Research

New Phytologist

Rapid mini-chromosome divergence among fungal isolates causing wheat blast outbreaks in Bangladesh and Zambia

Sanzhen Liu[1]*, Guifang Lin[1]*, Sowmya R. Ramachandran[2], Lidia Calderon Daza[1], Giovana Cruppe[1], Batiseba Tembo[3], Pawan Kumar Singh[4], David Cook[1], Kerry F. Pedley[2] and Barbara Valent[1]

[1]Department of Plant Pathology, Kansas State University, Manhattan, KS, 66506-5502, USA; [2]United States Department of Agriculture–Agricultural Research Service (USDA-ARS), Foreign Disease-Weed Science Research Unit, Ft. Detrick, MD, 21702-9253, USA; [3]Zambia Agricultural Research Institute, Mt. Makulu Central Research Station, Lusaka, 10101, Zambia; [4]International Maize and Wheat Improvement Center (CIMMYT), Texcoco, 56237, Mexico

Authors for correspondence:
Sanzhen Liu
Email: liu3zhen@ksu.edu

Barbara Valent
Email: bvalent@ksu.edu

Received: *20 October 2023*
Accepted: *27 October 2023*

New Phytologist (2024) **241:** 1266–1276
doi: 10.1111/nph.19402

Key words: effector, *Magnaporthe*, outbreak, supernumerary mini-chromosome, wheat blast.

Summary

- The fungal pathogen, *Magnaporthe oryzae Triticum* pathotype, causing wheat blast disease was first identified in South America and recently spread across continents to South Asia and Africa. Here, we studied the genetic relationship among isolates found on the three continents.
- *Magnaporthe oryzae* strains closely related to a South American field isolate B71 were found to have caused the wheat blast outbreaks in South Asia and Africa. Genomic variation among isolates from the three continents was examined using an improved B71 reference genome and whole-genome sequences.
- We found strong evidence to support that the outbreaks in Bangladesh and Zambia were caused by the introductions of genetically separated isolates, although they were all close to B71 and, therefore, collectively referred to as the B71 branch. In addition, B71 branch strains carried at least one supernumerary mini-chromosome. Genome assembly of a Zambian strain revealed that its mini-chromosome was similar to the B71 mini-chromosome but with a high level of structural variation.
- Our findings show that while core genomes of the multiple introductions are highly similar, the mini-chromosomes have undergone marked diversification. The maintenance of the mini-chromosome and rapid genomic changes suggest the mini-chromosomes may serve important virulence or niche adaptation roles under diverse environmental conditions.

Introduction

Magnaporthe oryzae (synonym of *Pyricularia oryzae*) is the ascomycetous fungus that causes blast diseases on a wide variety of grass species, including important grain crops (Gladieux *et al.*, 2018a; Valent *et al.*, 2020; Ristaino *et al.*, 2021). The *M. oryzae Oryza* pathotype (MoO), the lineage causing blast disease on rice (*Oryza sativa*), is reported to have arisen via host shift from a *Setaria* population around the time of rice domestication in China *c.* 7000 yr ago (Couch *et al.*, 2005). MoO subsequently dispersed world-wide through movement in infected seed, either through human movement or seed/grain exchange (Saleh *et al.*, 2014; Gladieux *et al.*, 2018b). By contrast, wheat blast disease, caused by the *M. oryzae Triticum* pathotype (MoT), emerged in Brazil in 1985 and has spread within South America for < 4 decades (Cruz & Valent, 2017; Ceresini *et al.*, 2018). Wheat blast jumped continents for the first time to Bangladesh in South Asia in 2016 (Islam *et al.*, 2016; Malaker *et al.*, 2016), presumably through importation of contaminated wheat seed/grain from South America (Ceresini *et al.*, 2018). The disease is now established in Bangladesh, posing a threat to food security in Asia, including India and China, two major wheat-producing and wheat-consuming countries. In 2018, wheat blast was identified in Zambia, Africa (Tembo *et al.*, 2020), sounding the alarm for world-wide spread of this recently-emerged disease.

Phylogenetic analyses of > 2500 genes and of whole-genome single nucleotide polymorphism (SNP) data showed that MoT isolates from Bangladesh were genetically close to the South American MoT isolate B71 collected in Bolivia in 2012 (Gladieux *et al.*, 2018a). Studies using genotyping data of 84 SNPs indicated that the MoT isolates closely related to B71 were responsible for the wheat blast outbreaks in both Bangladesh and Zambia (Latorre & Burbano, 2021; Win *et al.*, 2021). We previously generated the reference genome sequence for the B71 isolate, which has seven indispensable core chromosomes and one supernumerary dispensable mini-chromosome (Peng *et al.*, 2019). Supernumerary chromosomes, those which are not required for normal physiological growth and development, and which are often nonuniformly present in a population, exist in many plants, animals, and other fungi (Bertazzoni *et al.*, 2018;

*These authors contributed equally to this work.

Soyer *et al.*, 2018; Yang *et al.*, 2020). Supernumerary chromosomes are also called B-chromosomes, dispensable chromosomes, extra chromosomes, or accessory chromosomes. In *M. oryzae*, supernumerary chromosomes are typically smaller than the core chromosomes, consisting of a few megabases of DNA sequences, and they were therefore referred to as mini-chromosomes (Talbot *et al.*, 1993; Orbach *et al.*, 1996; Peng *et al.*, 2019; Langner *et al.*, 2021). The supernumerary chromosome is hypothesized to be an accelerator for fungal adaptive evolution (Coleman *et al.*, 2009; Croll *et al.*, 2013). Compared with core chromosomes, supernumerary mini-chromosomes in *M. oryzae* have a lower gene density and higher levels of repetitive sequences. The repeat-rich mini-chromosome provides abundant intrachromosomal homology for DNA duplication, loss, and rearrangements, which can serve as a reservoir to accelerate fungal genome evolution (Peng *et al.*, 2019; Huang *et al.*, 2023). Indeed, sequence comparisons among mini-chromosomes from different *M. oryzae* strains showed that they are highly variable (Peng *et al.*, 2019; Langner *et al.*, 2021). For Bangladeshi and Zambian isolates that are genetically similar to B71, it is unknown whether they carry mini-chromosomes and, if so, how similar they would be compared with the B71 mini-chromosome.

Strains in the wheat-adapted MoT lineage are still evolving and becoming more aggressive pathogens of wheat (Cruppe *et al.*, 2020; Valent *et al.*, 2021). The MoT population is still in the initial stages of moving from its center of origin into additional wheat regions of the world (Islam *et al.*, 2016; Malaker *et al.*, 2016; Tembo *et al.*, 2020). Understanding and contrasting the evolutionary potential of MoT pathotype strains in the South American center of diversity and in the recently-emerged bottleneck populations in Bangladesh and Zambia is a critical step to understand the movement and adaptation of wheat blast. In this study, we analyzed both SNPs and genomic structural variation among isolates from Bangladesh, Zambia, and South America. Furthermore, Oxford Nanopore long-read sequencing and assembly were conducted for a Zambian isolate and used for the genome comparison between the Zambian strain and B71. Our results show that all the analyzed isolates contained at least one mini-chromosome, and we document the high-speed creation of genetic variability within the mini-chromosomes from closely related isolates.

Materials and Methods

Genome assembly of B71

Nanopore long-read data generated previously (He *et al.*, 2020) were recalled using Basecaller Guppy (v.3.4.4) with default parameters (http://nanoporetech.com/community). Reads were assembled with Canu (v.1.9) with parameters (genomeSize = 45 m minReadLength = 5000 minOverlapLength = 1000 corOut Coverage = 80). The previous assembly version (B71Ref1.6) was used to add missing chromosome ends (He *et al.*, 2020). A contig containing ribosomal DNA sequences was added to the beginning of chromosome 1. The resulting assembly was polished using Nanopolish with Nanopore raw FAST5 data and Pilon with Illumina reads (SRA accession: SRR6232156; Walker *et al.*, 2014; Loman *et al.*, 2015).

