

# SCIENTIFIC REPORTS



OPEN

## Transcriptome analysis reveals the host selection fitness mechanisms of the *Rhizoctonia solani* AG1IA pathogen

Yuan Xia<sup>1,2</sup>, Binghong Fei<sup>1,2</sup>, Jiayu He<sup>1,2</sup>, Menglin Zhou<sup>1,2</sup>, Danhua Zhang<sup>1,2</sup>, Linxiu Pan<sup>1,2</sup>, Shuangcheng Li<sup>1,3</sup>, Yueyang Liang<sup>1,3</sup>, Lingxia Wang<sup>1,3</sup>, Jianqing Zhu<sup>1,3</sup>, Ping Li<sup>1,3</sup> & Aiping Zheng<sup>1,2</sup>

*Rhizoctonia solani* AG1IA is a major generalist pathogen that causes sheath blight. Its genome, which was the first to be sequenced from the *Rhizoctonia* genus, may serve as a model for studying pathogenic mechanisms. To explore the pathogen-host fitness mechanism of sheath-blight fungus, a comprehensive comparative transcriptome ecotype analysis of *R. solani* AG1IA isolated from rice, soybean and corn during infection was performed. Special characteristics in gene expression, gene ontology terms and expression of pathogenesis-associated genes, including genes encoding secreted proteins, candidate effectors, hydrolases, and proteins involved in secondary metabolite production and the MAPK pathway, were revealed. Furthermore, as an important means of pathogenic modulation, diverse alternative splicing of key pathogenic genes in *Rhizoctonia solani* AG1IA during infections of the abovementioned hosts was uncovered for the first time. These important findings of key factors in the pathogenicity of *R. solani* AG1IA ecotypes during infection of various hosts explain host preference and provide novel insights into the pathogenic mechanisms and host-pathogen selection. Furthermore, they provide information on the fitness of *Rhizoctonia*, a severe pathogen with a wide host range.

The soil-borne Basidiomycete fungus, *Rhizoctonia solani* Kühn, is a destructive and widespread plant pathogen that can infect approximately 200 species of crops<sup>1</sup>. According to its anastomosis criterion, ecology and host range, *R. solani* is divided into 14 anastomosis groups (AG1 to AG13 and AGBI)<sup>2</sup>. *R. solani* AG1 consists of six primary intraspecific subgroups (AG1IA, AG1IB, AG1IC, AG1ID, AG1IE and AG1IF)<sup>3,4</sup>. The AG1IA subgroup is considered the main pathogen that causes sheath blight in rice (*Oryza sativa* L.), maize (*Zea mays* L.) and soybean (*Glycine max*)<sup>5-7</sup>. As one of the most destructive diseases of cultivated rice, rice sheath blight caused by *R. solani* AG1IA can result in yield losses of 50% in Asia<sup>8</sup>. Although resistance to sheath blight has been reported<sup>9</sup>, completely resistant cultivars or immune varieties have not been found<sup>10</sup>. However, most research has been conducted on genus-species and anastomosis groups<sup>11</sup>, and few studies have investigated pathogenicity at the molecular level<sup>12,13</sup> or its genetic variability<sup>14-17</sup>.

RNA-sequencing (RNA-Seq) has become a very popular method for elucidating the molecular mechanisms involved in specific biological and disease processes<sup>18</sup>. Using RNA-Seq, much research has been conducted on pathogen-host interactions (PHIs) in major plant diseases, including *Magnaporthe oryzae*<sup>19</sup>, *Ustilago maydis*<sup>20</sup> and *Phytophthora infestans*<sup>21</sup>. Transcriptomics technology has also played an important role in the study of various pathogens such as *Postia placenta*<sup>22</sup>, *Colletotrichum*<sup>23</sup>, *Moniliophthora perniciosa*<sup>24</sup>, *Xanthomonas*<sup>25</sup>, and *Fusarium*<sup>26</sup> and their hosts.

