High-quality genome assembly-based and functional analyses reveal the pathogenesis mechanisms and evolutionary landscape of wheat sharp eyespot Rhizoctonia cerealis

Lin Lu
Institute of Crop Sciences, National Key Facility for Crop Gene Resources and Genetic Improvement, Chinese Academy of Agricultural Sciences, Beijing, China

Feilong Guo
Institute of Crop Sciences, National Key Facility for Crop Gene Resources and Genetic Improvement, Chinese Academy of Agricultural Sciences, Beijing, China

Zhichao Zhang
Department of Plant Pathology/The Key Laboratory of Plant Immunity, Nanjing Agricultural University, Nanjing, China

Lijun Pan
Institute of Crop Sciences, National Key Facility for Crop Gene Resources and Genetic Improvement, Chinese Academy of Agricultural Sciences, Beijing, China

Yu Hao
Institute of Crop Sciences, National Key Facility for Crop Gene Resources and Genetic Improvement, Chinese Academy of Agricultural Sciences, Beijing, China

Xiuliang Zhu
Institute of Crop Sciences, National Key Facility for Crop Gene Resources and Genetic Improvement, Chinese Academy of Agricultural Sciences, Beijing, China

Jinfeng Yu
Shandong Agricultural University

Wenwu Ye
Nanjing Agricultural University https://orcid.org/0000-0001-7347-8935

Zengyan Zhang (zhangzengyan@caas.cn)
Institute of Crop Sciences, National Key Facility for Crop Gene Resources and Genetic Improvement, Chinese Academy of Agricultural Sciences, Beijing, China

Article

Keywords: wheat, Rhizoctonia cerealis, wheat sharp eyespot, R. cerealis Rc207

Posted Date: March 11th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-152320/v1

Abstract

Wheat (*Triticum aestivum*) is one of the most important staple crops. The necrotrophic binucleate fungus *Rhizoctonia cerealis* is the causal agent for the devastating disease wheat sharp eyespot and additional diseases of other agricultural crops and bioenergy plants. In this study, we present the first high-quality genome assembly of *R. cerealis* Rc207, a highly aggressive strain isolated from wheat. The genome encodes expand and diverse sets of virulence-related proteins, especially secreted effectors, carbohydrate-active enzymes (CAZymes), metalloproteases, Cytochrome P450 (CYP450), and secondary metabolite-associated enzymes. Many of these genes, in particular those encoding secretory proteins and CYP450, showed markedly up-regulation during infection in wheat. Of 831 candidate secretory effectors, ten up-regulated secretory proteins, such as CAZymes, metalloproteases and antigens, were functionally validated as virulence factors required for the fungal infection in wheat. Further intra-species and inter-species comparative genomics analyses showed that repeat sequences, accounting for 17.87% of the genome, are the major driving force for the genome evolution, and frequently intraspecific gene duplication contributes to expansion of pathogenicity-related gene families. This is the first genome-scale investigation elucidating the pathogenesis mechanisms and evolutionary landscape of *R. cerealis*. Our results provide essential tools for further development of effective disease control strategies.

Introduction

Wheat (*Triticum aestivum* L.) is the most-extensively cultivated staple crop (~ 17% of the total cultivated area) in the world[1]. It serves as a leading food for human consumption as an essential source of starch, proteins, vitamins, dietary fiber, and phytochemicals[2]. Wheat crops worldwide are severely affected by sharp eyespot, a devastating culm/stem-base disease that is primarily caused by the necrotrophic and filamentous basidiomycete fungus *Rhizoctonia cerealis* van der Hoeven (teleomorph *Ceratobasidium cereal* Murray & Burpee)[3–5]. This disease can cause severe yield losses and a reduction in wheat quality[6,7]. In China, sharp eyespot has been a very highly economically important wheat disease for many decades[6,7], with a total of 6.67–9.33 million hectares of wheat plants being affected by the disease every year[6–8] (http://www.agri.cn/V20/bchqb). *R. cerealis* has a broad range of host species and a necrotrophic lifestyle, surviving both in the soil and on infected crop residues[5,7–11]. Besides causing wheat sharp eyespot, *R. cerealis* is also responsible for sharp eyespot disease in other cereal crops (such as barley, oat, and rye), yellow patch in grasses, and root rot disease in cotton, potato, sugar beet and several species of legumes[4,5,8,10,11]. Hence, it is very important to control the infection of this pathogen and urgent to ensure food security and the quality of life across the globe. However, current management of wheat sharp eyespot is not effective due to a poor understanding of the pathogenic mechanisms of this pathogenic fungus.

The *Rhizoctonia* genus is a heterogeneous and very complex group of filamentous fungi of basidiomycete, and includes uninucleate, binucleate, and multinucleate *Rhizoctonia* species based on young cell nuclear number. According to reproductively incompatible anastomosis, the binucleate

Rhizoctonia species can be divided into 7 anastomosis groups (AG A to AG Q), while the multinucleate *Rhizoctonia* species are classified into 14 distinct groups (AG1 to AG13, and AGBI)[9]. Among these, the *Rhizoctonia* anastomosis group AG D subgroup I (AG-DI) is responsible for sharp eyespot disease in wheat, and for yellow patch on turfgrasses[4,6,7,9]. In contrast to the multinucleate *R. solani, R. cerealis* contains narrower hyphae and two nuclei within a single hyphal cell, and is able to grow at a relatively slowarate[3,7,12,13]. Previous studies on *R. cerealis* mainly focused on the identification of the fungal pathogen and its life cycle, the symptoms and geographical distribution of the disease, and the genetic structure of populations[3–13]. The binucleate *R. cerealis* exists primarily as vegetative mycelia and sclerotia, does not form asexual spores, and its sexual stage is probably rare in nature[3,9,12]. Hence, it is especially difficult to establish effective methods that allow for the stable transformation and study pathogenesis mechanisms in *R. cerealis*. The whole genome assembly, using Illumina next-generation sequencing (NGS) or long-read sequencing technologies, provides an important tool for enhancing current knowledge on the original pathogenic mechanisms and evolution in pathogens[14–18]. However, no study addressing whole genome and transcriptomic sequencing has ever been reported for *R. cerealis* during infection of host plants. Furthermore, the current knowledge regarding pathogenesis and evolution mechanisms of *R. cerealis* remains very scarce.

In this study, we aimed to elucidate the pathogenesis and evolutionary mechanisms of the necrotrophic fungus *R. cerealis*, and the molecular basis of host-pathogen interactions. In order to achieve this, we employed both Oxford nanopore and Illumina NGS technologies to generate a high-quality whole-genome assembly for *R. cerealis* strain Rc207, a highly aggressive isolate of the wheat sharp eyespot in China. Subsequently, we observed the process of fungal infection on and within the cells of wheat leaf sheaths through scanning electron microscopy (SEM), and profiled gene expression patterns of the fungal pathogenic determinants during wheat infection. We identified 831 (including 29 novel) candidate secretory effectors with diverse activity, and validated the functions of ten up-regulated candidate secreted effector proteins, including: 5 carbohydrate-active enzymes (CAZymes), 1 metalloprotease, 1 tripeptidy peptidase, 2 antigens and 1 guanyl-specific ribonuclease, which have all been confirmed as effectors or virulence factors. Most importantly, to uncover mechanisms explaining the adaptive genomic evolution of *R. cerealis*, we also completed the whole genome assembly of *R. cerealis* Rc301 (R0301, another strain is virulent to wheat) by using Illumina NGS and assembling technologies, then performed intraspecific and interspecific comparative genome analyses and identified abundant potential mobile chromosomes or plastic regions driven by repeat sequences. Beyond uncovering the evolutionary landscape of the *R. cerealis* genome, this study also reveals genome-scale insights onto the pathogenicity and adaptation mechanisms, and provides effective control strategies.

Results

2.1 A high-quality whole-genome assembly and annotation of R. cerealis Rc207

We generated 178,680 clean reads with mean length of 34.1 kb and N_{50} length of 44.8 kb by using Oxford Nanopore long-read sequencing. After a *de novo* assembly by using the filtered and polished subreads of the Nanopore sequencing, and a further polishing/correcting by using ~ 32 million Illumina NGS reads (from six libraries), we obtained the final high-quality *R. cerealis* Rc207 genome assembly consisted of 55 scaffolds with a total length of 56.36 Mb, a max scaffold length of 3.52 Mb, an N_{50} scaffold length of 1.68 Mb, and a GC content of 48.63% (Table 1). Benchmarking Universal Single-Copy Orthologs (BUSCO) analysis revealed that the genome assembly contained 91.03% (264/290) complete BUSCOs, of which 90.53% (239/264) were single-copy (**Table S1**). Mapping with the Illumina short reads showed that 90.16% of the reads were properly mapped (**Table S1**).

Table 1
Summary of *R. cerealis* RC207 genome assembly and annotation

Item	Result
Nanopore clean read	
- read number	178,680
- total length (Gp)	6.09
- mean length (bp)	34,095
- N50 length (bp)	44,799
Genome assembly	
- assembly size (Mb)	56.36
- scaffold number	55
- scaffold N50 (Mb)	1.68
- scaffold N90 (Mb)	0.53
- scaffold Max (Mb)	3.52
- GC content (%)	48.63
- Repeat sequence (%)	17.87
Protein-coding genes	
- gene number	14,433
- average gene length (kb)	2.2
- average CDS length (kb)	1.54
- average intron number	6.32
Non-coding RNAs	
- rRNA number	60
- rRNA family	4
- tRNA number	138
- tRNA family	46
- other ncRNA number	42
- other ncRNA family	13

Approximately 17.87% of the sequences of the assembly were identified as repeat sequences, including 16.12% Class I (retro-transposons) and 1.49% Class II (DNA transposons) transposable elements (TEs, **Table S2**). A total of 14,433 protein-coding genes were predicted, with an average length of 2.20 kb, and an average intron number of 6.32 per gene (Table 1). A total of 10,222 (70.82%) of the predicted proteins could be mapped into an ortholog group using the OrthoMCL database, and 95.07% (13,721) of them could be annotated in different databases, including GO (15.31%), KEGG (26.87%), KOG (46.22%), Pfam (64.04%), Swissprot (51.02%), TrEMBL (94.75%), and NCBI-nr (94.91%). Hence, as much as ~ 5% of these proteins might be specific to the *R. cerealis* Rc207 (**Table S3**).