Public Illumina sequencing data

Publicly available data of MoT strains were downloaded from Sequence Read Archive (SRA). Accession numbers are available in Supporting Information Table S1. These data were used for constructing phylogenetic trees and examining genomic copy number variation among strains.

Illumina whole-genome sequencing

All MoT strains are maintained in frozen storage and manipulated under Biosafety Level 3 (BSL3) laboratory conditions in the USDA-ARS Foreign Disease and Weed Science Research Unit (FDWSRU) in Fort Detrick, MD, USA, and in the Biosecurity Research Institute (BRI) at Kansas State University in Manhattan, KS, USA, as authorized by PPQ permits from the USDA-Animal and Plant Health Inspection Services. *Magnaporthe* isolates were cultured on oatmeal agar (OMA) plates seeded with dried Whatman filter paper containing mycelia and conidia (Valent *et al.*, 1986). The culture plates were incubated under continuous light at room temperature for 7–12 d. Liquid cultures were seeded with pieces of OMA culture, total area 1 cm^2, from the actively growing region of the plate into complete medium (3 g Casamino acids, 3 g yeast extract, and 6 g sucrose in 1 l H_2O) at 24°C in a stationary flask. After 7 d, mycelial mats from each sample were collected, blotted dry on paper towels, lyophilized for 24 h, and stored at room temperature. DNA extraction was performed by finely grinding 200 mg of lyophilized mycelium in liquid nitrogen. Five hundred microliters of extraction buffer (1% CTAB, 0.7 M NaCl, 100 mM Tris (pH 7.5), 10 mM EDTA, 1% 2-Mercaptoethanol, 0.3 mg ml^{-1} Proteinase K) was added and mixed thoroughly by shaking and tube inversion. The mixture was incubated at 65°C for 30 mins and then allowed to cool to room temperature. One milliliter of phenol : chloroform : isoamyl alcohol (25 : 24 : 1) was added and mixed by shaking and inverting tubes. The tubes were centrifuged for 10 min at 13 000 ***g*** in a benchtop centrifuge. The aqueous phase was removed, and the DNA was precipitated by mixing in 0.54 volumes of room temperature isopropanol, followed by centrifugation at 13 000 ***g*** for 10 min. The pellets were rinsed with 70% ethanol and then allowed to air dry. Pellets were dissolved in 50 ml of TE buffer containing 1 mg ml^{-1} RNase. Paired-end sequencing was performed on the isolates using the Illumina platform at Novogene (Sacramento, CA, USA).

Nanopore WGS for *de novo* genome assemblies of ZM2-1

Genomic DNA was extracted from the mycelial powder using a modified CTAB method (He, 2000). To filter small-size DNA fragments, BluePippin Gel Cassette (Cat. #BLF7510; Sage Science, Beverly, MA, USA) was used to perform a > 20 kb High-Pass size selection. Around 1 µg of the selected DNA was used to build a library using SQK-LSK110 sequencing kit (Cat.

#SQK-LSK110; Oxford Nanopore, UK). Twelve microliters of the DNA library were loaded to a R9.4.1 flow cell (Cat. #FLO-106D; Oxford Nanopore), and whole-genome sequencing (WGS) was performed on a MinION Sequence Device (Cat. #MinION Mk1B; Oxford Nanopore, Oxford, UK). The Nanopore raw data stored in FAST5 format were converted to FASTQ using Guppy Basecaller (v.6.0.1, https://community.nanoporetech.com). Genomes were assembled using Canu (v.2.2, https://github.com/marbl/canu) (Koren *et al.*, 2017) with the parameters of 'minReadLength=5000 minOverlapLength=1500 corOutCoverage=80 correctedErrorRate=0.1'. The contigs in Canu assemblies were aligned to B71Ref2 to determine the chromosome number and the orientation using NUCmer (v.4.0.0, https://mummer4.github.io/) with the parameters of '-L10000 -I 90' (Marçais *et al.*, 2018). The resulting assembly was polished using Nanopolish (v.0.13.3) for two runs and Pillon (v.1.24) for two runs (Walker *et al.*, 2014; Loman *et al.*, 2015). For the strain ZM2-1 Nanopore sequencing, we noticed the contamination of *Neurospora crassa* based on the result from data analysis. Therefore, contigs matching the *N. crassa* reference genome (NC_026501.1) were removed after Nanopolishing.

Estimation of base errors in genome assemblies

Base errors were estimated based on software KAD (v.0.1.7) with the parameter of '--minc 5 --klen 35' (He *et al.*, 2020). The script running KAD is available on GitHub (https://github.com/PlantG3/B71branch).

Annotation of transposable elements

Extensive *de novo* TE Annotator (EDTA, v.2.0.0) was used for transposable element annotation (Ou *et al.*, 2019). The script running EDTA is available on GitHub (https://github.com/PlantG3/B71branch).

Genome annotation

Pipeline funannotate (v.1.8.8) was used for genome annotation. RNA-Seq data from both plate culture and *in planta* infection were used as expression evidence (SRA accessions: SRR9126640, SRR9127597–SRR9127602). Protein data included MG8 protein annotation from rice isolate 70–15 (Dean *et al.*, 2005), UniProtKB/Swiss-Prot protein database released in March 2021, and an effector collection (https://raw.githubusercontent.com/liu3zhenlab/collected_data/master/Magnaporthe/known.effectors.db01.fasta).

Contour-clamped homogeneous electric field karyotyping of ZM2-1

Contour-clamped homogeneous electric field (CHEF) karyotyping was conducted per the procedure established previously (Peng *et al.*, 2019). Briefly ZM2-1 protoplasts were mixed with 1.5% low melting-temperature agarose, followed by loading into plug molds and lysis with proteinase K. A CHEF Mapper XA System was used for the core- and mini-chromosome separation on the CHEF gel of 1% agarose in 0.5 × TBE. The CHEF electrophoresis was carried out using the initial switch interval of 120 s and the end switch interval of 3600 s. The electrophoresis was run at the voltage of 2 V cm^{-1} for 96 h at the temperature of 14°C, and with the angle parameter at 120°.

SNP identification

Paired-end Illumina sequencing reads were trimmed with Trimmomatic (v.0.38; Bolger *et al.*, 2014). The resulting paired-end reads were aligned to the reference (B71Ref2) with BWA (v.0.7.17-r1188; Li & Durbin, 2010). Alignments were filtered to retain confident alignments with at least 60-base pairs (bp) matches, > 95% identity, and > 95% coverage per read. GATK (v.4.1.0.0) was used for SNP discovery and filtering (McKenna *et al.*, 2010). We only retained biallelic SNP sites with the B71 isolate matching the B71 reference genome.

Construction of phylogenetic trees

Phylogenetic trees were constructed using IQ-TREE (v.1.6.12; Nguyen *et al.*, 2015). Briefly, two steps were performed. The first step identified the optimized model, which was then used in the second step for tree construction. Two phylogenetic trees were constructed, including a larger tree of 57 MoT strains and a smaller tree of MoT strains in the B71 branch and an outgroup strain. The minor allele frequency of each SNP of at least 0.1 and the missing data of at most 20% were required for each input SNP. In total, 141 081 and 1416 SNPs were used to construct the larger and the smaller trees, respectively. The optimized models for both phylogenetic trees were the same, namely 'TVM + F + ASC + R3'. The final constructed trees were visualized in iTOL (Letunic & Bork, 2019).

