

## Article

# Genome-Wide Characterization Reveals Variation Potentially Involved in Pathogenicity and Mycotoxins Biosynthesis of *Fusarium proliferatum* Causing Spikelet Rot Disease in Rice

Ling Wang <sup>\*</sup>, Shuailing Ge, Wenhao Liang, Weiyang Liao, Wen Li, Gui'ai Jiao, Xiangjin Wei , Gaoneng Shao, Lihong Xie, Zhonghua Sheng, Shikai Hu, Shaoqing Tang and Peisong Hu <sup>\*</sup>

State Key Laboratory of Rice Biology, China National Center for Rice Improvement, China National Rice Research Institute, Hangzhou 311401, China

<sup>\*</sup> Correspondence: wangling03@caas.cn (L.W.); hupeisong@caas.cn (P.H.)

**Abstract:** *Fusarium proliferatum* is the primary cause of spikelet rot disease in rice (*Oryza sativa* L.) in China. The pathogen not only infects a wide range of cereals, causing severe yield losses but also contaminates grains by producing various mycotoxins that are hazardous to humans and animals. Here, we firstly reported the whole-genome sequence of *F. proliferatum* strain Fp9 isolated from the rice spikelet. The genome was approximately 43.9 Mb with an average GC content of 48.28%, and it was assembled into 12 scaffolds with an N<sub>50</sub> length of 4,402,342 bp. There is a close phylogenetic relationship between *F. proliferatum* and *Fusarium fujikuroi*, the causal agent of the bakanae disease of rice. The expansion of genes encoding cell wall-degrading enzymes and major facilitator superfamily (MFS) transporters was observed in *F. proliferatum* relative to other fungi with different nutritional lifestyles. Species-specific genes responsible for mycotoxins biosynthesis were identified among *F. proliferatum* and other *Fusarium* species. The expanded and unique genes were supposed to promote *F. proliferatum* adaptation and the rapid response to the host's infection. The high-quality genome of *F. proliferatum* strain Fp9 provides a valuable resource for deciphering the mechanisms of pathogenicity and secondary metabolism, and therefore shed light on development of the disease management strategies and detoxification of mycotoxins contamination for spikelet rot disease in rice.

**Keywords:** *Fusarium proliferatum*; spikelet rot disease; rice; genome; pathogenicity; mycotoxins

**Key Contribution:** The detailed characterization of genome of *F. proliferatum* revealed the extensive variability involved in pathogenicity and mycotoxins biosynthesis in *F. proliferatum*. The genome sequence represents a useful platform not only for dissecting the genetic determinants of plant–pathogen interactions, but also for reduction of mycotoxins exposure.



**Citation:** Wang, L.; Ge, S.; Liang, W.; Liao, W.; Li, W.; Jiao, G.; Wei, X.; Shao, G.; Xie, L.; Sheng, Z.; et al. Genome-Wide Characterization Reveals Variation Potentially Involved in Pathogenicity and Mycotoxins Biosynthesis of *Fusarium proliferatum* Causing Spikelet Rot Disease in Rice. *Toxins* **2022**, *14*, 568.

<https://doi.org/10.3390/toxins14080568>

Received: 19 July 2022

Accepted: 17 August 2022

Published: 19 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Spikelet rot disease in rice (*Oryza sativa* L.) is disease that has been emerging over recent years in China. It is widespread along the middle and lower reaches of the Yangtze River, due to the susceptibility of cultivated *japonica* varieties and *indica-japonica* hybrids. The initial symptoms of the disease are characterized by reddish-brown or rust-red spots on the glumes at the flowering and milking stage, then become blackish-brown at the ripening stage. Affected spikelets are partially filled or aborted and therefore do not produce edible grains. The average yearly occurrence of spikelet rot disease in rice affects more than 800,000 hectares in China. In warm and rainy regions, the prevalence of spikelet rot disease dominates, and yield losses can reach up to 30% in the field. In order to decrease damage of spikelet rot disease in rice, the application of fungicides is a primary control measure when weather conditions are conducive to fungal infection. However, the effectiveness of

fungicides varies and is strongly dependent on different factors, such as the resistance of the rice cultivar, disease severity, and spraying technology.

The filamentous ascomycete *Fusarium proliferatum* (Matsushima) Nirenberg [teleomorph: *Gibberella fujikuroi* (Kuhlman) Samuels], belonging to *Fusarium* section *Liseola*, was reported for the first time in 2011 as the primary cause of spikelet rot disease in rice in China [1]. The fungus is a notorious phytopathogen, not only resulting in devastating diseases on diverse agricultural crops, such as wheat, rice, barley, and maize [2], but also producing a wide spectrum of toxic secondary metabolites (SMs) or mycotoxins, including fumonisins, moniliformin, fusaric acid, fusarin C, and beauvericin, which can accumulate in food and feed crops and therefore pose risks to human and animal health. Especially the production of fumonisin mycotoxins has received greater attention than other mycotoxins. High levels of fumonisins are epidemiologically associated with adverse health effects, including leukoencephalomalacia in horses, pulmonary edema in swine, hepatotoxicity in rats, nephrotoxicity in rodents, and cancer of the esophagus as well as neural tube defects in humans [3]. Some evidence founded that fumonisins production by *F. proliferatum* as a potential virulence factor was required for development of disease symptoms on hosts [4,5]. To minimize human and animal exposure to mycotoxins produced by *F. proliferatum*, including fumonisins, moniliformin, beauvericin, and fusarins, regulatory organizations have established maximum permissible levels in cereals and the products by the European Union, the joint Food and Agriculture Organization (FAO), and World Health Organization (WHO) [6,7].

Despite the potential importance concerning economic impact or human health, the lifestyle of *F. proliferatum* remains poorly documented and is still required to be extensively investigated. Genome decoding is the first step to decipher the range of molecular basis driving genetic variability, host–pathogen interactions, and the regulation of secondary metabolism. However, our knowledge of the complete gene set of *F. proliferatum* still remains incomplete. To date, only a few whole-genome sequences are publicly available of *F. proliferatum* strains, the strain ITEM 2341 from date palm [8], the strain A8 from onion [9], strain ET1 from orchid, and strain NRRL 62905 from maize [10]. Further expansion of the genomic resources of *F. proliferatum* provides a basis for our understanding of the pathogenesis and toxin biosynthesis of this pathogen.

As rice is an important crop used for human consumption, its safety is of concern. Moreover, increasing evidence suggested that fungicides might not be efficient at reducing mycotoxins contamination [11]. Hence, comprehensive genetic analyses are needed to provide knowledge focusing on morphology, pathology, and toxigenicity from the fungal aspect to develop the practical intervention strategies of mycotoxins decontamination. In this study, we first report a high-quality draft genome for *F. proliferatum* isolated from rice. We then approximated the repertoire of the genes and characterized structural features of the genome. Then, we evaluated the genetic components involved in pathogenicity and secondary metabolism, with a particular emphasis on the main gene families and the biosynthetic gene clusters of mycotoxins. The results not only gave insight into the genetic architecture of *F. proliferatum* but also contributed to strategies for reduction of crop diseases and mycotoxin contamination, which could also impact positively on food safety and security.

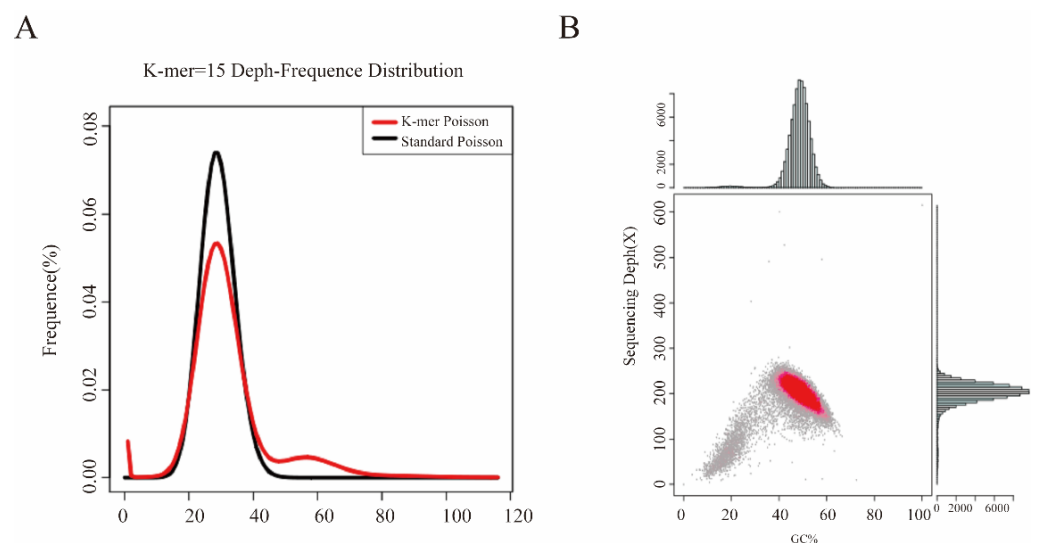
## 2. Results

### 2.1. Genome Sequence Assembly

The monoconidial strain Fp9 of *F. proliferatum* was isolated from an infected spikelet of rice (cv. Xiushui 09). The strain Fp9 was confirmed as *F. proliferatum* based on the concatenated alignment of the nucleotide sequences of *ITS*, *TEF-1 $\alpha$* , and  *$\beta$ -TUB* genes with MEGA (Figure S1).

The genome of the strain Fp9 was sequenced using massively-parallel-sequencing Illumina technology. One paired-end library with an average insert size of 400 bp was established to obtain clean reads of 5953 Mb, and two mate-pair libraries with insert

