Zheng et al. BMC Genomics (2018) 19:927
https://doi.org/10.1186/s12864-018-5222-8

BMC Genomics

RESEARCH ARTICLE Open Access

Comparative genomic analysis revealed rapid differentiation in the pathogenicity-related gene repertoires between *Pyricularia oryzae* and *Pyricularia penniseti* isolated from a *Pennisetum* grass

Huakun Zheng[1,2†], Zhenhui Zhong[2,3†], Mingyue Shi[2,3], Limei Zhang[2,3], Lianyu Lin[2,4], Yonghe Hong[2,4], Tian Fang[2,3], Yangyan Zhu[2,3], Jiayuan Guo[2,3], Limin Zhang[2,3], Jie Fang[2,4], Hui Lin[1], Justice Norvienyeku[2,4], Xiaofeng Chen[5], Guodong Lu[1,2*], Hongli Hu[2,3*] and Zonghua Wang[2,3,4*]

Abstract

Background: A number of *Pyricularia* species are known to infect different grass species. In the case of *Pyricularia oryzae* (syn. *Magnaporthe oryzae*), distinct populations are known to be adapted to a wide variety of grass hosts, including rice, wheat and many other grasses. The genome sizes of *Pyricularia* species are typical for filamentous ascomycete fungi [~ 40 Mbp for *P. oryzae*, and ~ 45 Mbp for *P. grisea*]. Genome plasticity, mediated in part by deletions promoted by recombination between repetitive elements [Genome Res 26:1091-1100, 2016, Nat Rev Microbiol 10:417-430,2012] and transposable elements [Annu Rev Phytopathol 55:483-503,2017] contributes to host adaptation. Therefore, comparisons of genome structure of individual species will provide insight into the evolution of host specificity. However, except for the *P. oryzae* subgroup, little is known about the gene content or genome organization of other *Pyricularia* species, such as those infecting *Pennisetum* grasses.

Results: Here, we report the genome sequence of *P. penniseti* strain P1609 isolated from a *Pennisetum* grass (JUJUNCAO) using PacBio SMRT sequencing technology. Phylogenomic analysis of 28 Magnaporthales species and 5 non-Magnaporthales species indicated that P1609 belongs to a *Pyricularia* subclade, which is genetically distant from *P. oryzae*. Comparative genomic analysis revealed that the pathogenicity-related gene repertoires had diverged between P1609 and the *P. oryzae* strain 70–15, including the known avirulence genes, other putative secreted proteins, as well as some other predicted *Pathogen-Host Interaction* (*PHI*) genes. Genomic sequence comparison also identified many genomic rearrangements relative to *P. oryzae*.

Conclusion: Our results suggested that the genomic sequence of the *P. penniseti* P1609 could be a useful resource for the genetic study of the *Pennisetum*-infecting *Pyricularia* species and provide new insight into evolution of pathogen genomes during host adaptation.

Keywords: *Pennisetum*, *Pyricularia*, PacBio sequencing, Pathogenesis-related genes, Comparative genomic analysis

* Correspondence: gdlufafu@163.com; huhongli7905@gmail.com; zonghuaw@163.com
†Huakun Zheng and Zhenhui Zhong contributed equally to this work.
[1]National Engineering Research Center of JUNCAO Technology, College of Life Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China
[2]State Key Laboratory of Ecological Pest Control for Fujian and Taiwan Crops, College of Plant Protection, Fujian Agriculture and Forestry University, Fuzhou 350002, China
Full list of author information is available at the end of the article

Background

Pyricularia was established by Saccardo to accommodate a type of fungal species based on pyriform conidia when the first species of this pathogen, *Pyricularia grisea*, was isolated from crabgrass (*Digitaria sanguinalis* L.) [1]. *Pyricularia* became an important research focus due to rice blast disease and now wheat blast caused by *Pyricularia oryzae* (syn. *Magnaporthe oryzae*) [2–4]. To date, 100 plant genera comprising 256 species have been documented as hosts of *Pyricularia* species (https://nt.ars-grin.gov/fungaldatabases/), among which 54 genera belong to the Poaceae family. Seven *Pyricularia* species (including one unidentified species) have been isolated from *Pennisetum* spp., a large genus in the Poaceae family, and more than one *Pyricularia* species can be found on the same *Pennisetum* species. For instance, 4 *Pyricularia* species, namely, *P. penniseti*, *P. penniseticola*, *P. setariae* and *Pyricularia sp.* have been found on *P. typhoides* [5, 6].

The genome sequence of *P. oryzae* strain 70–15, a strain produced by an initial cross of rice-infecting isolate 104–3 and the weeping love grass isolate AR4 [7]. A series of selections from crosses, including three generations crossing to the rice isolate Guy11 generated fertile rice infecting strains that have facilitated studies of developmental and pathogenic mechanisms of the blast fungus and helped *P. oryzae* to become one of the most important fungal models of plant pathogenesis [8, 9]. Since the publication of the *P. oryzae* strain 70–15 genome, more field isolates were sequenced and assembled, including Ina168, HN9311, FJ81278, Y34, P131, 98–06 and Guy11 [10–14]. Comparative genomic analyses and functional studies of these strains revealed genome plasticity and the involvement of the lineage specific genes in pathogenicity [12, 14]. More recently, facilitated by the fast development of sequencing technologies, a number of field isolates from rice, as well as isolates from different grass and cereal hosts, were sequenced and subjected to population-level analyses, revealing host immunity as a major force driving specialization [2, 15–18]. However, genomes of most of the species of the *Pyricularia* complex remain unexplored. For example, among the 7 identified *Pyricularia* species isolated from *Pennisetum* grasses [5], only *P. pennisetigena* was recently sequenced [2].

Here, we reported the whole-genome sequence of *P. penniseti* [19] isolated from a *Pennisetum* grass JUJUNCAO (*Pennisetum giganteum* Z. X. Lin). JUJUNCAO was originally developed as a biomass crop for the cultivation of edible mushrooms by Lin et al, and later became a versatile grass that is used as forage for cattle and sheep, material for biofuel production, and a tool for the prevention of soil erosion [20–22]. We have recently isolated a fungus, producing pyriform-shaped conidia from leaf spots of JUJUNCAO. One isolate, P1609, caused a typical blast fungal disease symptom on JUJUNCAO, showing small, round or elliptical lesions as an initial symptom with spindle shaped, grayish to tan necrotic lesion centers, and yellow halos at a later disease stage. The morphologic and phylogenetic analyses distinguished P1609 from other identified *Pyricularia* species, but it was not clearly distinguishable from the *P. penniseti* reported in 1970 in India [5, 23]. We therefore identified this fungus as *P. penniseti* [19]. In this study, we performed genome sequencing of this strain, aiming for a proper classification of this fungus in the *Pyricularia* population and identification of genes that may be involved in the adaptation of this fungus to JUJUNCAO.

Results

Genome sequencing and assembly

We sequenced the P1609 genome with the long-read PacBio technology. In total, 312,061 reads with 8.6 Kb average lengths were obtained, representing about 60-fold coverage of the genome (Fig. 1a). The genome sequence was assembled with the HAGP pipeline [24], resulting in a total assembly space of 41.82 Mb (Table 1), similar to assemblies of other isolates sequenced by PacBio [10]. The assembly contains 53 contigs, with the N50 of 3.4 Mb and the largest contig of 7.56 Mb (Fig. 1b). Contigs > 1 Mb cover 89.7% and contigs > 100 Kb cover 98.5% of the genome (Fig. 1b; Table 1), indicating long-continuity of the assembly. The GC content of the assembly is 50.3%, similar to genomes of *Pyricularia* isolates from different host plants, which range from 48.6 to 51% [18]. Genome annotation identified 13,102 genes with average gene size of 1758 bp, fairly evenly dispersed on contigs (Table 2, Fig. 1c, track b).