Further functional annotation showed that the *R. cerealis* Rc207 genome contained an arsenal of potentially diverse pathogenicity factors, such as: 1,080 secreted proteins, 822 carbohydrate-active enzymes (CAZymes), 461 proteases, 511 protein kinases, 38 mitogen-activated protein kinase (MAPK) pathway genes, 281 transporters including 89 ABC transporters, 227 transcription factors, 225 nucleases, 55 GTPases, 8 G-protein-coupled receptors (GPCRs), 89 lipases, 3,659 pathogen-host interaction (PHI) proteins, 197 cytochrome P450 (CYP450) proteins, and 196 fungal virulence factors (FVFs) (**Table S4**). Also, we predicted 165 genes that belong to 15 secondary metabolite-associated gene clusters, including one non-ribosomal peptide synthetases (NRPS) type cluster, six terpene synthases type clusters, and eight NRPS-like type clusters (**Table S5**). All of these proteins may play significant roles in pathogenesis of *R. cerealis*, and might also be associated with the exclusive necrotrophic lifestyle and the adaptation of this pathogen to an unique ecological niche.

2.2 SEM and transcriptomic analysis during R. cerealis Rc207 infection in wheat

We used scanning electron microscopy to track the hyphae infection of *R. cerealis* Rc207 strain (on and) inside the leaf sheaths of the susceptible wheat cultivar Wenmai 6 at the tillering stage. At 18 hai, the hyphae started to pierce plant cell walls (Fig. 1A, **Fig. S1**). At 36 hai, the hyphae penetrated into the plant cell walls and spread into the cytoplasm of the infected wheat sheaths (Fig. 1A, **Fig. S1**). At 72 hai, the hyphae markedly proliferated on the surface of plant cells, and colonized and grew inside the invaded wheat cells (Fig. 1A). At the same time, small brown lesions were first visible at the surface of the infected leaf sheaths. At 96 hai, the thriving fungal hyphae occupied the whole plant cells and destroyed the cell walls (Fig. 1A), while the dark-brown lesions on the wheat leaf sheaths had expanded and continued to develop (**Fig. S2**). At 240 hai, the fungal hyphae massively proliferated inside the colonizing plant cells and continued to destroy the host cells, while the sharp eyespot symptoms on the inoculated wheat sheaths and stems became more severe (**Fig. S2**).

To investigate how the fungal transcriptional reprogramming occurs during wheat infection, we performed deep RNA-sequencing to investigate the global gene transcript dynamics of the *R. cerealis* Rc207 strain at five infection time-points (18, 36, 72, 96, and 240 hai) and *in vitro* mycelia. The transcriptomes showed that a total of 12,706 genes were expressed, and that 7,212, 6,931, 8,841, 12,071 and 12,085 genes were expressed at 18, 36, 72, 96, and 240 hai, respectively (**Table S6**). Compared with *in vitro* mycelia, a total of 70, 78, 495, 661 and 663 genes were significantly up-regulated ($\log_2$fold-change >

1, FDR $P < 0.05$) at 18, 36, 72, 96 and 240 hai, respectively (**Table S7**, Fig. 1B). The corresponding proteins of the 912 genes that were significantly up-regulated during infection belonged to a total of 529 OrthoMCL-annotated ortholog groups and 40 paralog groups, but 58 proteins could not be grouped. When compared to the whole genome, we found that 13 ortholog groups were significantly over-represented among the 912 genes (**Table S8**). Nine of these groups contained the CAZymes of Glycoside Hydrolase (GH) families 3 (OG6_100201), 6 (OG6_109747), 10 (OG6_102896), 32 (OG6_101843), 43 (OG6_110274), 51 (OG6_109774), and 61 (now is Auxiliary Activity family 9 or AA9; OG6_118077), carbohydrate-binding module (CBM) family 5 (OG6_100121), and AA7 (OG6_113017). Moreover, three of the groups contained aromatic peroxygenases (OG6_118111), malate dehydrogenase (OG6_123068), and sterol 14-demethylase (Cytochrome P450 family 51; OG6_103408). One of the groups contained stress response-related proteins (OG6_111600). Almost the entire set of protein groups belonged to the predicted secretome (**Table S8**), indicating that significantly up-regulated secreted proteins may play important roles during the interactions between *R. cerealis* and wheat, and that these secreted proteins may be major pathogenesis determinants .

2.3 Up-regulated secretory CAZymes contribute to the infection of R. cerealis

The *R. cerealis* Rc207 genome contained 822 genes encoding CAZymes that potentially degrade the structure of the plant cell wall and modify the fungal cell wall **(Table S9)**. Among these, 404 CAZymes belonged to secretory proteins. Compared with 8 previously sequenced pathogenic fungi (*Cryptococcus gattii*, *Fusarium graminearum*, *Melampsora larici-populina*, *Postia placenta*, *Puccinia graminis*, *R. solani* AG1 IA, *R. solani* AG8, and *Ustilago maydis*), *R. cerealis* seems to be well equipped with more expanded gene families of CAZymes, and has the highest number in categories GH, carbohydrate esterase (CE) and AA (**Table S10**). Moreover, relatively to the rice sheath blight pathogen *R. solani* AG1 IA (399 CAZyme-genes), the *R. cerealis* RC207 genome contained a more abundant number of cellulose-degradation enzyme GH5 genes (52 *vs.* 15), hemicellulose-degradation enzyme GH16 genes (33 *vs.* 15), pectin-degradation enzyme GH28 genes (15 *vs.* 6), xylan-degradation enzyme GH43 genes (18 *vs.* 9), cutinase CE5 genes (14 *vs.* 1), chitin deacetylase CE4-coding genes (15 *vs.* 8), and AA9-coding genes (34 *vs.* 10) (**Fig. S3, Table S11)**.

RNA-seq based expression profile analysis showed that, relative to the *in vitro* mycelia, 9, 15, 135, 208, and 202 CAZyme-coding genes were significantly up-regulated at 18, 36, 72, 96, and 240 hai, respectively (Fig. 2A, **Table S12, Fig. S4**). The transcript levels of 46 cellulose-targeted enzymes (GH5, GH7, and AA9), 34 hemicellulose-targeted enzymes (GH16, GH10 and GH3), 37 pectin degradation enzymes (GH28, PL1 and PL3), and 14 xylan-degradation enzymes (GH43 and GH51) were significantly up-regulated during the infection process (**Fig. S5A-Fig. S5D**). Additionally, 5 chitin deacetylase CE4-coding genes were significantly up-regulated during advanced infection stages (72 to 240 hai).

To further test an association between the up-regulated CAZyme-coding genes and pathogenesis, we examined the transcript profiles of 13 genes by qRT-PCR analyses. The results showed that the transcript levels of these 13 CAZymes-coding genes, including 1 *GH28*, 3 *GH5*, 1 *GH6*, 3 *GH10*, 1 *GH51*, 1 *AA9*, 2

CBM1, and 1 *CE5*, were significantly induced during infection in wheat (18, 36, 72, 96, and 240 hai) compared to the *in vitro* fungal mycelia (**Fig. S6**). Further, the functional validation results showed that the heterologously-expressed proteins of two up-regulated and cysteine-rich secretory CAZymes, including RcGH6-1 (Rc_00803.1) and RcGH28 (Rc_03586.1) **(**Fig. 2B**)**, were able to trigger plant cell death/necrosis by infiltration into the leaves of the susceptible wheat cv. Wenmai 6 **(**Fig. 2C**).** Importantly, through assay the infection size on the filtrated wheat leaves plus *R. cerealis* liquid mycelia inoculation, results showed that the heterologously-expressed RcGH6-1 and RcGH28 proteins were further verified to promote the fungal infection **(**Fig. 2D, **Fig. S7)**.

2.4 Secretory effectors in R. cerealis Rc207 contribute to the infection

Although previous research annotated the candidate effectors with a signal peptide-containing protein pipeline[17], experimentally-validated effectors and virulence factors tend to be cysteine-rich (≥ 4) or up-regulated secreted proteins[15,19–21]. The *R. cerealis* Rc207 secretome is comprised by 1,080 secreted proteins (**Table S13**). Transcriptome analysis showed that 300 genes coding for secreted proteins were significantly up-regulated during fungal infection in wheat (**Table S14**, Fig. 3A). Of these, we found that 177 and 105 genes coding even and odd cysteine-containing secreted proteins, respectively, were significantly up-regulated. Furthermore, the *R. cerealis* Rc207 secretome includes 755 cysteine-rich (the number of cysteine ≥ 4) secretory proteins, which likely function as candidate effectors (**Fig. S8A**). Taken together, a total of 831 candidate effectors, including 755 cysteine-rich secretory proteins and 282 up-regulated secretory proteins, were identified from the *R. cerealis* Rc207 ssecretome and classified according to their functional annotation (**Table S14**). Among them, those containing even cysteines outnumbered those containing odd cysteine numbers.

This set comprised 29 novel candidate effectors, including: alkaline phosphatases, antigens (carboxypeptidase Y inhibitor), cytochrome b2, choline transporter, glycerophosphoryl diester phosphodiesterase family, fruiting body protein SC7, tyrosinase, distantly related to plant expansin, allergen, alkaline nonlysosomal ceramidase, endoplasmic reticulum protein, FAD-dependent oxidoreductase, guanyl-specific ribonuclease, tripeptidyl peptidase, signal peptide-containing isochorismatase, and other 2 uncharacterized proteins that have obvious characteristics of effectors (**Table S15**). In addition, the *R. cerealis* Rc207 secretome also contained 1 cytochrome C oxidase (Rc_02038.1) with 13.21% amino acid identity to the oxidase domain of *R. solani* AG1 IA assembly protein CtaG/cox (AG1IA_05310, a novel effector reported)[14], and three peptidase inhibitors I9 (Rc_08365.1, Rc_04605.1 and Rc_10139.1) with 28.57%, 25.94%, and 26.32% identities to the peptidase inhibitor domain of the *R. solani* AG1 IA peptidase inhibitor I9 (AG1IA_07795, a novel effector reported)[14].

Previous studies showed that during necrotrophic fungal pathogen infection to host plants, secretory effectors function as essential determinants of pathogenicity or virulence through induction of plant cell death and necrosis[22]. We cloned and validated the functions of 11 up-regulated (qRT-PCR results, **Fig. S6, Fig. S9**, Fig. 4A), cysteine-rich candidate effectors with diverse activity during fungal infection in wheat, including: 5 CAZymes [RcGH5 (Rc_08801.1), RcGH6-1 (Rc_00803.1), RcGH6-2 (Rc_07778.1), RcGH28

(Rc_03586.1), and RcAA9 (Rc_05515.1)], 1 GDSL-lipase RcLP (Rc_08199.1), 1 metalloprotease RcFL1 (Rc_11192.1), 1 tripeptidyl peptidase RcTP (Rc_02278.1), 2 antigens RcOV16-1 and RcOV16-2 (Rc_05066.1 and Rc_11267.1), and 1 guanyl-specific ribonuclease RcRNase (Rc_06868.1). The results showed that except that the heterologously-expressed Rclipase (Rc_08199.1) did not obviously induce necrosis in infiltrated wheat leaves (**Fig. S10**), the remaining 10 secretory proteins were able to induce cell death in the treated leaves of wheat or *N. benthamiana*, and thus promoted *R. cerealis* Rc207 infection in host plants (Fig. 2C-2D, Fig. 3B-3D, Fig. 4C-4E).

