quantified using the Qubit2.0 Fluorometer and then diluted to $1.5\ \text{ng}/\mu\text{L}$. qRT-PCR was used to accurately quantify the effective concentration of the library to ensure the quality of the library. The library preparations were sequenced d by the Illumina NovaSeq 6000.

The raw reads were preprocessed by removing low-quality reads and adaptor reads using SOAPnuke (60). All the high-quality reads were aligned against the reference genome of strain FCN23 using HISAT v2.0.5 (61). The data of RNA-seq has been deposited in NCBI Short Read Archive database (BioProject number [PRJNA795051](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA795051)).

Gene expression and quantitative RT-PCR. According to the position information of gene alignment on the reference genome, the number of reads covered by each gene from start to end was counted using featureCounts v1.5.0-p3 (62). Reads with a comparison quality value lower than 10, reads on non-comparison pairs, and reads in multiple regions of the genome were filtered out. Gene expression levels were estimated by the fragments per kilobase per million mapped fragments (FPKM) for each sample. The differential expression analysis was performed using the DESeq2 R package v1.20.0 (63). Genes with an adjusted $\text{padj} \leq 0.05$ and $|\log_2(\text{foldchange})| \geq 1$ were assigned to be differentially expressed. Gene Ontology (GO) enrichment analysis of the differentially expressed genes (DEGs) was implemented by the clusterProfiler R package v3.8.1, in which gene length bias was corrected (64). GO terms with corrected P values less than 0.05 were considered significantly enriched by DEGs. Then we used clusterProfiler software to test the statistical enrichment of differential expression genes in KEGG pathways.

For quantitative RT-PCR, inoculated samples were collected at 0, 6, 12, 24, and 36 h postinoculation (hpi). Reactions were performed on Bio-Rad CFX96 real-time PCR detection system with the SYBR qPCR Master Mix (Transgen Biotech, China). For each gene, three biological replicates were prepared for qPCR

analysis and the average threshold cycle (Ct) was calculated. The relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method. The translation elongation factor 1- α (*EF-1 α*) gene of FCN23 was chosen as the endogenous reference gene. The F23a002499 gene was amplified with 499F (5'ACGACGACTGCTGGAA3') and 499R (5'TAAACACCGCCGCTCTTG3'), while the *EF-1 α* gene was amplified with EF1nF (5'TCAGGGTGCCGCTTCT3') and EF1nR (5'CTTGACGATGGCGGAGT3').

Transient expression analysis of candidate effectors in *Nicotiana benthamiana*. The candidate effector genes (without signal peptide regions) in the FCN23 genome were amplified from the cDNA library using KOD FX DNA polymerase (Toyobo, Osaka, Japan). The primer sets were listed in Table S12. The expression vector pBinGFP was linearized by digestion with KpnI and XbaI. Purified PCR product was cloned into pBinGFP by using ClonExpress MultiS One-Step Cloning Kit (Vazyme, China). Then these constructs were transformed into the *Agrobacterium tumefaciens* strain GV3101 using the freeze-thaw method.

The verified individual colonies of *A. tumefaciens* were cultured in LB medium containing rifampicin (10 μ g/mL) and kanamycin (50 μ g/mL) for 24 h. Then bacterial cells were centrifuged at 5000 rpm for 5 min, washed three times with 10 mM MgCl₂, and resuspended in MMA buffer (10 mM MgCl₂, 10 mM MES, 150 μ M acetosyringone, pH 5.6). The OD₆₀₀ was adjusted to 0.5 and left in a dark at room temperature for 3 h. Leaves of 4-week-old *N. benthamiana* plants were inoculated by needleless syringes as previously described (65). The pBinGFP was used as the negative control, while the addition of the proapoptotic protein BAX was used as the positive control that induced cell death. Cell death symptoms were evaluated and photographed at 3 days after infiltration. The experiments were repeated at least three times.

Data Availability. The entire genome of *Fusarium commune* FCN23 has been deposited at DDBJ/ENA/GenBank under the accession number [JAJTCY000000000](https://www.ncbi.nlm.nih.gov/nuccore/JAJTCY000000000). The data of RNA-seq has been deposited in NCBI Short Read Archive database (BioProject number: [PRJNA795051](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA795051)).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, XLSX file, 1.3 MB.

SUPPLEMENTAL FILE 2, PDF file, 0.9 MB.