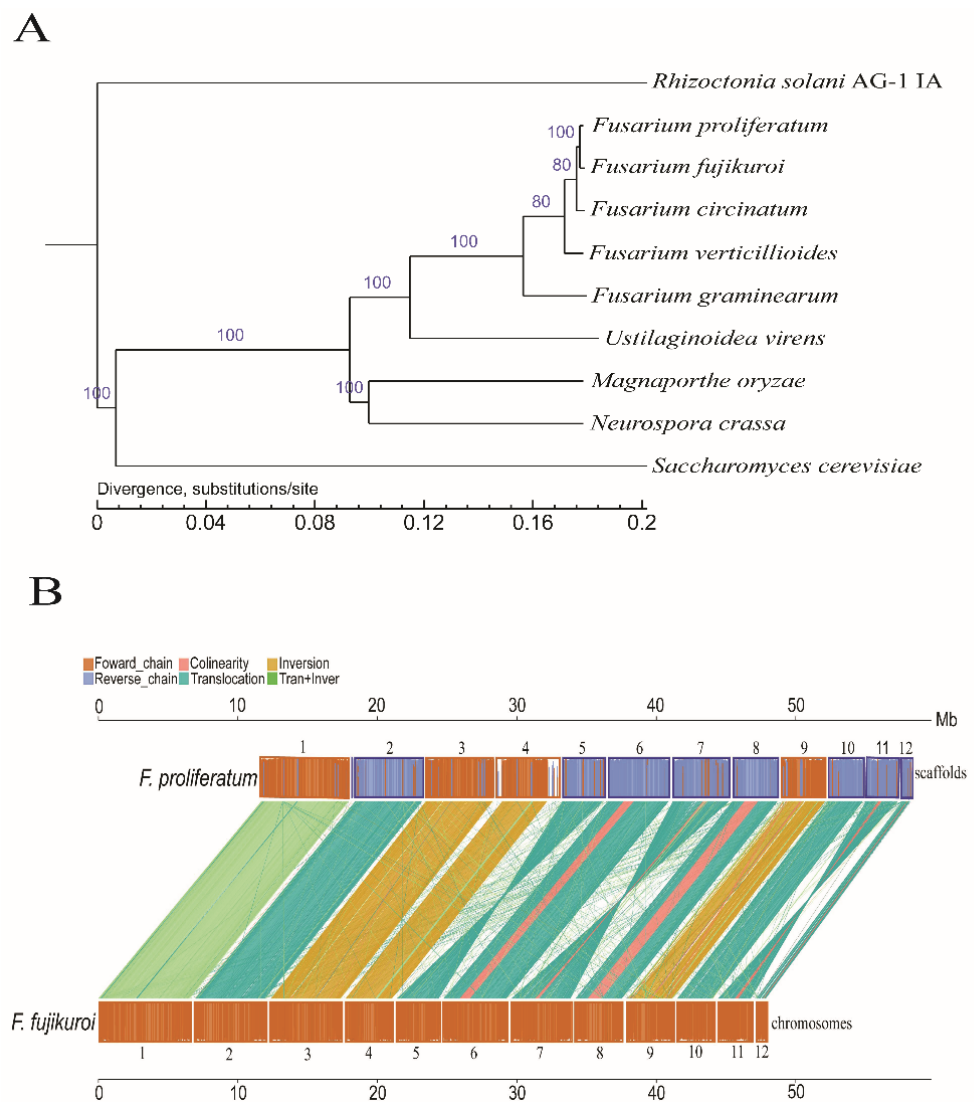
sizes of 2 kb and 6 kb were constructed to obtain clean reads of 3906 Mb and 4156 Mb, respectively. The error-corrected genome sequences were generated by mapping Illumina short reads to long reads. After discarding low-quality reads and adaptor reads, a total of 43,919,396 reads were obtained corresponding to approximately 120× sequence depth (Figure 1). The filtered sequences were assembled by SOAPdenovo into 190 contigs with N<sub>50</sub> length of 730,687 bp and N<sub>90</sub> length of 210,543 bp, respectively. The contigs were further assembled into 12 scaffolds with N<sub>50</sub> length of 4,402,342 bp and N<sub>90</sub> length of 2,484,728 bp, respectively. Accordingly, the estimated genome size was 43.9 Mb with an average GC content of 48.28%. The draft genome assembly had 98.4% and 99.2% completeness with the ascomycota\_odb10 dataset ( $n = 1706$ ) and fungi\_odb10 dataset ( $n = 758$ ), respectively, according to the evaluations with BUSCO, which suggests a high degree of completeness for the genome assembly. The genome sequence had been deposited in the GenBank database under the accession number WKFO00000000 (PRJNA517537 for BioProject and SAMN10829574 for BioSample).



**Figure 1.** The analyses of k-mer distribution and GC depth distribution for Illumina sequencing data in *F. proliferatum*. (A) Fifteen k-mer depth distribution of whole-genome Illumina reads. (B) The relationship between GC content and sequencing depth of the genome data used for assembly. The average sequencing depth was detected at near 120-fold coverage of the genome.

## 2.2. Genome Comparison and Phylogenetic Relationship

Phylogenetic analysis for the evolutionary relationship between the *F. proliferatum* strain Fp9 and other fungal species (eight ascomycota and one basidiomycota outgroup) showed that *F. proliferatum* was most closely related to *Fusarium fujikuroi*, the causal agent of the bakanae disease of rice (Figure 2A). Despite the shared genomic synteny in both genera (Figure 2B), it is noteworthy that the large fragments containing species-specific genes were significantly enriched in the subtelomeric regions, e.g., 56.2-kb and 24.6-kb, and inserts were present in far end of the left and right arms of scaffold IV in *F. proliferatum* with respect to *F. fujikuroi* (Table S1). Functional enrichment analysis using the MIPS ‘FunCat’ program revealed that subcategories of the ‘oxidation-reduction process’ was remarkably enriched in *F. proliferatum* and underrepresented in *F. fujikuroi*, whilst genes annotated as being integral to membrane, nucleic acid binding, zinc ion binding, regulation of transcription, and metabolic process were significantly higher in *F. fujikuroi* than in *F. proliferatum* (Table S2). To validate structural variants (SVs) between *Fusarium proliferatum* and *Fusarium fujikuroi*, 20 SVs were randomly selected for performing PCR amplification and paired-end Sanger sequencing. Each amplification gave the product size expected from the event inferred in silico (Table S3, Figure S2). The sequencing of amplified products likewise confirmed the inferred events in each case (Additional File S1).



**Figure 2.** The phylogenomic relationship of *F. proliferatum* with other filamentous fungi. **(A)** Phylogenetic tree of *F. proliferatum* and other nine selected fungal species (eight Ascomycota and one Basidiomycota). Basidiomycota fungus *Rhizoctonia solani* AG-I IA was used as the outgroup. Bootstrap values were calculated from 1000 replicates and marked in each node. **(B)** A whole-genome alignment between *F. proliferatum* and *F. fujikuroi*. The synteny of nucleotide sequences was indicated by vertical lines. In the upper and lower axes, the brown frame represented the genome forward chain, the blue frame represented the genome reverse chain, and the height of the color filled in the frame represented the similarity degree of the comparison. The color comparison types of the linked graph between the upper and lower axes were as follows: Collinear (red), Translocation (green), Inversion (yellow), and Tran + Inver (light green).

### 2.3. Expansion of Prominent Gene Families

A total of 14,054 protein-coding genes were predicted for the genome of *F. proliferatum* strain Fp9 by GeneMarkS with integrated model. For the functional annotations of predicted genes, all genes were interrogated into the known databases. Among the predicted coding genes, 13,798 (98.18%), 4269 (30.38%), 2226 (15.84%), 2085 (14.84%), 3279 (23.33%), and 13,248 (94.26%) genes had homologs with known functions in the Non-Redundant Protein Database databases (NR), Kyoto Encyclopedia of Genes and Genomes (KEGG), Clusters of Orthologous Groups (COG), Gene Ontology (GO), Swiss-Prot, and TrEMBL databases, respectively.

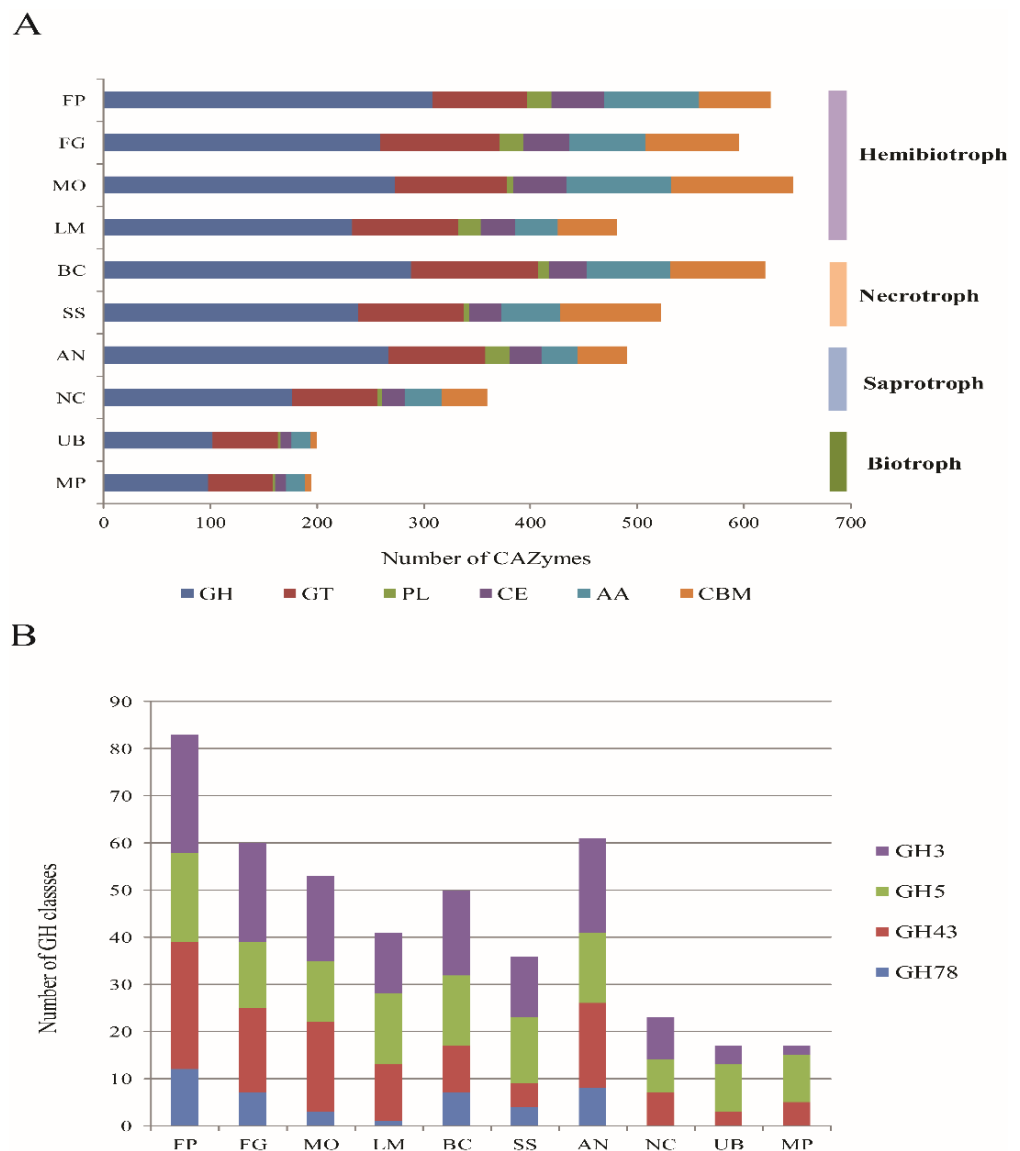
As a fungal plant pathogen, 1591 (11.32%) genes were considered to be involved in interaction between the pathogen and host, according to the best-matched with the pathogen–host interaction database (PHI-base). A total of 1222 (8.70%) genes harboring N-terminal signal peptides were predicted to produce the secreted proteins using SignalP 5.0. The secreted protein-encoding genes were thought to be one of the essential components of aggressiveness, which randomly distributed across the genome of *F. proliferatum*, with no evidence of clustering on particular chromosomes as had been reported for *Fusarium oxysporum* [12]. We found that 1402 (9.98%) genes were predicted to be membrane transporters in *F. proliferatum* as indicated by Transporter Classification Database, which surpassed the number of most other ascomycetes (Table S4). *F. proliferatum* encoded considerably major facilitator superfamily (MFS), which was the largest classes of membrane transporters.

Additionally, 625 (4.45%) genes encoding putative carbohydrate-active enzymes (CAZymes) were identified with the CAZymes database, of which, 308 glycosyl hydrolases (GHs), 89 glycoside transferases (GTs), 23 polysaccharide lyases (PLs), 49 carbohydrate esterases (CEs), 89 auxiliary activities enzymes (AAs), and 67 carbohydrate-binding modules (CBMs) (Table S4). The CAZymes repertoires of *F. proliferatum* were similar to those of the hemibiotrophic and necrotrophic fungi, i.e., *Magnaporthe oryzae*, *Fusarium graminearum*, and *F. fujikuroi*, but more expanded than those found in the saprotrophic and biotrophic fungi (Figure 3A). Particularly, some families of carbohydrate and non-carbohydrate substrates (GHs) and pectin-based components (PLs), such as GH3, GH5, GH43, and GH78, were enriched in *F. proliferatum* compared with other fungi representing various nutritional lifestyles (Figure 3B, Table S5).