Analysis of copy number variation via CGRD

Pipeline CGRD (v.0.3.5) was used to find large copy number variation (e.g. presence/absence variation and duplication) using Illumina data with the parameters (--knum --adj0 --cleanup --groupval '-5 -0.4 0.4 0.8'; Lin *et al.*, 2021). The script running CGRD is available on GitHub (https://github.com/PlantG3/B71branch).

Analysis of structural variation via Syri

Genome sequences of B71 and ZM2-1 were aligned with NUCmer (v.4.0.0) with the parameters (--maxmatch -c 500 -b 500 -l 20; Marçais *et al.*, 2018). Alignments were filtered with 'delta-filter' in NUCmer with the parameters (-m -i 90 -l 500). Filtered alignments were then input to SyRI (v.1.5) for structural variation analysis (Goel *et al.*, 2019). The in-house R module 'mgplot' was employed for displaying chromosomal alignments and large structural variation.

Read coverages of *PWL2* and *BAS1*

Strains that are in the B71 branch and have at least 8 million pairs of reads were used for examining read coverages of *PWL2* and *BAS1*. T25, which has no mini-chromosomes and did not contain either of these two genes (Peng *et al.*, 2019), was used as the control. Briefly, trimmed Illumina reads were aligned to each gene using BWA (v.0.7.17-r1188; Li & Durbin, 2010). Alignments with at least 60 match, 98% identity, and 98% coverage were retained, and read coverage per base pair position was determined and plotted.

Results

Establishment of a finished reference genome of South American isolate B71

The previous reference B71 genome (B71Ref1), assembled from PacBio long reads, contained 22 gaps in core chromosomes and five scaffolds from the mini-chromosome (Peng *et al.*, 2019). To generate a complete genome assembly, we produced >250× Oxford Nanopore long reads with an N50 of 29 kb and the longest read of 166 kb (Fig. S1). The resulting assembly, produced using the CANU assembler and polishing with both Nanopore reads and Illumina reads, resulted in telomere-to-telomere assemblies of chromosomes 2–7, chromosome 1 with telomere repeat sequence on one end, a telomere-to-telomere mini-chromosome, and a circularized mitochondrial genome (Fig. 1; Table S2; Walker *et al.*, 2014; Koren *et al.*, 2017). Telomeric retrotransposons MoTeRs with various copy numbers were found in all telomeric regions, including on mini-chromosomes (Starnes *et al.*, 2012). Notably, one MoTeR insertion was identified *c.* 759 kb away from the end of chromosome 3. The unfinished assembly at the beginning of chromosome 1 is presumably due to the high copy number of ribosomal DNA repeats (rDNA). The accuracy of assembly sequences is estimated to be >99.99% (Table S3; He *et al.*, 2020). The length of the finished mini-chromosome is 1.9 Mb, consisting of 51.7% repeats. Core chromosomes contain less repetitive sequences and different core chromosomes vary in their repetitive sequence content. Chromosomes 2, 4, and 5 contain repeats ranging from 3.7% to 4.5%, while chromosomes 1, 3, 6, and 7 contain higher levels of repeat content, ranging from 9.8% to 17.9% (Table S2). Centromere sequences can be inferred from all chromosomes, including the mini-chromosome (Fig. 1; Yadav *et al.*, 2019). Genome annotation with the updated reference genome results in 11 864 protein-coding genes, which contain homologs of 25 known effector genes (Dataset S1).

The B71 branch contains isolates from Bangladesh, Zambia, Bolivia, and Brazil

We collected publicly available WGS data of MoT strains, including Bangladeshi isolates collected from 2016 to 2020, Zambian isolates from 2018 to 2020, and isolates from South America. Phylogenetic analysis confirmed that all Bangladeshi and Zambian isolates cluster closely with B71 (Fig. 2a; Latorre & Burbano, 2021; Win *et al.*, 2021). We also identified a Brazilian strain, 12.1.181, collected in 2012 (Castroagudín *et al.*, 2016), in the same phylogenetic clade as B71 and isolates from Bangladesh and Zambia (Fig. 2a). Here, the clade is referred to as the B71 branch. Within the clade, Bangladeshi and Zambian isolates were separately clustered (Figs 2b, S2–S5). The Brazilian strain 12.1.181 is more similar to Bangladeshi isolates and the Bolivian strain B71 is more similar to Zambian isolates. These results indicate that the wheat blast outbreaks on the Asian and African continents were likely caused by different sub-branches of isolates. We identified 41 SNPs with all Bangladeshi isolates sharing one genotype, and 12.1.181 and B71 sharing another genotype (Dataset S2). None of the Zambian strains harbored these Bangladeshi SNP genotypes, indicating that Zambian isolates spread from South America rather than from Bangladesh.

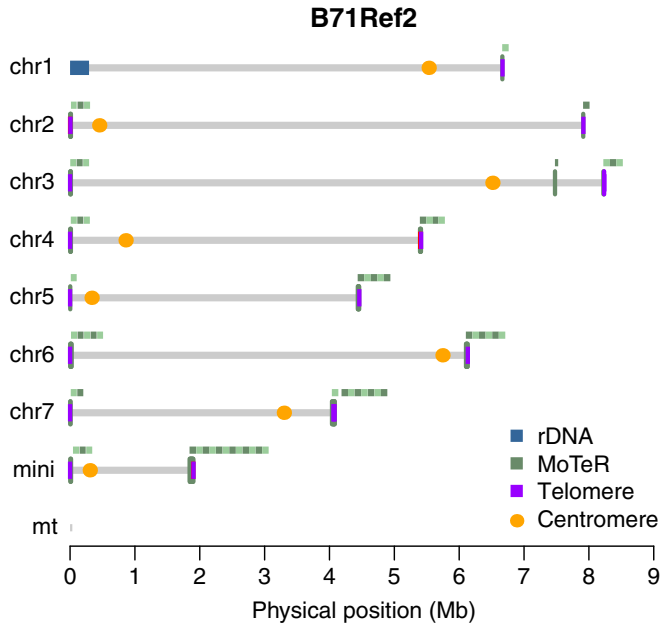

Fig. 1 Chromosomal ends and centromeres of B71Ref2. The finished B71 reference genome (B71Ref2) consists of seven core chromosomes, one supernumerary mini-chromosome, and a circularized mitochondrial genome. Positions of ribosomal DNA (rDNA) repeats, transposable element MoTeR repeats, telomere repeats, and inferred centromeres are highlighted. The close-up of each MoTeR repeat locus is shown above the locus with two green colors to indicate the copy number of MoTeR repeats.

Genomic structural variation among B71-related isolates

We performed Comparative Genomics Read Depth (CGRD) analysis against the new B71 assembly to estimate copy number variation using WGS read depth data (Lin *et al.*, 2021). Given the relatively low sequence divergence between the isolates, as measured by SNPs, we were interested in assessing if copy number variation, including presence/absence variation and duplications, had taken place among the B71-related isolates. A few sequence segments that were present in B71 but absent in some other isolates were observed. This included a 155.5-kb segment at the end of chromosome 3 (8082–8238 kb) and a 22.6-kb

