



# The Evolution of Mini-chromosomes in the Fungal Genus *Colletotrichum*

Haoming Wang<sup>1</sup>, Rong Huang<sup>2</sup>, Jingyi Ren<sup>1</sup>, Lihua Tang<sup>2</sup>, Suiping Huang<sup>2</sup>, Xiaolin Chen<sup>2</sup>, Jun Fan<sup>1</sup>, Qinhu Wang<sup>1</sup>, Tom Hsiang<sup>3</sup>, Huiquan Liu<sup>1</sup>, Qili Li<sup>2\*</sup>

<sup>1</sup> State Key Laboratory of Crop Stress Biology for Arid Areas, College of Plant Protection, Northwest A&F University, Yangling, Shaanxi, China

<sup>2</sup> Institute of Plant Protection, Guangxi Academy of Agricultural Sciences, Key Laboratory of Green Prevention and Control on Fruits and Vegetables

in South China Ministry of Agriculture and Rural Affairs and Guangxi Key Laboratory of Biology for Crop Diseases and Insect Pests, Nanning, Guangxi, China

<sup>3</sup> School of Environmental Sciences, University of Guelph, Guelph, Ontario, Canada

\* E-mail: liqili@gxaas.net

**Abstract:** Anthracnose diseases caused by *Colletotrichum* species are among the most common fungal diseases. Genome sequencing of several species shows the presence of mini-chromosomes. These are thought to contribute to virulence, but their formation and activity remain to be fully elucidated. Here, we assembled 17 *Colletotrichum* genomes (16 isolated from mango plus one from persimmon) through PacBio long-read sequencing. Half of the assembled scaffolds had telomeric repeats at both ends indicating full-length chromosomes. Based on comparative genomics analysis at interspecies and intraspecies levels, we identified extensive chromosomal rearrangements events. We analyzed mini-chromosomes of *Colletotrichum* spp. and found large variation among close relatives. In *C. fructicola*, homology between core chromosomes and mini-chromosomes suggested that some mini-chromosomes were generated by recombination of core chromosomes. In *C. musae* GZ23-3, we found 26 horizontally transferred genes arranged in clusters on mini-chromosomes. In *C. asianum* FJ11-1, several potential pathogenesisrelated genes on mini-chromosomes were upregulated, especially in strains with highly pathogenic phenotypes. Mutants of these upregulated genes showed obvious defects in virulence. Our findings provide insights into the evolution and potential relationships to virulence associated with mini-chromosomes.