Currently, the genomes of *R. solani* AG1IA, AG1IB (isolate 7/3/14), AG3, AG8, and AG2-2IIB have been sequenced<sup>27-31</sup> and the genomes of *R. solani* AG1IA, AG1IB, AG3, and AG8 have been compared<sup>32</sup>. In particular, whole-genome sequencing of *R. solani* AG1IA has enabled prediction of pathogenesis-associated genes and

<sup>1</sup>Rice Research Institute, Sichuan Agricultural University, Chengdu, 611130, China. <sup>2</sup>Key Laboratory of Sichuan Crop Major Diseases, Sichuan Agricultural University, Chengdu, 611130, China. <sup>3</sup>Key Laboratory of Southwest Corp Gene Resource and Genetic Improvement of Ministry of Education, Sichuan Agricultural University, Ya'an, 625014, China. Correspondence and requests for materials should be addressed to A.Z. (email: [apzh0602@gmail.com](mailto:apzh0602@gmail.com))

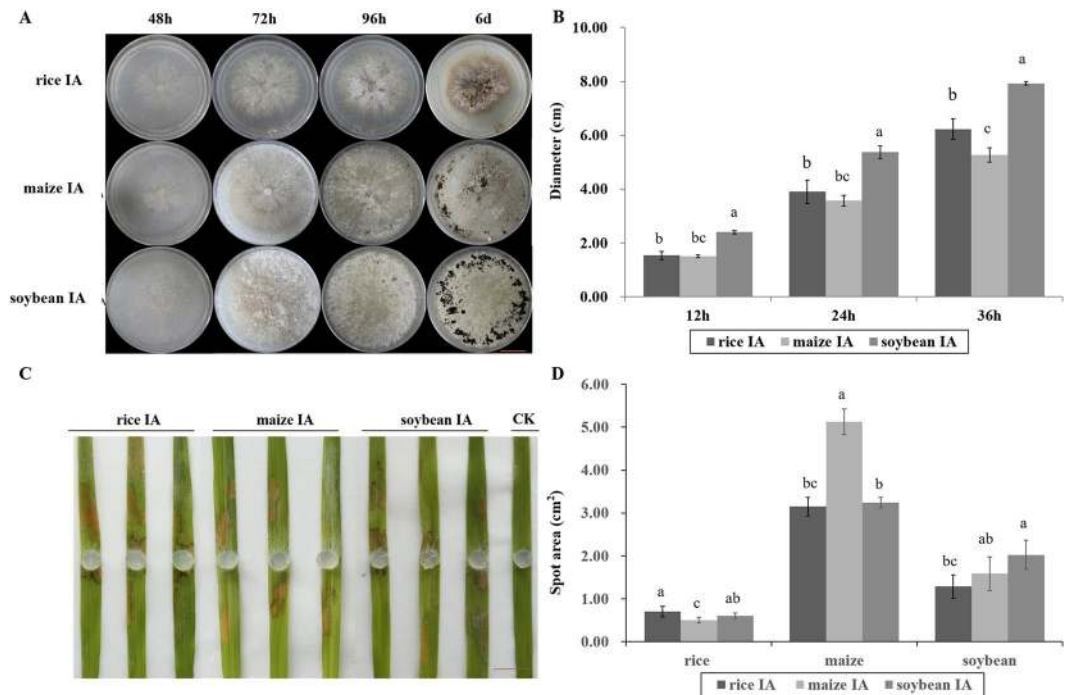
effectors, which could lead to novel insights into pathogenic mechanisms<sup>28</sup>. The *R. solani* proteome has increased our understanding of the mechanisms by which *R. solani* adapts to different plant hosts<sup>33</sup>. Furthermore, the transcriptome of *R. solani* AG1IA during early invasion of *Zoysia japonica* roots<sup>34</sup>, and changes in the transcriptome change after infection with *R. solani*, were recently reported<sup>35</sup>. Genome sequencing and annotation have provided a global view of *R. solani* IA genes, and analysis of the transcriptome at six stages of infection has improved our understanding of the mechanism by which *R. solani* AG1 infects rice. Despite considerable advances in our understanding of *R. solani* AG1IA in the past several years, progress in genetic studies and our understanding of pathogenic factors has been limited due to the multinucleate and heterokaryotic nature of *R. solani*, which has hindered our ability to study the host selection fitness mechanisms of sheath-blight fungus. To date, the majority of studies on *R. solani* AG1IA PHIs have focused on comparisons of pathogenicity and biological characteristics in different hosts<sup>36,37</sup>. Based on this, we provide the first comparative transcriptome analysis of *R. solani* AG1IA gene expression during infection of different hosts and provide novel insights into the host selection fitness mechanisms during infection by sheath-blight fungus. High-throughput sequencing and bioinformatics were used to profile the transcriptomes of three strains of different ecotypes of *R. solani* AG1IA under different infection conditions. Many differentially expressed, pathogenesis-associated genes—including secreted proteins, candidate effectors, carbohydrate-active enzymes (CAZymes), and secondary metabolites involved in the pathogenesis of *R. solani* AG1IA in various hosts were identified. Moreover, alternative splicing of key pathogenesis-related genes in *R. solani* AG1IA during the infection of different hosts was evaluated. By comparing the transcriptomes of *R. solani* AG1IA ecotypes cultivated in the same host and original/non-original hosts to those cultivated on vegetative mycelia, we hope to improve our understanding of the infection mechanism of the sheath blight fungus, *R. solani* AG1IA.