#### 2.4. Diversification of the Biosynthetic Gene Clusters for Mycotoxins

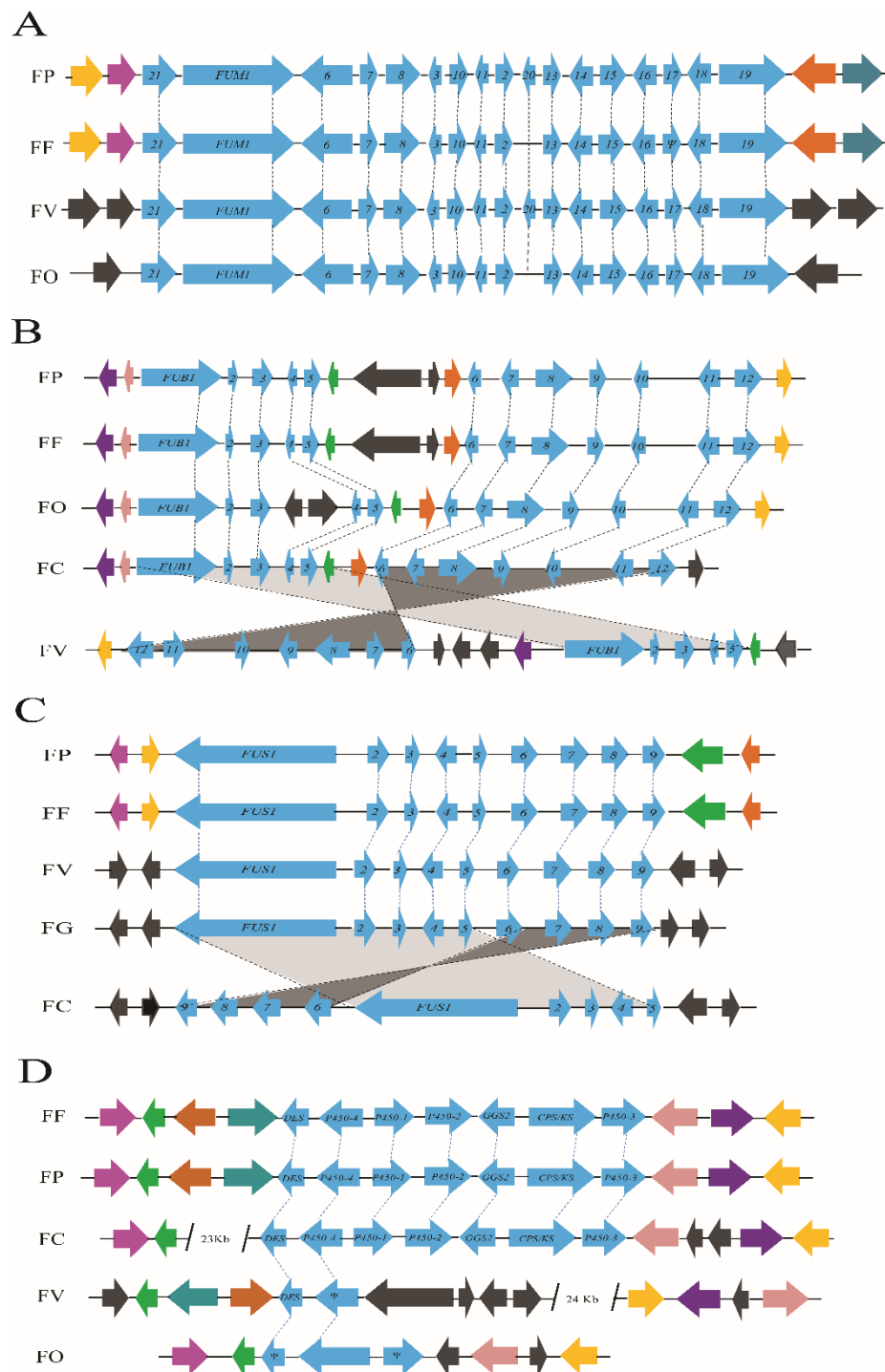
To estimate the genetic potential of *F. proliferatum* to produce SMs, genes encoding key classes of SM-associated enzymes were predicted. Altogether, there were 778 genes present in 57 gene clusters in *F. proliferatum*, which catalyzed synthesis of molecules that serve as core structures for secondary metabolites, i.e., fifteen polyketide synthases (PKSs), eleven nonribosomal peptide synthetases (NRPSs), five hybrid PKS-NRPS, fourteen terpenes, one lantipeptide, as well as eleven other clusters. A genome-wide comparison of gene clusters among *Fusarium* genomes showed that most of the gene clusters of *F. proliferatum* were typically conserved compared to other *Fusarium* species, whilst some of SM genes had marked differences in gene content of homologous clusters.

The content and arrangement of the fumonisin gene (*FUM*) cluster in *F. proliferatum* was very similar to that in *F. fujikuroi* and *Fusarium verticillioides*, and the *FUM* cluster was arranged in a highly collinear manner in the three fusaria (Figure 4A). Although the gene order and orientation of *FUM* cluster in *F. proliferatum* were the same as those in *F. verticillioides*, there was striking difference in the cluster flanking genes between them. The fusaric acid gene (*FUB*) cluster showed collinearity among fusaric acid (FA)-producing fusaria, but the homologous segments were separated by three or four genes which were not involved in FA biosynthesis, and the *FUB* genes exhibited a discontinuous distribution (Figure 4B). The genomic locations of orthologous genes of *FUB* cluster in *F. verticillioides* were inverted in comparison to *F. proliferatum*. In contrast to the *FUB* cluster, the biosynthetic gene clusters of the bikaverin and fusarubin, which are responsible for production of the PKS-derived perithecial pigmented naphthoquinone, were almost perfectly maintained with high levels of similarity of gene content, and flanking genes at least one side were largely collinear amongst the members of *Gibberella fujikuroi* species complex (GFC) (data not shown). The biosynthetic gene cluster of apicidin F was present in the genome sequence of *F. fujikuroi* but absent from that of the highly-related species *F. proliferatum* and the other *Fusarium* species examined, e.g., *F. verticillioides*, *F. oxysporum*, or *Fusarium circinatum*.



**Figure 3.** Carbohydrate-active enzymes (CAZymes) significantly enriched in *F. proliferatum* relative to other fungi with various lifestyles. **(A)** Number of fungal CAZyme families of *F. proliferatum* and other fungal species. Hemibiotroph: FP, *Fusarium proliferatum*; FG, *Fusarium graminearum*; MO, *Magnaporthe oryzae*; LM, *Leptosphaeria maculans*. Necrotroph: BC, *Botrytis cinerea*; SS, *Sclerotinia sclerotiorum*; Saprotroph: AN, *Aspergillus nidulans*; NC, *Neurospora crassa*. Biotroph: UB, *Ustilago bromivora*, MP, *Melanopsichium pennsylvanicum*. Families: GH, glycoside hydrolases; GT, glycosyl transferase; PL, polysaccharide lyase; CE, carbohydrate esterase; AA, auxiliary activities enzymes; CBM, carbohydrate binding module. **(B)** Comparison of GH repertoires of *F. proliferatum* and other fungal species. Families GH3 and GH5 encode cellulose-degrading enzymes, family GH43 encodes both cellulose and pectin-degrading enzymes, and family GH78 encodes pectin-degrading enzymes.

The flanking regions of the fusarin C gene (*FUS*) cluster do not exhibit homology among *Fusarium* species, and *FUS* genes were reversed in *F. circinatum* compared to *F. proliferatum* (Figure 4C). The flanking genes of the *FUS* cluster were arranged markedly different among *Fusarium* genomes examined. The homologs of the entire gene cluster for gibberellin (GA) biosynthesis in *F. fujikuroi* were also found in *F. proliferatum* (Figure 4D). *F. circinatum* had one homolog of the GA cluster, whereas the gene content of the flanking regions was absent. In contrast, *F. verticillioides* and *F. oxysporum* lacked GA clusters but had partial, nonfunctional homologs that were present in a similar syntenic region as the GA cluster in *F. fujikuroi*.



**Figure 4.** Comparison of the biosynthetic gene clusters of secondary metabolites in *F. proliferatum* with homologous regions in other *Fusarium* species. (A) Content and arrangement of the fumonisins biosynthetic gene (*FUM*) cluster in *Fusarium* species. (B) Content and arrangement of the fusaric acid gene (*FUB*) cluster in *Fusarium* species. (C) Content and arrangement of the fusarin C gene (*FUS*) cluster in *Fusarium* species. (D) Content and arrangement of the gibberellin gene (*GA*) cluster in *Fusarium* species. Organisms: FP, *Fusarium proliferatum*; FF, *Fusarium fujikuroi*; FV, *Fusarium verticillioides*; FO, *Fusarium oxysporum*; FG, *Fusarium graminearum*; FC, *Fusarium circinatum*. The genes considered to be part of the gene clusters were represented by horizontal blue arrows. The same colored arrows represented homologous genes in two or more fungi. Black arrows indicated that the gene did not have a closely-related homolog in the SM cluster of the species examined. The Greek letter Ψ indicated a pseudogene. The arrowheads indicated the direction of transcription of the gene.

### 3. Discussion

In recent years, spikelet rot disease in rice has been considered an emerging disease even where it was previously absent or scarce in China. Management efforts using resistant varieties and/or fungicides were inadequate for disease control due to mycotoxins contamination. Given its economic and public health importance, the causal agent *F. proliferatum* has been attracted greater attention. The availability of genome sequences has facilitated in the understanding of genome composition and host virulence as well as the biosynthetic capacity of natural compounds. Although there were several published genome databases of *F. proliferatum*, these were not sampled from rice fields. Here, we presented a genome of *F. proliferatum* strain Fp9 collected from rice spikelet in the fields, which allowed us to make the comprehensive investigation of the genetic bases of *F. proliferatum*.

In the current study, next-generation sequencing (NGS) technology was applied to generate the *F. proliferatum* reference genome. The high quality genome sequence of strain Fp9 that was assembled into twelve scaffolds was used for structural annotation and comparative analysis. However, NGS technologies (such as Illumina, 454, and Ion-torrent) have shortcomings in their capability to investigate repetitive elements, structural variants, or native base modifications [12,13]. Several of the limitations can be overcome by third-generation sequencing technologies (TGS). The advent of TGS technologies, such as the Pacific Biosciences (PacBio) and Oxford Nanopore machines, provides new possibilities for contig assembly, scaffolding, and high-performance computing in bioinformatics due to its long reads [14]. The high level of error can be reduced to <1% with the PacBio technology with circular consensus long read (CCS) approach [15]. NanoVar generated by Nanopore technologies exhibits higher structural variant calling accuracy using low-depth simulated datasets [16]. For projects where both genome continuity and consensus accuracy matters, a combination of longer reads and transcriptome analysis may be needed to obtain the accurate and realistic genome of *F. proliferatum* in the future studies.

The phylogenetic analysis revealed the high synteny between *F. proliferatum* and *F. fujikuroi*, the causative agent of the bakanae disease of rice. Abundant species-specific genes were found near the subterminal regions. Variation in the proximal telomere was also characterized in *Aspergillus* species by the expansion and deletion of the unique genes with a bias towards subtelomeric locations [17]. Such phenomenon is common in fungal pathogens and more generally in eukaryotic genomes. In fungi, gene loss and changes in genomic architecture are important adaptation processes, which may be a consequence from their parasitic lifestyle that pushes their need to occupy an ecological niche [18]. Without selective pressure, the higher speed of diversification in telomeric-like and subtelomeric-like regions had been already reported in the genome of *F. graminearum* [19]. Codon usage indices suggested that subtelomeric genes use a lower frequency of optimal codons and display a lower codon adaptation index with a higher effective number of codons [20]. Our analyses provided the support for the hypothesis that species-specific genes were concentrated in subtelomeric regions that led to genetic divergence between *F. proliferatum* and *F. fujikuroi*. The most abundant category of *F. proliferatum*-specific genes with a known function that was associated with oxidoreductase, which are likely be related to several important facets of fungal biology, such as cell differentiation, growth, and reproduction [21]. As a follow up to this study, we will focus on specific genes that benefit potential contribution by conferring adaptive metabolic advantages in the evolution of *F. proliferatum*.