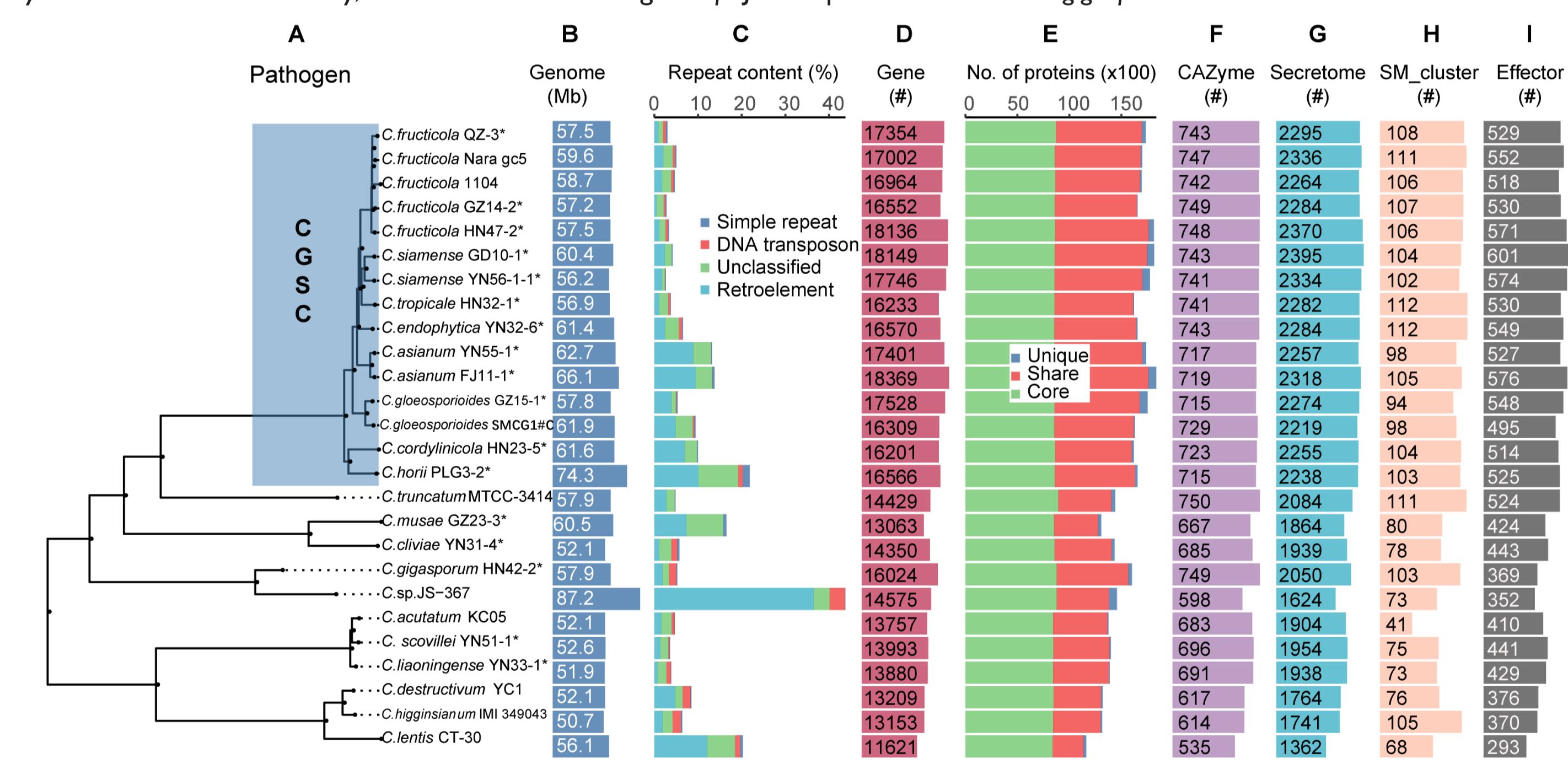
## 1 Seventeen high-quality *Colletotrichum* genomes

After genome assembly of the 17 strains, the numbers of scaffolds ranged from 9 to 21 and average scaffold N50 was higher than 4.39 Mb. These strains had an average of 49% scaffolds with telomere repeats "TTAGGG" at both ends. According to previous studies, the number of chromosomes of *Colletotrichum* is usually 10–13, which is close to the number of scaffolds assembled in this study for each genome, and hence the assemblies of these genomes are considered chromosomal level.