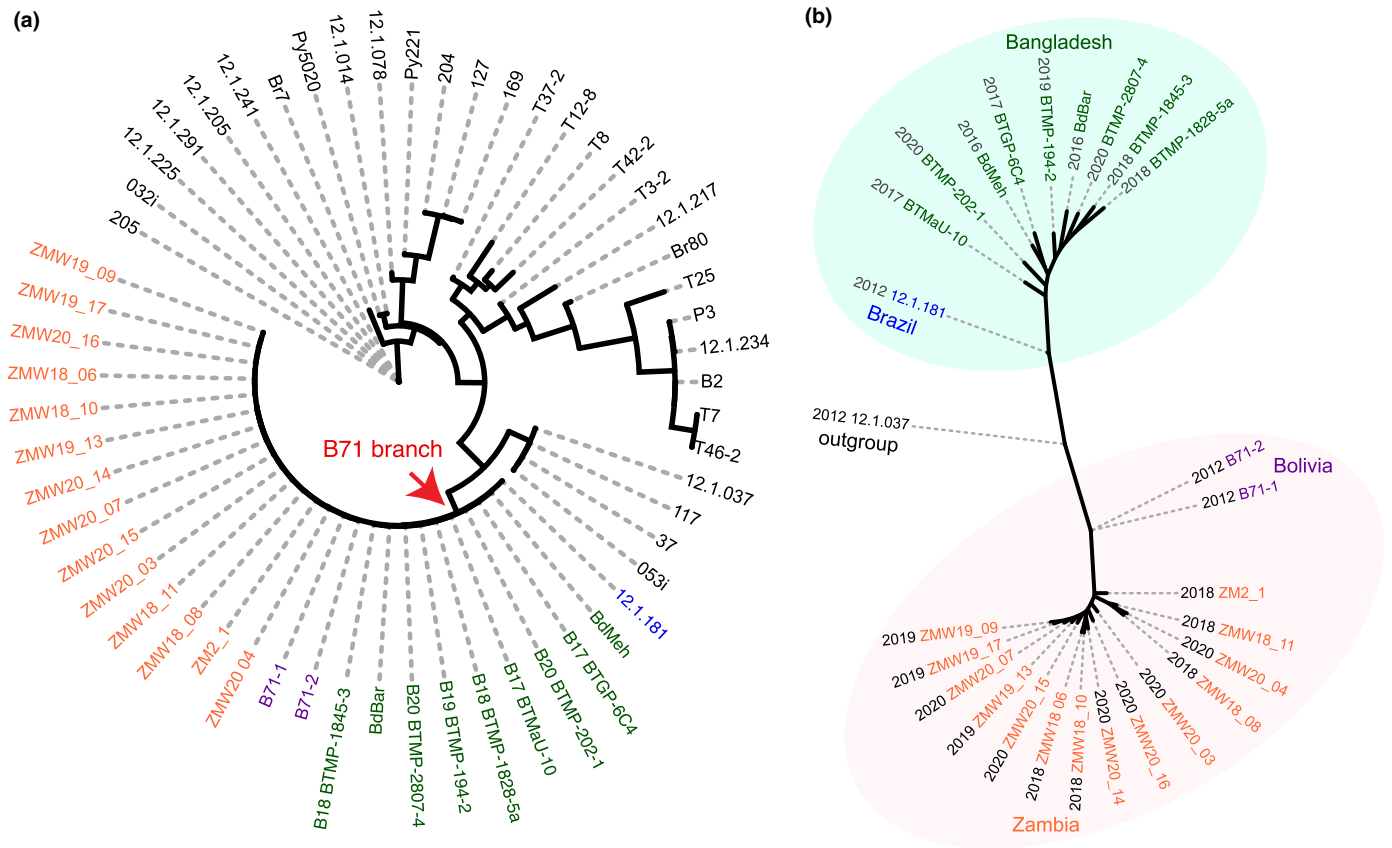

Fig. 2 Phylogenetic tree of *Magnaporthe oryzae Triticum* (MoT) isolates. Publicly available whole-genome sequencing data of MoT isolates were used to identify single nucleotide polymorphisms (SNPs) for the phylogeny construction using the maximum likelihood approach. (a) Scaled tree of MoT isolates. B71-1 and B71-2 represent two independent samples of B71. (b) Scaled tree of isolates from the B71 branch. Brazilian isolate 12.1.037 served as the outgroup isolate. The year and country isolated for each isolate are labeled.

segment at the beginning of chromosome 4 (16–22.6 kb) that were both absent in all Bangladeshi isolates and the Brazilian isolate 12.1.181 (Fig. 3a,b). Both segments were present in all Zambian isolates. The results indicate the identified presence/absence variants were standing variation in South America before the Bangladesh introduction. The results also support the hypothesis that Zambian isolates are from a different sub-branch than those causing the Bangladesh outbreak.

In contrast to core chromosomes, pervasive structural variation on the mini-chromosome was observed among isolates in the B71 branch (Fig. 3c). All analyzed isolates from the branch contained most sequences of the B71 mini-chromosome, indicating that all these isolates carry a mini-chromosome similar to the one originally described in B71. As a control, we mapped reads from isolate 12.1.037, which was collected from Brazil, but is more genetically distinct from B71. As expected, we found that 12.1.037 has little sequence homology to the B71 mini-chromosome (Fig. 3c). A few sequence regions of the B71 mini-chromosome were absent in the Bangladeshi isolates and 12.1.181 (Fig. 3c). Two regions of the B71 mini-chromosome, *c.* 6 and 86 kb at mini-chromosome positions 1.173–1.179 and 1.364–1.450 Mb, respectively, were repeatedly found to be absent in genomes of multiple Bangladeshi isolates. The former region is also absent in the 12.1.181 genome. Only one region, from *c.* 100 to 159 kb, was absent in two of 14 Zambian isolates, namely ZM2-1 and ZMW20_04. CGRD analysis indicates that all Bangladeshi isolates carry one mini-chromosome, while Zambian isolates carry one or multiple mini-chromosomes (Fig. 3c). More than one mini-chromosome was observed in Zambian strains isolated from Mpika in 2018, from Mount Makulu in 2019, and from multiple areas in 2020 (Fig. 3c). The isolates containing either a single mini-chromosome or multiple mini-chromosomes were not clustered separately in the phylogenetic tree using whole-genome SNP markers (Fig. 2b).

Nanopore long reads data of 13/14 Zambian strains are publicly available, among which three strains (ZMW18_06, ZMW19_09, and ZMW20_04) had sufficient data for *de novo* genome assembly. We assembled genomes of these three strains and identified a contig that was highly similar to the B71 mini-chromosome in each draft genome assembly, which was referred to as the mini-chromosome contig. Note that, with the result from the CGRD analysis, ZMW18_06 and ZMW19_09 were evidenced to have > 1 mini-chromosome, while ZMW20_04 was indicated to have one mini-chromosome (Fig. 3c). Analysis of read depth supporting each contig found that each of the mini-chromosome contigs of ZMW18_06 and ZMW19_09 had roughly twice depths as compared with contigs from core chromosomes. By contrast, the mini-chromosome contig of ZMW20_04 had a similar depth