2.5 Secretory metalloprotease RcFL1 is required for fungal infection in wheat

Compared to the closely-related fungus *R. solani* AG1 IA (307 protease-coding genes), the *R. cerealis* Rc207 genome contained more genes coding for proteases (461 genes), particularly richer in aspartic proteases, including pepsin A1As (44 *vs.* 23), cysteine proteases (99 *vs.* 64), metalloproteases (116 *vs.* 100), serine proteases such as prolyl aminopeptidase S33s (178 *vs.* 103), and threonine proteases (24 *vs.* 17) **(Fig. S11, Table S16**). Among the 116 metalloproteases (**Fig. S12**), 6, 4, 13, 19 and 19 metalloprotease-coding genes were significantly up-regulated during wheat infection at 18, 36, 72, 96 and 240 hai, respectively (**Fig. S13, Table S17)**.

In order to further investigate an association between the up-regulated metalloproteases with pathogenesis, we examined the gene transcript profiles and the pathogenic functions of the M36 domain-containing metalloprotease (fungalysin) RcFL1 (Rc_11192.1) and M43 domain-containing metalloprotease RcMEP1[23], both with signal peptides, during *R. cerealis* infection in wheat. The qRT-PCR analysis results showed that the transcript abundances of *RcFL1* and *RcMEP1*[23] were markedly up-regulated during the infection process in wheat (18, 36, 72, 96, and 240 hai) compared to *in vitro* fungal mycelia (Fig. 4A). Our previous functional analyses showed that the up-regulated secretory *RcMEP1* acted as virulence factor[23]. Hence, we further investigated functional role of the *RcFL1* during fungal infection to wheat. *Agrobacterium tumefaciens* mediated transient expression assays in *Nicotiana benthamiana* leaves showed that the RcFL1 was able to secrete into the apoplasts and to trigger plant cell death, but that the signal peptide-deleting mutant lost both of these activities in *planta* (Fig. 4B). Furthermore, compared to His-TF (CK), the heterologously-expressed proteins His-TF-RcFL1 (expressing the full RcFL1 protein) and His-TF-RcFL1-M36 (expressing the M36 domain of RcFL1) were able to trigger necrosis and plant cell death on the infiltrated leaves of susceptible wheat cv. Wenmai 6 or of *N. benthamiana* (Fig. 4C, **Fig. S14A**, **Fig. S14B**). Notably, both the RcFL1 and RcFL1-M36 proteins were able to increase the fungal infection size after *R. cerealis* liquid mycelia inoculation on the wheat leaves **(**Fig. 4D-4E**)**. Additionally, the RcFL1 protein significantly repressed the expression of wheat defense genes encoding *chitinases* (*TaChit3* and *TaChitIV*) and the receptor-like kinase CERK1 (*TaCERK1*) in the infiltrated leaves (Fig. 4F). These results demonstrated that the secretory metalloprotease RcFL1, acting as a virulence factor, is required for the fungal infection into wheat, and that both the signal peptide and the M36 domain are necessary to ensure the virulence role of RcFL1.

2.6 Repeat sequences and frequently-intraspecies gene duplication contribute to the genome evolution of R. cerealis and expansion of pathogenicity-related gene families

A phylogenetic tree including *R. cerealis* Rc207, ten previously sequenced basidiomycotina fungi (*Coprinopsis cinerea, Cryptococcus gattii*, *Laccaria bicolor, M. larici-populina*, *Phanerochaete carnosa*, *Postia placenta*, *Puccinia graminis*, *R. solani* AG1 IA, *R. solani* AG8, and *U. maydis*), and the ascomycete *F. graminearum* as an outgroup, was constructed based on their core orthologs. The tree showed that *R. cerealis* Rc207 is more closely related to *R. solani* AG1 IA (and *R. solani* AG8) than to the other fungi in the basidiomycotina family. The divergence between *R. cerealis* Rc207 and *R. solani* AG1 IA/ *R. solani* AG8 occurred approximately 140 million years ago (MYA; **Fig. S15**).

To explore the major mechanisms underlying the genomic evolution of *R. cerealis*, we completed the genome assembled of *R. cerealis* Rc301 (R0301, another strain is also virulent to wheat isolated from Nanjing, China) by using Illumina reads (**Table S18**), then performed intra-species gene collinearity analysis between *R. cerealis* Rc207 and *R. cerealis* Rc301, and inter-species gene collinearity analyses between *R. cerealis* and *R. solani* AG1 IA/ *R. solani* AG8. These analyses showed there were more collinear gene clusters identified from the longer scaffolds of *R. cerealis* Rc207 genome assembly (nanopore sequencing), indicating that these scaffolds include conserved regions. In contrast, more *R. cerealis*-specific the genes were present in the short scaffolds (Fig. 5A-5B). In addition, we identified 3,010 intra-species syntenic (IS) genes (20.86%) in 53 syntenic gene clusters within the *R. cerealis* Rc207 genome (**Table S19**). These IS genes exhibited a high frequency among the short scaffolds and a low frequency among the long scaffolds (Fig. 6), and the similar pattern was also observed in *R. cerealis* Rc301 genome (Fig. 5B). Interestingly, the distribution pattern of IS genes was similar to that of repeat sequences that usually accompany the genomic structures for a rapid evolution[16,24,25]. These findings suggest that there may be abundant duplicated genes and gene clusters within the *R. cerealis* genome, and that these genes are frequently located at repeat sequences-rich genomic regions that may be undergoing rapid evolution.

According to the lengths of the scaffolds, from long to short, and to the frequency of IS genes, with 4.6%, 21.6% and 40.3%, we divided the scaffolds into three groups to perform preliminary comparisons. Specifically, group1, group2 and group3 included the top five longer scaffolds, the 13 medium-length scaffolds, and the remaining 37 short scaffolds, respectively (Fig. 5B; Fig. 6, **Table S21**). In a manner consistent with the observations above on the repeat sequences, the IC genes, the secretome- and CAZymes-encoding genes and the species-specific genes, the unexpressed and low-expressed genes also exhibited an increasing frequency trend from group1 to group3. In contrast, an opposite trend was observed for gene density, GC content, the frequency of conserved genes and that of genes with collinearity to Rc301 (Fig. 5B, Fig. 6; **Table S20**). In addition, the IC genes, the species-specific genes, the unexpressed and low-expressed genes, and the repeat sequences, concentrated on the edge of scaffolds; while the conserved genes, the highly expressed genes, the majority of the secretome, and the genes with collinearity to RC301 concentrated on the central regions of scaffolds (Fig. 5B, Fig. 6A-C; **Table S21**). These analyses further revealed that repeat sequences likely drive a rapid evolution in specific regions of

the *R. cerealis* genome, which may be in the forms of mobile chromosomes or plastic regions, such as the edge of the chromosomes.

Discussion

In this study, using an Illumina NGS platform, we generated a draft genome assembly for the *R. cerealis* Rc207 strain comprising 41.49 Mb in size, which is slightly bigger than the multinucleate *R. solani* AG1 IA (36.94 Mb) and the *R. solani* AG8 (39.8 Mb) genome assemblies[14,15]. Phylogenetic analysis indicated that the *R. cerealis* Rc207 is more closely related to the *R. solani* AG1 IA/*R. solani* AG8 than to other pathogens in the basidiomycotina family, generating important clues about their orgins and lifestyles. The divergence time of *R. cerealis* Rc207 with *R. solani* AG1 IA/*R. solani* AG8 was estimated at ~ 140 MYA, which is close to the separation time between monocots and dicots[26]. Repeat sequences or TEs are key contributors to the genomic plasticity and virulence variation in some fungal plant-pathogens[16,24,25]. However, it is difficult to effectively perform an assembly of repeat sequence-rich genomic regions using the Illumina NGS reads. Hence, we completed a high-quality *R. cerealis* Rc207 genome assembly using long-read sequencing technology. This yielded a 35.81% larger genome size, a 4.63-fold increased contig N_{50} length with an N50 size of 1638.05 Kb, an additional 2,616 genes, and 5.37-fold repeat sequences when compared to the Illumina NGS assembly. As the only sequenced plant fungal pathogen in the binucleate *Rhizoctonia*, the complete genome assembly of *R. cerealis* will provide the essential genomics-based resource for further comparative genomics, including genomic features and evolution, functional genomics, and for uncovering the mechanisms behind pathogenicity, virulence variation and adaptation to this important pathogenic-fungus.

Based on the high-quality genome assembly, interspecific and intraspecific comparative genomic analyses revealed the core and plastic genomic scaffolds and regions in *R. cerealis*. In particular, the repeat sequence-rich scaffolds/regions contained a relatively higher number of secreted proteins- and CAZymes-encoding genes (validated or candidate virulence factors), more species-specific genes and more unexpressed and low-expressed genes. These findings suggest that these genes may undergo rapid evolution, maintain the pathogen unique genome and niche adaption, and are responsible for the variation in virulence in *R. cerealis*. Many repeat sequence-rich scaffolds/regions are part of the short scaffolds or distributed on the edge of scaffolds in the *R. cerealis* genome, which is consistent with the existence of mobile chromosomes and plastic chromosome regions previously reported in the genomes of some other fungal plant pathogens[16,24,26]. Repeat sequences likely drive a rapid evolution in *R. cerealis* specific genomic regions and virulence variation between *R. cerealis* and *R. solani* or between the *R. cerealis* isolates. Furthermore, these comparative genomic analyses identified that 20.86% of *R. cerealis* genes formed 53 syntenic gene clusters, which included abundant virulence-related genes. Hence, frequently intraspecific gene duplication may play a major mechanistic role explaining the expansion of several pathogenicity-related gene families in *R. cerealis*. Even though *R. cerealis* is a binucleate organism, we obtained a very low estimation of genome-wide heterozygosity (1.06%) based on k-mer

analysis using the Illumina NGS reads. Thus, the majority of the short scaffolds, or candidate duplicated genomic regions, are not likely the assemblies of the second copies of different allelic fragments.