Strain	Species	Host	Collection location	Species complex
YN55-1	<i>C. asianum</i>	<i>Mangifera indica</i>	Yunnan Province	<i>gloeosporioides</i>
FJ11-1	<i>C. asianum</i>	<i>Mangifera indica</i>	Fujian Province	<i>gloeosporioides</i>
HN47-2	<i>C. fructicola</i>	<i>Mangifera indica</i>	Hainan Province	<i>gloeosporioides</i>
QZ-3	<i>C. fructicola</i>	<i>Mangifera indica</i>	Guangxi Province	<i>gloeosporioides</i>
GD10-1	<i>C. siamense</i>	<i>Mangifera indica</i>	Guangdong Province	<i>gloeosporioides</i>
YN56-1-1	<i>C. siamense</i>	<i>Mangifera indica</i>	Yunnan Province	<i>gloeosporioides</i>
GZ15-1	<i>C. gloeosporioides</i>	<i>Mangifera indica</i>	Guizhou Province	<i>gloeosporioides</i>
GZ14-2	<i>C. fructicola</i>	<i>Mangifera indica</i>	Guizhou Province	<i>gloeosporioides</i>
HN32-1	<i>C. tropicale</i>	<i>Mangifera indica</i>	Hainan Province	<i>gloeosporioides</i>
HN42-2	<i>C. gigasporum</i>	<i>Mangifera indica</i>	Hainan Province	<i>gigasporum</i>
YN31-4	<i>C. clavigae</i>	<i>Mangifera indica</i>	Yunnan Province	<i>orchidearum</i>
YN32-6	<i>C. endophytica</i>	<i>Mangifera indica</i>	Yunnan Province	<i>gloeosporioides</i>
YN33-1	<i>C. liaoingense</i>	<i>Mangifera indica</i>	Yunnan Province	<i>magnum</i>
YN51-1	<i>C. scovillei</i>	<i>Mangifera indica</i>	Yunnan Province	<i>acutatum</i>
HN23-5	<i>C. cordylinicola</i>	<i>Mangifera indica</i>	Hainan Province	<i>gloeosporioides</i>
GZ23-3	<i>C. musae</i>	<i>Mangifera indica</i>	Guizhou Province	<i>gloeosporioides</i>
PLG3-2	<i>C. horii</i>	<i>Diospyros</i> sp.	Guangxi Province	<i>gloeosporioides</i>

## 2 Genome and gene features

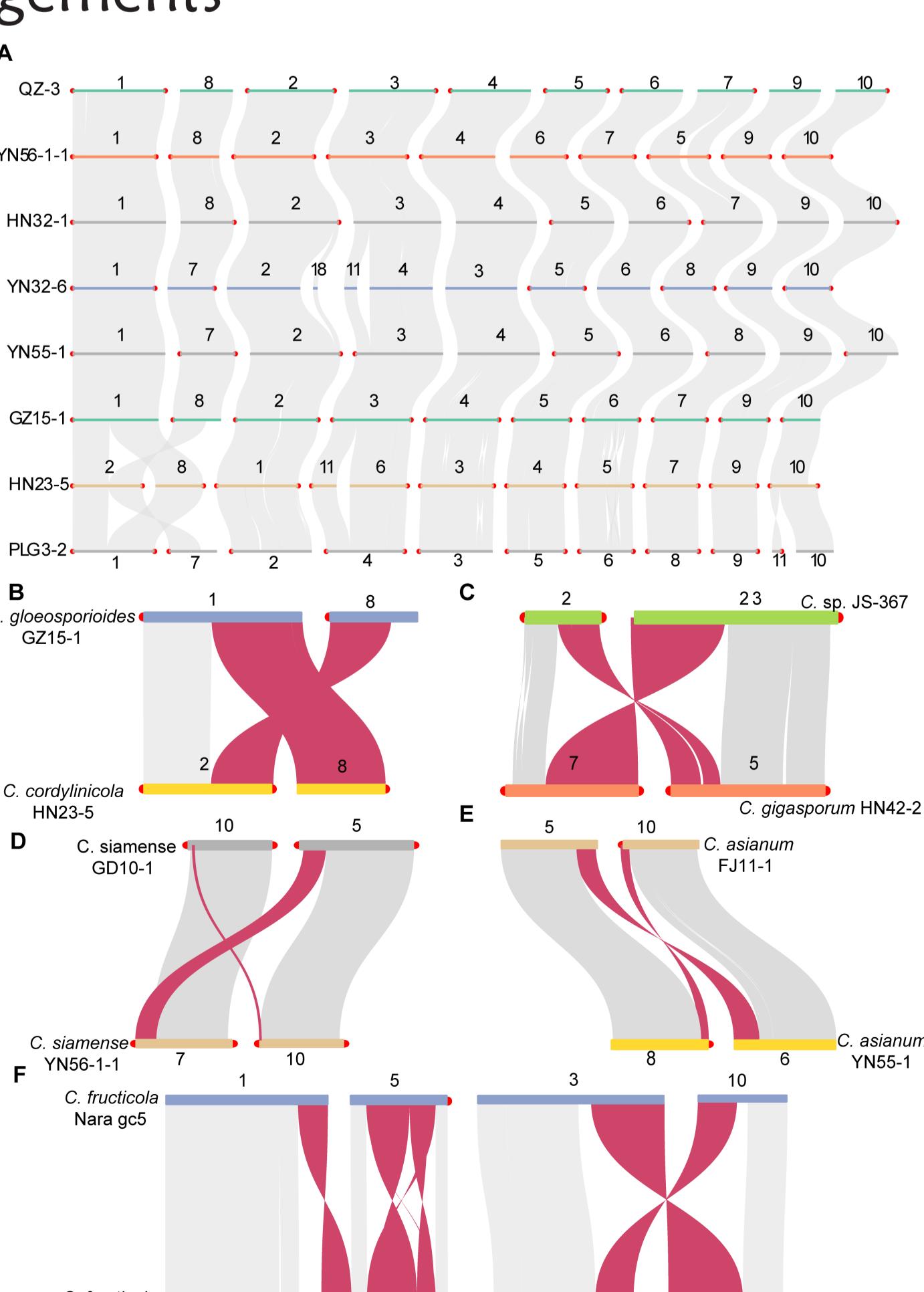
To better understand the characteristics of genomes of *Colletotrichum*, we obtained sequences of an additional nine genomes from the NCBI genome database. In total, 26 genomes of *Colletotrichum* strains (19 species) were used for comparison. The average size of the nuclear genome sequence of these 26 genomes was 59.04 Mb. Among them, *C. sp. JS-367* had the largest genome at 87.2 Mb and *C. higginsianum* had the smallest genome at 50.7 Mb. These results indicate that genomes of different *Colletotrichum* strains may vary greatly. We performed phylogenetic analysis of these 26 strains based on all 399,364 genes belonging to 21,943 orthologous. Previous studies have shown that *C. higginsianum*, *C. destructivum*, and *C. lenti* belong to the destructivum species complex. *C. scovillei* YN51-1 belongs to the same clade as *C. acutatum*. *C. liaoingense* YN33-1 belongs to the magnum species complex which is closer to *C. acutatum* species complex. *C. truncatum* MTCC-3414 belongs to the truncatum species complex. The precise identity of *C. sp. JS-367*, an endophyte obtained from mulberry, is still unknown. Although *C. sp. JS-367* placed closest to *C. gigasporum* HN42-2.