De novo repeat sequence analysis identified 7.67% repeat sequences, among which 4.27% were Gypsy and Copia, the two long terminal repeats (LTR)-type retrotransposons. Although the overall content of repeat sequences in P1609 is less than that of the sequenced *P. oryzae* strains [8, 10, 12], the high proportion of Gypsy and Copia in repeat sequence is similar to that of the reported *P. oryzae* strains [25]. The TEs are not evenly distributed along the contigs, with some contigs being highly enriched with TEs (Fig. 1c, track c), suggesting that the P1609 genome also underwent transposon expansion as observed in other plant pathogens [26].

Comparative and phylogenetic analysis

To understand the genetic relationship of P1609 with other fungal phytopathogens, we generated a phylogenetic tree of P1609 with *Botrytis cinerea*, *Colletotrichum gloeosporioides*, *Fusarium graminearum*, *Neurospora crassa*, *Pyricularia grisea* (Pg), *Pyricularia oryzae* (Po), *Sclerotinia sclerotiorum*, *Trichoderma reesei* and *Ustilago*

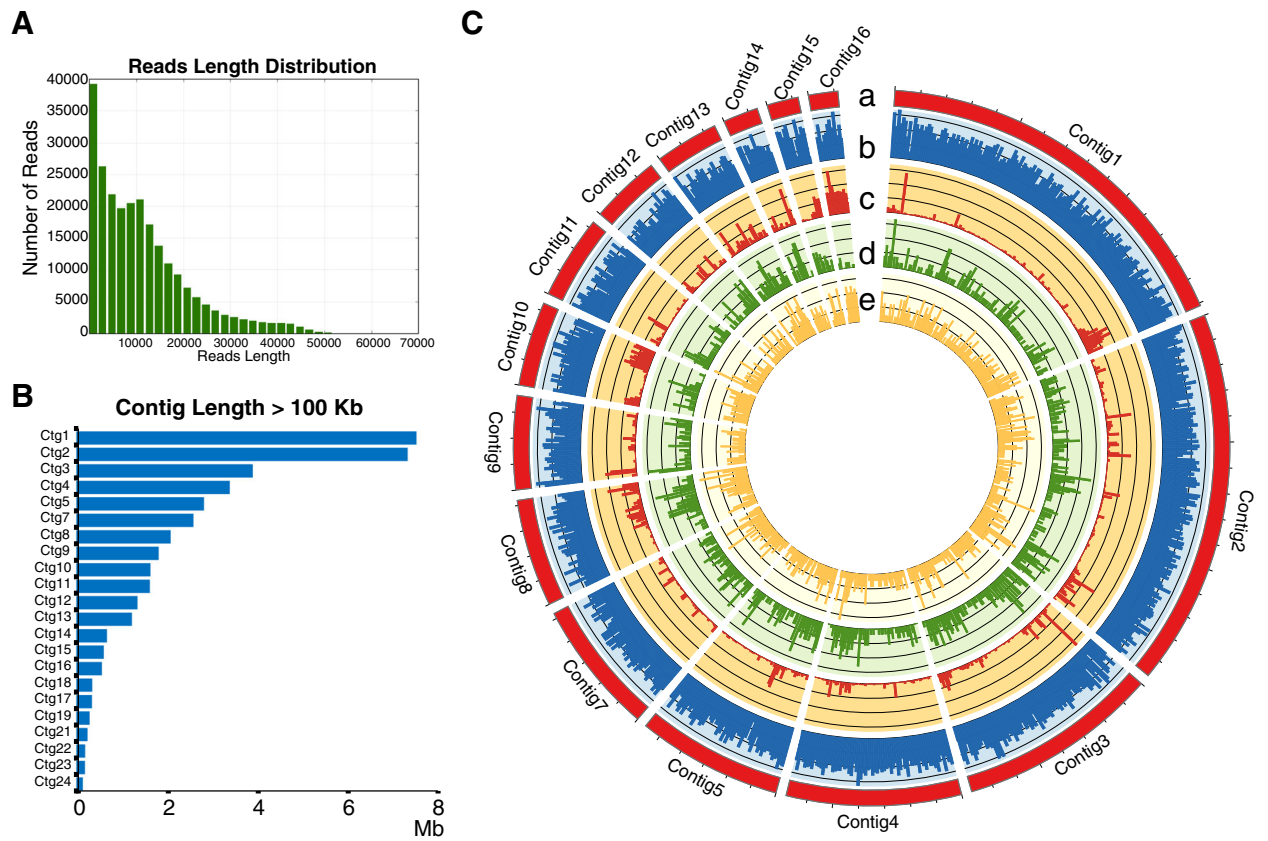

Fig. 1 PacBio sequencing and genome assembly of P1609. **a** Reads length distribution. **b** Contig length of assembled contig > 100 Kb. **c** Overview of P1609 genome. (Track a) Contig1 to contig16 of P1609, (track b) gene density, (track c) Transposon elements density, (track d) secreted proteins density, (track e) unique gene (compared with *P. oryzae*, *N. crassa*, *F. graminearum* and *C. gloeosporioides*) density of P1609 per 50 Kb

maydis. Since P1609 showed a close morphological relationship with *P. oryzae* isolates, we included *Pyricularia* isolates collected from different host plants, including *Oryza sativa* (PoOs), *Triticum aestivum* (PoTa), *Digitaria sanguinalis* (PgDs), *Setaria viridis* (PoSv), and *Eleusine indica* (MoEi). In total, 2051 single-copy genes shared by all the examined genomes were selected to infer phylogeny [27]. The resulting phylogenetic tree indicated that P1609 is more distantly related to PoOs, PoTa, PoSv, and PoEi (*P. oryzae*) than PgDs (*P. grisea*) (Fig. 2a). We then estimated divergence time of P1609 and *Pyricularia* isolates by assuming a constant molecular clock calibrated in a previous study [28, 29], which estimated the divergence of *Neurospora* and *Pyricularia*

Table 1 Details of sequencing reads and genome assembly of P1609

Feature	Value
Number of Subreads	312,061
Total Length of Subreads (bp)	2,688,966,115
Mean Coverage	59
Polished Contigs	53
Maximum Contig Length (bp)	7,555,856
N50 Contig Length (bp)	3,411,241
Sum of Contig Lengths (Mb)	41.82
GC level	50.30%
length > 1 Mb	89.70%
length > 100 Kb	98.50%

Table 2 Details of genome annotation of P1609

Category	Value
Total TE	7.67%
LINEs	1.50%
LTR elements	4.27%
DNA elements	0.63%
Unclassified	1.27%
Gene Number	13,102
Average Gene Length (bp)	1758
Secreted Protein Number	1409