MFS transporters are currently the largest characterized superfamily of transmembrane secondary transport proteins, which are responsible for nutrient uptake, metabolite extrusion, and resistance to various toxic compounds, including not only secondary metabolites but also fungicides and antibiotics [22]. Most significantly, MFS transporters were enriched in *F. proliferatum*, that likely cooperate and collectively contribute to mycotoxin efflux, as the self-defense of MFS transporters against deoxynivalenol (DON) in *F. graminearum* [23]. Intriguingly, MFS transporters could function directly to avoid the toxicity of chemicals by pumping them out of the cells, thus provided defense against toxic oxidants and acted as a virulence factor by protecting the fungi from toxic compounds [24].



On the other hand, MFS transporters play a role in nutrient availability for the survival, including the transport of lipids, ions, and small metabolites, as previously implicated in *Botrytis cinerea* [25]. Therefore, the large number of MFS transporters in the genome of *F. proliferatum* might involve in nutrients uptake from hosts to survive in the first site of infection, and detoxification in counteracting the physiological impact of antimicrobial host defense compounds to thrive in stressful host niche colonization sites.

In order to breach plant barriers for invasion and colonization, the phytopathogenic fungi were equipped with abundant and diverse enzymes for the degradation of plant polysaccharide materials. In this study, a comparison of the abundance of the full repertoires of CAZymes was performed among saprotrophic, necrotrophic, hemibiotrophic, and biotrophic fungi. We found that hemibiotrophic and necrotrophic fungi tended to have more CAZymes than biotrophic and saprotrophic fungi. As a hemibiotroph, *F. proliferatum* undergoes a shift in lifestyle from biotrophy to necrotrophy during the infection process. The CAZymes families related to the hydrolytic activity of celluloses (GHs) and the oxidative degradation of pectins (PLs) were more prevalent in *F. proliferatum* than other test fungi with different nutritional modes. As mentioned previously, GHs and PLs are often known as cell wall-degrading enzymes (CWDEs) due to their important roles in plant biomass decomposition by fungal pathogens [26]. Hemibiotrophic fungi obtain nutrition for growth and development mainly by the localized degradation of cell walls with a variety of CWDEs. Such expansion of CWDEs could explain, at least in part, the capacity of *F. proliferatum* for the penetration and successful infection of the hosts due to high cellulolytic, hemicellulolytic, and pectinolytic abilities. As a matter of fact, epidemic spreading across multiple hosts were caused by introductions of pathogen from restricted geographic regions, as reported for *Phytophthora infestans* [27]. The broad host-range pathogens were obviously more prone to rapidly respond to the environmental condition of a new habitat [28]. For spikelet rot disease in rice, the disease prevalence was likely relevant to multiple host adaptability of *F. proliferatum*. Clarifying the specific biological functions of CAZymes participating in the interactions with hosts, which likely provide information for a better understanding of its pathogenic process and host colonization.

Hemibiotrophy includes the two sequential stages of infection, biotrophy, and necrotrophy, in a series of steps that involve the participation of different virulence factors [29]. In *Colletotrichum higginsianum* infecting *Arabidopsis thaliana* and *Colletotrichum graminicola* infecting maize, comparative genomics showed that both fungi had large sets of pathogenicity-related genes, but families of genes encoding secreted effectors, pectin-degrading enzymes, secondary metabolism enzymes, transporters, and peptidases were expanded [30]. *Ustilaginoidea virens* (Cooke) Takah is an ascomycetous fungus that causes rice false smut, possessed reduced gene inventories for polysaccharide degradation, nutrient uptake, and secondary metabolism, which may play essential roles in a biotrophic lifestyle and adaptation to the specific floret infection [31]. In *Penicillium digitatum*, genes encoding plasma membrane transporters that may be required for assimilating major polysaccharides in the host wall, for example, oligopeptides, amino acids, and sugars, are also induced [32]. The rich sources of MFS transporters and CWDEs in *F. proliferatum*, suggested the pathogen required a maximum capacity for host manipulation during intracellular colonization, and biotrophic hyphae could provide the uptake of sugars and amino acids, which might be an adaptation to exploit a broad range of host plants by its multistage hemibiotrophic infection strategy. Major hemibiotrophic plant pathogens, such as the rice blast fungus *Magnaporthe oryzae* and charcoal rot fungus *Macrophomina phaseolina*, undergo major transformations in cell morphology and infection mode when switching from growth on the plant surface to intracellular biotrophy and from biotrophy to necrotrophy [33,34]. In this case, *F. proliferatum* invades the host cell in the biotrophic stage, then produces cellulolytic enzymes to hijack host metabolic pathways for a better uptake of plant nutrients by transporters during the necrotrophic stage, and finally, the fungus differentiates fast-growing hyphae that kill and destroy host tissues. MFS transporters and CWDEs in *F. proliferatum*

may thus be particularly important for the ability of the fungus to rapidly adapt to changing environments and occupy a wide range of habitats, including a diversity of crop plants.

In filamentous fungi, SMs are usually regarded as not essential for fungal growth and development, but they fulfill various functions, such as virulence factors, chemical signals for communication, defense against competitors, and symbiotic interactions [35]. Fungal SM biosynthetic genes were typically located in clusters for common gene regulation. Several studies had explored the evolutionary drivers of divergence of SM genes, including point or indel null mutations in the core genes, gene-conversion, inversion, and duplication/deletion, gene rearrangements and HGT events [36–38]. The species of the fungus *Fusarium* were known for the potential to synthesize structurally distinct SMs and some were known to be involved in playing roles in triggering host cell death and disease development [39,40]. The comparison of the genome of *F. proliferatum* to those of other fusaria revealed that most of SMs biosynthetic genes were relatively conserved. However, several gene clusters displayed multiple and disparate patterns of distribution, which exhibited the divergence of metabolites in the genus *Fusarium*. A similar situation was found for the aflatoxin gene cluster in the genus *Aspergillus* section Flavi, some of the species produced aflatoxin (e.g., *A. flavus*, *A. parasiticus*, and *A. nomius*), whereas others had lost the ability for aflatoxin production (e.g., *A. oryzae* and *A. sojae*) [41]. It appeared that structure and distribution of SM gene clusters contributed to metabolic diversity of the fusaria. However, only few gene clusters responsible for the natural products were already well characterized, probably due to the fact that most are not or only minimally expressed under standard laboratory conditions [42]. Furthermore, comprehensive analyses are required to fully understand the regulation of secondary metabolism in the fusaria.

The genus *Fusarium* can potentially synthesize harmful mycotoxins deriving from secondary metabolism, as exemplified by fumonisins, trichothecenes, zearalenone, moniliformin, and beauvericin. The economic impact of *Fusarium*-induced mycotoxin contamination via its effects on health of humans and livestock as well as on international trade was estimated to be in the hundreds of millions of US dollars every year [43]. The examination of the genetics of mycotoxins biosynthesis is motivated by the need to limit mycotoxin contamination of food and feed and to understand how mycotoxins impact pathogen-infected crops. In the current study, the composition and arrangement of biosynthetic gene clusters involved in the known mycotoxins production of *F. proliferatum* and other fusaria were analyzed in more detail.

The analyses of cluster flanking regions provided evidence that the functional homologs responsible for fumonisin production were located at different genomic position in *F. proliferatum* relative to *F. verticillioides*, which supported the hypothesis that the *FUM* cluster was acquired by horizontal gene transfer (HGT) from a progenitor homolog [44]. In *F. verticillioides*, fungi can expand its metabolic capabilities in such a way that, in turn, enhanced the adaptive strategies [45]. It had been shown that an intact *FUM* cluster also existed in *Aspergillus niger* [46]. Synonymous site divergence between the *A. niger* and *F. verticillioides* highlighted how the *FUM* cluster can diverge through the rearrangement of gene positions or by addition or loss of genes within *FUM* cluster [47]. Nevertheless, the presence of the *FUM* gene cluster did not always correlate with the ability to produce fumonisins. In contrast to the high amounts of fumonisin production presenting in the cracked maize kernel culture of *F. proliferatum* strain ITEM 2287 and *F. verticillioides* strain NRRL 20956, the main fumonisins producers among the genus *Fusarium*, no or relatively small amounts of fumonisins was observed in *F. fujikuroi* strains GL24 and GL28 [44]. Possible reasons could be that *FUM20* was absent and *FUM17* was present as a pseudogene that caused a premature truncation in *F. fujikuroi*.

The gene cluster responsible for fusaric acid exhibited a discontinuous distribution among *Fusarium* species, displaying gene insertion/deletion, inversion, or translocations. A possible explanation for this finding was that multiple independent events of genome rearrangements occurred within *FUB* gene clusters as species diverged from each other. The presence of the non-*FUB* genes within *FUB* cluster among *Fusarium* genera could

lead to analogs of metabolites that had the similar core structure but differ markedly in modifications to the structure [48], as was the case for structural diversity of trichothecenes in *Fusarium* [49]. However, trichothecene biosynthetic (*TRI*) genes were organized into three segments, the *TRI*-gene segments differed from the *FUB*-gene segments in that they were distinct loci, and in *F. graminearum*, each of these *TRI* loci was on a different chromosome [50]. In *F. verticillioides*, the *FUB1* to *FUB5* and *FUB6* to *FUB12* segments of the *FUB* cluster underwent inversion relative to the homologous segments in other *Fusarium* species, which suggested that genome rearrangements were another potential cause of the phylogenetic incongruence and the deep divergence of the cluster sequences of fusaric acid.

The most studied NRPSs in *Fusarium* species is the one responsible for the biosynthesis of Apicidin F [51]. Most members of *Fusarium* species exhibited small or large deletions in the biosynthesis genes that were responsible for apicidin F. It was most likely that an ancestor of the genus *Fusarium* maintained the entire gene cluster of apicidin F in genome [52], which had been lost or retained remnants of cluster paralogues in some species during evolutionary history. Although the biosynthetic gene cluster of apicidin F was present in the genome sequence of *F. fujikuroi*, additional genome sequences of *F. fujikuroi* showed that this cluster was only present in strain B14 but not present in other strains of this species [20].

The first fungal PKS-NRPS to be characterized was the fusarin C synthetase [53]. Although the order and orientation of the *FUS* cluster responsible for fusarin C production was largely conserved among the *Fusarium* species examined, the flanking regions were arranged markedly different, which likely reflected that the *FUS* cluster had undergone multiple independent rearrangement events associated with translocation and inversion, e.g., via HGT events. A functional full-length homolog of *FUS* cluster was required for synthesis of fusarin mycotoxins. However, the deletion events of *FUS* cluster had likely occurred. For example, *F. proliferatum* from a variety of hosts and geographic locations indicated that 33 of the strains had an intact *FUS* cluster, and only three isolates had the partially deleted *FUS* cluster [54]. A possible explanation for this finding was that the partial deletion of the *FUS* cluster occurred in a common ancestor of *F. proliferatum* in such a way that it resulted in two alleles of the *FUS* cluster.

Gibberellins (GAs) belong to a large family of tetracyclic diterpenoid carboxylic acids, and some members of them are phytohormones. GAs were first isolated as secondary metabolites of the pathogenic fungus *F. fujikuroi* in rice seedlings. Complete and partial homologs of the GA clusters existed in some *Fusarium* species, although no GA or GA-like compounds had been reported in these species [42]. The lack of GA production in other fusaria could be caused by a number of factors, including mutations that leave ORFs intact but render enzymes nonfunctional, the reduced transcription of GA genes, improper GA transcript processing, and/or altered translation [42]. *F. proliferatum* lost the ability to produce GAs due to mutations in *P450-1* and *P450-4* genes that lead to premature termination of the predicted ORF [55]. In *F. avenaceum*, a GA gene cluster with homologs of six of the seven GA biosynthetic genes was found, the homologs lacked a *P450-3* homolog encoding the 13-hydroxylase [56]. In *F. fujikuroi*, the *P450-3* enzyme catalyzed the last step of the conversion of GA7 to GA3 in GA biosynthetic pathway [57]. It is noteworthy that although GA biosynthetic genes were present in some related species, the ability to produce Gas and express GA biosynthetic genes was limited to *F. fujikuroi* [58].