## Results

**Biological characteristics of strains from different hosts.** The three strains, rice IA, maize IA, and soybean IA, were randomly isolated from heavily infected rice (*O. sativa* L.), maize (*Z. mays* L.) and soybean (*G. max*) plants in Sichuan province in areas of different altitudes (Xichang, Chengdu, and Yaan, respectively). To explore the biological differences in strains from different host, a comparative analysis including hyphal anastomosis, growing situation and pathogenicity were conducted. Hyphal anastomosis among rice IA, maize IA, and soybean IA showed somatic incompatibility and no fusion of hyphae, and strains from self-pairings were compatible and exhibited no demarcation between paired ecotypes (Supplementary Fig. S1). Cultures of three strains all could anastomose with AG1IA and all ecotypes from rice IA, maize IA, and soybean IA hosts showed identical ribosomal DNA ITS (Internal Transcribed Spacer Identification). We found that some genetic differences still existed after long-term co-evolution of the strains and their host plants.

For colony morphology on potato-dextrose agar (PDA) (Fig. 1A), the strains exhibited differences in sclerotium formation. Soybean IA formed the most aerial hyphae, followed by maize IA and rice IA. The sclerotia of *R. solani* are the primary source of sheath blight infection, and the number of sclerotia is a determinant of disease severity. The sclerotia formed by soybean IA were relatively large and distributed at the edge of the plate. Sclerotia formed by maize IA were not connected and distributed at the edge of the plate, with few located near the inoculation point. In contrast, rice IA formed a sclerotium near the inoculation point. The mycelium diameters of rice IA and maize IA, were not significantly different at the early stage (from 0 to 24 h after inoculation), but soybean IA showed significantly faster growth than the other two strains, from 12 to 36 h (Fig. 1B). To explore the differences in pathogenicity among the three *R. solani* AG1 IA strains isolated from rice, maize, and soybean, a detached-leaf inoculation method was used, that allows for convenient inoculation, spot amplification, and rapid disease development<sup>38</sup>. Differences in the area of spots were observed (Fig. 1C). Spots from rice leaves showed that the disease caused by rice IA was significantly more severe than that caused by maize IA but showed no significant difference with soybean IA, and there was no significant difference between maize IA and soybean IA at 48 h post-inoculation (hpi) (Fig. 1D). Furthermore, the same experiment was performed using maize and soybean as host plants at 48 h post-inoculation (hpi). Spot areas from maize leaves showed that the disease caused by maize IA was significantly more severe than that caused by rice IA and soybean IA and spots from soybean leaves also indicated that the disease caused by soybean IA was significantly more severe than rice IA. However, no significant difference was found between soybean IA and maize IA (Fig. 1D, Supplementary Fig. S2). These results show that *R. solani* AG1IA strains isolated from different host plants exhibit different levels of pathogenicity and show greater pathogenicity to their original hosts.

**Transcriptome sequencing and analysis.** To gain insight into the pathogenesis of *R. solani* AG1IA in different hosts, transcriptomes of the abovementioned strains with the same infection stage (infection structures formed) in different hosts were sequenced using Illumina NGS technology; strains without host infection were used as controls. In total, 104.97 million paired-end reads with Q30 > 85%, which indicated raw data of high quality, were obtained (Supplementary Table S1). Of the short clean reads, 68.94–79.31% were mapped onto the *R. solani* AG1IA reference genome<sup>28</sup>, and mapped reads were further classified and annotated using TopHat<sup>39</sup> (Supplementary Table S2). Overall, 9,077, 9,106, 8,850, 8,878, 8,923, 8,997, 8,943 and 8,913 genes were identified in R-R (rice IA during rice infection), R0 (rice IA grown on PDA), M-M (maize IA during maize infection), M0 (maize IA grown on PDA), M-R (maize IA during rice infection), S-S (soybean IA during soybean infection), S0 (soybean IA grown on PDA), and S-R (soybean IA during rice infection) (FPKM value > 0.5), respectively. In addition, 1,840 novel genes named Rice\_newGene\_999 were discovered and assembled using Cufflinks<sup>40</sup> after filtering out short encoded peptides (<50 amino acid residues) and genes containing only a single exon. Among the novel genes, 11% were successfully annotated in KEGG (Kyoto Encyclopedia of Genes and Genomes), 26.9% in gene ontology (GO), 27% in COG, 41.5% in SWISS-PROT, and 85% in NR (Supplementary Table. S3). Twelve