ACKNOWLEDGMENTS

This research was funded by the Key Research and Development Program of Jiangxi Province of China, grant number 20161ACF60014 and The Educational Commission of Jiangxi Province of China, grant numbers GJJ190166 and GJJ170294. We are grateful to Jing Zhang from the Jiangxi Agricultural University for providing the vectors for the study.

We declare no conflict of interest.

REFERENCES

- Mukherjee PK, Mukherjee D, Maji AK, Rai S, Heinrich M. 2009. The sacred lotus (*Nelumbo nucifera*)-phytochemical and therapeutic profile. *J Pharm Pharmacol* 61:407–422. <https://doi.org/10.1211/jpp/61.04.0001>.
- Guo HB. 2009. Cultivation of lotus (*Nelumbo nucifera* Gaertn. ssp. *nucifera*) and its utilization in China. *Genet Resour Crop Evol* 56:323–330. <https://doi.org/10.1007/s10722-008-9366-2>.
- Wang Z, Yong C, Zeng M, Wang Z, Fang Q, Wang Y, Jie C, He Z. 2021. Lotus (*Nelumbo nucifera* Gaertn.) leaf: a narrative review of its phytoconstituents, health benefits and food industry applications. *Trends Food Sci Tech* 112:631–650. <https://doi.org/10.1016/j.tifs.2021.04.033>.
- Chen KL, Kirschner R. 2018. Fungi from leaves of lotus (*Nelumbo nucifera*). *Mycol Progress* 17:275–293. <https://doi.org/10.1007/s11557-017-1324-y>.
- Cui RQ, Sun XT. 2012. First report of *Curvularia lunata* causing leaf spot on lotus in China. *Plant Dis* 96:1068–1068. <https://doi.org/10.1094/PDIS-03-12-0234-PDN>.
- Koyama Y, Thar SP, Kizaki C, Toyota K, Sawada E, Abe N. 2013. Development of specific primers to *Hirschmanniella* spp. causing damage to lotus and their economic threshold level in Tokushima prefecture in Japan. *Nematol* 15:851–858. <https://doi.org/10.1163/15685411-00002723>.
- Skovgaard K, Rosendahl S, O'Donnell K, Nirenberg HI. 2003. *Fusarium commune* is a new species identified by morphological and molecular phylogenetic data. *Mycologia* 95:630–636. <https://doi.org/10.2307/3761939>.
- Dong C, Wang R, Zheng X, Zheng X, Jin L, Wang H, Chen S, Shi Y, Wang M, Liu D, Yang Y, Hu Z. 2018. Integration of transcriptome and proteome analyses reveal molecular mechanisms for formation of replant disease in *Nelumbo nucifera*. *RSC Adv* 8:32574–32587. <https://doi.org/10.1039/c8ra06503a>.
- Edel-Hermann V, Lecomte C. 2019. Current status of *Fusarium oxysporum* formae speciales and races. *Phytopathology* 109:512–530. <https://doi.org/10.1094/PHYTO-08-18-0320-RVW>.
- Nisikado Y, Watanabe K. 1953. On the rhizome rot of lotus, *Nelumbo nucifera* Gaertn., caused by a new *Fusarium*, *F. bulbigenum* Wr. *nelumbicolium* Nis. et Wat. *J Berichte Des Ohara Instituts Für Landwirtschaftliche Forschungen* 10:1–8.
- Minamikawa K, Nishizawa T, Saito H. 1959. Studies on the edible East Indian lotus, *Nelumbo nucifera* Gaertn. II: On the control of the rhizome rot of lotus. *Engei Gakkai Zasshi* 28:241–256. <https://doi.org/10.2503/jjshs.28.241>.
- Tang H, Zheng L, Yan SL, Wang QZ, Li J. 2017. First report of rot disease on post-harvest lotus roots caused by *Fusarium oxysporum* in China. *J Plant Pathol* 99:534.
- Deng S, Ma X, Chen Y, Feng H, Zhou D, Wang X, Zhang Y, Zhao M, Zhang J, Daly P, Wei L. 2022. LAMP assay for distinguishing *Fusarium oxysporum* and *Fusarium commune* in lotus (*Nelumbo nucifera*) rhizomes. *Plant Dis* 106:231–246. <https://doi.org/10.1094/PDIS-06-21-1223-RE>.
- Ma L-J, Geiser DM, Proctor RH, Rooney AP, O'Donnell K, Trail F, Gardiner DM, Manners JM, Kazan K. 2013. *Fusarium* Pathogenomics. *Annu Rev Microbiol* 67:399–416. <https://doi.org/10.1146/annurev-micro-092412-155650>.
- Zuriegat Q, Zheng Y, Liu H, Wang Z, Yun Y. 2021. Current progress on pathogenicity-related transcription factors in *Fusarium oxysporum*. *Mol Plant Pathol* 22:882–895. <https://doi.org/10.1111/mpp.13068>.
- Schmidt SM, Lukasiewicz J, Farrer R, Dam PV, Bertoldo C, Rep M. 2016. Comparative genomics of *Fusarium oxysporum* f. sp. *melonis* reveals the secreted protein recognized by the *Fom-2* resistance gene in melon. *New Phytol* 209:307–318. <https://doi.org/10.1111/nph.13584>.