On the whole, the biosynthesis of mycotoxins was complex among members of the *Fusarium* species. Mutations, HGT events, and genome reorganization, as well as gene loss and gene gain, could directly influence the re-shuffling of biosynthesis genes of mycotoxins, which probably led to synthesize putatively novel derivatives or lack the corresponding metabolic capability. Unique biosynthetic clusters of mycotoxins in *F. proliferatum* suggested sources for an exclusive metabolite that might be at least partially beneficial to the lifestyle specific to the fungus.

## 4. Conclusions

In summary, our study presented the high-quality genome sequence of the strain Fp9 of *F. proliferatum*, the causative agent of spikelet rot disease in rice. Despite the close phylogenetic relationships of *F. proliferatum* related to *F. fujikuroi*, there were sets of genes that were unique to *F. proliferatum* near the subtelomeric regions. The expansion of genes encoding MFS transporters and CWDEs was apparent in *F. proliferatum* compared with other fungi with different nutritional lifestyles, which seemed to be a specific feature of *F. proliferatum*. Although the high level of gene conservation and the collinearity of biosynthetic gene clusters responsible for mycotoxins was present among *F. proliferatum* and other fusaria, there were several differences in gene content and SM production, which provided the potential to contribute to the divergence of mycotoxins in the *Fusarium* species. To the best of our knowledge, this was the first whole-genome sequence of *F. proliferatum* isolated from rice. The genomic resources of the species foster the functional analysis of molecular components involved in pathogenicity, its interactions with host plants, and mycotoxins biosynthesis.

## 5. Materials and Methods

### 5.1. Fungal Strains and Culture Conditions

The *Fusarium proliferatum* strain Fp9 was isolated from naturally infected rice spikelets collected in Hangzhou, Zhejiang province, China [59]. Pathogenicity testing showed that the strain Fp9 caused the typical symptoms of spikelet rot in rice by artificial inoculation [59]. The strain was cultured in potato dextrose broth (PDB) for 5 to 7 days at 28 °C on an orbital shaker at 120 rpm, and mycelia were harvested by filtration. The strain was grown in YEPD medium (0.5% yeast extract, 1% peptone, and 2% glucose) to obtain spores, and stored in 20% glycerol stocks at −80 °C for long-term storage.

### 5.2. Species Identification

Genomic DNA was extracted from the mycelia using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The species was identified based on colony morphology and the PCR amplification of three genes encoding internal transcribed spacer (*ITS*) region, translation elongation factor (*TEF*), and beta-tubulin ( $\beta$ -*TUB*) using the primer pairs ITS1/ITS4 [60], EF1/EF2 [61], and T1/T22 [62], respectively. The PCR cycling parameters were as follows: an initial denaturation step at 94 °C for 5 min, 30 cycles at 94 °C for 15 s, 55 °C for 30 s, and 72 °C for 1 min, followed by a final extension step at 72 °C for 10 min. Each PCR reaction yielded a single band detected in a 1.0% (*w/v*) agarose gel stained with GelRed. PCR products were purified using the Wizard<sup>®</sup> SV Gel and PCR Clean-Up System kit (Promega, Madison, WI, USA) and were directly sequenced in both directions using Sanger dideoxy chain termination in an ABI 3730 DNA sequencer at Sangon Company (Shanghai, China). A phylogeny tree was constructed using the neighbor-joining method implemented in MEGA [63] based on the concatenated alignment of the nucleotide sequences of genes with 1000 bootstrap replicates.

### 5.3. Genome Sequencing and De Novo Assembly

The genome was sequenced with massively parallel sequencing (MPS) Illumina technology. A paired-end library with 500-bp insert and two mate-pair libraries with 2-kb and 5-kb inserts were constructed and sequenced using an Illumina HiSeq2500 with PE125 strategy. Library construction and sequencing were performed at the Beijing Novogene Bioinformatics Technology Co., Ltd. (Beijing, China). The quality control of both paired-end and mate-pair reads was performed with an in-house program. Raw sequencing data were obtained through quality-filtering by removing the adapters and low-quality reads. The filtered reads were assembled into scaffolds by the software SOAPdenovo v2, Ruibang Luo, Hong Kong, China [64]. Genome size was evaluated based on Illumina short reads using the K-mer analysis toolkit [65]. The completeness of the genome assembly was assessed

using BUSCO (benchmarking universal single copy ortholog) v5.2.2 with parameters ‘-m OGS’ [66].

#### 5.4. Gene Prediction, Annotation and Protein Classification

Protein-encoding genes were predicted by the GeneMarkS programs [67]. Gene annotations were performed with BLAST search ( $E$ -value  $\leq 10^{-5}$ , minimal alignment length percentage  $\geq 40\%$ ) against the following databases: NR [68], KEGG [69], COG [70], GO [71], Swiss-Prot, and TrEMBL databases [72]. Protein domains were defined in the Pfam [73] and Interpro [74] databases. Potential virulence-related proteins were identified by searching against the PHI-base [75]. Potential secreted proteins with signal peptide were predicted by SignalP 5.0 [76]. Membrane transporters were predicted by Phobius [77] and TMHMM [78], then classified based on the Transporter Classification Database [79]. The CAZymes were classified into different sub-families with the CAZyme database [80]. Biosynthetic gene clusters of SMs were analyzed using antiSMURF 4.0 with the Hidden Markov Model [81].

#### 5.5. Orthology and Phylogenomic Analysis

To evaluate the phylogenomic position of *F. proliferatum*, the genomes of the other fungi were retrieved from NCBI database (Table S6). These fungi included the *Fusarium* species, *F. proliferatum* (*Gibberella fujikuroi* MP-D), *F. fujikuroi* (MP-C), *F. circinatum* (MP-H), *F. verticillioides* (MP-A), *F. graminearum* (non-GFC species), rice important fungal pathogens, *Magnaporthe oryzae*, *Ustilaginoidea virens*, *Rhizoctonia solani* AG-I IA, two model fungal species, *Neurospora crassa*, and *Saccharomyces cerevisiae*. Core/Pan genes were clustered by the CD-HIT [82] rapid clustering of similar proteins with a threshold of 50% pairwise identity and 0.7 length difference cutoff in amino acids. A genome-scale phylogenetic tree was constructed by the program TreeBeST using the maximum likelihood method of PhyML v3.0 [83], and the setting of bootstraps was 1000 calculated based on the protein sequences of the orthologous genes. Synteny analysis was performed using MUMmer v3.0 [84] package with a cluster length of exact matches of at least 100 nt and, at most, 500 nt mismatches between two exact matches.

#### 5.6. Enrichment Analysis

Functional enrichment analysis was performed using the MIPS Functional Catalogue (FunCat) program [85]. FunCat categories with a false discovery rate (FDR) under 0.01 were defined as significantly enriched. The FDR correction of each annotated category was calculated using the Benjamini–Hochberg method. The enrichment analysis of gene families was conducted using the Gene Ontology Enrichment tools proposed online by the Eukaryotic Pathogen genome resource (EuPathDB) project [86] using biological ontology and InterPro predictions. The chi-squared test was used to compare the difference of the number of genes in the genome of *F. proliferatum* with the number of genes in the genome of other fungi. Over and underrepresentation were accepted at the significance threshold of  $p$ -value  $< 0.001$ .

#### 5.7. Comparison of the Biosynthetic Gene Clusters of SMs

The content and arrangement of the biosynthetic gene cluster of SMs in *F. proliferatum* were compared to publicly available data and homologous sequences in other *Fusarium* species (Table 1). The genomes of *Fusarium* species examined were obtained from NCBI database (Table S6). The ortholog data of the biosynthetic gene cluster were used to determine the level of collinearity among *Fusarium* species. A fumonisin gene (*FUM*) cluster had been described in *F. verticillioides* [44]. A fusaric acid gene (*FUB*) cluster had been reported in *F. verticillioides* [87]. A bikaverin gene (*Bik*) cluster was characterized in *F. fujikuroi* [88]. A fusarubin gene (*FSR*) cluster was previously identified in *F. fujikuroi* [89]. An apicidin F gene (*APF*) cluster was characterized in *F. fujikuroi* [51]. A fusarin C gene

(FUS) cluster was characterized in *F. fujikuroi* [53]. A gibberellins (GAs) gene cluster had been demonstrated in *F. fujikuroi* [90].

**Table 1.** Descriptions of genes coding key biosynthetic enzymes for secondary metabolites (SMs) in *Fusarium* species.

Species	Secondary Metabolite	Gene	Predicted Function		
<i>Fusarium verticillioides</i>	fumonisin	<i>FUM21</i>	Zn(II)2Cys6-type transcription factor		
		<i>FUM1</i>	polyketide synthase		
		<i>FUM6</i>	cytochrome P450 monooxygenase		
		<i>FUM7</i>	dehydrogenase		
		<i>FUM8</i>	aminotransferase		
		<i>FUM3</i>	dioxygenase		
		<i>FUM10</i>	fatty acyl-CoA synthetase		
		<i>FUM11</i>	tricarboxylate transporter		
		<i>FUM2</i>	cytochrome P450 monooxygenase		
		<i>FUM20</i>	unknown		
		<i>FUM13</i>	short-chain dehydrogenase/reductase		
		<i>FUM14</i>	peptide synthetase		
		<i>FUM15</i>	cytochrome P450 monooxygenase		
		<i>FUM16</i>	fatty acyl-CoA synthetase		
		<i>FUM17</i>	longevity assurance factor		
		<i>FUM18</i>	longevity assurance factor		
		<i>FUM19</i>	ABC transporter		
		<i>Fusarium verticillioides</i>	fusaric acid	<i>FUB1</i>	polyketide synthase
				<i>FUB2</i>	hypothetical protein of unknown function
<i>FUB3</i>	amino acid kinase				
<i>FUB4</i>	hydrolase				
<i>FUB5</i>	acetyltransferase				
<i>FUB6</i>	dehydrogenase				
<i>FUB7</i>	sulfhydrylase				
<i>FUB8</i>	dehydrogenase				
<i>FUB9</i>	oxidase				
<i>FUB10</i>	C6 transcription factor				
<i>FUB11</i>	major facilitator superfamily transporter				
<i>FUB12</i>	C6 transcription factor				
<i>Fusarium fujikuroi</i>	bikaverin	<i>Bik1</i>	polyketide synthase		
		<i>Bik2</i>	putative FAD-dependent monooxygenase		
		<i>Bik3</i>	O-methyltransferase		
		<i>Bik4</i>	putative NmrA-like transcriptional regulator		
		<i>Bik5</i>	Zn(II)2Cys6-type transcription factor		
		<i>Bik6</i>	major facilitator superfamily transporter		

Table 1. Cont.