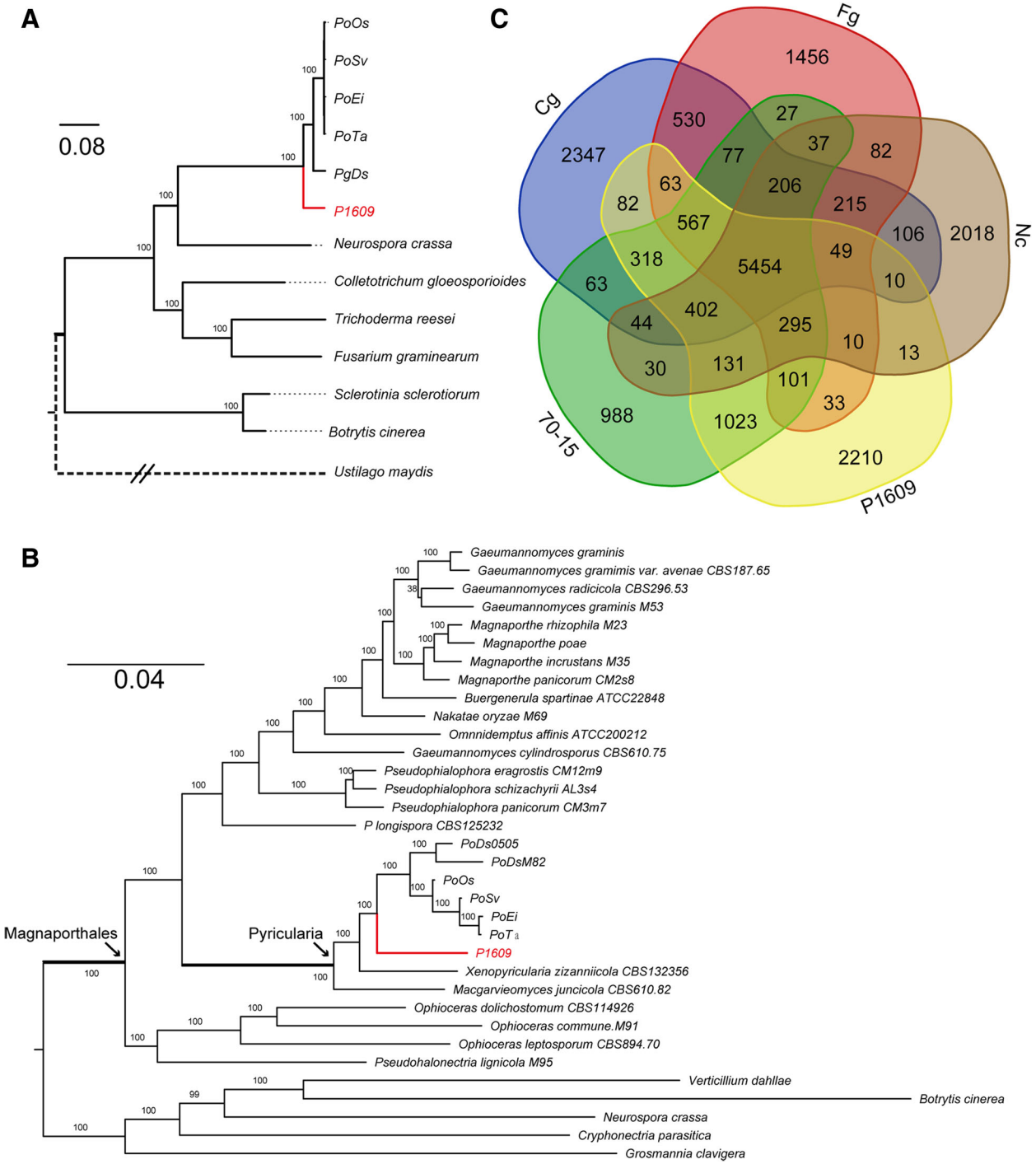

Fig. 2 Phylogenetic and comparative genomic study of P1609. **a** Phylogenomic tree of P1609 with *Botrytis cinereal*, *Colletotrichum gloeosporioides*, *Fusarium graminearum*, *Neurospora crassa*, *Sclerotinia sclerotiorum*, *Trichoderma reesei* and *Ustilago maydis* as well as *Pyricularia* isolates collected from *O. sativa* (70–15), *T. aestivum* (PoTa), *D. sanguinalis* (PgDs), *S. viridis* (PoSv), and *E. indica* (PoEi) based on 2051 single copy genes. The values of all of the branches are 100. **b** Maximum likelihood tree of P1609 and 28 Magnaporthales species, as well as 5 Sordariomycetes used as outgroup species based on 82,715 amino acid positions derived from 226 genes. **c** Venn diagram showed an overlap of gene families among P1609, *Pyricularia* rice isolates (PoOs), *C. gloeosporioides* (Cg), *F. graminearum* (Fg) and *N. crassa* (Nc)

at about 200 million years ago (MYA). The estimation indicated that P1609 and *Pyricularia* isolates diverged at about 31 MYA, earlier than the divergent time of rice- and *S. viridis*-infecting isolates (about 10,000 years ago) [28, 30]. The phylogenetic tree also indicated that P1609 is a member of Magnaporthales, but is distantly related to the *P. oryzae*. To further determine exactly where P1609 localized in Magnaporthales, we also generated a phylogenomic tree of P1609 with 28 Magnaporthales species and 5 non-Magnaporthales species using amino acid sequences of 226 conserved orthologous genes as described [31]. The result showed that P1609 was localized in the *Pyricularia* subclade, and was closer to *P. oryzae* than *Xenopyricularia zizaniicola* (Fig. 2b), indicating that P1609 is a *Pyricularia* species.

We next conducted a comparative genomic study of P1609 with 70–15 (the reference isolate of *P. oryzae*), *N. crassa* (Nc), *F. graminearum* (Fg) and *C. gloeosporioides* (Cg) using OrthoFinder [27]. The five organisms share 5454 of gene families with each other, which covers 50.2% of the gene set of P1609 (Fig. 2c). Notably, the number of P1609 specific genes in this comparison was 2210. This is about twice the number of 70–15 specific genes. Although most of these P1609 specific genes had no functional annotations, Pfam annotation indicated that some encode carbohydrate-active enzymes (CAZymes) involved in polysaccharide metabolism pathways (Additional file 1: Figure S1; Additional file 2: Table S1).

To identify genes under positive selection, Ka/Ks ratios of the orthologous genes in P1609 versus 70–15 were calculated, with the assumption that Ka/Ks ratio > 1 suggested a positive selection of the gene [11]. Among the 5991 pairs of orthologs of P1609 and 70–15 scanned, 6 genes with Ka/Ks > 1 were identified in P1609 (Table 3), suggestive of positive selection on these genes in P1609 during the adaptation to JUJUNCAO. These genes are involved in different secondary metabolic pathways (See Discussion).

Gene categories involved in pathogenicity

Plant pathogenic fungi employed diverse gene repertoires to invade host plants and subvert host immune systems, which include effectors, carbohydrate-active enzymes (CAZymes), other secreted enzymes and fungal secondary metabolisms [32, 33]. To understand the differences between *P. penniseti* and *P. oryzae* secretomes, we examined the 1409 putative secreted proteins of P1609 predicted by SignalP [34], and found 236 were unique in P1609 (Additional file 3: Table S2). By contrast, 165 putative effectors identified in the genome of 70–15, the rice blast reference fungal genome, were absent in the P1609 genome (Additional file 4: Table S3). Notably, all known avirulence genes (*AVRs*) from the rice-infecting isolate genomes were absent in P1609 genome.

Given the important roles of CAZymes in enabling plant pathogens to break down the plant cell wall [35], we next compared CAZymes between P1609 and 70–15. Our BLASTp search results showed that P1609 contains more predicted CAZyme-coding genes than 70–15 (Additional file 2: Table S1). Detailed analysis showed that the P1609 genome encodes five unique CAZymes belonging to five families, namely CBM61, GH117, GH35, GH65 and PL24, respectively (Additional file 2: Table S1). While it has six copies of GH28 pectinases (three copies in *P. oryzae and P. grisea*; Additional file 5: Figure S2). The high representation of carbohydrate-active enzymes may be related to the adaptation to the host *P. giganteum*. We then further analyzed the distribution of annotated PHI genes in P1609. In total, we identified 1692 potential PHI genes belonging to 1154 gene families (Additional file 6: Table S4). Interestingly, we found that several PHI genes exhibited great expansion in P1609 genome. For instance, *MGG_12656*, a gene involved in virulence in *P. oryzae*, has 107 homologs in P1609, and ChLae1 contributing to toxin production and virulence in the maize pathogen *Cochliobolus heterostrophus* has 17 homologs in P1609 [36, 37]. Compared with 70–15, 35 PHI genes were unique in P1609 (Table 4), most of which had highly similar homologs in *Fusarium* (*Gibberella*).