## 3 Extensive chromosomal rearrangements

Syntenic analysis of the genomes of the CGSC showed that 10 core chromosomes were found in different species. We observed extensive chromosomal rearrangement events on core chromosomes. For example, in strain HN23-5 (*C. cordylinicola*), chromosomes #2 and #8 have been rearranged and contain portions homologous to chromosomes #1 and #8 of *C. gloeosporioides* GZ15-1 strain. Outside of the CGSC, chromosomes #2 and #23 of *C. sp. JS-367* showed homology with chromosomes #5 and #7 of HN42-2 (*C. gigasporum*). Furthermore, a chromosomal fusion event occurred in the genome of YN31-4 (*C. clavigae*), which resulted in only nine core chromosomes.

We also performed intraspecies genomic synteny analysis, and results showed that chromosomal rearrangement events were observed in *C. siamense*, *C. asianum*, and *C. fructicola*. For example, in *C. siamense*, chromosomes #10 and #5 of GD10-1 showed homology to chromosomes #7 and #10 of YN56-1-1. In *C. asianum*, chromosomes #5 and #10 of FJ11-1 showed homology with chromosomes #6 and #8 of YN55-1. In *C. fructicola*, chromosomal rearrangements occurred more frequently than for other species, especially for strain Nara gc5 which showed many rearrangements compared to other strains of this species.