The high-quality genome assembly presented here reveals that the number of CAZymes, secreted proteins, PHI factors, proteases including metalloproteases, CYP450 and secondary metabolism enzymes present in the wheat sharp eyespot pathogen *R. cerealis* Rc207 markedly exceeds those observed in the rice sheath blight pathogen *R. solani* AG1 IA[14]. This suggests unique infection mechanisms and host adaption strategies for each of these pathogens. Many of these genes are expressed during fungal infection in wheat. For example, 46 cellulose-targeted enzymes (GH5, GH7 and AA9), 34 hemicellulose-targeted enzymes (GH16, GH10 and GH3), 37 pectin degradation enzymes (GH28, PL1 and PL3), 14 xylan-degradation enzymes and 5 polysaccharide (chitin) deacetylase CE4 displayed significantly up-regulation. Recently, the GH16 ENG1, which was uncovered in *F. graminearum*[21], the AA9-type effector MoCDIP4, which was found in the rice blast fungus *Magnaporthe oryzae*[27], the secretory polysaccharide deacetylase VdPDA1, which was described in the soil-borne fungus *Verticillium dahlia*[28], the *R. cerealis* cutinase RcCUT1[29] and hemicellulose-digested enzyme GH10 RcXYN1[30], were all shown to function as virulence factors. Here, we investigated the pathogenic functions of the five up-regulated CAZyme-encoding genes, namely *RcGH5*, *RcGH6-1*, *RcGH6-2, RcGH28* and *RcAA9* with 78.33% sequence identity to the fungal-type cellulose-binding domain of MoCDIP4, and demonstrated that these proteins were able to trigger plant cell death and promote fungal infection in wheat.

Compared to *R. solani* AG1 IA, *R. cerealis* Rc207 appears to require more proteases. The current research firstly verified that the secretory tripeptidyl peptidase RcTP (Rc_02278.1) functions as a virulence factor of *R. cerealis* Rc207. Our previous research revealed that the secretory M43 domain-containing metalloprotease RcMEP1acted as virulence factor of *R. cerealis*[23]. In the maize anthracnose pathogen fungus *Colletotrichum graminicola*, the M36 domain-containing metalloprotease Cgfl was demonstrated to act as virulence factor through inhibiting host chitinases[31]. Here, our functional analyses demonstrated that the M36 domain-containing metalloprotease RcFL1 could induce plant cell-death and contribute to fungal infection by repressing the expression of the wheat *chitinases* and *TaCERK*, whose ortholog in *Arabidopsis* has been shown to be essential for chitin elicitor signaling[32]. The recent paper reported that VdPDA1, functioning as a secretory effector, contributes to virulence by preventing CERK-mediated chitin-triggered immunity during *V. dahliae* infection[32]. Thus, RcFL1 likely prevented the host chitin-triggering immunity, which shed light on a novel mechanism underlying the virulence roles of the kind of proteases in necrotrophic fungal pathogens.

Moreover, a total of 831 diverse candidate effectors (up-regulated, cysteine-rich secretory proteins) were identified from the *R. cerealis* Rc207 secretome. Interestingly we found 29 novel effectors, such as tripeptidyl peptidase, antigen, guanyl-specific ribonuclease, cytochrome b2, and one signal peptide-containing isochorismatase. However, in the *R. cerealis* Rc207 genome, there are no RXLR- and Crinkler-type effectors widely present in the oomycete species[33], suggesting a significant difference in virulence mechanisms between *R. cerealis* Rc207 and oomycete species. Previous studies reported that genes

coding effectors were often up-regulated during host infection and required for plant cell-death and pathogenicity in pathogens [14,21,31,34]. Here, our functional experiments verified that in the necroptrophic *R. cerealis* Rc207, 10 up-regulated, or cysteine-rich secreted candidate effectors, including RcGH5, RcGH6-1, RcGH6-2, RcGH28, RcAA9, RcFL1, RcTP, RcOV16-1, RcOV16-2 and RcRNase act as virulence factors through the induction of plant cell death and the promotion of infection in wheat. To our knowledge, guanyl-specific ribonuclease, tripeptidyl peptidase and antigen (carboxypeptidase Y inhibitor) have been, for the first time, confirmed as virulence factors involved in pathogen-host plant interaction. The antigen RcOV16 belongs to the phosphatidyl ethanolamine-binding protein (PEBP) family likely through the inhibition of serine proteinases[35], which may be widely useful in crop protection and biotechnology[36]. Even though some resistant quantitative trait loci were identified in wheat, there are no crop cultivars with high resistance to *R. cerealis*. The important virulence factors validated here, including RcMEP6, RcGH28, RcGH6-1, RcOV16, and RcAA9, may be knocked-out/down in host plants by host-induced gene silencing in order to improve *R. cerealis*-resistance in wheat and other economically important cereal crops and bioenergy plants.

In conclusion, this study reports a high-quality whole-genome assembly of the soil-borne basidiomycete pathogen fungi *R. cerealis* and fills gaps in our knowledge of plant pathogen biology. In addition to providing the first binucleate *Rhizoctonia* genome assembly, our study focused on a genome-wide pathogenic mechanisms and functional assay, along with comparative genomic and genome evolution analyses. This study undoubtedly deepens our understanding on the necrotrophic–plant interactions and provides valuable information and resources about the genomic features, adaptation strategies, fungal taxa, evolution, and pathogenicity mechanisms of this fungal pathogen. The novel effectors or virulence factors identified enrich the pathways associated with pathogenesis and fungal pathogen-plant interactions. Moreover, they constitute a vital tool for the study of PHI mechanisms and for molecular breeding. The strategy for impairing the important virulence factors validated, i.e. metalloproteases, cytochrome C oxidase, antigen and peptidase inhibitors, may be applied to control sharp eyespot of cereals including wheat and even diverse diseases affecting host plants.

Materials And Methods

R. cerealis **strain, culture conditions and wheat infection.** The binucleate fungus *R. cerealis* strain Rc207 is a highly aggressive strain with strong virulence in wheat crops from Northern-China, and was collected from the wheat field of Taian, Shandong. *R. cerealis* Rc301 (R0301), another isolate is also virulent to wheat isolated from Nanjing, Jiangsu,China, was kindly provided by Prof. Huigu Chen. The strains were grown in a potato dextrose broth medium at 25 ℃ for 10 days, in the dark and with vigorous shaking (100 r.p.m.), and was subsequently washed with sterile H_2O, frozen in liquid N_2 and freeze dried for genome sequencing.

The wheat cultivar Wenmai6 is highly susceptible to *R. cerealis* infection. Wheat plants were grown in a greenhouse under a cycle of 13-h light (~ 25°C) and 11-h dark (~ 15°C). The *R. cerealis* isolate Rc207 mycelia were inoculated between second base sheaths and stems of wheat plants at their tillering stage

using the toothpick inoculation method. Wheat sheaths at five different *R. cerealis* Rc207 infection time points (18, 36, 72, 96, and 240 hai) were sampled and used for observation by scanning electron microscope, and for RNA-seq analyses.

NGS genome sequencing and assembly. Genomic DNA was extracted with a modified CTAB method. The harvested DNA was detected by agarose gel electrophoresis and quantified using Qubit. Six Illumina NGS libraries, including one 500 bp, one 2,193 bp, two 6,111 bp, and two 10,612 bp, were constructed and sequenced using the Illumina Hiseq 2500 technology with a PE125 strategy. Whole-genome sequencing and assembly were performed at the Beijing Novogene Bioinformatics Technology Co., Ltd.

Evolution analysis. A phylogenetic tree, including Rc207 and eleven other fungal species (*C. cinerea, C. gattii*, *F. graminearum*, *L. bicolor, M. larici-populina*, *P. carnosa*, *P. placenta*, *P. graminis*, *R. solani* AG1 IA, *R. solani* AG8, and *U. maydis*), was constructed based on conserved/core orthologs using a maximum likelihood model in TreeBest at the Beijing Novogene Bioinformatics Technology Co., Ltd. The divergence times were estimated based on the average time observed in basidiomycetes[37].

Oxford nanopore sequencing and assembly. The genomic DNA sample of *R. cerealis* Rc207 was purified by Oxford nanopore tech. (ONT) and sequenced and assembled at Beijing Biomarker Technologies. The Pilon[38] software was used to correct the assembled genome with second-generation data in order to obtain the final genome with higher accuracy. The final *R. cerealis* Rc207 assembly was generated using two pieces of evidence: 1) the evaluation of Illumina NGS data returns ratio and 2) assessment of the integrity of the fungal genome assembly using the BUSCO v2.0 software. *De novo* prediction, homologous protein prediction and transcriptomic prediction was used to predict gene structure, after which the three predictions were integrated. The predicted genes were blasted to different databases, including KOG, KEGG, Swiss-PROt, TrEMBL and NCBI-nr, in order to obtain the gene function annotation.

Genes involved in pathogenicity and virulence. The CAZyme genes were identified using the dbCAN2 meta server (http://bcb.unl.edu/dbCAN2). Cytochrome P450 genes were annotated in the CYP450 database with a threshold *E*-value of less than $1e^{-50}$ (http://drnelson.uthsc.edu/cytochrome P450.html). PHI and FVF were identified with an *E*-value less than $1e^{-50}$ and a minimal alignment length percentage larger than 40%. SignalP v5.0[39] and TMHMM v2.0[40] were used to detect protein sequences with signal peptides and transmembrane helices, respectively. Proteins with signal peptide and without transmembrane domains were identified as secreted proteins. Secondary metabolite-associated gene clusters were predicted using the antiSMA0 software[41]. Proteases were predicted using the MEROPS database (release 12.3)[42].

Comparative genomic analysis. A pairwise gene synteny comparison between genome assemblies was identified using LAST and the MCScanX software (Python version)[43] using predicted proteins as input. The gene synteny regions were visualized as dot plots and a macro-synteny plot using the JCVI graphics functions. Genome scaffolds, gene density and other genomic features using the synteny regions as input were plotted using Circos v0.69–8[44].

RNA-sequencing and transcriptomic analyses. Fungal complementary DNA libraries were constructed from the *R. cerealis* Rc207 strain by infecting healthy wheat leaf sheaths for 18, 36, 72, 96 and 240 hours, and from *in vitro* mycelia stage. The libraries were sequenced on an Illumina Hiseq 2500 platform and 125 bp paired-end reads were generated in the Beijing Novogene Bioinformatics Technology Co., Ltd. The RNA expression analysis was based on the predicted genes of *R. cerealis* Rc207. The index was built and clean reads were aligned to the reference genome using HISAT2 v2.2.1[45]. The StringTie v1.3.5[46] software was used to construct and identify both known and novel transcripts from the HISAT2 alignment results. The DESeq2 v1.30.0[47] software was used to perform read count normalization and differential gene expression (DGE) analysis using fold-change greater than or equal to 2.0 and a FDR *P*-value < 0.05.