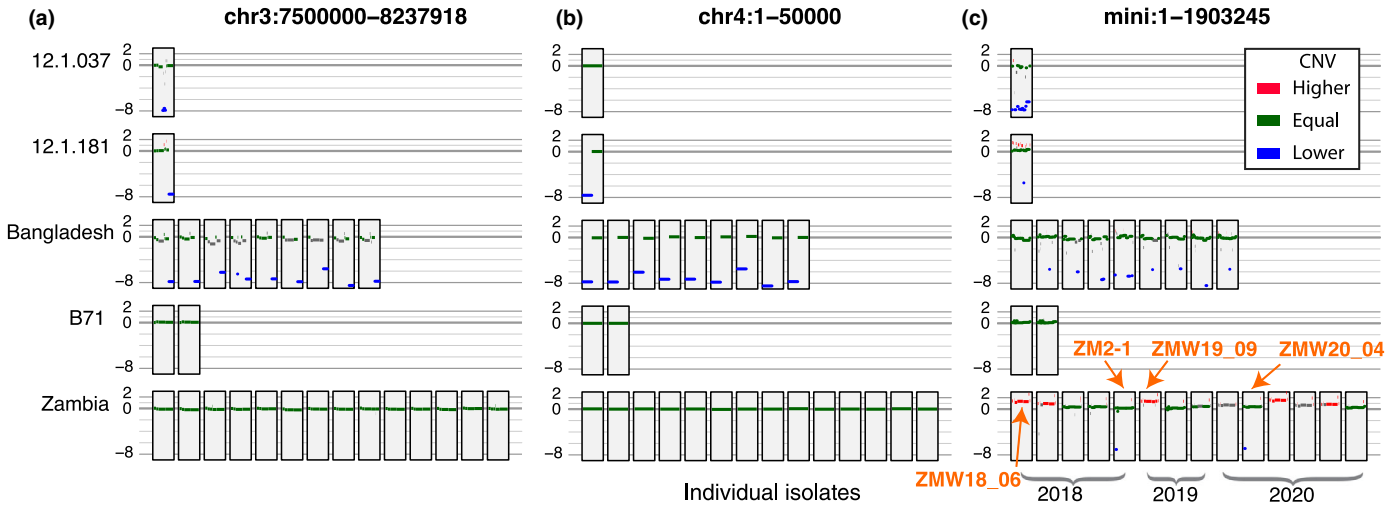

Fig. 3 Regional results of Comparative Genomics Read Depth (CGRD). Each block represents the CGRD result at a region indicated on the top of each plot for an *Magnaporthe oryzae* isolate. The CGRD result in each block was from the comparison between whole-genome sequencing data of an isolate and B71. Results from two small regions on chromosome 3 (a) and chromosome 4 (b), as well as the whole mini-chromosome (c) were displayed. The Y-axis values represent $\log_2$ values of ratios of read depths of an isolate to B71, signifying copy number variation (CNV). A value > 0.8, colored as red, indicates a higher copy number in the isolate as compared to B71. A value < −5, colored as blue, indicates the absence of the B71 sequence in the isolate. The value between −0.4 and 0.4, colored as green, is deemed to be an equal copy number between the two genomes. Other values are colored as gray to indicate ungrouped regions. Years for Zambian strains are labeled in (c). Orange arrows point at strains subjected to *de novo* genome assembly.

to contigs from core chromosomes (Fig. S6). This result, along with the CGRD analysis, strongly supported that some Zambian strains carried multiple near-identical mini-chromosomes, which were collapsed to single contigs in the genome assemblies.

Rapid sequence divergence in the mini-chromosome of Zambian isolates

To further examine genomic variation in Zambian isolates, we conducted Oxford Nanopore long-read sequencing of a Zambian isolate, ZM2-1. The isolate was collected from the Mpika district of Muchinga Province, Zambia in 2018, which was previously termed as Zambia2.1 (ZM2-1; Tembo *et al.*, 2020). The assembly resulted in a nearly finished assembly, consisting of *c.* 45 Mb with seven core chromosomes and a mini-chromosome that was a telomere-to-telomere sequence (Fig. S7; Table S4).

Core chromosomes of the ZM2-1 genome are highly similar to the B71 core chromosomes with a few large structural variants (Fig. 4). The presence of the mini-chromosome was confirmed by CHEF electrophoresis (Fig. S8). On chromosome 2 of ZM2-1, a 170-kb region at *c.* 677–847-kb and a 100-kb region at the chromosomal end were absent in B71 (Fig. 4; Table S5). In addition, an inversion of > 65 kb was found at *c.* 3.3 Mb on chromosome 6 as compared to B71 (Table S6). The comparison confirmed that ZM2-1 had no deletions at the end of chromosome 3 or at the beginning of chromosome 4 identified in Bangladesh strains and 12.1.181 via CGRD. Overall, we estimated that structural variants impacted < 1% of core chromosomes and largely occurred in gene-poor regions, possibly mediated by repetitive sequences.

Consistent with our analysis using short reads of Zambian strains, long-read genome assemblies of ZM2-1 have one mini-chromosome that is similar to the B71 mini-chromosome (Fig. 3). The B71-like mini-chromosome of ZM2-1 showed the absence of a large region (DEL_1, from 120.2 to 192.8 kb) of B71 and multiple duplications (Fig. 4; Table S5). DEL_1 largely overlapped with the ZM2-1 deletion on the mini-chromosome identified in CGRD (Fig. 3b) and the boundary of DEL_1 was expected to be more accurate due to the limitation of short reads used in CGRD. Four inversion events were identified, indicative of frequent inversion occurring within the B71 branch (Fig. 4; Table S6). Overall, structural variants consisted of > 15% of the mini-chromosome. Although large structural variation exists between mini-chromosomes of B71 and ZM2-1, both mini-chromosomes contained two effector genes *BAS1* and *PWL2* in a syntenic region (Fig. 5). The *BAS1* and *PWL2* segment was also identified on the mini-chromosome contigs in the draft genome assemblies of ZMW18_06, ZMW19_09, and ZMW20_04 (Fig. 5). Using whole-genome short-read sequencing data of isolates with sufficient sequencing depths, reads with a nearly full coverage on both *BAS1* and *PWL2* were identified from all examined isolates of the B71 branch (Fig. S9).

Discussion

The wheat blast population that has been expanding in South America since before its first discovery in 1985 is genetically complex. Indeed, a recent analysis showed that individual MoT genomes are composed of admixture contributions of chromosome segments derived from five different host-adapted populations of *M. oryzae* (Rahnama *et al.*, 2023). Rahnama *et al.* suggest the present-day population in South America is structured according to chromosomal haplotypes that were defined when various progeny inherited chromosome segments from different donors through sexual recombination. This is in sharp contrast to the simple population structures in Bangladesh and Zambia established by

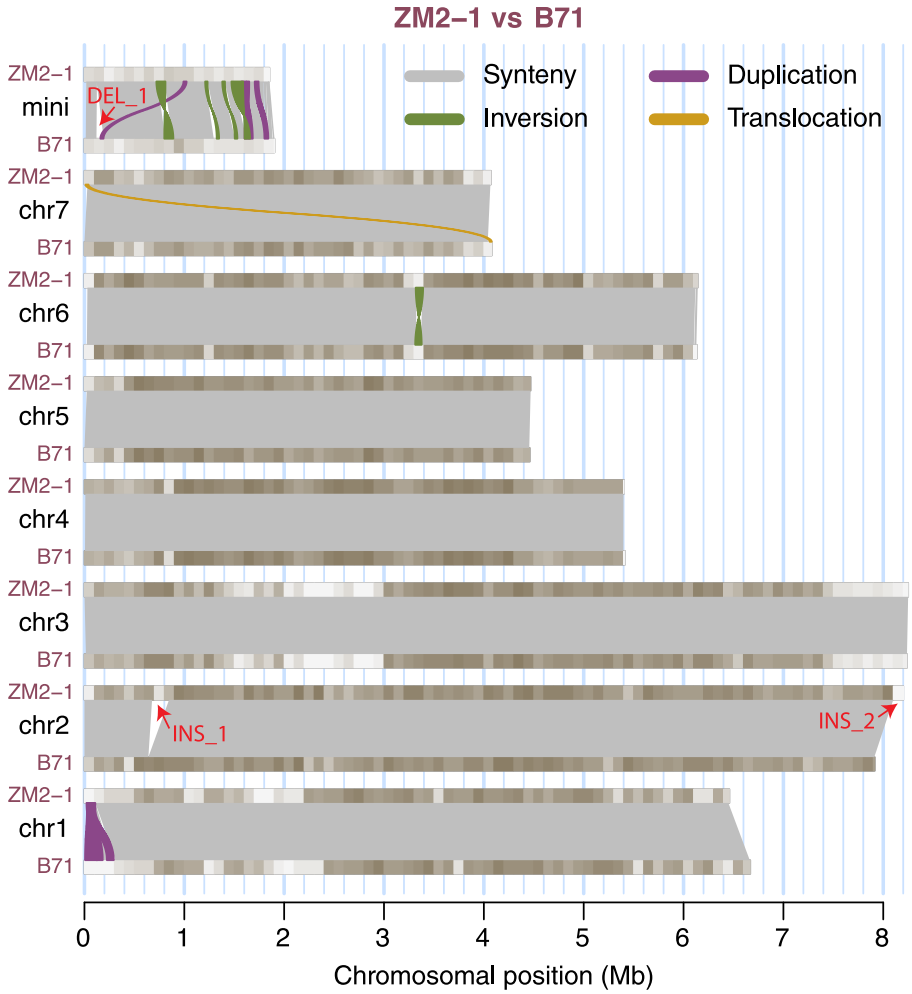

Fig. 4 Genomic comparison between ZM2-1 and B71. Chromosomal comparisons based on Syri analysis of structural variation. Inversion, duplication, and translocation events with at least 20 kb are color-coded. The gradient colors from tan to light gray on chromosomes signify gene density from high to low. Red arrows point at large insertions (INS_1 and INS_2) and a large deletion (DEL_1). The duplication events at the beginning of chromosome 1, the ribosome DNA repetitive locus, were probably artifacts due to incomplete assembly of the ribosome DNA repeats.