## 4 The mini-chromosomes are widespread and highly specific

In addition to the conserved core chromosomes, CGSC also had some mini-chromosomes (also known as accessory chromosomes). The numbers of these mini-chromosomes ranged from three to eight per genome. Many mini-chromosomes had telomeric repeats at both ends, indicating that they were complete chromosomes. NonCGSC strains also had mini-chromosomes, except for YN31-4, JS-367, HN42-2, YN33-1, or MTCC-3414. The strains with the most mini-chromosomes were YN32-6 (*C. endophytica*) and HN23-5 (*C. cordylinicola*), both of which had eight mini-chromosomes.

We then compared the GC content and sequence similarity of mini and core chromosomes. We found that the mini-chromosomes had more repetitive sequences, with a lower proportion of GC content and fewer conserved sequences. The functions of most genes found on mini-chromosomes are still unknown. dN/dS analysis showed that these genes were under strong selection pressure, indicating that genes on mini-chromosomes evolved more rapidly than genes on core chromosomes.

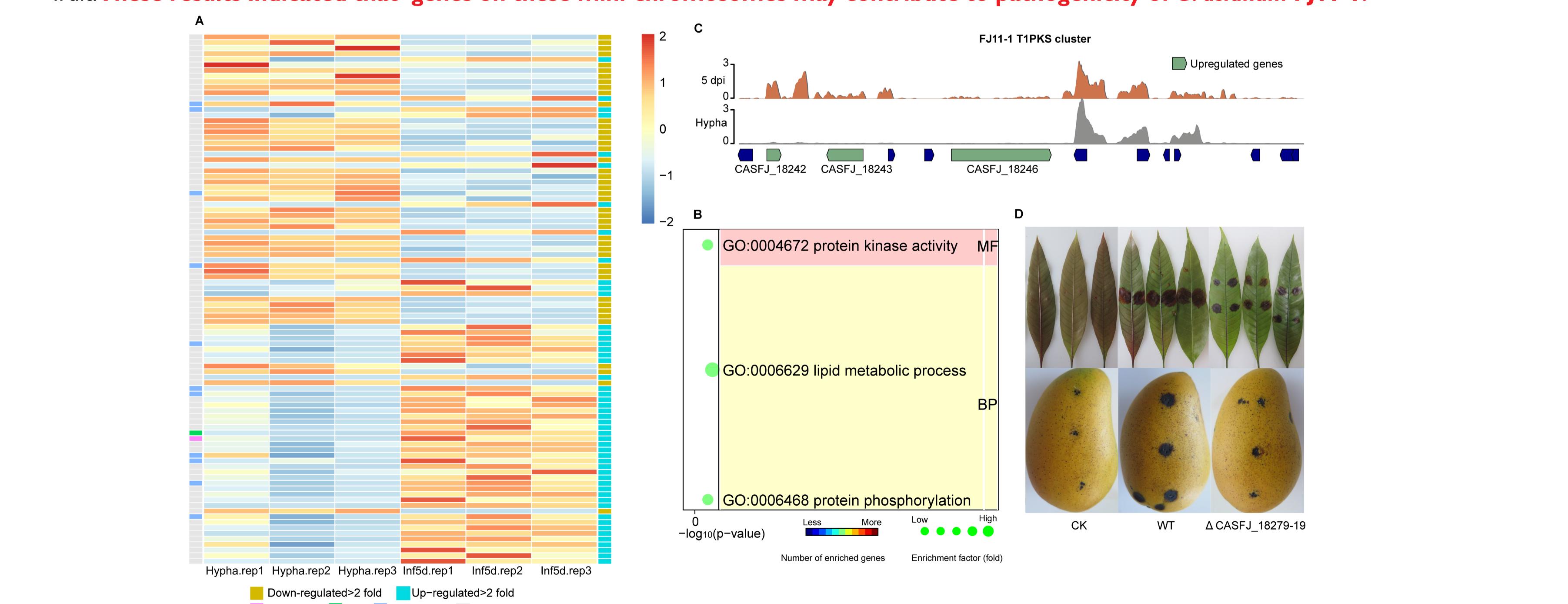
## 5 Generation of mini-chromosomes

We conducted intraspecies genomic synteny analysis and found a sequence of approximately 590 kb on the mini-chromosomes #11 of *C. fructicola* QZ-3 to be highly similar to a region of core chromosome #3 of *C. fructicola* Nara gc5. This region consisted of 184 genes. To determine their origin, we aligned this sequence against the core chromosomes of every strain and found that the region from the mini-chromosome had the highest similarity (99.4%) to positions 3,548,799–4,135,098 of chromosome #5 of QZ-3 itself. This implied that the mini-chromosome fragment was derived from a core chromosome in the same genome. In QZ-3, a short fragment near the telomeric repeat of mini-chromosome #11 was similar to a fragment of core chromosome #11 of Nara gc5. This fragment happened to be missing in core chromosome #6 of QZ-3. Mini-chromosome #11 of QZ-3 was assembled in a complete sequence without any breakpoints. These results provided evidence that mini-chromosome #11 of QZ-3 was generated by the recombination of its own core chromosomes #5 and #6.

We found that 162 genes were most similar to genes in non-*Colletotrichum* species, thus suspected they were obtained by horizontal gene transfer (HGT). *C. musae* GZ23-3 had the largest number of potentially HGT genes on its mini-chromosomes, with 37 genes, 26 of which matched those from bacterial genomes, and the rest showed homology to other fungal genomes. Further analysis of the 26 genes in GZ23-3 showing homology to bacterial