Species	Secondary Metabolite	Gene	Predicted Function
<i>Fusarium fujikuroi</i>	fusarubin	<i>FSR1</i>	$\beta$ -ketoacyl synthase
		<i>FSR2</i>	O-methyltransferase
		<i>FSR3</i>	monooxygenase
		<i>FSR4</i>	alcohol dehydrogenase
		<i>FSR5</i>	short-chain dehydrogenase/reductase
		<i>FSR6</i>	Zn(II)2Cys6-type transcription factor
<i>Fusarium fujikuroi</i>	apicidin	<i>APF1</i>	non-ribosomal peptide synthetase
		<i>APF2</i>	transcription factor
		<i>APF3</i>	$\Delta$ 1-Pyrroline-5-carboxylate reductase
		<i>APF4</i>	aminotransferase
		<i>APF5</i>	fatty acid synthase
		<i>APF6</i>	O-methyltransferase
		<i>APF7</i>	cytochrome P450 monooxygenase
		<i>APF8</i>	cytochrome P450 monooxygenase
		<i>APF9</i>	FAD-dependent monooxygenase
		<i>APF11</i>	major facilitator superfamily transporter
		<i>APF12</i>	cytochrome b5-like
		<i>Fusarium fujikuroi</i>	fusarin C
<i>FUS2</i>	$\alpha/\beta$ hydrolase		
<i>FUS3</i>	glutathione S-transferase		
<i>FUS4</i>	peptidase		
<i>FUS5</i>	serine hydrolase		
<i>FUS6</i>	MFS transporter		
<i>FUS7</i>	aldehyde dehydrogenase		
<i>FUS8</i>	cytochrome P450 monooxygenase		
<i>FUS9</i>	methyltransferase		
<i>Fusarium fujikuroi</i>	gibberellins	<i>GGs2</i>	geranylgeranyl diphosphate synthase
		<i>DES</i>	GA4 desaturase
		<i>CPS/KS</i>	bifunctional ent-copalyl diphosphate/ent-kaurene synthase
		<i>P450-1</i>	cytochrome P450 monooxygenase
		<i>P450-2</i>	cytochrome P450 monooxygenase
		<i>P450-3</i>	cytochrome P450 monooxygenase
		<i>P450-4</i>	cytochrome P450 monooxygenase

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/toxins14080568/s1>, Figure S1: Neighbor-joining tree inferred from the concatenated nucleotide sequences of *ITS*, *TEF-1 $\alpha$*  and  *$\beta$ -TUB* of the strains of *F. proliferatum* and other *Fusarium* species; Figure S2: Gel electrophoresis images of PCR validation of structural variants (SVs) between *F. proliferatum* and *F. fujikuroi*; Table S1: Specific regions in the subtelomeres are present in *F. proliferatum* but absent in *F. fujikuroi*; Table S2: GO enrichment analysis of genes are present in the genome of *F. proliferatum* and absent in that of *F. fujikuroi*; Table S3: Summary of PCR validations of structural variants (SVs) between *F. proliferatum* and *F. fujikuroi*. Table S4: Comparison of genes coding for membrane transporters in *F. proliferatum* with that in other fungi; Table S5: Characterization of carbohydrate-active enzymes (CAZymes) of *F. proliferatum* and other fungi; Table S6: The protein-encoding genes and genomes of species used in this study; Additional file S1: Sanger sequencing results of structural variants (SVs) based on PCR validation for *F. proliferatum* and *F. fujikuroi*.

**Author Contributions:** Conceptualization, L.W., S.T. and P.H.; data curation, L.W.; formal analysis, L.W.; funding acquisition, P.H.; methodology, L.W., S.G., W.L. (Wenhao Liang), W.L. (Weiyang Liao), W.L. (Wen Li) and G.J.; project administration, S.T. and P.H.; resources, S.T. and P.H.; software, L.W., X.W., G.S., L.X., Z.S. and S.H.; supervision, S.T. and P.H.; visualization, P.H.; writing—original draft, L.W.; writing—review and editing, S.T. and P.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by National Natural Science Foundation of China (31800133, 32188102), Natural Science Foundation of Zhejiang Province, China (LQ18C140005), Key Research and Development Program of Zhejiang Province, China (2022C02011, 2021C02063, 2021C02056), Science and Technology Major Project of Anhui Province, China (202203c08020006), and the Agricultural Sciences and Technologies Innovation Program of Chinese Academy of Agricultural Sciences (CAAS-ZDRW202001).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank National Natural Science Foundation of China (31800133, 32188102), Natural Science Foundation of Zhejiang Province, China (LQ18C140005), Key Research and Development Program of Zhejiang Province, China (2022C02011, 2021C02063, 2021C02056), Science and Technology Major Project of Anhui Province, China (202203c08020006), and the Agricultural Sciences and Technologies Innovation Program of Chinese Academy of Agricultural Sciences (CAAS-ZDRW202001) for supporting this study.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Huang, S.-W.; Wang, L.; Liu, L.-M.; Tang, S.-Q.; Zhu, D.-F.; Savary, S. Rice Spikelet Rot Disease in China—1. Characterization of Fungi Associated with the Disease. *Crop Prot.* **2011**, *30*, 1–9. [[CrossRef](#)]
- Cendoya, E.; Pinson-Gadais, L.; Farnochi, M.C.; Ramirez, M.L.; Chéreau, S.; Marcheguay, G.; Ducos, C.; Barreau, C.; Richard-Forget, F. Abiotic Conditions Leading to *FUM* Gene Expression and Fumonisin Accumulation by *Fusarium proliferatum* Strains Grown on a Wheat-Based Substrate. *Int. J. Food Microbiol.* **2017**, *253*, 12–19. [[CrossRef](#)]
- Lin, Y.; Totsuka, Y.; Shan, B.; Wang, C.; Wei, W.; Qiao, Y.; Kikuchi, S.; Inoue, M.; Tanaka, H.; He, Y. Esophageal Cancer in High-Risk Areas of China: Research Progress and Challenges. *Ann. Epidemiol.* **2017**, *27*, 215–221. [[CrossRef](#)] [[PubMed](#)]
- Sun, L.; Chen, X.; Gao, J.; Zhao, Y.; Liu, L.; Hou, Y.; Wang, L.; Huang, S. Effects of Disruption of Five *FUM* Genes on Fumonisin Biosynthesis and Pathogenicity in *Fusarium proliferatum*. *Toxins* **2019**, *11*, 327. [[CrossRef](#)]
- Xie, L.; Wu, Y.; Wang, Y.; Jiang, Y.; Yang, B.; Duan, X.; Li, T. Fumonisin B1 induced Aggressiveness and Infection Mechanism of *Fusarium proliferatum* on Banana Fruit. *Environ. Pollut.* **2021**, *288*, 117793. [[CrossRef](#)] [[PubMed](#)]
- Kamle, M.; Mahato, D.K.; Devi, S.; Lee, K.E.; Kang, S.G.; Kumar, P. Fumonisin: Impact on Agriculture, Food, and Human Health and their Management Strategies. *Toxins* **2019**, *11*, 328. [[CrossRef](#)] [[PubMed](#)]
- Awuchi, C.G.; Ondari, E.N.; Ogbonna, C.U.; Upadhyay, A.K.; Baran, K.; Okpala, C.O.R.; Korzeniowska, M.; Guiné, R.P.F. Mycotoxins Affecting Animals, Foods, Humans, and Plants: Types, Occurrence, Toxicities, Action Mechanisms, Prevention, and Detoxification Strategies—a Revisit. *Foods* **2021**, *10*, 1279. [[CrossRef](#)]
- Almiman, B.F.; Shittu, T.A.; Muthumeenakshi, S.; Baroncelli, R.; Sreenivasaprasad, S. Genome Sequence of the Mycotoxigenic Crop Pathogen *Fusarium proliferatum* Strain ITEM 2341 from Date Palm. *Microbiol. Resour. Announc.* **2018**, *7*, e00964-18. [[CrossRef](#)]