recent introductions of strains that are highly similar to aggressive MoT strain B71 (Malaker *et al.*, 2016; Gladieux *et al.*, 2018a; Latorre & Burbano, 2021; Win *et al.*, 2021). Here, we examined genomic variation among isolates of the B71 branch using an improved B71 reference genome. We constructed phylogenetic trees using SNPs and performed structural variation analyses using WGS Illumina reads and genome assemblies with Nanopore long reads. Our results support that different sub-branches from South America were independently introduced to Bangladesh and Zambia. Genome sequencing data indicated that at least one mini-chromosome was maintained in all the analyzed isolates, providing the opportunity for examination of genomic variation of both core- and mini-chromosomes that diversified in a short evolutionary period.

Isolates of different sub-branches caused outbreaks in Bangladesh and Zambia

Phylogenetic analysis using SNP genotyping data showed that Bangladeshi and Zambian isolates are closely related to MoT isolates B71 from Bolivia and 12.1.181 from Brazil. The phylogeny showed that all Bangladeshi isolates are closer to 12.1.181 and all Zambian isolates are closer to B71, indicative of separate sub-branches causing the outbreaks in Bangladesh and Zambia. We identified the 41 SNPs with a genotype conserved in all Bangladeshi strains, which were isolated from multiple cities in multiple years, and consistently different in all Zambian strains examined. The SNPs were located on all chromosomes, providing evidence that the introduction of wheat blast into Zambia occurred from South America rather than from Bangladesh. This finding is consistent with the conclusions from a recent independent analysis (Latorre *et al.*, 2023). The hypothesis of independent wheat blast introductions to the two continents from South America was further supported by presence/absence polymorphisms on chromosomes 3 and 4. Importantly, for these two loci, both presence and absence genotypes existed in strains from South America, implying that the diversification of the ancestors of Bangladeshi and Zambian isolates occurred in South America. In conclusion, our genomic analyses strongly support that isolates from different sub-branches from South America caused the first outbreaks of wheat blast disease in Bangladesh and Zambia.

The mini-chromosome has been maintained, but is rapidly diversifying

Our results indicate that all field isolates from the B71 branch carry at least one mini-chromosome, and some Zambian

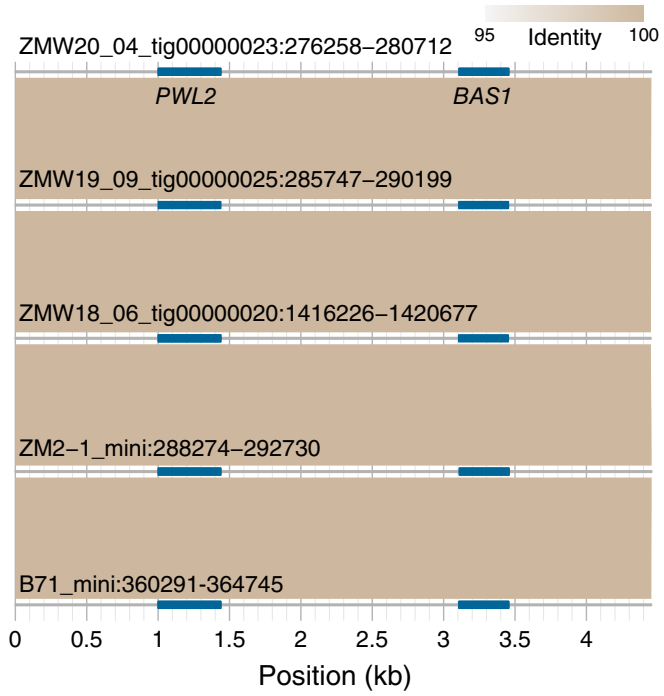

Fig. 5 Syntenic alignments of the regions containing *PWL2* and *BAS1*. Both effector genes *PWL2* and *BAS1* are present in B71 and Zambian strains subjected to *de novo* genome assembly. The regions harboring these two genes are nearly identical. Blue color bars signify the locations of the two genes and the bisque color indicates alignments between regions whose coordinates on the B71 mini-chromosome are listed.

strains carry more than one. The maintenance of the mini-chromosome in all examined isolates indicates that depletion of the mini-chromosome rarely occurred and/or selection pressure existed for retaining the mini-chromosome. The mini-chromosome harbors >50% repetitive sequences, including a high proportion of retrotransposons and DNA transposons, while core chromosomes contain *c.* 10% repetitive sequences and a higher density of genes. What selection could be acted upon to maintain these otherwise dispensable sequences remains to be determined. We observed pervasive large structural variation among mini-chromosomes from different isolates in the same branch, in contrast to the relatively infrequent structural changes found between core chromosomes. Although there is extensive structural variation, some genes in mini-chromosomes appear to be retained in all strains. Whole-genome sequencing data showed that all examined strains of the B71 branch contain sequences of the effector genes *BAS1* and *PWL2* (Sweigard *et al.*, 1995; Mosquera *et al.*, 2009; Peng *et al.*, 2019), indicating the genomic region containing the two genes is relatively stable despite so far only being localized to highly dynamic mini-chromosomes in the MoT population. Additionally, *BAS1* and *PWL2* co-occur in 6% of wheat isolates collected between 1986 and 1989, and in 91% of wheat isolates collected in 2017 and 2018, indicating that the overall frequency of wheat field isolates containing these two genes has increased through time (Navia-Urrutia *et al.*, 2022). These results suggest that *BAS1* and *PWL2* might be playing a role in the enhanced aggressiveness of more recent MoT strains towards wheat. Due to the presence of *BAS1* and *PWL2* on mini-chromosomes, their presence in all examined strains implies that the frequency of mini-chromosomes in field isolates has increased through time.

Multiple mini-chromosomes were found in individual Zambian strains isolated in multiple years (2018, 2019, and 2020). Our analysis did not provide evidence that Zambian strains with >1 mini-chromosome have increased in frequency, but this result is limited due to the limited number of Zambian strains analyzed to date. It is intriguing to observe mini-chromosome variation occurring in a natural population in a matter of years. While copy number variation is a well-documented means to increase biotic and abiotic stress tolerance (Cook *et al.*, 2012; Maron *et al.*, 2013), most examples are linked to copy number changes in the core genome. To our knowledge, there are few examples of rapid sequence divergence of noncore genomic compartments in natural populations. Recent analysis of naturally occurring herbicide-resistant *Amaranthus palmeri*, an economically important weed impacting crop production, found significant expansion of the coding sequence for 5-enolpyruvylshikimate-3-phosphate synthase (*EPSPS*) occurring on extrachromosomal circular DNA (eccDNA; Koo *et al.*, 2018). The importance of accessory chromosomes is clear in plant pathogenic fungi, where they can be transferred and affect the host range and/or habitat diversity of individual strains of *Fusarium* spp. (Coleman *et al.*, 2009; Ma *et al.*, 2010). The biological impact of the documented mini-chromosome divergence described here remains to be determined. Likewise, it is not clear whether there is a genetic or environmental factor contributing to the mini-chromosome divergence found in strains collected in Zambia, or whether this is a matter of ascertainment bias and represents a common phenomenon in *M. oryzae*. Analysis of additional strains from this pandemic branch in coming years will help further address whether the mini-chromosome can be depleted in field strains and how multiple mini-chromosomes in one strain arose.