17. Wang B, Yu H, Jia Y, Dong Q, Steinberg C, Alabouvette C, Edel-Hermann V, Kistler H, Ye K, Ma L, Guo L. 2020. Chromosome-scale genome assembly of *Fusarium oxysporum* strain Fo47, a fungal endophyte and biocontrol agent. *Mol Plant Microbe Interact* 33:1108–1111. <https://doi.org/10.1094/MPMI-05-20-0116-A>.
18. Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman PM, Kang S, Shim W-B, Woloshuk C, Xie X, Xu J-R, Antoniw J, Baker SE, Bluhm BH, Breakspear A, Brown DW, Butchko RAE, Chapman S, Coulson R, Coutinho PM, Danchin EGJ, Diener A, Gale LR, Gardiner DM, Goff S, Hammond-Kosack KE, Hilburn K, Hua-Van A, Jonkers W, Kazan K, Kodira CD, Koehrsen M, Kumar L, Lee Y-H, Li L, Manners JM, Miranda-Saavedra D, Mukherjee M, Park G, Park J, Park S-Y, Proctor RH, Regev A, Ruiz-Roldan MC, Sain D, et al. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* 464:367–373. <https://doi.org/10.1038/nature08850>.
19. Brakhage AA. 2013. Regulation of fungal secondary metabolism. *Nat Rev Microbiol* 11:21–32. <https://doi.org/10.1038/nrmicro2916>.
20. Dodds PN, Rathjen JP. 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. *Nat Rev Genet* 11:539–548. <https://doi.org/10.1038/nrg2812>.
21. Gassmann W, Bhattacharjee S. 2012. Effector-triggered immunity signaling: from gene-for-gene pathways to protein-protein interaction networks. *Mol Plant Microbe Interact* 25:862–868. <https://doi.org/10.1094/MPMI-01-12-0024-IA>.
22. Fokkens L, Guo L, Dora S, Wang B, Ye K, Sánchez-Rodríguez C, Croll D. 2020. A Chromosome-scale genome assembly for the *Fusarium oxysporum* strain Fo5176 To establish a model *Arabidopsis*-Fungal Pathosystem. *G3 (Bethesda)* 10:3549–3555. <https://doi.org/10.1534/g3.120.401375>.
23. Garcia-Maceira FI, Pietro AD, Huertas-Gonzalez MD, Ruiz-Roldan MC, Roncero MI. 2001. Molecular characterization of an endopolygalacturonase from *Fusarium oxysporum* expressed during early stages of infection. *Appl Environ Microbiol* 67:2191–2196. <https://doi.org/10.1128/AEM.67.5.2191-2196.2001>.
24. Zhang L, Kars I, Essenstam B, Liebrand T, Wagemakers L, Elberse J, Tagkalaki P, Tjoitang D, van den Ackerveken G, Kan J. 2014. Fungal endopolygalacturonases are recognized as microbe-associated molecular patterns by the *Arabidopsis* receptor-like protein RESPONSIVENESS TO BOTRYTIS POLYGALACTURONASES1. *Plant Physiol* 164:352–364. <https://doi.org/10.1104/pp.113.230698>.
25. Kachewar NR, Gupta V, Ranjan A, Patel HK, Sonti RV. 2019. Overexpression of OsPUB41, a rice E3 ubiquitin ligase induced by cell wall degrading enzymes, enhances immune responses in rice and *Arabidopsis*. *BMC Plant Biol* 19:1–17. <https://doi.org/10.1186/s12870-019-2079-1>.
26. Ma Z, Song T, Zhu L, Ye W, Wang Y, Shao Y, Dong S, Zhang Z, Dou D, Zheng X, Tyler BM, Wang Y. 2015. A *Phytophthora sojae* glycoside hydrolase 12 protein is a major virulence factor during soybean infection and is recognized as a PAMP. *Plant Cell* 27:2057–2072. <https://doi.org/10.1105/tpc.15.00390>.
27. Gui Y-J, Chen J-Y, Zhang D-D, Li N-Y, Li T-G, Zhang W-Q, Wang X-Y, Short DPG, Li L, Guo W, Kong Z-Q, Bao Y-M, Subbarao KV, Dai X-F. 2017. *Verticillium dahliae* manipulates plant immunity by glycoside hydrolase 12 proteins in conjunction with carbohydrate-binding module 1. *Environ Microbiol* 19:1914–1932. <https://doi.org/10.1111/1462-2920.13695>.
28. Zhang L, Yan J, Fu Z, Shi W, Ninkuu V, Li G, Yang X, Zeng H. 2021. FoEG1, a secreted glycoside hydrolase family 12 protein from *Fusarium oxysporum*, triggers cell death and modulates plant immunity. *Mol Plant Pathol* 22: 522–538. <https://doi.org/10.1111/mpp.13041>.
29. de Wit P, Mehrabi R, Burg H, Stergiopoulos I. 2009. Fungal effector proteins: past, present and future. *Mol Plant Pathol* 10:735–747. <https://doi.org/10.1111/j.1364-3703.2009.00591.x>.
30. Jones DA, Bertazzoni S, Turo CJ, Syme RA, Hane JK. 2018. Bioinformatic prediction of plant-pathogenicity effector proteins of fungi. *Curr Opin Microbiol* 46:43–49. <https://doi.org/10.1016/j.mib.2018.01.017>.
31. Kettles GJ, Bayon C, Sparks CA, Canning G, Kanyuka K, Rudd JJ. 2018. Characterization of an antimicrobial and phytotoxic ribonuclease secreted by the fungal wheat pathogen *Zymoseptoria tritici*. *New Phytol* 217:320–331. <https://doi.org/10.1111/nph.14786>.