Quantitative RT-PCR assay. qRT-PCR reactions were performed using a SYBR Premix Ex Taq kit (TaKaRa, Japan) in an ABI 7500 real time PCR system/instruction (Applied Biosystems, USA). The qRT-PCR data were analyzed using the comparative $2^{-\Delta\Delta CT}$ method. The *R. cerealis* actin gene *RcActin* was used as an internal reference for testing the fungal genes. Each treatment included three independent technical replicates. The primer pairs used in this section are listed in **Table S21**.

Functional validation of candidate effectors. Among the 831 candidate effectors, we selected 10 candidate virulence genes showing up-regulation for functional verification. To test the cell death-induction of the candidate effectors or virulence factors, these proteins were heterologously expressed in *E. coli* Transetta (DE3) and purified, and then infiltrated into leaves of wheat or of *N. benthamiana*[14,30]. To test their infection-promotion activity, Rc207 liquid mycelia were inoculated onto the wheat leaves infiltrated with the heterologously expressed proteins for 6 h and observed/photographed for 3 d[48]. *A. tumefaciens* mediated transient expression of RcFL1 and RcRNase in *N. benthamiana* leaves, and plant cell death-induction assays were performed as described by Ma *et al.*[48] and Yang *et al.*[34]. The GFP localization signals were assayed and photographed for 2 d post agro-infiltration under a confocal microscope (Zeiss LSM 700, Heidenheim, Germany).

Declarations

Acknowledgments

This study was funded by the National Key Project for Research on Transgenic Biology, China (Grant no. 2016ZX08002001 to Zengyan Zhang). Authors are grateful to Porf. Huigu Chen for providing *R. cerealis* Rc301, Dr. Yilong Yang (University of New Hamshire, USA) for his nice suggestion for genome sequencing, Prof. Lingjian Ma (Northwest A&F University) and Ms. Fangdi Shen (Ningbo Polytechnic) for help on scanning electron microscopy observation.

Author contributions

Z.-Y.Z. conceive and supervis the research project, designed the experiments and revised the manuscript. W.-W. Y and Z.-C. Z. performed main bioinformatic and the genome evolution analyses. L.L.

and F.-L.G. perfromed the majority of the experiments. L.-J. P., Y.H., and X.-L.Z., contribueted to RcFL1 main function, SEM observation, and sampling for RNA-seq. J.-F. Y. isolated and provided the *R. cerealis* Rc207 strain. Z.-Y.Z. and W.-W. Y. wrote the draft manuscript.

Competing financial interests: The authors declare no competing financial interests

References

1. Zhou, Y. *et al.* Triticum population sequencing provides insights into wheat adaptation. *Nat Genet* **52**, 1412–1422 (2020).
2. Shewry, P.R. & Hey, S.J. The contribution of wheat to human diet and health. *Food Energy Secur* **4**, 178–202 (2015).
3. Boerema, G.H. & Verhoeven, A.A. Check-list for scientific names of common parasitic fungi. Series 2b: Fungi on field crops: cereals and grasses. *Neth J Plant Pathol* **83**(1977).
4. Lipps, P.E. & Herr, l.j. Etiology of *Rhizoctonia cerealis* in sharp eyespot of wheat. *Phytopathology* **72**(1982).
5. Van der Hoeven, E.P. & Bollen, G.J. Effect of benomyl on soil fungi associated with rye. 1. Effect on the incidence of sharp eyespot caused by *Rhizoctonia cerealis*. *Neth J Plant Pathol* **86**(1980).
6. Li, W., Guo, Y., Zhang, A. & Chen, H. Genetic Structure of Populations of the Wheat Sharp Eyespot Pathogen *Rhizoctonia cerealis* Anastomosis Group D Subgroup I in China. *Phytopathology* **107**, 224–230 (2017).
7. Hamada, M.S., Yin, Y., Chen, H. & Ma, Z. The escalating threat of *Rhizoctonia cerealis*, the causal agent of sharp eyespot in wheat. *Pest Manag Sci* **67**, 1411–9 (2011).
8. Tomaso-Peterson, M. & Trevathan, L. Characterization of *Rhizoctonia*-Like Fungi Isolated from Agronomic Crops and Turfgrasses in Mississippi. *Plant Dis* **91**(2007).
9. Sharon, M., *et al.* The advancing identification and classification of *Rhizoctonia* spp. using molecular and biotechnological methods compared with the classical anastomosis grouping. *Mycoscience* **47**, 299–316 (2006).
10. Reinecke, P.F., H. Infection experiments with *Rhizoctonia cerealis* Van der Hoeven on cereals. *J Plant Dis Prot* **86**, 241–246 (1979).
11. Burpee, L. *Rhizoctonia cerealis* causes yellow patch of turfgrasses. *Plant Dis* **64**, 1114–1116 (1980).
12. Ji, L., Liu, C., Zhang, L., Liu, A. & Yu, J. Variation of rDNA Internal Transcribed Spacer Sequences in *Rhizoctonia cerealis*. *Curr Microbiol* **74**, 877–884 (2017).
13. Li, W., Sun, H., Deng, Y., Zhang, A. & Chen, H. The heterogeneity of the rDNA-ITS sequence and its phylogeny in *Rhizoctonia cerealis*, the cause of sharp eyespot in wheat. *Curr Genet* **60**, 1–9 (2014).
14. Zheng, A. *et al.* The evolution and pathogenic mechanisms of the rice sheath blight pathogen. *Nat Commun* **4**, 1424 (2013).

15. Hane, J.K., Anderson, J.P., Williams, A.H., Sperschneider, J. & Singh, K.B. Genome sequencing and comparative genomics of the broad host-range pathogen *Rhizoctonia solani* AG8. *PLoS Genet* **10**, e1004281 (2014).
16. Bao, J. *et al.* PacBio Sequencing Reveals Transposable Elements as a Key Contributor to Genomic Plasticity and Virulence Variation in *Magnaporthe oryzae*. *Mol Plant* **10**, 1465–1468 (2017).
17. Müller, M.C. *et al.* A chromosome-scale genome assembly reveals a highly dynamic effector repertoire of wheat powdery mildew. *New Phytol* **221**, 2176–2189 (2019).
18. Moss, E.L., Maghini, D.G. & Bhatt, A.S. Complete, closed bacterial genomes from microbiomes using nanopore sequencing. *Nat Biotechnol* **38**, 701–707 (2020).
19. Oliva, R. *et al.* Recent developments in effector biology of filamentous plant pathogens. *Cell Microbiol* **12**, 705–15 (2010).
20. Kale, S.D. Oomycete and fungal effector entry, a microbial Trojan horse. *New Phytol* **193**, 874–81 (2012).
21. Zhang, X.W. *et al.* In planta stage-specific fungal gene profiling elucidates the molecular strategies of *Fusarium graminearum* growing inside wheat coleoptiles. *Plant Cell* **24**, 5159–76 (2012).
22. Liu, Z. *et al.* The cysteine rich necrotrophic effector SnTox1 produced by *Stagonospora nodorum* triggers susceptibility of wheat lines harboring Snn1. *PLoS Pathog* **8**, e1002467 (2012).
23. Pan, L., et al. The M43 domain-containing metalloprotease RcMEP1 in *Rhizoctonia cerealis* is a pathogenicity factor during the fungus infection to wheat. *J Intergr Agric* **19**, 2044–2055 (2020).
24. Möller, M. & Stukenbrock, E.H. Evolution and genome architecture in fungal plant pathogens. *Nat Rev Microbiol* **15**, 756–771 (2017).
25. Dong, S., Raffaele, S. & Kamoun, S. The two-speed genomes of filamentous pathogens: waltz with plants. *Curr Opin Genet Dev* **35**, 57–65 (2015).
26. Fouché, S. *et al.* Stress-Driven Transposable Element De-repression Dynamics and Virulence Evolution in a Fungal Pathogen. *Mol Biol Evol* **37**, 221–239 (2020).
27. Xu, G. *et al.* A fungal effector targets a heat shock-dynamin protein complex to modulate mitochondrial dynamics and reduce plant immunity. *Sci Adv* **6**(2020).
28. Gao, F. *et al.* Deacetylation of chitin oligomers increases virulence in soil-borne fungal pathogens. *Nat Plants* **5**, 1167–1176 (2019).
29. Lu, L., Rong, W., Massart, S. & Zhang, Z. Genome-Wide Identification and Expression Analysis of Cutinase Gene Family in *Rhizoctonia cerealis* and Functional Study of an Active Cutinase RcCUT1 in the Fungal-Wheat Interaction. *Front Microbiol* **9**, 1813 (2018).
30. Lu, L., Liu, Y. & Zhang, Z. Global Characterization of GH10 Family Xylanase Genes in *Rhizoctonia cerealis* and Functional Analysis of Xylanase RcXYN1 During Fungus Infection in Wheat. *Int J Mol Sci* **21**(2020).
31. Sanz-Martín, J.M. *et al.* A highly conserved metalloprotease effector enhances virulence in the maize anthracnose fungus *Colletotrichum graminicola*. *Mol Plant Pathol* **17**, 1048–62 (2016).

32. Erwig, J. *et al.* Chitin-induced and CHITIN ELICITOR RECEPTOR KINASE1 (CERK1) phosphorylation-dependent endocytosis of *Arabidopsis thaliana* LYSIN MOTIF-CONTAINING RECEPTOR-LIKE KINASE5 (LYK5). *New Phytol* **215**, 382–396 (2017).
33. Lévesque, C.A. *et al.* Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. *Genome Biol* **11**, R73 (2010).
34. Yang, B. *et al.* Fg12 ribonuclease secretion contributes to *Fusarium graminearum* virulence and induces plant cell death. *J Integr Plant Biol* (2020).
35. Mima, J. *et al.* Structure of the carboxypeptidase Y inhibitor IC in complex with the cognate proteinase reveals a novel mode of the proteinase-protein inhibitor interaction. *J Mol Biol* **346**, 1323–34 (2005).
36. Sabotič, J. & Kos, J. Microbial and fungal protease inhibitors–current and potential applications. *Appl Microbiol Biotechnol* **93**, 1351–75 (2012).
37. Krings, M., Dotzler, N., Galtier, J. & Taylor, T.N. Oldest fossil basidiomycete clamp connections. *Mycoscience* **52**, 18–23 (2011).
38. Walker, B.J. *et al.* Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* **9**, e112963 (2014).
39. Almagro Armenteros, J.J. *et al.* SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**, 420–423 (2019).
40. Chen, Y., Yu, P., Luo, J. & Jiang, Y. Secreted protein prediction system combining CJ-SPHMM, TMHMM, and PSORT. *Mamm Genome* **14**, 859–65 (2003).
41. Blin, K. *et al.* antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* **47**, W81-w87 (2019).
42. Rawlings, N.D. *et al.* The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res* **46**, D624-d632 (2018).
43. Tang, H. *et al.* Synteny and collinearity in plant genomes. *Science* **320**, 486–8 (2008).
44. Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome Res* **19**, 1639–45 (2009).
45. Kim, D., Paggi, J.M., Park, C., Bennett, C. & Salzberg, S.L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907–915 (2019).
46. Pertea, M., Kim, D., Pertea, G.M., Leek, J.T. & Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc* **11**, 1650–67 (2016).
47. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol* **11**, R106 (2010).
48. Ma, Z. *et al.* A *Phytophthora sojae* Glycoside Hydrolase 12 Protein Is a Major Virulence Factor during Soybean Infection and Is Recognized as a PAMP. *Plant Cell* **27**, 2057–72 (2015).