Chromosome rearrangements

Chromosome rearrangements were broadly documented and had been reported to play important roles in fungal evolution, host adaptation and pathogenicity of phytopathogens [38, 39]. To explore genome collinearity and rearrangement between P1609 and 70–15, the identified

Table 3 Positively selected genes in P1609

Protein ID	Ka	Ks	Ka/Ks	*P*-Value (Fisher)	Description
P1609_7184	0.707979	0.440748	1.60631	0.00697146	Unknown
P1609_5032	0.550905	0.343609	1.60329	1.98443E-06	Isoamyl alcohol oxidase
P1609_3360	0.183224	0.118517	1.54597	0.00293936	Glycerol uptake protein 1
P1609_791	0.302231	0.199786	1.51277	9.75093E-06	Folylpolyglutamate synthase
P1609_7869	0.491873	0.328109	1.49911	0.0296442	Thioredoxin reductase
P1609_1006	0.40206	0.275456	1.45962	0.0288585	Spore surface glycoprotein BclB

Table 4 Unique Pathogen Host Interaction (PHI) genes in P1609

Protein ID	PHI ID	Reference Organism	Phenotypes
P1609_2497_494aa	PHI:115	*Cochliobolus carbonum*	Unaffected pathogenicity
P1609_11506_262aa, P1609_12781_266aa, P1609_4501_233aa, P1609_683_218aa	PHI:698	*Vibrio cholerae*	Reduced virulence
P1609_11592_161aa, P1609_4614_892aa	PHI:1284	*Fusarium graminearum*	Unaffected pathogenicity
P1609_11335_610aa	PHI:1803	*Fusarium graminearum*	Unaffected pathogenicity
P1609_9374_326aa	PHI:2394	*Fusarium graminearum*	Increased virulence
P1609_1007_600aa	PHI:225	*Fusarium solani*	Reduced virulence
P1609_13007_833aa	PHI:1714	*Fusarium graminearum*	Lethal
P1609_3609_435aa, P1609_5044_426aa, P1609_8181_460aa, P1609_960_144aa	PHI:1455	*Fusarium graminearum*	Unaffected pathogenicity
P1609_11548_383aa, P1609_5843_397aa	PHI:1823	*Fusarium graminearum*	Unaffected pathogenicity
P1609_2665_64aa	PHI:1421	*Fusarium graminearum*	Lethal
P1609_2607_206aa, P1609_2608_820aa	PHI:3418	*Staphylococcus saprophyticus*	Mixed outcome
P1609_11380_688aa	PHI:1809	*Fusarium graminearum*	Lethal
P1609_7767_257aa	PHI:317	*Aspergillus fumigatus*	Reduced virulence
P1609_6736_447aa	PHI:323	*Verticillium fungicola*	Reduced virulence
P1609_13_169aa	PHI:1147	*Fusarium pseudograminearum*	Unaffected pathogenicity
P1609_2376_825aa	PHI:1871	*Fusarium graminearum*	Lethal
P1609_10320_326aa	PHI:1522	*Fusarium graminearum*	Lethal
P1609_11485_355aa, P1609_2130_433aa	PHI:3116	*Pseudomonas syringae*	Mixed outcome
P1609_10203_738aa	PHI:1815	*Fusarium graminearum*	Unaffected pathogenicity
P1609_11600_2486aa	PHI:2628	*Salmonella enterica*	Reduced virulence, 663
P1609_11388_1000aa	PHI:3076	*Candida parapsilosis*	Mixed outcome
P1609_11869_369aa	PHI:2375	*Alternaria brassicicola*	Mixed outcome
P1609_5035_383aa	PHI:233	*Cochliobolus carbonum*	Reduced virulence
P1609_5702_488aa	PHI:1271	*Fusarium graminearum*	Unaffected pathogenicity
P1609_11864_596aa	PHI:2380	*Alternaria brassicicola*	Mixed outcome

collinear gene blocks that linked with different chromosomes in 70–15 (Fig. 3a) were visualized in the contigs (ctgs) of P1609 (Fig. 3b). P1609 and 70–15 overall displayed high genome collinearity. Ctg1, ctg3, ctg5 and ctg7 in P1609 correspond to chr.2, chr.6, chr.1 and chr.5 of 70–15, respectively. We found that the second largest contig in the P1609, ctg2, is a combination of chr.4 and chr.7 of 70–15 (Fig. 3b). The joining region was spanned by multiple single PacBio long reads (Fig. 3d), excluding the possibility that the rearrangement was an artifact due to assembly errors. Meanwhile, notably, the contig end regions of P1609 showed a higher level of chromosome deletion and rearrangement when compared to 70–15 (Fig. 3c). For instance, ctg4 and ctg8 were merged with blocks homologous to 70–15 chr.1 and chr.3 at the end of the contig, while ctg9 was merged with blocks from chr.3 and chr.5 at the end of the contig (Fig. 3b).

Discussion

Here we sequenced the genome of *P. penniseti* using PacBio SMRT technology. Comparative genomic analysis revealed several chromosome rearrangement events relative to *P. oryzae* and large differences in pathogenicity-related gene repertoires between the *P. penniseti* strain P1609 and the *P. oryzae* strain 70–15.

Chromosome fission and fusion

Long-read sequencing greatly improved genome assembly and thus provided more valuable details on genome structures at a chromosome level. By comparing the chromosome structure between P1609 and 70–15, we found several chromosome splitting and rearrangement events. Chromosomal rearrangements have been reported to be associated with virulence evolution in several pathogens due to the loss of *AVR* gene(s) [26, 40, 41]. In addition to rearrangements, the telomere regions were also rearranged relative to *P. oryzae*. Further study is required to investigate the role of the chromosome recombination during the adaptation to JUJUNCAO as well as the variation that occurs between individuals in the *P. penneseti* population.