32. Pedersen C, Ver Loren van Themaat E, McGuffin LJ, Abbott JC, Burgis TA, Barton G, Bindschedler LV, Lu X, Maekawa T, Wessling R, Cramer R, Thordal-Christensen H, Panstruga R, Spanu PD. 2012. Structure and evolution of barley powdery mildew effector candidates. *BMC Genomics* 13: 694. <https://doi.org/10.1186/1471-2164-13-694>.
33. Pliego C, Nowara D, Bonciani G, Gheorghe DM, Xu R, Surana P, Whigham E, Nettleton D, Bogdanove AJ, Wise RP, Schweizer P, Bindschedler LV, Spanu P. 2013. Host-induced gene silencing in barley powdery mildew reveals a class of ribonuclease-like effectors. *Mol Plant Microbe Interact* 26:633–642. <https://doi.org/10.1094/MPMI-01-13-0005-R>.
34. Hao Z, Li Y, Jiang Y, Xu J, Li J, Luo L. 2021. Genome sequence analysis of the fungal pathogen *Fusarium graminearum* using Oxford Nanopore Technology. *JoF* 7:699. <https://doi.org/10.3390/jof7090699>.
35. Yang B, Wang Y, Tian M, Dai K, Zheng W, Liu Z, Yang S, Liu X, Shi D, Zhang H, Wang Y, Ye W, Wang Y. 2021. Fg12 ribonuclease secretion contributes to *Fusarium graminearum* virulence and induces plant cell death. *J Integr Plant Biol* 63:365–377. <https://doi.org/10.1111/jipb.12997>.
36. Mbofung GY, Hong SG, Pryor BM. 2007. Phylogeny of *Fusarium oxysporum* f. sp. *lactucaae* inferred from mitochondrial small subunit, elongation factor 1- α , and nuclear ribosomal intergenic spacer sequence data. *Phytopathology* 97:87–98. <https://doi.org/10.1094/PHYTO-97-0087>.
37. O'Donnell K, Kistler HC, Cigelnik E, Ploetz RC. 1998. Multiple evolutionary origins of the fungus causing Panama disease of banana: concordant evidence from nuclear and mitochondrial gene genealogies. *Proc Natl Acad Sci U S A* 95:2044–2049. <https://doi.org/10.1073/pnas.95.5.2044>.
38. Stewart J, Abdo Z, Dumroese R, Klopfenstein N, Kim M-S. 2012. Virulence of *Fusarium oxysporum* and *F. commune* to Douglas-fir (*Pseudotsuga menziesii*) seedlings. *Forest Pathol* 42:220–228. <https://doi.org/10.1111/j.1439-0329.2011.00746.x>.
39. Mishra PK, Fox RT, Culham A. 2003. Development of a PCR-based assay for rapid and reliable identification of pathogenic *Fusaria*. *FEMS Microbiol Lett* 218:329–332. <https://doi.org/10.1111/j.1574-6968.2003.tb11537.x>.
40. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
41. Marçais G, Kingsford C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* 27:764–770. <https://doi.org/10.1093/bioinformatics/btr011>.
42. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
43. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007. Rfam: consistent and rapid annotation. *Nucleic Acids Res* 35: 3100–3108. <https://doi.org/10.1093/nar/gkm160>.
44. Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res* 44:W54–57. <https://doi.org/10.1093/nar/gkw413>.
45. Nawrocki EP, Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29:2933–2935. <https://doi.org/10.1093/bioinformatics/btt509>.
46. Mario S, Burkhard M. 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res* 33: 465–467.
47. Birney E, Clamp M, Durbin R. 2004. GeneWise and Genomewise. *Genome Res* 14:988–995. <https://doi.org/10.1101/gr.1865504>.
48. Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. 2008. Automated eukaryotic gene structure annotation using EvidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol* 9:R7. <https://doi.org/10.1186/gb-2008-9-1-r7>.
49. Seppy M, Manni M, Zdobnov EM. 2019. BUSCO: Assessing genome assembly and annotation completeness. *Methods Mol Biol* 1962:227–245. https://doi.org/10.1007/978-1-4939-9173-0_14.
50. Xu L, Dong Z, Fang L, Luo Y, Wei Z, Guo H, Zhang G, Gu YQ, Devin CD, Xia Q, Wang Y. 2019. OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res* 47:52–58.
51. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. 2020. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant* 13:1194–1202. <https://doi.org/10.1016/j.molp.2020.06.009>.
52. Zhang H, Tanner Y, Huang L, Sarah E, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. 2018. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 46:95–101.
53. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach M, Weber T, Takano E, Breitling R. 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* 39: 339–346.