Figures

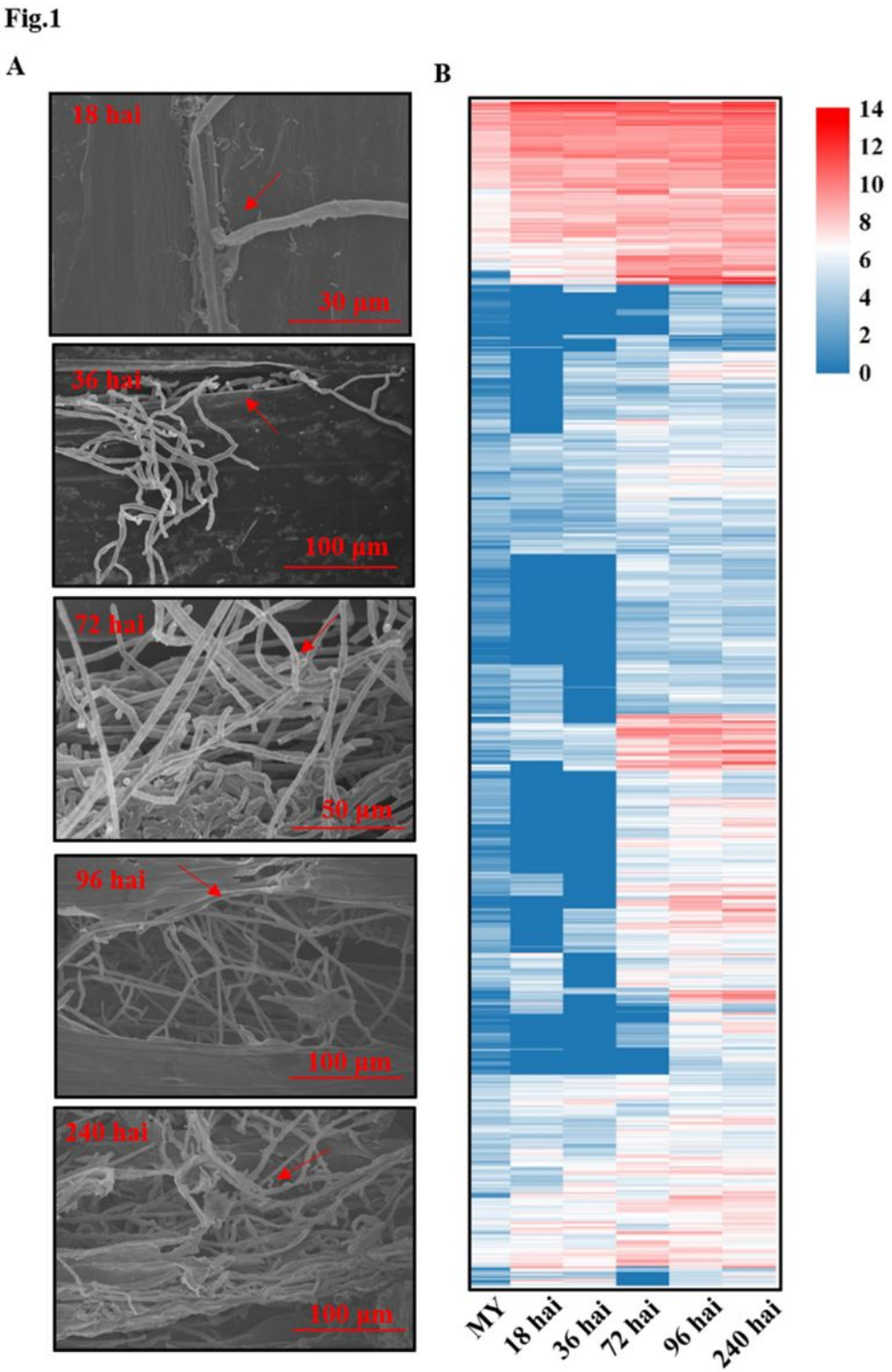


Figure 1

R. cerealis Rc207 during the infection process in wheat. (A) Scanning electron microscopy images of R. cerealis Rc207 infection at the time points listed above the figure. The red arrow represents the hyphae.

(B) Heap map showing the expression levels of up-regulated genes of R. cerealis Rc207 during infection in wheat. The color bar represents the log2 of (FPKM+1) value, ranging from 0 (blue) to 14 (red).

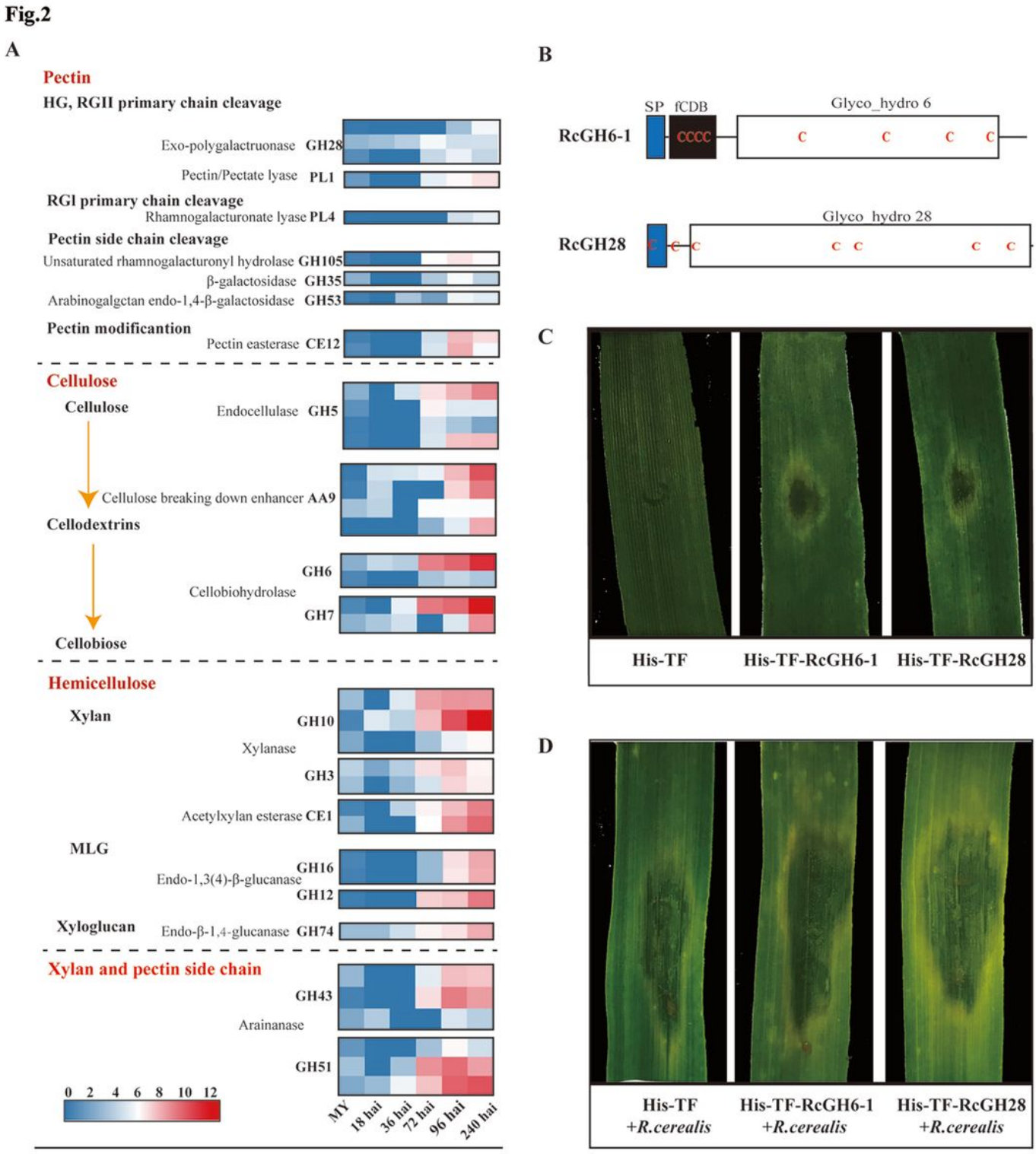


Figure 2

CAZymes expression patterns of R. cerealis Rc207 and CAZymes functional validation. (A) Expression patterns of preferentially cell wall-related genes that were grouped based on their putative targeted substrates. The color bar represents the log2 of (FPKM+1) value, ranging from 0 (blue) to 12 (red). (B)

Schematic representation of RcGH6-1 (Rc_00803.1) and RcGH28 (Rc_03586.1). (C) RcGH6-1 and RcGH28 proteins induce cell death in infiltrated leaves of susceptible wheat cv. Wenmai 6. Each heterologously-expressed protein was infiltrated in 5 μM. (D) RcGH6-1 and RcGH28 promote infection of the R. cerealis Rc207 mycelia in leaves of wheat cv. Wenmai 6. Photographs of the lesions caused by R. cerealis were taken three days after inoculation with the fungal liquid mycelia on wheat leaves infiltrated heterologously-expressed RcGH6-1 or RcGH28 (5 μM).