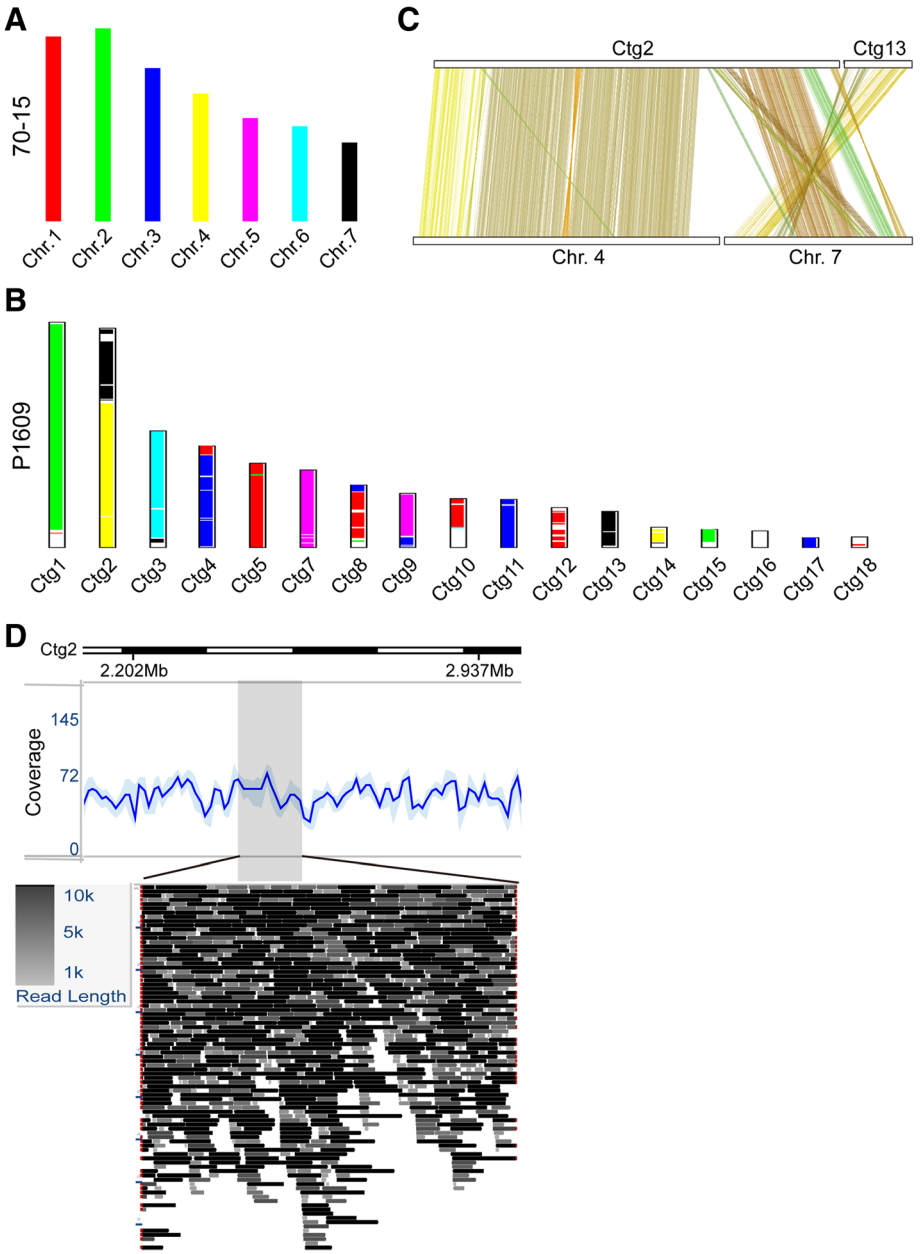

Fig. 3 Chromosome rearrangement and splitting between P1609 and 70–15. **a** Bar plot showing the chromosomes in 70–15. **b** Bar plot showing the assembled contigs of P1609. Colinear chromosomes of 70–15 and contigs of P1609 are indicated by the same color. **c** Dual synteny plot showing splitting of Ctg2 of P1609 into chr. 4 and chr. 7 in 70–15. **d** PacBio long-read coverage from 2.02 to 2.94 Mb of Ctg 2. Color of reads indicate different read lengths.

Unique genes and positively selected genes

In this study, we found 2210 unique genes in *P. penniseti* that were not present in the 70–15 genome. The P1609 genome encodes more CAZymes than 70–15 genome. For example, P1609 contains twice the number of genes encoding glycosyl hydrolase family 28 (GH28) pectinases in 70–15. Generally, necrotrophic plant pathogens possess more GH28 enzymes than biotrophic and non-pathogens fungi [42]. These results suggested that P1609 might heavily rely on CAZymes in the interaction with JUJUNCAO.

To identify genes under positive selection, the Ka/Ks ratios in the 5991 orthologous gene pairs between P1609 and 70–15 were calculated, and only 6 genes with Ka > Ks in P1609 were identified. Functional annotation showed that most of these genes appear to be involved in secondary metabolic pathways. For example, P1609_5032 encodes an isoamyl alcohol oxidase that turns isoamyl alcohol into

isovaleraldehyde [43]. In *Saccharomyces cerevisiae*, isoamyl alcohol could induce filament formation [44]. P1609_5032 might play a role in fungal development through controlling the level of isoamyl alcohol. P1609_791 encodes a folylpolyglutamate synthase involved in biosynthesis of folate, required for protein synthesis in bacteria, mitochondria, and chloroplasts, and biosynthetic pathways for purines, dTMP, methionine, and formyl-methionyl-tRNA [45]. P1609_3360 encodes a glycerol uptake protein, which is an important intracellular osmolyte participating in osmotic stress response [46] and critical for the function of appressorium in *P. oryzae* [47]. P1609_7869 encodes a homolog of a spore surface glycoprotein shown to be involved in spore adhesion to hydrophobic surfaces in several *Colletotrichum* species [48–50]. Adhesion of spore tip mucilage is important for infection in *P. oryzae* [51]. P1609_1006 encodes a BclB glycoprotein (collagen-like protein). Its homologs in *Bacillus anthracis* are important components of the infection-associated structure exosporium [52–54], although its role in filamentous pathogens remains unknown. The positive selection on these genes suggested that they might play roles in the interaction between P1609 and JUJUNCAO.

Putative secreted proteins

Increasingly, studies of pathogen populations support the view that plant immunity is the major force driving the specialization of adapted pathogens [2, 15–18]. It was previously proposed by Schulze-Lefert and Panstruga that effector-triggered immunity (ETI) is the major force driving the host specificity of pathogens [55]. Our previous study focusing on *AVR* gene evolution of the *Pyricularia* species also revealed that directional selection exerted by host plants is the direct force driving host specificity in *Pyricularia* species [18]. Recent studies in the interaction between *M. oryzae* and rice revealed that both ETI and pathogen-associated molecular patterns (PAMPs)-triggered immunity (PTI) are involved in determining effector repertoires and specialization to the two subspecies of *P. oryzae* [15–17]. This opinion was supported by the results that the japonica-isolate genomes contain many effectors to overcome the strong basal immunity posed by japonica rice, but prevent them from infection of indica rice through the triggering of ETI. By contrast, the indica-isolates deserted most of the effectors to avoid recognition by *R* genes encoded by indica plants, but are disabled in conquering the elevated basal immunity in japonica rice [16]. Our comparative genomic analysis showed that P1609 contains a large number of unique effector candidates compared with 70–15, but lost (or never possessed) many putative effectors found in the 70–15 genome, including all known AVR effectors. One possible explanation is that the JUJUNCAO harbors a high level of basal immunity, as well as an arsenal of resistance genes, which drove P1609 to gain host-specific effectors to overcome the robust basal immunity posed by JUJUNCAO. The large difference in effector/Avr gene content between *P. penniseti* and *P. oryzae* suggest the selection driving evolution of these fungi can help direct investigation of the molecular mechanisms underlying the interaction between *Pyricularia* species and their hosts.

Conclusion

Pyricularia species are pathogens of grasses, many of which are either food or forage grasses. The model fungus *P. oryzae* had been well studied. However, there are only a few whole-genome sequences available for other species, such as those from *Pennisetum* grasses. Here, we generated long-read PacBio reads and produced a assemblage with long-continuity contig sequences. The phylogenomic and comparative genomic analysis suggested that P1609 is a *Pyricularia* species genetically distant from *P. oryzae*, and may have employed diverse mechanisms during the adaptation to JUJUNCAO, including the pathways associated with secondary metabolics (positively selected genes), CAZymes and the putative secreted proteins, as well as other PHI proteins. In summary, the P1609 assembly and genome annotation represents the few available *Pyricularia* genome resources for studying the pathogenic mechanism of this fungus towards *Pennisetum* grasses.

Methods

Isolation of the fungal strain

The *Pennisetum*-infecting strain P1609 was isolated from the leaf spot lesion of JUJUNCAO (*Pennisetum giganteum* Z. X. Lin), in the nursery of National Engineering Research Center of JUNCAO Technology, Fujian Agriculture and Forestry University located at No. 63 Xiyuangong Road, Minhou County, Fuzhou, Fujian Province, China.

DNA extraction, amplification and sequencing

To prepare the genomic DNA for sequencing, the P1609 isolate was cultured in the liquid complete medium (CM) in a 110-rpm shaker at 25 °C for 3 to 4 days. The mycelia were then collected for the preparation of genomic DNA using a CTAB method as previously described [18]. Sequencing libraries were prepared using the SMRTbellTM Template Prep Kit 1.0 (PACBIO) and sequenced using PacBio Sequel platform (NovoGene, China).