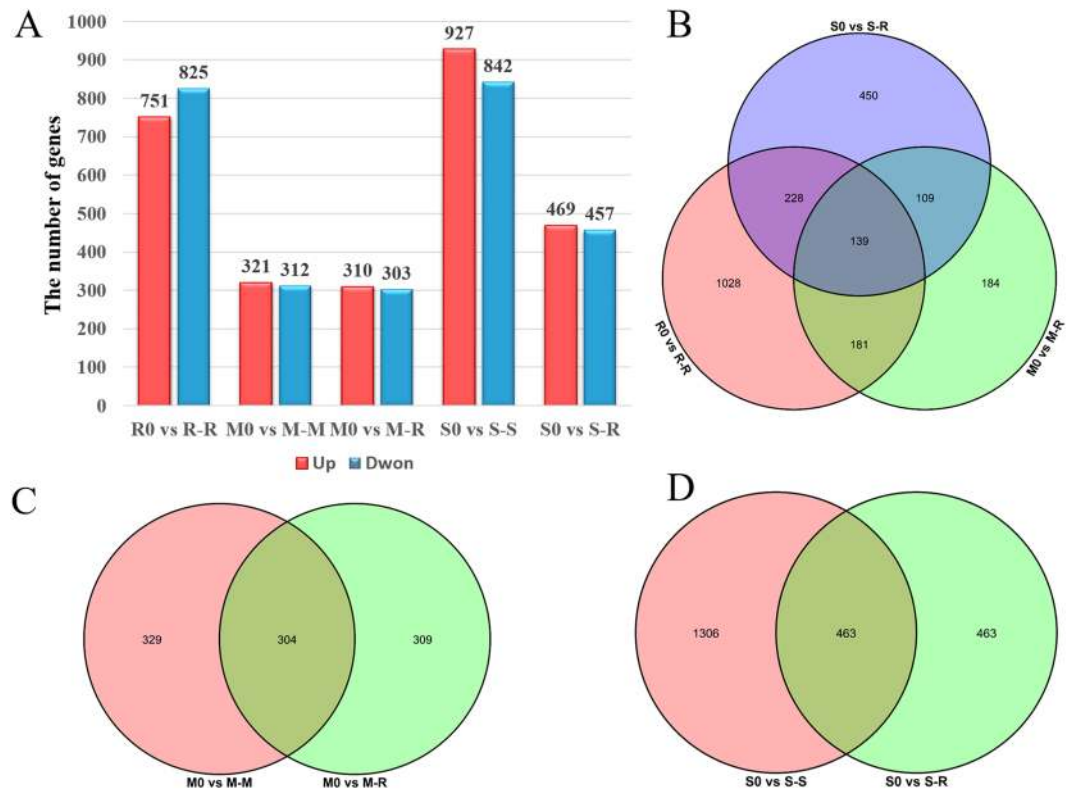


**Figure 1.** Biological characteristics of strains from different hosts. **(A)** Colony morphology of rice IA, maize IA and soybean IA after culture on PDA for 48, 72 and 96 h and 6 days. Scale bars, 2 cm. **(B)** Diameter of mycelia of rice IA, maize IA and soybean IA cultured on PDA at 12, 24 and 36 h post-inoculation (hpi). Error bars,  $\pm$  standard error (SE) of means of raw data. The least significant difference (LSD) method was used; different lower case letters (a–c) indicate significant differences among three strains at  $P < 0.05$ . **(C)** Pathogenicity of rice IA, maize IA and soybean IA. Four-week-old rice cultivar 93–11 leaves were inoculated onto a block of PDA containing fresh medium (diameter 0.5 cm); controls were inoculated with fresh PDA. Typical leaves were photographed at 48 hpi. Scale bars, 1 cm. **(D)** Histogram depicting spot area caused by rice IA, maize IA and soybean IA under infection of rice, maize and soybean leaves at 48 hpi. Error bars,  $\pm$  SE of means of raw data. The least significant difference (LSD) method was used; different lower case letters (a–c) indicate significant differences among three strains at  $P < 0.05$ .

genes were selected randomly for quantitative polymerase chain reaction (qPCR) analysis, and most of the qPCR results were consistent with those of RNA-Seq data (Supplementary Fig. S3). Thus, the RNA-Seq data are reliable.