Fig.3

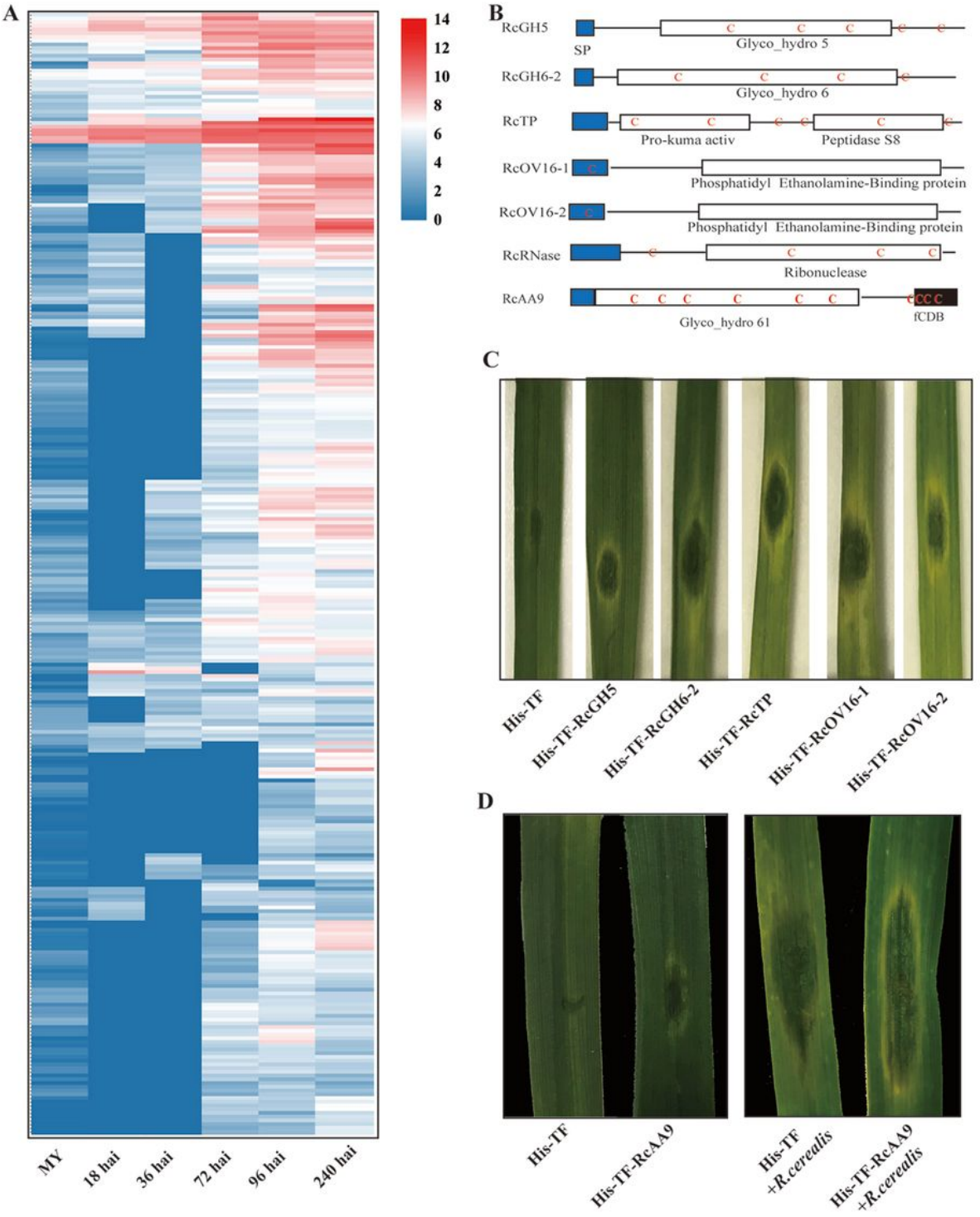


Figure 3

The functional validation of candidate effectors. (A) Heap map showing the expression levels of up-regulated candidate effectors of R. cerealis Rc207 during infection in wheat. The color bar represents the log2 of (FPKM+1) value, ranging from 0 (blue) to 14 (red). (B) Schematic representations of RcGH5 (Rc_08801.1), RcGH6-2 (Rc_07778.1), RcAA9 (Rc_05515.1), RcTP (Rc_02278.1), RcOV16-1 (Rc_05066.1), RcOV16-2 (Rc_11267.1) and 1 RcRNase (Rc_06868.1) proteins. (C) RcGH5, RcGH6-2, RcTP, RcOV16-1 and RcOV16-2 induce cell death in infiltrated leaves of the susceptible wheat cv. Wenmai 6. Each heterologously-expressed protein was infiltrated in 5 μM. (D) RcAA9 contributes to the infection of R. cerealis Rc207 mycelia in leaves of the wheat cv. Wenmai6. Photograph of the infection/lesions caused by R. cerealis was taken three days after inoculation with the fungal liquid mycelia on wheat leaves infiltrated with the heterologously-expressed protein.

Fig.4

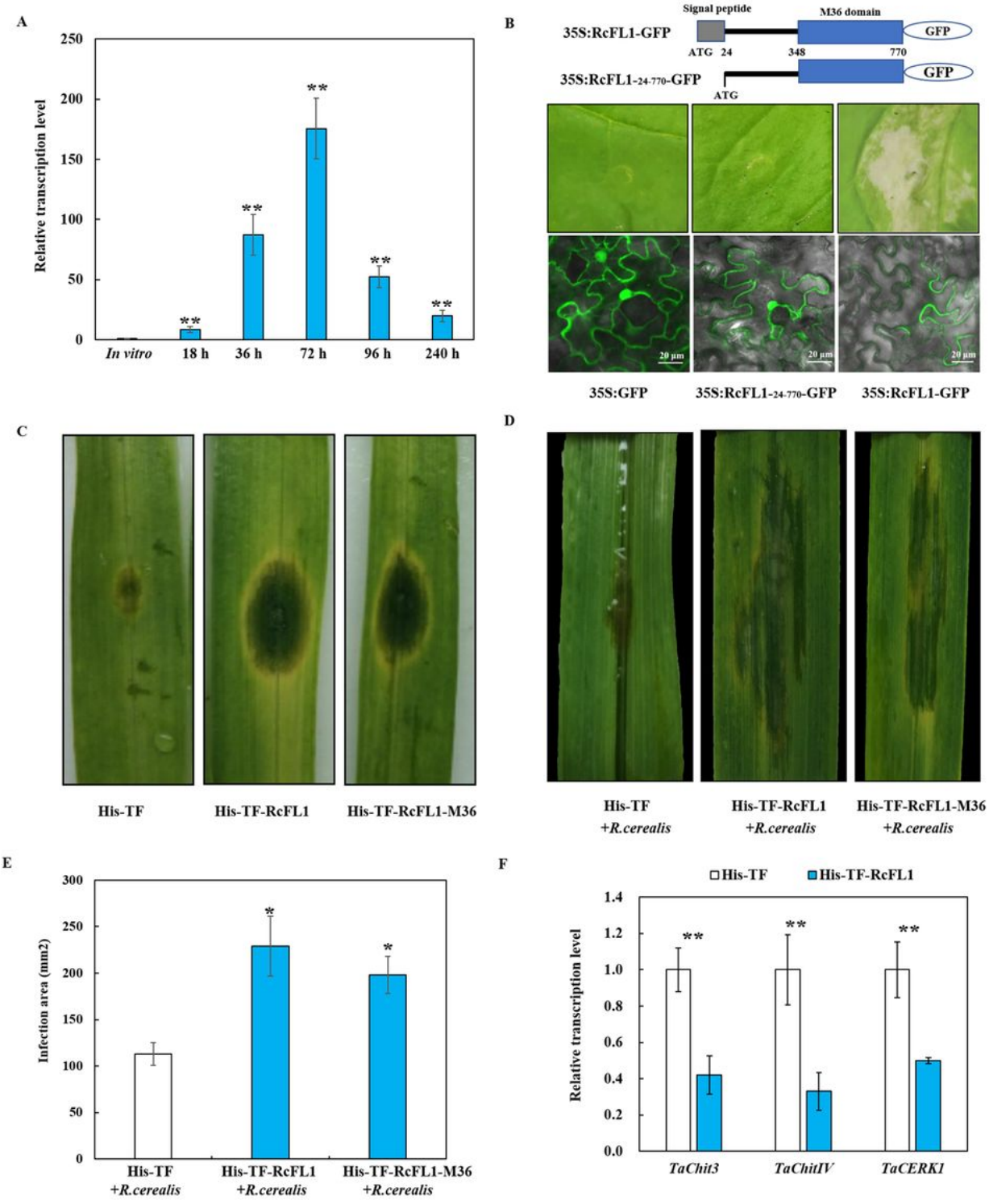


Figure 4

The secretory metalloprotease RcFL1 is required for R. cerealis infection in wheat. (A) qRT-PCR assay showing the expression levels of the RcFL1 gene in R. cerealis Rc207 during different infection stages. The R. cerealis Actin gene was used as an internal control to normalize the data. Standard error (SE) bar was calculated based on three technical replicates and significance was assessed using Student's t-tests (*, P < 0.05; **, P < 0.01). (B) A. tumefaciens-mediated transient expression in N. benthamiana showed

that the cell death-induced activity of RcFL1 depended on the signal peptide function. Photographs were taken five days post agro-infiltration. Confocal microscopy images of 35S: RcFL1-GFP, 35S: RcFL1-24-770-GFP or 35S: GFP in these agro-infiltrated leaves of N. benthamiana. These GFP signals were observed under a confocal microscope. Bars= 20 μm. Photographs were taken three days post agro-infiltration. (C) RcFL1 and its M36 domain induce cell death in infiltrated leaves of the susceptible wheat cv. Wenmai 6. Each heterologously-expressed protein was infiltrated in 5 μM. (D) RcFL1 and its M36 domain-containing peptide RcFL1-M36 promote the fungal infection in the leaves of wheat cv. Wenmai6. The photograph of the infection lesions caused by R. cerealis was taken three days post inoculation with fungus liquid mycelia on wheat leaves infiltrated heterologously-expressed protein (5 μM). (E) The lesion area/infection measured by R. cerealis liquid mycelia inoculation for three days on leaves infiltrated with heterologously-expressed His-TF-RcFL1 (5 μM), or His-TF-RcFL1-M36 (5 μM) or His-TF (5 μM). (F) Heterologously-expressed RcFL1 protein decreases the gene transcript levels of wheat chitinases and receptor-like kinase in infiltrated leaves. Both samples were derived from the wheat cv. Wenmai 6 leaves at three days post inoculation with the His-TF-RcFL1 and the His-TF (CK) at a concentration of 5 μM. The wheat Actin gene was used as an internal control to normalize the data. SE was calculated based on three technical replicates. * indicates significant differences (** $p < 0.01$ or * $p < 0.05$, t-test) between His-TF-RcFL1 and the His-TF (CK) infiltrated leaves.