Assembly and annotation

De novo sequence assembly was conducted by SMRTLink v. 5.0.1.10424, HGAP 4 pipeline provided by Pacific Bioscience Company. In HGAP 4 pipeline, the expected

genome size was set as 45 Mb based on the reported size of *Pyricularia* genomes, and default settings were used for other parameters. Gene prediction was conducted using Fgenesh from SoftBerry (MolQuest II v2.4.5.1135, http://linux1.softberry.com/berry.phtml) with *Pyricularia* additional variants as training organism. Gene functional domain annotation was conducted by InterproScan (version 4.8, http://www.ebi.ac.uk/ interpro/), and PfamScan [56]. Pathogen-Host Interaction (PHI) genes were predicted by performing a whole genome blastp analysis against the PHI database ($E < 10^{-10}$) [57, 58]. Putative carbohydrate-active enzymes (CAZymes) were identified using the HMMER 3.1b1 by searching annotated HMM profiles of CAZyme families downloaded from the dbCAN database in protein sequences of P1609 [59].

Repeat analysis

De novo repeat sequence identification was analyzed by using RepeatModeler (version 1.0.8) with default settings. Repeat sequences obtained from RepeatModeler have been used to search for repeat sequences in the P1609 genome by RepeatMasker (version 3.3.0) (http://www.repeatmasker.org/) [60].

Phylogenetic analysis and comparative genomic analysis

Phylogenomic tree of P1609 and *B. cinerea* [61], *C. gloeosporioides* [62], *F. graminearum* [63], *N. crassa* [64], *P. oryzae* [8], *S. sclerotiorum* [61],*T. reesei* [65] and *U. maydis* [66] was built based on single copy orthologs from clustering result of OrthoFinder (v0.6.1) [27]. 2051 single copy genes have been selected out from 13 organisms in total (see Fig. 2a) and aligned with MAFFT (mafft-linsi-anysymbol) [67]. The phylogenetic tree was constructed using FastTree based on the alignments of single-copy orthologs with approximately-maximum-likelihood model and 100 bootstrap iterations [68]. For divergence time estimation, the phylogenetic analysis was conducted using r8s (version 1.81), and the divergence time of *Pyricularia* and *Neurospora* (200 MYA) was used as a reference [28, 29]. Clustering result of 13 genomes was also used for unique gene identification and comparative genomic analysis. The comparative genomic study of homologs among P1609 and 70–15 (the reference isolate of *P. oryzae*), *N. crassa*, *F. graminearum* and *C. gloeosporioides* was conducted using the OrthoFinder [27]. The identification of positively selected genes was performed as described previously [11]. Genes were aligned in pairs between P1609 and 70–15. Codeml tool in PAML suite was used to calculate Ka/Ks ratio, with the assumption that Ka/Ks ratio > 1 suggested the gene was under positive selection. For the prediction of secreted proteins, SignalP 4.1 was implemented to predict signal peptides and TMHMM 2.0 have been used to predict the transmembrane domain [34, 69]. Proteins with a signal peptide cleavage site, amino acid length smaller than 400 amino acids, and no transmembrane domain after the region signal peptide cleavage site were defined as putative secreted proteins in this study. We used MCScanX to identified syntenic blocks between P1609 and 70–15. To detect the conserved synteny blocks, the reciprocal best-match paralogs of P1609 and 70–15 were conducted by all-against-all BLASTP comparison, with E-value < 10^{-10} [70].

Additional files

Additional file 1: Figure S1. Gene copy numbers of CAZymes in Botrytis cinereal, Colletotrichum gloeosporioides, Fusarium graminearum, Neurospora crassa, Sclerotinia sclerotiorum, Trichoderma reesei and *Ustilago maydis* as well as Pyricularia isolates collected from *O. sativa* (PoOs), *T. aestivum* (PoTa), *D. sanguinalis* (PgDs), *S. viridis* (PoSv), and *E. indica* (PoEi). Log2 Copy Number presents variation of copy number with increased red color means increased number of CAZymes. (XLSX 20 kb)

Additional file 2: Table S1. The CAZymes identified in P1609 and other fungi. (XLS 47 kb)

Additional file 3: Table S2. P1609_vs_7015_unique_secreted. (XLSX 16 kb)

Additional file 4: Table S3. 70–15 VS P1609 unique secreted proteins. (XLSX 71 kb)

Additional file 5: Figure S2. GH28 of P1609 (P1609_11576, P1609_5879, P1609_2497, P1609_5781, P1609_680 and P1609_5514)) and PoOs (MGG_09608, MGG_08752 and MGG_08938), PgDs (Ds0505_9820). Extra copies of GH28 in P1609 is marked by blue. (JPG 237 kb)

Additional file 6: Table S4. Predicted PHI in P1609. (JPG 113 kb)

Abbreviations

AVR gene: Avirulence gene; *B. cinerea*: *Botrytis cinerea*; CAZymes: Carbohydrate-active enzymes; CBM: Carbohydrate-binding module family; Cg: *Colletotrichum gloeosporioides* (*C. gloeosporioides*); chr: Chromosome; CM: Complete medium; CTAB: Cetyltrimethyl ammonium bromide; ctg: Contig; ETI: Effector-triggered immunity; Fg: *Fusarium graminearum* (*F. graminearum*); GH: Glycosyl hydrolase family pectinases; LTR: Long terminal repeats; MYA: Million years ago; Nc: *Neurospora crassa* (*N. crassa*); *P. giganteum*: *Pennisetum giganteum*; *P. grisea*: *Pyricularia grisea*; *P. penniseti*: *Pyrucularia penniseti*; *P. penniseticola*: *Pyricularia penniseticola*; *P. setariae*: *Pyricularia setariae*; *P. typhoides*: *Pennisetum typhoides*; PgDs: *Pyricularia* strain isolated from *Digitaria sanguinalis*; PHI: Pathogen-Host Interaction; PL: Polysaccharide Lyase Family; Po: *Pyricularia oryzae* (*P. oryzae*); PoEi: *Pyricularia* strain isolated from *Eleusine indica*; PoOs: *Pyricularia* strain isolated from *Oryza sativa*; PoSv: *Pyricularia* strain isolated from *Setaria viridis*; PoTa: *Pyricularia* strain isolated from *Triticum aestivum*; PTI: Pathogen-associated molecular pattern (PAMPs)-triggered immunity; *R* genes: Resistance gene; *S. Sclerotiorum*: *Sclerotinia sclerotiorum*; SMRT: Single-Molecule Real-Time; *T. reesei*: *Trichoderma reesei*; TE: Transposable elements; *U. maydis*: *Ustilago maydis*

Acknowledgments

We would like to thank Dr. Sanzhen Liu at Kansas State University and Dr. Daniel Ebbole at Texas A&M University for critically reading the manuscript.

Funding

This work was supported by grants from the Natural Science Foundations of China to Z.W (31770156), the Scientific Research Foundation of the Graduate School of FAFU to Z.Z, and the State Key Laboratory of Ecological Pest Control for Fujian and Taiwan Crops to H.Z (SKB2017002).

Availability of data and materials

Genome assembly and PacBio reads are available in GenBank under BioProject PRJNA416656. The Whole Genome sequence has been deposited at GenBank under the accession PELF00000000.

Authors' contributions
The study was conceived and designed by ZW and GL. The initial collection and culturing of the strain was performed by HZ, XC, HH, MS, LiZ, YH, TF, YZ, JG, LnZ, JF and HL. Bioinformatics was performed by ZZ, and LL. ZZ, HZ, JN, GL and ZW wrote, revised and approved the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note

Author details
[1]National Engineering Research Center of JUNCAO Technology, College of Life Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China. [2]State Key Laboratory of Ecological Pest Control for Fujian and Taiwan Crops, College of Plant Protection, Fujian Agriculture and Forestry University, Fuzhou 350002, China. [3]College of Plant Protection, Fujian Agriculture and Forestry University, Fuzhou 350002, China. [4]College of life science, Fujian Agriculture and Forestry University, Fuzhou 350002, China. [5]Institute of Oceanography, Minjiang University, Fuzhou 350108, China.