9. Armitage, A.D.; Taylor, A.; Hulin, M.T.; Jackson, A.C.; Harrison, R.J.; Clarkson, J.P. Draft Genome Sequence of an Onion Basal Rot Isolate of *Fusarium proliferatum*. *Microbiol. Resour. Announc.* **2019**, *8*, e01385-18. [[CrossRef](#)]
10. Niehaus, E.M.; Munsterkotter, M.; Proctor, R.H.; Brown, D.W.; Sharon, A.; Idan, Y.; Oren-Young, L.; Sieber, C.M.; Novak, O.; Pencik, A.; et al. Comparative “Omics” of the *Fusarium fujikuroi* Species Complex Highlights Differences in Genetic Potential and Metabolite Synthesis. *Genome Biol. Evol.* **2016**, *8*, 3574–3599. [[CrossRef](#)]
11. Alberts, J.F.; van Zyl, W.H.; Gelderblom, W.C.A. Biologically Based Methods for Control of Fumonisin-Producing *Fusarium* Species and Reduction of the Fumonisin. *Front. Microbiol.* **2016**, *7*, 548. [[CrossRef](#)] [[PubMed](#)]
12. Treangen, T.J.; Salzberg, S.L. Repetitive DNA and Next-Generation Sequencing: Computational Challenges and Solutions. *Nat. Rev. Genet.* **2011**, *13*, 36–46. [[CrossRef](#)]
13. Gu, W.; Zhou, A.; Wang, L.; Sun, S.; Cui, X.; Zhu, D. SVLR: Genome Structural Variant Detection Using Long-Read Sequencing Data. *J. Comput. Biol.* **2021**, *28*, 774–788. [[CrossRef](#)] [[PubMed](#)]
14. Qi, W.; Colarusso, A.; Olombrada, M.; Parrilli, E.; Patrignani, A.; Tutino, M.L.; Toll-Riera, M. New Insights on *Pseudoalteromonas haloplanktis* TAC125 Genome Organization and Benchmarks of Genome Assembly Applications Using Next and Third Generation Sequencing Technologies. *Sci. Rep.* **2019**, *9*, 16444. [[CrossRef](#)]
15. Goodwin, S.; McPherson, J.D.; McCombie, W.R. Coming of Age: Ten Years of Next-Generation Sequencing Technologies. *Nat. Rev. Genet.* **2016**, *17*, 333–351. [[CrossRef](#)]
16. Tham, C.Y.; Tirado-Magallanes, R.; Goh, Y.; Fullwood, M.J.; Koh, B.T.; Wang, W.; Ng, C.H.; Chng, W.J.; Thiery, A.; Tenen, D.G.; et al. NanoVar: Accurate Characterization of Patients’ Genomic Structural Variants Using Low-Depth Nanopore Sequencing. *Genome Biol.* **2020**, *21*, 56. [[CrossRef](#)] [[PubMed](#)]
17. Zhao, S.; Gibbons, J.G. A Population Genomic Characterization of Copy Number Variation in the Opportunistic Fungal Pathogen *Aspergillus fumigatus*. *PLoS ONE* **2018**, *13*, e0201611. [[CrossRef](#)] [[PubMed](#)]
18. Gladieux, P.; Ropars, J.; Badouin, H.; Branca, A.; Aguilera, G.; Vienne, D.M.; de la Vega, R.C.R.; Branco, S.; Giraud, T. Fungal Evolutionary Genomics Provides Insight into the Mechanisms of Adaptive Divergence in Eukaryotes. *Mol. Ecol.* **2014**, *23*, 753–773. [[CrossRef](#)] [[PubMed](#)]
19. Laurent, B.; Moinard, M.; Spataro, C.; Ponts, N.; Barreau, C.; Foulongne-Oriol, M. Landscape of Genomic Diversity and Host Adaptation in *Fusarium graminearum*. *BMC Genom.* **2017**, *18*, 203. [[CrossRef](#)] [[PubMed](#)]
20. Chiara, M.; Fanelli, F.; Mulè, G.; Logrieco, A.F.; Pesole, G.; Leslie, J.F.; Horner, D.S.; Toomajian, C. Genome Sequencing of Multiple Isolates Highlights Subtelomeric Genomic Diversity within *Fusarium fujikuroi*. *Genome Biol. Evol.* **2015**, *7*, 3062–3069. [[CrossRef](#)]
21. Zhang, L.-B.; Feng, M.-G. Antioxidant Enzymes and Their Contributions to Biological Control Potential of Fungal Insect Pathogens. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 4995–5004. [[CrossRef](#)] [[PubMed](#)]
22. Yan, N. Structural Biology of the Major Facilitator Superfamily Transporters. *Annu. Rev. Biophys.* **2015**, *44*, 257–283. [[CrossRef](#)] [[PubMed](#)]
23. Wang, Q.; Chen, D.; Wu, M.; Zhu, J.; Jiang, C.; Xu, J.R.; Liu, H. MFS Transporters and GABA Metabolism Are Involved in the Self-Defense Against DON in *Fusarium graminearum*. *Front. Plant Sci.* **2018**, *9*, 438. [[CrossRef](#)] [[PubMed](#)]
24. Lin, H.-C.; Yu, P.-L.; Chen, L.-H.; Tsai, H.-C.; Chung, K.-R. A Major Facilitator Superfamily Transporter Regulated by the Stress-Responsive Transcription Factor Yap1 Is Required for Resistance to Fungicides, Xenobiotics, and Oxidants and Full Virulence in *Alternaria alternata*. *Front. Microbiol.* **2018**, *9*, 2229. [[CrossRef](#)]
25. Vela-Corcía, D.; Srivastava, D.A.; Dafa-Berger, A.; Rotem, N.; Barda, O.; Levy, M. MFS Transporter from *Botrytis cinerea* Provides Tolerance to Glucosinolate-Breakdown Products and Is Required for Pathogenicity. *Nat. Commun.* **2019**, *10*, 2886. [[CrossRef](#)]
26. Zhao, Z.; Liu, H.; Wang, C.; Xu, J.-R. Comparative Analysis of Fungal Genomes Reveals Different Plant Cell Wall Degrading Capacity in Fungi. *BMC Genom.* **2013**, *14*, 274. [[CrossRef](#)]
27. Kamoun, S.; Furzer, O.; Jones, J.D.G.; Judelson, H.S.; Ali, G.S.; Dalio, R.J.D.; Roy, S.G.; Schena, L.; Zambounis, A.; Panabières, F.; et al. The Top 10 *Oomycete* Pathogens in Molecular Plant Pathology. *Mol. Plant Pathol.* **2014**, *16*, 413–434. [[CrossRef](#)]
28. Engering, A.; Hogerwerf, L.; Slingenberg, J. Pathogen-Host Environment Interplay and Disease Emergence. *Emerg. Microbes Infect.* **2013**, *2*, e5. [[CrossRef](#)]
29. Horbach, R.; Navarro-Quesada, A.R.; Knogge, W.; Deising, H.B. When and How to Kill a Plant Cell: Infection Strategies of Plant Pathogenic Fungi. *J. Plant Physiol.* **2011**, *168*, 51–62. [[CrossRef](#)]
30. O’Connell, R.J.; Thon, M.R.; Hacquard, S.; Amyotte, S.G.; Kleemann, J.; Torres, M.F.; Damm, U.; Buiate, E.A.; Epstein, L.; Alkan, N.; et al. Lifestyle Transitions in Plant Pathogenic *Colletotrichum* Fungi Deciphered by Genome and Transcriptome Analyses. *Nat. Genet.* **2012**, *44*, 1060–1065. [[CrossRef](#)]
31. Zhang, Y.; Zhang, K.; Fang, A.; Han, Y.; Yang, J.; Xue, M.; Bao, J.; Hu, D.; Zhou, B.; Sun, X.; et al. Specific Adaptation of *Ustilagoideae virens* in Occupying Host Florets Revealed by Comparative and Functional Genomics. *Nat. Commun.* **2014**, *5*, 3849. [[CrossRef](#)] [[PubMed](#)]
32. de Ramón-Carbonell, M.; Sánchez-Torres, P. *Penicillium digitatum* MFS Transporters Can Display Different Roles During Pathogen-Fruit Interaction. *Int. J. Food Microbiol.* **2021**, *337*, 108918. [[CrossRef](#)] [[PubMed](#)]
33. Ma, X.; Duan, G.; Chen, H.; Tang, P.; Su, S.; Wei, Z.; Yang, J. Characterization of Infected Process and Primary Mechanism in Rice Acute Defense Against Rice Blast Fungus, *Magnaporthe oryzae*. *Plant Mol. Biol.* **2022**, online ahead of print. [[CrossRef](#)]
34. Chowdhury, S.; Basu, A.; Kundu, S. Biotrophy-Necrotrophy Switch in Pathogen Evoke Differential Response in Resistant and Susceptible Sesame Involving Multiple Signaling Pathways at Different Phases. *Sci. Rep.* **2017**, *7*, 17251. [[CrossRef](#)]

35. Keller, N.P. Fungal secondary metabolism: Regulation, function and drug discovery. *Nat. Rev. Microbiol.* **2019**, *17*, 167–180. [[CrossRef](#)]
36. Lind, A.L.; Wisecaver, J.H.; Lameiras, C.; Wiemann, P.; Palmer, J.M.; Keller, N.P.; Rodrigues, F.; Goldman, G.H.; Rokas, A. Drivers of Genetic Diversity in Secondary Metabolic Gene Clusters Within a Fungal Species. *PLoS Biol.* **2017**, *15*, e2003583. [[CrossRef](#)]
37. Villani, A.; Proctor, R.H.; Kim, H.-S.; Brown, D.W.; Logrieco, A.F.; Amatulli, M.T.; Moretti, A.; Susca, A. Variation in Secondary Metabolite Production Potential in the *Fusarium incarnatum-equiseti* Species Complex Revealed by Comparative Analysis of 13 Genomes. *BMC Genom.* **2019**, *20*, 314. [[CrossRef](#)]
38. Massonnet, M.; Morales-Cruz, A.; Minio, A.; Figueroa-Balderas, R.; Lawrence, D.P.; Travadon, R.; Rolshausen, P.E.; Baumgartner, K.; Cantu, D. Whole-Genome Resequencing and Pan-Transcriptome Reconstruction Highlight the Impact of Genomic Structural Variation on Secondary Metabolite Gene Clusters in the Grapevine Esca Pathogen *Phaeoacremonium minimum*. *Front. Microbiol.* **2018**, *9*, 1784. [[CrossRef](#)]
39. López-Díaz, C.; Rahjoo, V.; Sulyok, M.; Ghionna, V.; Martín-Vicente, A.; Capilla, J.; Di Pietro, A.; López-Berges, M.S. Fusaric Acid Contributes to Virulence of *Fusarium oxysporum* on Plant and Mammalian Hosts. *Mol. Plant Pathol.* **2017**, *19*, 440–453. [[CrossRef](#)]
40. Jia, L.-J.; Tang, H.-Y.; Wang, W.-Q.; Yuan, T.-L.; Wei, W.-Q.; Pang, B.; Gong, X.-M.; Wang, S.-F.; Li, Y.-J.; Zhang, D.; et al. A Linear Nonribosomal Octapeptide from *Fusarium graminearum* Facilitates Cell-to-Cell Invasion of Wheat. *Nat. Commun.* **2019**, *10*, 922. [[CrossRef](#)]
41. Mohammed, A.; Faustinelli, P.C.; Chala, A.; Dejene, M.; Fininsa, C.; Ayalew, A.; Ojiewo, C.O.; Hoisington, D.A.; Sobolev, V.S.; Martínez-Castillo, J.; et al. Genetic Fingerprinting and Aflatoxin Production of *Aspergillus* Section Flavi Associated with Groundnut in Eastern Ethiopia. *BMC Microbiol.* **2021**, *21*, 239. [[CrossRef](#)]
42. Wiemann, P.; Sieber, C.M.K.; von Bargen, K.W.; Studt, L.; Niehaus, E.-M.; Espino, J.J.; Huß, K.; Michielse, C.B.; Albermann, S.; Wagner, D.; et al. Deciphering the Cryptic Genome: Genome-wide Analyses of the Rice Pathogen *Fusarium fujikuroi* Reveal Complex Regulation of Secondary Metabolism and Novel Metabolites. *PLoS Pathog.* **2013**, *9*, e1003475. [[CrossRef](#)]
43. Brown, D.W.; Proctor, R.H. Insights into Natural Products Biosynthesis from Analysis of 490 Polyketide Synthases from *Fusarium*. *Fungal Genet. Biol.* **2016**, *89*, 37–51. [[CrossRef](#)] [[PubMed](#)]
44. Proctor, R.H.; Van Hove, F.; Susca, A.; Stea, G.; Busman, M.; van der Lee, T.; Waalwijk, C.; Moretti, A.; Ward, T.J. Birth, Death and Horizontal Transfer of the Fumonisin Biosynthetic Gene Cluster During the Evolutionary Diversification of *Fusarium*. *Mol. Microbiol.* **2013**, *90*, 290–306. [[CrossRef](#)] [[PubMed](#)]
45. Gao, S.; Gold, S.E.; Wisecaver, J.; Zhang, Y.; Guo, L.; Ma, L.-J.; Rokas, A.; Glenn, A.E. Genome-Wide Analysis of *Fusarium verticillioides* Reveals Inter-Kingdom Contribution of Horizontal Gene Transfer to The Expansion of Metabolism. *Fungal Genet. Biol.* **2019**, *128*, 60–73. [[CrossRef](#)] [[PubMed](#)]
46. Frisvad, J.C.; Smedsgaard, J.; Samson, R.A.; Larsen, T.O.; Thrane, U. Fumonisin B2 Production by *Aspergillus niger*. *J. Agric. Food Chem.* **2007**, *55*, 9727–9732. [[CrossRef](#)]
47. Khaldi, N.; Wolfe, K.H. Evolutionary Origins of the Fumonisin Secondary Metabolite Gene Cluster in *Fusarium verticillioides* and *Aspergillus niger*. *Int. J. Evol. Biol.* **2011**, *2011*, 423821. [[CrossRef](#)]
48. Bohni, N.; Hofstetter, V.; Gindro, K.; Buyck, B.; Schumpp, O.; Bertrand, S.; Monod, M.; Wolfender, J.-L. Production of Fusaric Acid by *Fusarium* spp. in Pure Culture and in Solid Medium Co-Cultures. *Molecules* **2016**, *21*, 370. [[CrossRef](#)]
49. Proctor, R.H.; McCormick, S.P.; Kim, H.S.; Cardoza, R.E.; Stanley, A.M.; Lindo, L.; Kelly, A.; Brown, D.W.; Lee, T.; Vaughan, M.M.; et al. Evolution of Structural Diversity of Trichothecenes, A Family of Toxins Produced by Plant Pathogenic and Entomopathogenic Fungi. *PLoS Pathog.* **2018**, *14*, e1006946. [[CrossRef](#)]
50. Proctor, R.H.; McCormick, S.P.; Gutiérrez, S. Genetic Bases for Variation in Structure and Biological Activity of Trichothecene Toxins Produced by Diverse Fungi. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 5185–5199. [[CrossRef](#)]
51. Niehaus, E.M.; Janevska, S.; von Bargen, K.W.; Sieber, C.M.K.; Harrer, H.; Humpf, H.U.; Tudzynski, B. Apicidin F: Characterization and Genetic Manipulation of a New Secondary Metabolite Gene Cluster in the Rice Pathogen *Fusarium fujikuroi*. *PLoS ONE* **2014**, *9*, e103336. [[CrossRef](#)] [[PubMed](#)]
52. Witte, T.E.; Harris, L.J.; Nguyen, H.D.T.; Hermans, A.; Johnston, A.; Sproule, A.; Dettman, J.R.; Boddy, C.N.; Overy, D.P. Apicidin Biosynthesis Is Linked to Accessory Chromosomes in *Fusarium poae* Isolates. *BMC Genom.* **2021**, *22*, 591. [[CrossRef](#)] [[PubMed](#)]
53. Niehaus, E.-M.; Kleigrew, K.; Wiemann, P.; Studt, L.; Sieber, C.; Connolly, L.R.; Freitag, M.; Güldener, U.; Tudzynski, B.; Humpf, H.-U. Genetic Manipulation of the *Fusarium fujikuroi* Fusarin Gene Cluster Yields Insight into the Complex Regulation and Fusarin Biosynthetic Pathway. *Chem. Biol.* **2013**, *20*, 1055–1066. [[CrossRef](#)]
54. Proctor, R.H.; Desjardins, A.E.; Moretti, A. Biological and chemical complexity of *Fusarium proliferatum*. In *The Role of Plant Pathology in Food Safety and Food Security*; Strange, R.N., Lodovica Gullino, M., Eds.; Springer: Berlin, Germany, 2009; pp. 97–111.
55. Malonek, S.; Rojas, M.C.; Hedden, P.; Hopkins, P.; Tudzynski, B. Restoration of Gibberellin Production in *Fusarium proliferatum* by Functional Complementation of Enzymatic Blocks. *Appl. Environ. Microbiol.* **2005**, *71*, 6014–6025. [[CrossRef](#)] [[PubMed](#)]
56. Lysøe, E.; Harris, L.J.; Walkowiak, S.; Subramaniam, R.; Divon, H.H.; Rüser, E.S.; Llorens, C.; Gabaldón, T.; Kistler, H.C.; Jonkers, W.; et al. The Genome of the Generalist Plant Pathogen *Fusarium avenaceum* Is Enriched with Genes Involved in Redox, Signaling and Secondary Metabolism. *PLoS ONE* **2014**, *9*, e112703. [[CrossRef](#)]
57. Troncoso, C.; González, X.; Bömk, C.; Tudzynski, B.; Gong, F.; Hedden, P.; Rojas, M.C. Gibberellin biosynthesis and gibberellin oxidase activities in *Fusarium sacchari*, *Fusarium konzum* and *Fusarium subglutinans* strains. *Phytochemistry* **2010**, *71*, 1322–1331. [[CrossRef](#)]