Fig.5

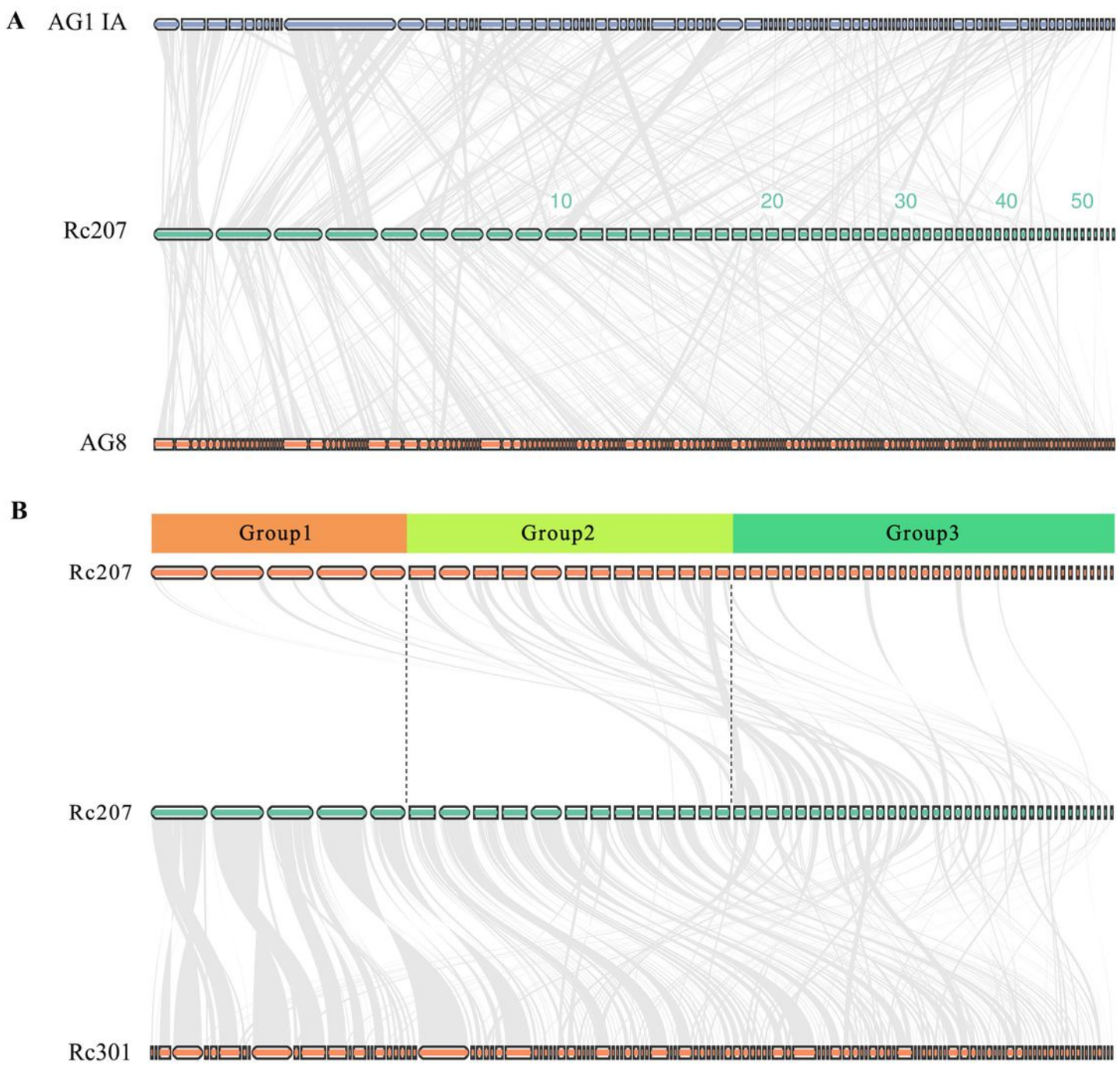


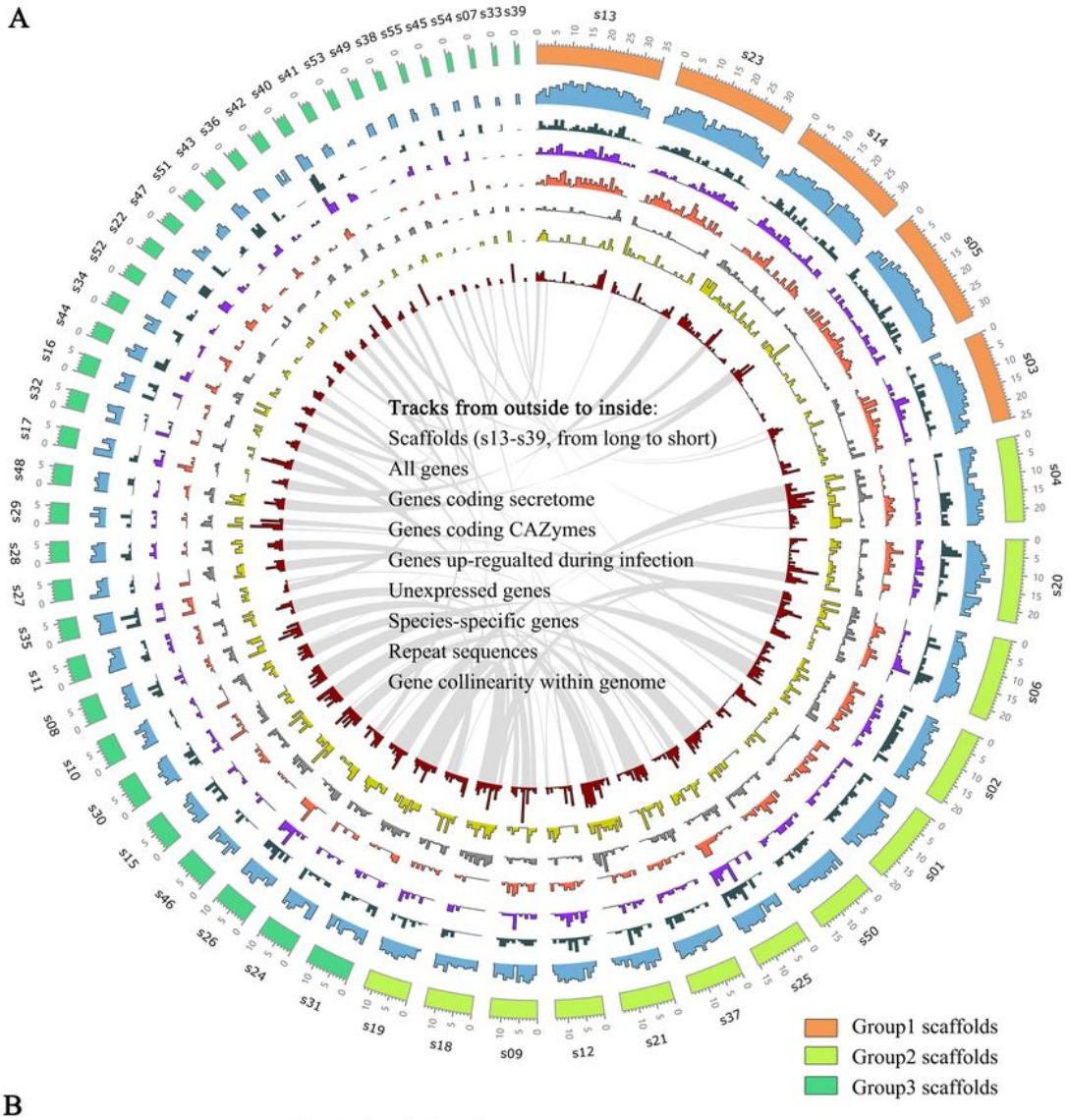
Figure 5

Inter- and intra-species gene collinear relationships of R. cerealis Rc207 and R. solani AG1-IA/AG8, or and R. cerealis Rc301. (A) Gene collinear relationships between R. cerealis Rc207 and R. solani AG1 IA, and between R. cerealis Rc207 and R. solani AG8. (B) Gene collinear relationship within R. cerealis Rc207 genome, and between R. cerealis Rc207 and R. cerealis Rc301 Group1, group2, and group3 include the

top five long scaffolds, the 13 medium length scaffolds, and the remaining 37 short scaffolds, respectively.

Fig.6

A



B

	Group1	Group2	Group3
Total number of scaffolds	5	13	37
Mean scaffold length (Mb)	3.2	1.7	0.5
Gene density (/Mb)	314.4	243.4	220.5
Repreat sequences	11.4%	20.2%	26.2%
GC content	49.1%	48.6%	48.2%
Genes coding conserved orthologs	77.5%	68.4%	65.8%
Species-specific genes	2.5%	5.9%	6.6%
Genes coding secretome	5.8%	8.2%	8.7%
Genes coding CAZymes	5.4%	5.7%	6.2%
Genes with collinearity within Rc207	4.6%	21.6%	40.3%
Mean gene number in homolog group	6.5	17.6	22.1
Mean expression level (log2FPKM)	24.3	16.9	12.7
Unexpressed genes	4.6%	14.0%	18.5%
Genes with collinearity to RC301	78.7%	58.4%	46.9%

C

	Relative distribution in a scaffold				
	0%-20%	20%-40%	40%-60%	60%-80%	80%-100%
All genes	(1.31)	0.01	0.82	1.11	(0.63)
Genes coding conserved orthologs	(1.28)	0.19	0.81	1.04	(0.77)
High-expressed genes	(1.43)	0.18	0.88	0.92	(0.55)
Genes coding secretome	0.01	(0.30)	0.37	1.33	(1.41)
Specific-specific genes	1.46	(0.34)	(1.04)	(0.62)	0.54
Genes with collinearity within Rc207	1.48	(0.87)	(0.39)	(0.78)	0.55
Low-expressed genes	1.35	(0.44)	0.98	(0.68)	0.75
Repeat sequences in group1	1.47	(0.53)	(0.58)	(0.94)	0.57
Repeat sequences in group2	(0.29)	0.31	(0.50)	(1.07)	1.56
Repeat sequences in group3	(1.38)	1.19	0.35	0.44	(0.60)
Genes with collinearity to RC301	(1.42)	0.27	0.77	0.98	(0.60)

Figure 6

Genome features of R. cerealis Rc207. (A) The circle diagram displays tracks of genome features, as indicated. (B) Comparison of genome features between the three scaffold groups. (C) Comparison of genome features according to relative distribution in a scaffold.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Fig.S1.jpg
- Fig.S2.jpg
- Fig.S3.jpg
- Fig.S4.jpg
- Fig.S5.jpg
- Fig.S6.jpg
- Fig.S7.jpg
- Fig.S8.jpg
- Fig.S9.jpg
- Fig.S10.jpg
- Fig.S11.jpg
- Fig.S12.jpg
- Fig.S13.jpg
- Fig.S14.jpg
- Fig.S15.jpg
- 2.22TableS1S21.xls