genes indicated that they showed high similarity to bacterial genes related to ATP binding, and were arranged in clusters on mini-chromosomes #12 and #13. Phylogenetic analysis showed that two examples (CMUGZ\_14714 and CMUGZ\_13716) were shared by the bacterial order Rhizobiales and were most similar to genes of Bradyrhizobium. Potentially HGT genes showed different codon usage preferences compared to genes of *Colletotrichum*.

## 6 Pathogenesis-related genes on mini-chromosomes are upregulated during infection

We conducted RNA-seq analysis of six dominant strains causing mango anthracnose (*C. asianum* FJ11-1, *C. asianum* YN55-1-1, *C. siamense* GD10-1, *C. siamense* YN56-1-1, *C. fructicola* HN47-2, and *C. gloeosporioides* GZ15-1) and analyzed their gene expression profile. Compared to the hyphal stage, an average of 2,375 genes were upregulated at 5 days after infection in each strain. Among these upregulated genes, an average of 15 genes were located on mini-chromosomes. Previous study showed that strains FJ11-1 and GD10-1 were more aggressive than their closest relative such as YN55-1 and YN56-1-1 in this study. Compared to isolates YN55-1 and YN56-1-1, FJ11-1 and GD10-1 had more potential pathogenic genes located on mini-chromosomes. Among upregulated genes located on mini-chromosomes of FJ11-1, eight genes encoded secreted proteins and one was an SM gene. GO enrichment analysis showed that upregulated genes were related to protein kinase activity and phosphorylation which are crucial for infection by pathogenic fungi. We further investigated the expression of TIPKS (type I polyketide synthase) genes on FJ11-1 mini-chromosome #16, which were highly expressed 5 days after inoculation. Among them, CASFJ\_18242, CASFJ\_18243, and CASFJ\_18246 were upregulated during the infection stage. The TIPKS cluster is specific to strain FJ11-1. To validate the function of genes on mini-chromosomes, we knocked out gene CASFJ\_18279 which was upregulated during the infection stage and was identified only on mini-chromosome of FJ11-1. All three mutants ( $\Delta$ CASFJ\_18279-19,  $\Delta$ CASFJ\_18279-22, and  $\Delta$ CASFJ\_18279-33) showed reduced aggressiveness compared to wild-type on both mango leaves and fruit. These results indicated that genes on these mini-chromosomes may contribute to pathogenicity of *C. asianum* FJ11-1.



**Acknowledgment:** The work was supported by the National Natural Science Foundation of China (Grant no. 32160622), Guangxi Key Laboratory of Biology for Crop Diseases and Insect Pests Foundation (2020KF01), and Guangxi Key Research and Development Program (Guixi AB20159041).