**Differentially expressed genes during host infection.** A detailed comparative analysis of differentially expressed genes (DEGs) was performed to investigate the mechanisms underlying the pathogenesis of *R. solani* AG11A during the infection of various hosts. A total of 2,319 DEGs were detected and comparison of the strains during rice infection resulted in the identification of 1,576, 613, and 926 DEGs in R0 vs. R-R, M0 vs. M-R, and S0 vs. S-R (Fig. 2A), respectively. Therefore, during infection of rice, more genes were involved in the regulation of rice IA than in maize IA and soybean IA. Surprisingly, few DEGs (139, 5.99%) were common to the three strains; 1,028, 184, and 450 genes were specifically regulated in R-R, M-R, and S-R, respectively, which accounts for 71.67% of the total (Fig. 2B). During infection of the same host (rice), the gene expression profiles of *R. solani* strains varied markedly, suggesting the co-evolution of strains with their original hosts. Next, DEGs in strains of different ecotypes during infection of the original host and non-original host (rice) were analysed. During infection of rice and maize by maize IA, a total of 942 DEGs were detected, 304 were commonly expressed, and 329 and 309 DEGs were specifically expressed in M-M and M-R, respectively (Fig. 2C). In contrast, 2,232 DEGs were detected in S-S and S-R; specifically, 1,306 and 463 genes were regulated in S-S and S-R, respectively and 463 DEGs were commonly expressed (Fig. 2D). Interestingly, the total number of DEGs in S-S and S-R was markedly higher than in M-R and M-M, suggesting that more DEGs participated in soybean IA infection. More DEGs were found upon infection of the original host than the non-original host in maize IA and soybean IA. These results suggest the gene expression pattern of *R. solani* AG11A differed according to the host plant and possesses considerable host-adaptation ability, which could likely explain its wide host range.

**Gene ontology enrichment analysis of DEGs.** Gene ontology (GO) enrichment analysis of cDNA libraries was performed to classify the function of DEGs during host infection and to gain insight into their involvement in biological, molecular, and cellular processes. To obtain a global view of differences in the transcriptome, the number of GO categories with increased and decreased mRNA levels was compared (Fig. 3). Among them, cellular processes, metabolic processes, single-organism processes, cell part, cell and catalytic activity accounted for the majority of DEGs in all comparisons. In addition, metabolic processes and catalytic



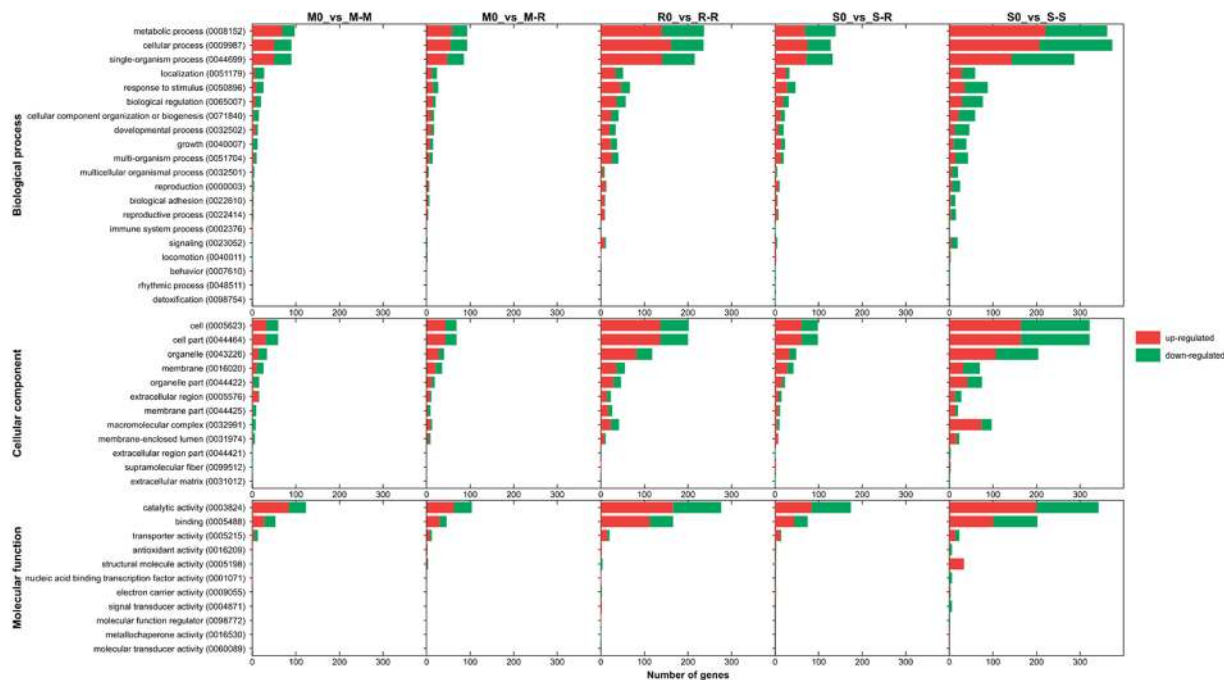
**Figure 2.** Overview of serial analysis of DEGs identified by pairwise comparison of eight transcriptomes. In total, 2,319 DEGs were found during different host infections with a  $|\log_2FC| \geq 1$  and  $FDR < 0.05$ . **(A)** Numbers of DEGs identified in R0 vs. R-R, M0 vs. M-M, M0 vs. M-R, S0 vs. S-S and S0 vs. S-R. **(B)** Venn diagrams of DEGs in three strains during infection of rice. **(C)** Venn diagrams of DEGs in maize IA during infection of rice and maize. **(D)** Venn diagrams of DEGs in soybean IA during infection of rice and soybean.