Received: 17 June 2018 Accepted: 5 November 2018
Published online: 13 December 2018

References
1. Saccardo P. Fungorum Extra-Europaeorum pugillus. Michelia. 1880;2(6):136–49.
2. Gladieux P, Condon B, Ravel S, Soanes D, Leodato Nunes Maciel J, Nhani J, Antonio N, Chen L, Terauchi R, Lebrun M-H, Tharreau D, et al. Gene flow between divergent cereal- and grass-specific lineages of the rice blast fungus *Magnaporthe oryzae*: mBio; 2017. p. e01219-01217. https://doi.org/10.1128/mBio.01219-01217.
3. Igarashi S, Utiamada C, Igarashi L, Kazuma A, Lopes R. *Pyricularia em trigo*. 1. Ocorrência de Pyricularia sp. no estado do Paraná. Fitopatol Bras. 1986;11:351–2.
4. Wilson RA, Talbot NJ. Under pressure: investigating the biology of plant infection by *Magnaporthe oryzae*. Nat Rev Microbiol. 2009;7(3):185–95.
5. Klaubauf S, Tharreau D, Fournier E, Groenewald JZ, Crous PW, de Vries RP, Lebrun MH. Resolving the polyphyletic nature of *Pyricularia* (Pyriculariaceae). Stud Mycol. 2014;79:85–120.
6. Lenne JM. A world list of fungal diseases of tropical pasture species. Australas Plant Pathol. 1991;20(3):122–4.
7. Chao CCT, Ellingboe AH. Selection for mating competence in *Magnaporthe grisea* pathogenic to rice. Can J Bot. 1991;69(10):2130–4.
8. Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu JR, Pan H, et al. The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature. 2005;434(7036):980–6.
9. Yan X, Talbot NJ. Investigating the cell biology of plant infection by the rice blast fungus *Magnaporthe oryzae*. Curr Opin Microbiol. 2016;34:147–53.
10. Bao J, Chen M, Zhong Z, Tang W, Lin L, Zhang X, Jiang H, Zhang D, Miao C, Tang H. PacBio sequencing reveals transposable element as a key contributor to genomic plasticity and virulence variation in *Magnaporthe oryzae*. Mol Plant. 2017. https://doi.org/10.1016/j.molp.2017.1008.1008.
11. Chen C, Lian B, Hu J, Zhai H, Wang X, Venu RC, Liu E, Wang Z, Chen M, Wang B, et al. Genome comparison of two *Magnaporthe oryzae* field isolates reveals genome variations and potential virulence effectors. BMC Genomics. 2013;14. https://doi.org/10.1186/1471-2164-1114-1887.
12. Dong Y, Li Y, Zhao M, Jing M, Liu X, Liu M, Guo X, Zhang X, Chen Y, Liu Y, et al. Global genome and transcriptome analyses of *Magnaporthe oryzae* epidemic isolate 98-06 uncover novel effectors and pathogenicity-related genes, revealing gene gain and lose dynamics in genome evolution. PLoS Pathog. 2015;11(4):e1004801. https://doi.org/10.1371/journal.ppat.1004801.
13. Xue M, Yang J, Li Z, Hu S, Yao N, Dean RA, Zhao W, Shen M, Zhang H, Li C, et al. Comparative analysis of the genomes of two field isolates of the rice blast fungus *Magnaporthe oryzae*. PLoS Genet. 2012;8(8):e1002869. https://doi.org/10.1371/journal.pgen.1002869.
14. Yoshida K, Saitoh H, Fujisawa S, Kanzaki H, Matsumura H, Tosa Y, Chuma I, Takano Y, Win J, Kamoun S, et al. Association genetics reveals three novel avirulence genes from the rice blast fungal pathogen *Magnaporthe oryzae*. Plant Cell. 2009;21(5):1573–91.
15. Gladieux P, Ravel S, Rieux A, Cros-Arteil S, Adreit H, Milazzo J, Thierry M, Fournier E, Terauchi R, Tharreau D. Coexistence of multiple endemic and pandemic lineages of the rice blast pathogen. MBio. 2018;9(2):e01806-01817. https://doi.org/10.01128/mBio.01806-01817.
16. Liao J, Huang H, Meusnier I, Adreit H, Ducasse A, Bonnot F, Pan L, He X, Kroj T, Fournier E, et al. Pathogen effectors and plant immunity determine specialization of the blast fungus to rice subspecies. Elife. 2016:5. https://doi.org/10.7554/eLife.19377.
17. Zhong Z, Chen M, Lin L, Han Y, Bao J, Tang W, Lin Y, Somai R, Lu L, Zhang W, et al. Population genomic analysis of the rice blast fungus reveals specific events associated with expansion of three main clades. ISME J. 2018;12(8):1867–78.
18. Zhong Z, Justice N, Chen M, Bao J, Lin L, Chen L, Lin Y, Wu X, Cai Z, Zhang Q. Directional selection from host plants is a major force driving host specificity in *Magnaporthe* species. Sc Rep. 2016;6:25591. https://doi.org/10.21038/srep25591.
19. Hu H, Zheng H, Yang T, Chen X, Ye W, Lu G, Lin Z, Wang Z. First report of *Pyricularia* leaf spot on *Pennisetum giganteum* (JUJUNCAO) in China. Plant Dis. 2018. https://doi.org/10.1094/PDIS-05-18-0898-PDN.
20. Lin X, Lin Z, Lin D, Lin H, Luo H, Hu Y, Lin C, Zhu C. Effects of planting *Pennisetum sp*. (Giant Juncao) on soil microbial functional diversity and fertility in the barren hillside. Acta Ecol Sin. 2014;34(15):4304–12.
21. Lin ZX, Lin DM, Su DW, Lin H, Jing L, Zheng D, Yu S-K. Effect of different salt-affected soils on biological characteristics of *Pennisetum sp*. Southwest Chin J Agr Sci. 2015;28(2):675–80.
22. Shi J, Lin ZX, Lin DM, De-Wei SU, Luo HL, Lin XS, Lin ZS, Zheng D, Chen JH, Yao JX. Enzymolysis conditions of *Pennisetum sp*. cellulose. Pratacultural Sci. 2014;31(4):760–5.
23. Prasada R, Goyal JP. A new species of *Pyricularia* on Bajra. Cur Sci. 1970; 39(12):287–8.
24. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods. 2013;10(6):563–9.
25. Muszewska A, Hoffmansommer M, Grynberg M. LTR retrotransposons in Fungi. PLoS One. 2011;6(12):e29425. https://doi.org/10.21371/journal.pone.0029425.
26. Raffaele S, Kamoun S. Genome evolution in filamentous plant pathogens: why bigger can be better. Nat Rev Microbiol. 2012;10(6):417–30.
27. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. 2015;16:157. https://doi.org/10.1186/s13059-13015-10721-13052.
28. Hedges SB. The origin and evolution of model organisms. Nat Rev Genet. 2002;3(11):838–49.
29. Sanderson MJ. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics. 2003; 19(2):301–2.
30. Couch BC, Fudal I, Lebrun MH, Tharreau D, Valent B, van Kim P, Notteghem JL, Kohn LM. Origins of host-specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. Genetics. 2005;170(2):613–30.
31. Luo J, Qiu H, Cai G, Wagner NE, Bhattacharya D, Zhang N. Phylogenomic analysis uncovers the evolutionary history of nutrition and infection mode in rice blast fungus and other Magnaporthales. Sci Rep. 2015;5:9448.
32. Chiapello H, Mallet L, Guerin C, Aguileta G, Amselem J, Kroj T, Ortega-Abboud E, Lebrun MH, Henrissat B, Gendrault A, et al. Deciphering genome content and evolutionary relationships of isolates from the fungus *Magnaporthe oryzae* attacking different host plants. Genome Biol Evol. 2015;7(10):2896–912.
33. Kubicek CP, Starr TL, Glass NL. Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. Annu Rev Phytopathol. 2014;52(1):427–51.
34. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 2011;8(10):785–6.
35. Zhao Z, Liu H, Wang C, Xu JR. Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. BMC Genomics. 2013;14:274. https://doi.org/10.1186/1471-2164-1114-1274.