Acknowledgements

We thank Dr Mark Farman from the Department of Plant Pathology at the University of Kentucky for valuable comments. SRR was supported in part by an appointment to the Agricultural Research Service (ARS) Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the US Department of Energy (DOE) and the US Department of Agriculture (USDA). Funding was provided by the USDA NIFA award (2018-67013-28511) to SL, the USDA NIFA award (2021-68013-33719) to BV, the USDA NIFA award (2021-67013-35724) to SL, BV, DC, and D. Koo, the NSF award (1741090 and 2311738) to SL, and NSF awards (2011500) to SL, BV, DC, and D. Koo. This research was supported by USDA-ARS projects 8044-22000-046-00D and 8044-22000-051-00D. This is contribution no. 22-309-J from the Kansas Agricultural Experiment Station, Manhattan, Kansas.

Competing interests

USDA is an equal opportunity provider and employer. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. SL is the co-founder of Data2Bio, LLC. Other authors claim no competing interests.

Author contributions

KFP, BV and SL conceptualized experiments. BT and PKS collected and provided fungal isolates. GL, SRR, GC and LCD conducted experiments. GL, DC, SRR and SL analyzed data. SL, GL, SRR, DC, KFP and BV wrote the manuscript. All authors reviewed and revised the manuscript. SL and GL contributed equally to this work.

ORCID

David Cook https://orcid.org/0000-0002-2719-4701
Sanzhen Liu https://orcid.org/0000-0002-9513-855X
Kerry F. Pedley https://orcid.org/0000-0002-6827-8231
Barbara Valent https://orcid.org/0000-0002-5088-3345

Data availability

Genomic sequencing and assembly data of B71 and ZM2-1 have been deposited in the Sequence Read Archive (SRA) database under accessions PRJNA355407 and PRJNA849580, respectively. Related scripts and data are available on GitHub (https://github.com/PlantG3/B71branch).

References

Bertazzoni S, Williams AH, Jones DA, Syme RA, Tan K-C, Hane JK. 2018. Accessories make the outfit: accessory chromosomes and other dispensable DNA regions in plant-pathogenic fungi. *Molecular Plant–Microbe Interactions* **31**: 779–788.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Castroagudín VL, Moreira SI, Pereira DAS, Moreira SS, Brunner PC, Maciel JLN, Crous PW, McDonald BA, Alves E, Ceresini PC. 2016. *Pyricularia graminis-tritici*, a new *Pyricularia* species causing wheat blast. *Persoonia* **37**: 199–216.

Ceresini PC, Castroagudín VL, Rodrigues FÁ, Rios JA, Eduardo Aucique-Pérez C, Moreira SI, Alves E, Croll D, Maciel JLN. 2018. Wheat blast: past, present, and future. *Annual Review of Phytopathology* **56**: 427–456.

Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, Schmutz J, Taga M, White GJ, Zhou S *et al.* 2009. The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. *PLoS Genetics* **5**: e1000618.

Cook DE, Lee TG, Guo X, Melito S, Wang K, Bayless AM, Wang J, Hughes TJ, Willis DK, Clemente TE *et al.* 2012. Copy number variation of multiple genes at Rhg1 mediates nematode resistance in soybean. *Science* **338**: 1206–1209.

Couch BC, Fudal I, Lebrun M-H, Tharreau D, Valent B, van Kim P, Nottéghem J-L, Kohn LM. 2005. 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* **170**: 613–630.

Croll D, Zala M, McDonald BA. 2013. Breakage-fusion-bridge cycles and large insertions contribute to the rapid evolution of accessory chromosomes in a fungal pathogen. *PLoS Genetics* **9**: e1003567.

Cruppe G, Cruz CD, Peterson GL, Pedley KF, Mohammad A, Fritz A, Calderon L, da Silva CL, Todd T, Kuhnem P *et al.* 2020. Novel sources of wheat head blast resistance in modern breeding lines and wheat wild relatives. *Plant Disease* **104**: 35–43.

Cruz CD, Valent B. 2017. Wheat blast disease: danger on the move. *Tropical Plant Pathology* **42**: 210–222.

Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu J-R, Pan H *et al.* 2005. The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* **434**: 980–986.

Gladieux P, Condon B, Ravel S, Soanes D, Maciel JLN, Nhani A Jr, Chen L, Terauchi R, Lebrun M-H, Tharreau D *et al.* 2018a. Gene flow between divergent cereal- and grass-specific lineages of the rice blast fungus *Magnaporthe oryzae*. *mBio* **9**: e01219.

Gladieux P, Ravel S, Rieux A, Cros-Arteil S, Adreit H, Milazzo J, Thierry M, Fournier E, Terauchi R, Tharreau D. 2018b. Coexistence of multiple endemic and pandemic lineages of the rice blast pathogen. *mBio* **9**: 1110.

Goel M, Sun H, Jiao W-B, Schneeberger K. 2019. SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome Biology* **20**: 277.

He C, Lin G, Wei H, Tang H, White FF, Valent B, Liu S. 2020. Factorial estimating assembly base errors using *k*-mer abundance difference (KAD) between short reads and genome assembled sequences. *NAR Genomics and Bioinformatics* **2**: lqaa075.

He YQ. 2000. An improved protocol for fungal DNA preparation. *Mycosystem* **19**: 434.

Huang J, Liu S, Cook DE. 2023. Dynamic genomes – mechanisms and consequences of genomic diversity impacting plant–fungal interactions. *Physiological and Molecular Plant Pathology* **125**: 102006.

Islam MT, Croll D, Gladieux P, Soanes DM, Persoons A, Bhattacharjee P, Hossain MS, Gupta DR, Rahman MM, Mahboob MG *et al.* 2016. Emergence of wheat blast in Bangladesh was caused by a South American lineage of *Magnaporthe oryzae*. *BMC Biology* **14**: 84.

Koo D-H, Molin WT, Saski CA, Jiang J, Putta K, Jugulam M, Friebe B, Gill BS. 2018. Extrachromosomal circular DNA-based amplification and transmission of herbicide resistance in crop weed *Amaranthus palmeri*. *Proceedings of the National Academy of Sciences, USA* **115**: 3332–3337.

Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Research* **27**: 722–736.

Langner T, Harant A, Gomez-Luciano LB, Shrestha RK, Malmgren A, Latorre SM, Burbano HA, Win J, Kamoun S. 2021. Genomic rearrangements generate hypervariable mini-chromosomes in host-specific isolates of the blast fungus. *PLoS Genetics* **17**: e1009386.

Latorre SM, Burbano HA. 2021. The emergence of wheat blast in Zambia and Bangladesh was caused by the same genetic lineage of *Magnaporthe oryzae*. *Zenodo*. doi: 10.5281/zenodo.4619405.

Latorre SM, Were VM, Foster AJ, Langner T, Malmgren A, Harant A, Asuke S, Reyes-Avila S, Gupta DR, Jensen C *et al.* 2023. Genomic surveillance uncovers a pandemic clonal lineage of the wheat blast fungus. *PLoS Biology* **21**: e3002052.

Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v.4: recent updates and new developments. *Nucleic Acids Research* **47**: W256–W259.

Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**: 589–595.

Lin G, He C, Zheng J, Koo D-H, Le H, Zheng H, Tamang TM, Lin J, Liu Y, Zhao M *et al.* 2021. Chromosome-level genome assembly of a regenerable maize inbred line A188. *Genome Biology* **22**: 175.

Loman NJ, Quick J, Simpson JT. 2015. A complete bacterial genome assembled *de novo* using only nanopore sequencing data. *Nature Methods* **12**: 733–735.

Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B *et al.* 2010.

Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* **464**: 367–373.

Malaker PK, Barma NC, Tiwary TP, Collis WJ, Duveiller EP, Singh K, Joshi AK, Singh RP, Braun H-J, Peterson GL *et al.* 2016. First report of wheat blast caused by *Magnaporthe oryzae* pathotype *Triticum* in Bangladesh. *Plant Disease* **100**: 2330.

Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018. MUMmer4: a fast and versatile genome alignment system. *PLoS Computational Biology* **14**: e1005944.

Maron LG, Guimarães CT, Kirst M, Albert PS, Birchler JA, Bradbury PJ, Buckler ES, Coluccio AE, Danilova TV, Kudrna D *et al.* 2013. Aluminum tolerance in maize is associated with higher MATE1 gene copy number. *Proceedings of the National Academy of Sciences, USA* **110**: 5241–5246.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M *et al.* 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**: 1297–1303.

Mosquera G, Giraldo MC, Khang CH, Coughlan S, Valent B. 2009. Interaction transcriptome analysis identifies *Magnaporthe oryzae* BAS1-4 as biotrophy-associated secreted proteins in rice blast disease. *Plant Cell* **21**: 1273–1290.

Navia-Urrutia M, Mosquera G, Ellsworth R, Farman M, Trick HN, Valent B. 2022. Effector genes in *Magnaporthe oryzae Triticum* as potential targets for incorporating blast resistance in wheat. *Plant Disease* **106**: 1700–1712.

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **32**: 268–274.

Orbach MJ, Chumley FG, Valent B. 1996. Electrophoretic karyotypes of *Magnaporthe grisea* pathogens of diverse grasses. *Molecular Plant–Microbe Interactions* **9**: 261–271.

Ou S, Su W, Liao Y, Chougule K, Agda JRA, Hellinga AJ, Lugo CSB, Elliott TA, Ware D, Peterson T *et al.* 2019. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biology* **20**: 275.

Peng Z, Oliveira-Garcia E, Lin G, Hu Y, Dalby M, Migeon P, Tang H, Farman M, Cook D, White FF *et al.* 2019. Effector gene reshuffling involves dispensable mini-chromosomes in the wheat blast fungus. *PLoS Genetics* **15**: e1008272.

Rahnama M, Condon B, Ascari JP, Dupuis JR, Del Ponte EM, Pedley KF, Martinez S, Valent B, Farman ML. 2023. Recent co-evolution of two pandemic plant diseases in a multi-hybrid swarm. *Nature Ecology & Evolution*. doi: 10.1038/s41559-023-02237-z

Ristaino JB, Anderson PK, Bebber DP, Brauman KA, Cunniffe NJ, Fedoroff NV, Finegold C, Garrett KA, Gilligan CA, Jones CM *et al.* 2021. The persistent threat of emerging plant disease pandemics to global food security. *Proceedings of the National Academy of Sciences, USA* **118**: e2022239118.

Saleh D, Milazzo J, Adreit H, Fournier E, Tharreau D. 2014. South-East Asia is the center of origin, diversity and dispersion of the rice blast fungus, *Magnaporthe oryzae*. *New Phytologist* **201**: 1440–1456.

Soyer JL, Balesdent M-H, Rouxel T, Dean RA. 2018. To B or not to B: a tale of unorthodox chromosomes. *Current Opinion in Microbiology* **46**: 50–57.

Starnes JH, Thornbury DW, Novikova OS, Rehmeyer CJ, Farman ML. 2012. Telomere-targeted retrotransposons in the rice blast fungus *Magnaporthe oryzae*: agents of telomere instability. *Genetics* **191**: 389–406.

Sweigard JA, Carroll AM, Kang S, Farrall L, Chumley FG, Valent B. 1995. Identification, cloning, and characterization of PWL2, a gene for host species specificity in the rice blast fungus. *Plant Cell* **7**: 1221–1233.

Talbot NJ, Salch YP, Ma M, Hamer JE. 1993. Karyotypic variation within clonal lineages of the rice blast fungus, *Magnaporthe grisea*. *Applied and Environmental Microbiology* **59**: 585–593.

Tembo B, Mulenga RM, Sichilima S, M'siska KK, Mwale M, Chikoti PC, Singh PK, He X, Pedley KF, Peterson GL *et al.* 2020. Detection and characterization of fungus (*Magnaporthe oryzae* pathotype *Triticum*) causing wheat blast disease on rain-fed grown wheat (*Triticum aestivum* L.) in Zambia. *PLoS ONE* **15**: e0238724.

Valent B, Cruppe G, Stack JP, Cruz CD, Farman ML, Paul PA, Peterson GL, Pedley KF. 2021. Recovery plan for wheat blast caused by *Magnaporthe oryzae* pathotype *Triticum*. *Plant Health Progress* **22**: 182–212.

Valent B, Singh PK, He X, Farman M, Tosa Y, Braun HJ. 2020. CHAPTER 13: blast diseases: evolution and challenges of a staple food crop fungal pathogen. In: Ristaino JB, Records A, eds. *Epidemiology. emerging plant diseases and global food security*. St. Paul, MN, USA: The American Phytopathological Society, 267–292.

Valent B, Crawford MS, Weaver CG, Chumley FG. 1986. Genetic studies of fertility and pathogenicity in *Magnaporthe grisea* (*Pyricularia oryzae*). *Iowa State Journal of Research* **60**: 569–594.

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK *et al.* 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **9**: e112963.

Win J, Malmgren A, Langner T, Kamoun S. 2021. A pandemic clonal lineage of the wheat blast fungus. *Zenodo*. doi: 10.5281/zendo.4618522.

Yadav V, Yang F, Reza MH, Liu S, Valent B, Sanyal K, Naqvi NI. 2019. Cellular dynamics and genomic identity of centromeres in cereal blast fungus. *mBio* **10**: e01581.

Yang H, Yu H, Ma L-J. 2020. Accessory chromosomes in *Fusarium oxysporum*. *Phytopathology* **110**: 1488–1496.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Dataset S1 List of known effectors in B71.

Dataset S2 Single nucleotide polymorphisms that Bangladesh isolates carrying a unique genotype.

Fig. S1 Histogram of lengths of Nanopore raw reads.

Fig. S2 Unscaled phylogenetic tree of isolates of the B71 branch.

Fig. S3 Numbers of whole-genome SNPs from pairwise comparisons between strains.

Fig. S4 Numbers of single nucleotide polymorphisms on core chromosomes from pairwise comparisons between strains.

Fig. S5 Numbers of single nucleotide polymorphisms on the mini-chromosome from pairwise comparisons between strains.

Fig. S6 Read depth of contigs in three genome assemblies.

Fig. S7 Genomic features of ZM2-1.

Fig. S8 CHEF gel of ZM2-1.

Fig. S9 Read coverages on *PWL2* and *BAS1*.

Table S1 Sequence read archive accessions of publicly available whole-genome sequencing data.

Table S2 Chromosome lengths and repeat contents of B71.

Table S3 Base accuracy of B71Ref2.

Table S4 Chromosome lengths and repeat contents of ZM2-1.

Table S5 Insertion or deletion events with > 40 kb in ZM2-1 as compared with B71.

Table S6 Large inversions between B71 and ZM2-1 strains.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.