54. Armenteros JJA, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 37:420–423. <https://doi.org/10.1038/s41587-019-0036-z>.
55. Emanuelsson O, Brunak S, Von Heijne G, Nielsen H. 2007. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2: 953–971. <https://doi.org/10.1038/nprot.2007.131>.
56. Sperschneider J, Dodds P. 2022. EffectorP 3.0: prediction of apoplastic and cytoplasmic effectors in fungi and oomycetes. *Mol Plant Microbe Interact* 35:146–156. <https://doi.org/10.1094/MPMI-08-21-0201-R>.
57. Sperschneider J, Gardiner DM, Dodds PN, Tini F, Covarelli L, Singh KB, Manners JM, Taylor JM. 2016. EffectorP: predicting fungal effector proteins from secretomes using machine learning. *New Phytol* 210:743–761. <https://doi.org/10.1111/nph.13794>.
58. Sperschneider J, Williams AH, Hane JK, Singh KB, Taylor JM. 2014. Evaluation of secretion prediction highlights differing approaches needed for oomycete and fungal effectors. *Front Plant Sci* 6:1168.
59. Selin C, de Kievit T, Belmonte MF, Fernando W. 2016. Elucidating the role of effectors in plant-fungal interactions: progress and challenges. *Front Microbiol* 7:600.
60. Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, Li Y, Ye J, Yu C, Li Z, Zhang X, Wang J, Yang H, Fang L, Chen Q. 2018. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and pre-processing of high-throughput sequencing data. *Gigascience* 7:1–6. <https://doi.org/10.1093/gigascience/gix120>.
61. Mortazavi A, Williams BA, Mccue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5: 621–628. <https://doi.org/10.1038/nmeth.1226>.
62. Yang L, Smyth GK, Wei S. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30:923–930. <https://doi.org/10.1093/bioinformatics/btt656>.
63. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>.
64. Yu G, Wang LG, Han Y, He QY. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16:284–287. <https://doi.org/10.1089/omi.2011.0118>.
65. De Oliveira M, Xu G, Li B, de Souza Vespoli L, Meng X, Chen X, Yu X, de Souza S, Intorne AC, de Manhães A, Musinsky A, Koiwa H, de Souza Filho G, Shan L, He P. 2016. Specific control of *Arabidopsis* BAK1/SERK4-regulated cell death by protein glycosylation. *Nat Plants* 2:15218. <https://doi.org/10.1038/nplants.2015.218>.