36. Jeon J, Park SY, Chi MH, Choi J, Park J, Rho HS, Kim S, Goh J, Yoo S, Park JY, et al. Genome-wide functional analysis of pathogenicity genes in the rice blast fungus. Nat Genet. 2007;39(4):561–5.
37. Wu D, Oide S, Zhang N, Choi MY, Turgeon BG. ChLae1 and ChVel1 regulate T-toxin production, virulence, oxidative stress response, and development of the maize pathogen *Cochliobolus heterostrophus*. PLoS Pathog. 2012;8(2): e1002542. https://doi.org/10.1371/journal.ppat.1002542.
38. Mehrabi R, Mirzadi Gohari A, Kema GHJ. Karyotype variability in plant-pathogenic Fungi. Annu Rev Phytopathol. 2017;55:483–503.
39. Faino L, Seidl MF, Shi-Kunne X, Pauper M, van den Berg GC, Wittenberg AH, Thomma BP. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. Genome Res. 2016;26(8):1091–100.
40. de Jonge R, Bolton MD, Kombrink A, van den Berg GC, Yadeta KA, Thomma BP. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. Genome Res. 2013;23(8):1271–82.
41. Hartmann FE, Sanchez-Vallet A, McDonald BA, Croll D. A fungal wheat pathogen evolved host specialization by extensive chromosomal rearrangements. ISME J. 2017;11(5):1189–204.
42. Sprockett DD, Piontkivska H, Blackwood CB. Evolutionary analysis of glycosyl hydrolase family 28 (GH28) suggests lineage-specific expansions in necrotrophic fungal pathogens. Gene. 2011;479(1–2):29–36.
43. Yamashita N, Motoyoshi T, Nishimura A. Purification and characterization of Isoamyl alcohol oxidase ("Mureka"-forming enzyme). Biosci Biotechnol Biochem. 1999;63(7):1216–22.
44. Kern K, Nunn CD, Pichova A, Dickinson JR. Isoamyl alcohol-induced morphological change in *Saccharomyces cerevisiae* involves increases in mitochondria and cell wall chitin content. FEMS Yeast Res. 2004;5(1):43–9.
45. DeSouza L, Shen Y, Bognar AL. Disruption of cytoplasmic and mitochondrial folylpolyglutamate synthetase activity in *Saccharomyces cerevisiae*. Arch Biochem Biophys. 2000;376(2):299–312.
46. Varela JC, Mager WH. Response of *Saccharomyces cerevisiae* to changes in external osmolarity. Microbiology. 1996;142(Pt 4):721–31.
47. Foster AJ, Ryder LS, Kershaw MJ, Talbot NJ. The role of glycerol in the pathogenic lifestyle of the rice blast fungus *Magnaporthe oryzae*. Environ Microbiol. 2017;19(3):1008–16.
48. Mercure EW, Kunoh H, Nicholson RL. Adhesion of *Colletotrichum graminicola* conidia to corn leaves, a requirement for disease development. Physiological & Molecular Plant Pathology. 1995;45(6):407–20.
49. Sela-Buurlage MB, Epstein L, Rodriguez RJ. Adhesion of ungerminated *Colletotrichum musae* conidia. Physiological & Molecular Plant Pathology. 1991;39(5):345–52.
50. Young DH, Kauss H. Adhesion of *Colletotrichum lindemuthianum* spores to Phaseolus vulgaris hypocotyls and to polystyrene. Appl Environ Microbiol. 1984;47(4):616–9.
51. Hamer JE, Howard RJ, Chumley FG, Valent B. A mechanism for surface attachment in spores of a plant pathogenic fungus. Science. 1988;239(4837): 288–90.
52. Thompson BM, Hoelscher BC, Driks A, Stewart GC. Assembly of the BclB glycoprotein into the exosporium and evidence for its role in the formation of the exosporium 'cap' structure in *Bacillus anthracis*. Mol Microbiol. 2012; 86(5):1073–84.
53. Thompson BMWLN, Fox KF. The BdB glycoprotein of *Bacillus anthracis* is involved in exosporium integrity. J Bacteriol. 2007;189(18):6704–13.
54. Wang Y, Jenkins SA, Gu C, Shree A, Martinez-Moczygemba M, Herold J, Botto M, Wetsel RA, Xu Y. *Bacillus anthracis* spore surface protein BclA mediates complement factor H binding to spores and promotes spore persistence. PLoS Pathog. 2016;12(6):e1005678. https://doi.org/10.1371/journal.ppat.1005678.
55. Schulze-Lefert P, Panstruga R. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. Trends Plant Sci. 2011;16(3):117–25.
56. Mistry J, Bateman A, Finn RD. Predicting active site residue annotations in the Pfam database. BMC Bioinformatics. 2007;8:298. https://doi.org/10.1186/1471-2105-1188-1298.
57. Urban M, Cuzick A, Rutherford K, Irvine A, Pedro H, Pant R, Sadanadan V, Khamari L, Billal S, Mohanty S, et al. PHI-base: a new interface and further additions for the multi-species pathogen-host interactions database. Nucleic Acids Res. 2017;45(D1):D604–10.
58. Winnenburg R, Baldwin TK, Urban M, Rawlings C, Kohler J, Hammond-Kosack KE. PHI-base: a new database for pathogen host interactions. Nucleic Acids Res. 2006;34:D459–64.
59. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 2012; 40(Web Server issue):445–51.
60. Tarailo-Graovac M, Chen N. UNIT 4.10Using RepeatMasker to identify repetitive elements in genomic sequences. 2009; Chapter 4(Unit 4):4.10.11–14.10.14. doi: https://doi.org/10.1002/0471250953.bi0410s25.
61. Amselem J, Cuomo CA, van Kan JA, Viaud M, Benito EP, Couloux A, Coutinho PM, de Vries RP, Dyer PS, Fillinger S, et al. Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. PLoS Genet. 2011;7(8):e1002230. https://doi.org/10.1371/journal.pgen.1002230.
62. Gan P, Ikeda K, Irieda H, Narusaka M, O'Connell RJ, Narusaka Y, Takano Y, Kubo Y, Shirasu K. Comparative genomic and transcriptomic analyses reveal the hemibiotrophic stage shift of *Colletotrichum* fungi. New Phytol. 2013; 197(4):1236–49.
63. Cuomo CA, Güldener U, Xu JR, Trail F, Turgeon BG, Pietro AD, Walton JD, Ma LJ, Baker SE, Rep M. The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science. 2007;317(5843):1400–2.
64. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, et al. The genome sequence of the filamentous fungus *Neurospora crassa*. Nature. 2003;422(6934):859–68.
65. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, et al. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nat Biotechnol. 2008;26(5):553–60.
66. Kamper J, Kahmann R, Bolker M, Ma LJ, Brefort T, Saville BJ, Banuett F, Kronstad JW, Gold SE, Muller O, et al. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. Nature. 2006;444(7115): 97–101.
67. Kato H, Yamamoto M, Yamaguchiozaki T, Kadouchi H, Iwamoto Y, Nakayashiki H, Tosa Y, Mayama S, Mori N. Pathogenicity, mating ability and DNA restriction fragment length polymorphisms of *Pyricularia populations* isolated from Gramineae, Bambusideae and Zingiberaceae plants. J Gen Plant Pathol. 2000;66(1):30–47.
68. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013;30(12):2725–9.
69. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001;305(3):567–80.
70. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49. https://doi.org/10.1093/nar/gkr1293.