activity were both pathogenicity-associated. Overall, 185 genes related to catalytic activity were up-regulated in R-R, almost three-fold higher than M-R, for which only 63 genes were up-regulated in this term. In addition, 199 genes of S-S were up-regulated in catalytic activity, almost two-fold the amount of S-R. Surprisingly, 220 genes for S-S were up-regulated in metabolic processes, while only 74 genes were up-regulated in S-R, similar to the pattern for other processes (e.g., cellular processes, single-organism processes, cell part, cell). Differential expression of up-regulated and down-regulated genes in pathogenicity-associated categories suggests more complex mechanisms of gene regulation during infection of the original host than the non-original host in rice IA and soybean IA.

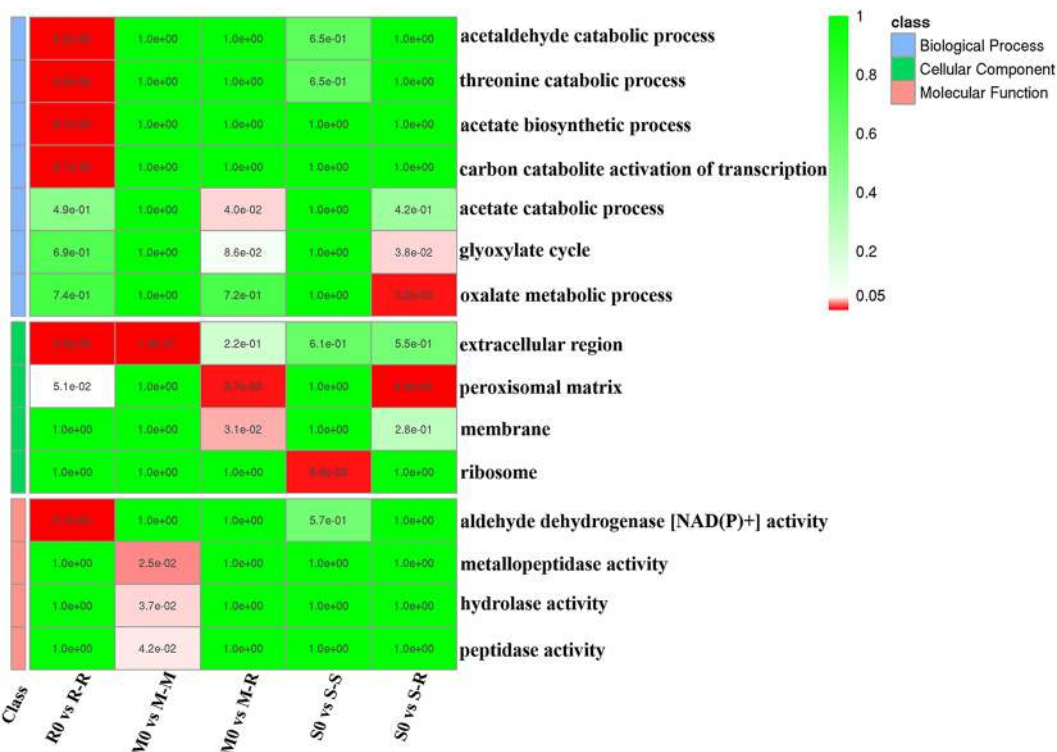
Analysis of R-R, M-R, and S-R during rice infection showed that the GO terms ‘acetaldehyde catabolic process’, ‘extracellular region’, and ‘aldehyde dehydrogenase activity’ were significantly ( $q < 0.05$ ) enriched in R-R (Fig. 4). However, ‘peroxisomal matrix’, which is closely related to the pathogenicity of filamentous fungi, was significantly enriched in M-R and S-R, and ‘oxalate metabolic process’ was only highly significantly enriched in S-R. The oxalate metabolic process is associated with pathogenicity because several pathogens (e.g., *M. oryzae* and *F. graminearum*) secrete oxalate to acidify host tissues and sequester calcium from host cell walls<sup>41,42</sup>. The GO term ‘extracellular region’ (associated with secretion) was highly significantly enriched in M-M compared to M-R, whereas ‘peroxisomal matrix’ was highly significantly enriched in M-R compared to M-M (Fig. 4). The GO terms of soybean IA during infection of soybean and rice were compared. The results showed that ‘ribosome’ (associated with thermotolerance and virulence in pathogenic fungi<sup>43</sup>) was highly significantly enriched in S-S, and ‘oxalate metabolic process’ and ‘peroxisomal matrix’ were highly significantly enriched in S-R (Fig. 4). Therefore, ‘peroxisomal matrix’ was the most abundant term in all strains during rice infection. In contrast, ‘extracellular region’ was the most abundant term used during maize infection. These results show that, although all three strains belong to the AGIIA subgroup, the GO terms (particularly those associated with pathogenicity) enriched in DEGs differ considerably according to host plant.