58. Albermann, S.; Linnemannstöns, P.; Tudzynski, B. Strategies for Strain Improvement in *Fusarium fujikuroi*: Overexpression and Localization of Key Enzymes of the Isoprenoid Pathway and Their Impact on Gibberellin Biosynthesis. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 2979–2995. [[CrossRef](#)] [[PubMed](#)]
59. Sun, L.; Wang, L.; Liu, L.M.; Hou, Y.X.; Li, Q.Q.; Huang, S.W. Screening for Strains of Rice Spikelet Rot Disease Pathogenic Fungus with High Fumonisin Production and Strong Pathogenicity. *Chin. J. Rice Sci.* **2018**, *32*, 610–616.
60. White, T.J.; Bruns, T.; Lee, S.; Taylor, J. Amplification and Direct Sequencing of Fungal Ribosomal RNA Genes for Phylogenetics. In *PCR Protocols: A Guide to Methods and Applications*; Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J., Eds.; Academic Press: New York, NY, USA, 1990; pp. 315–322.
61. O'Donnell, K.; Kistler, H.C.; Cigelnik, E.; Ploetz, R.C. Multiple Evolutionary Origins of the Fungus Causing Panama Disease of Banana: Concordant Evidence from Nuclear and Mitochondrial Gene Genealogies. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 2044–2049. [[CrossRef](#)]
62. O'Donnell, K.; Cigelnik, E. Two Divergent Intragenomic rDNA ITS2 Types within a Monophyletic Lineage of the Fungus *Fusarium* Are Nonorthologous. *Mol. Phylogenetics Evol.* **1997**, *7*, 103–116. [[CrossRef](#)]
63. Kumar, S.; Stecher, G.; Li, M.; Nknyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis Across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)] [[PubMed](#)]
64. Luo, R.; Liu, B.; Xie, Y.; Li, Z.; Huang, W.; Yuan, J.; He, G.; Chen, Y.; Pan, Q.; Liu, Y.; et al. SOAPdenovo2: An Empirically Improved Memory-Efficient Short-Read De Novo Assembler. *GigaScience* **2012**, *1*, 18. [[CrossRef](#)] [[PubMed](#)]
65. Mapleson, D.; Accinelli, G.G.; Kettleborough, G.; Wright, J.; Clavijo, B.J. KAT: A K-Mer Analysis Toolkit to Quality Control NGS Datasets and Genome Assemblies. *Bioinformatics* **2017**, *33*, 574–576. [[CrossRef](#)]
66. Manni, M.; Berkeley, M.R.; Seppey, M.; Simão, F.A.; Zdobnov, E.M. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Mol. Biol. Evol.* **2021**, *38*, 4647–4654. [[CrossRef](#)] [[PubMed](#)]
67. Besemer, J.; Lomsadze, A.; Borodovsky, M. GeneMarkS: A Self-Training Method for Prediction of Gene Starts in Microbial Genomes. Implications for Finding Sequence Motifs in Regulatory Regions. *Nucleic Acids Res.* **2001**, *29*, 2607–2618. [[CrossRef](#)]
68. Deng, Y.Y.; Li, J.Q.; Wu, S.F.; Zhu, Y.P.; Chen, Y.; He, F.C.; Li, J.; Wu, S.; He, F. Integrated Nr Database in Protein Annotation System and Its Localization. *Comput. Eng.* **2006**, *32*, 71–74.
69. Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a Reference Resource for Gene and Protein Annotation. *Nucleic Acids Res.* **2016**, *44*, D457–D462. [[CrossRef](#)]
70. Galperin, M.Y.; Wolf, Y.I.; Makarova, K.S.; Alvarez, R.V.; Landsman, D.; Koonin, E.V. COG Database Update: Focus on Microbial Diversity, Model Organisms, and Widespread Pathogens. *Nucleic Acids Res.* **2020**, *49*, D274–D281. [[CrossRef](#)]
71. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene Ontology: Tool for the Unification of Biology. *Nat. Genet.* **2000**, *25*, 25–29. [[CrossRef](#)]
72. Magrane, M.; UniProt Consortium. UniProt Knowledgebase: A Hub of Integrated Protein Data. *Database* **2011**, *2011*, bar009. [[CrossRef](#)]
73. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G.A.; Sonnhammer, E.L.L.; Tosatto, S.C.; Paladin, L.; Raj, S.; Richardson, L.J.; et al. Pfam: The Protein Families Database in 2021. *Nucleic Acids Res.* **2021**, *49*, D412–D419. [[CrossRef](#)] [[PubMed](#)]
74. Blum, M.; Chang, H.-Y.; Chuguransky, S.; Grego, T.; Kandasaamy, S.; Mitchell, A.; Nuka, G.; Paysan-Lafosse, T.; Qureshi, M.; Raj, S.; et al. The InterPro Protein Families and Domains Database: 20 Years on. *Nucleic Acids Res.* **2021**, *49*, D344–D354. [[CrossRef](#)]
75. Urban, M.; Czuzik, A.; Seager, J.; Wood, V.; Rutherford, K.; Venkatesh, S.Y.; De Silva, N.; Martinez, M.C.; Pedro, H.; Yates, A.D.; et al. PHI-base: The Pathogen–Host Interactions Database. *Nucleic Acids Res.* **2020**, *48*, D613–D620. [[CrossRef](#)]
76. Almagro Armenteros, J.J.; Tsirigos, K.D.; Sønderby, C.K.; Petersen, T.N.; Winther, O.; Brunak, S.; Von Heijne, G.; Nielsen, H. SignalP 5.0 Improves Signal Peptide Predictions Using Deep Neural Networks. *Nat. Biotechnol.* **2019**, *37*, 420–423. [[CrossRef](#)] [[PubMed](#)]
77. Käll, L.; Krogh, A.; Sonnhammer, E. Advantages of Combined Transmembrane Topology and Signal Peptide Prediction—The Phobius Web Server. *Nucleic Acids Res.* **2007**, *35*, W429–W432. [[CrossRef](#)] [[PubMed](#)]
78. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L. Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes. *J. Mol. Biol.* **2001**, *305*, 567–580. [[CrossRef](#)]
79. Saier, M.H., Jr.; Reddy, V.S.; Tsu, B.V.; Ahmed, M.S.; Li, C.; Moreno-Hagelsieb, G. The Transporter Classification Database (TCDB): Recent Advances. *Nucleic Acids Res.* **2016**, *44*, D372–D379. [[CrossRef](#)]
80. Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P.M.; Henrissat, B. The Carbohydrate-Active Enzymes Database (CAZy) in 2013. *Nucleic Acids Res.* **2014**, *42*, D490–D495. [[CrossRef](#)]
81. Blin, K.; Wolf, T.; Chevrette, M.G.; Lu, X.; Schwalen, C.J.; Kautsar, S.A.; Suarez Duran, H.G.; de Los Santos, E.L.C.; Kim, H.U.; Nave, M.; et al. AntiSMASH 4.0—Improvements in Chemistry Prediction and Gene Cluster Boundary Identification. *Nucleic Acids Res.* **2017**, *45*, W36–W41. [[CrossRef](#)]
82. Huang, Y.; Niu, B.; Gao, Y.; Fu, L.; Li, W. CD-HIT Suite: A Web Server for Clustering and Comparing Biological Sequences. *Bioinformatics* **2010**, *26*, 680–682. [[CrossRef](#)]
83. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)] [[PubMed](#)]

84. Kurtz, S.; Phillippy, A.; Delcher, A.L.; Smoot, M.; Shumway, M.; Antonescu, C.; Salzberg, S.L. Versatile and Open Software for Comparing Large Genomes. *Genome Biol.* **2004**, *5*, R12. [[CrossRef](#)] [[PubMed](#)]
85. Ruepp, A.; Zollner, A.; Maier, D.; Albermann, K.; Hani, J.; Mokejcs, M.; Tetko, I.; Güldener, U.; Mannhaupt, G.; Münsterkötter, M.; et al. The FunCat, a Functional Annotation Scheme for Systematic Classification of Proteins from Whole Genomes. *Nucleic Acids Res.* **2004**, *32*, 5539–5545. [[CrossRef](#)]
86. Lu, J.; Salzberg, S.L. Removing Contaminants from Databases of Draft Genomes. *PLoS Comput. Biol.* **2018**, *14*, e1006277. [[CrossRef](#)]
87. Brown, D.W.; Lee, S.H.; Kim, L.H.; Ryu, J.G.; Lee, S.; Seo, Y.; Kim, Y.H.; Busman, M.; Yun, S.H.; Proctor, R.H.; et al. Identification of a 12-Gene Fusaric Acid Biosynthetic Gene Cluster in *Fusarium* Species Through Comparative and Functional Genomics. *Mol. Plant-Microbe Interact.* **2015**, *28*, 319–332. [[CrossRef](#)] [[PubMed](#)]
88. Wiemann, P.; Willmann, A.; Straeten, M.; Kleigrewe, K.; Beyer, M.; Humpf, H.U.; Tudzynski, B. Biosynthesis of the Red Pigment Bikaverin in *Fusarium fujikuroi*: Genes, Their Function and Regulation. *Mol. Microbiol.* **2009**, *72*, 931–946. [[CrossRef](#)]
89. Studt, L.; Wiemann, P.; Kleigrewe, K.; Humpf, H.-U.; Tudzynski, B. Biosynthesis of Fusarubins Accounts for Pigmentation of *Fusarium fujikuroi* Perithecia. *Appl. Environ. Microbiol.* **2012**, *78*, 4468–4480. [[CrossRef](#)]
90. Tudzynski, B. Gibberellin Biosynthesis in Fungi: Genes, Enzymes, Evolution, and Impact on Biotechnology. *Appl. Microbiol. Biotechnol.* **2005**, *66*, 597–611. [[CrossRef](#)]