**Alternative splicing.** A comprehensive analysis of alternative splicing events in eight *R. solani* AGIIA transcriptomes was performed using high-throughput RNA-seq. A total of 11,341(4,462), 11,030 (4,471), 10,871 (4,465), 10,986 (4,355), 11,126 (4,468), 10,619 (4,381), 11,020 (4,463) and 10,794 (4,402) alternative splicing events (genes) of six AS types (FPKM > 0.5) were detected in R-R, R0, M-M, M0, M-R, S-S, S0, and S-R, respectively (Supplementary Table S4). Moreover, an average of 49.5% of genes in *R. solani* AGIIA underwent alternative splicing in this study. To the best of our knowledge, this is the first study to report a large number of alternative splicing events and a high proportion of AS genes in *Rhizoctonia*, both of which are markedly higher



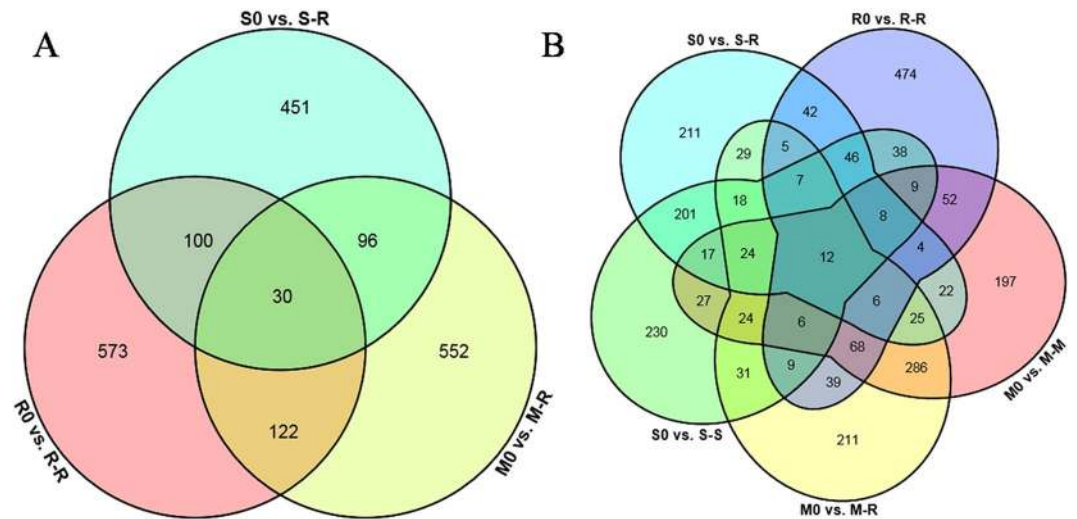


**Figure 3.** Numbers of GO categories with increased and decreased mRNA levels. Statistically significant data were grouped into broad GO categories (by TAIR). Bars represent number of up-regulated (red) and down-regulated (green) probe sets.



**Figure 4.** GO term enrichment analysis of strains during infection of different hosts. GO terms were enriched with a Hypergeometric test, and q-values were adjusted with FDR < 0.05. Scale represents p-value after homogenization.

than the equivalent values for *R. solani* AG11B<sup>32</sup>. The following six classes of alternative splicing were predicted among the eight transcriptomes: exon skipping, intron retention, alternative 5' splice site, alternative 3' splice site, alternative first exon and alternative last exon (Supplementary Fig. S4). Intron retention was the most abundant



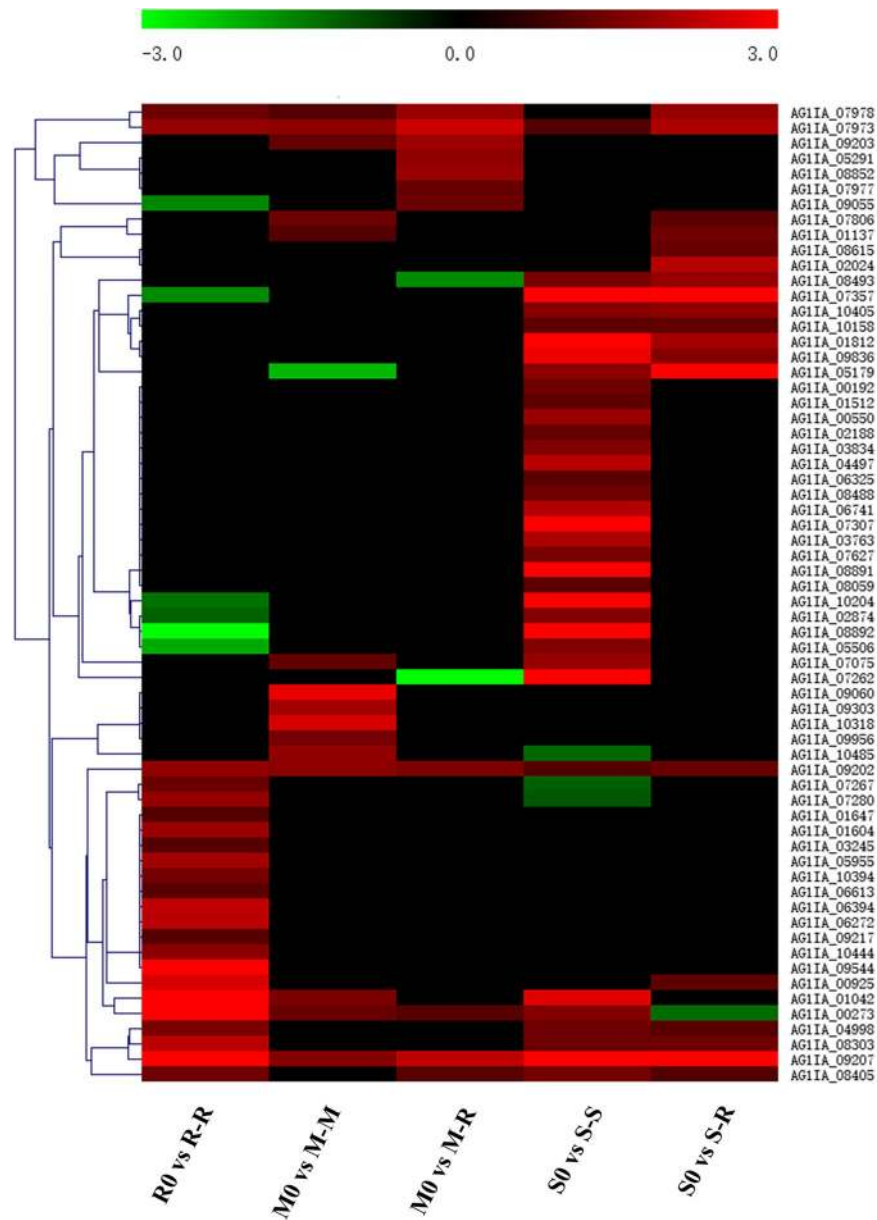
**Figure 5.** Genes underwent alternative splicing only during host infection compared to vegetative mycelia. (A) Venn diagrams of genes underwent alternative splicing only during infection of rice. (B) Venn diagrams of genes underwent alternative splicing during infection of rice, maize and soybean.

type of alternative splicing. For which *AG1IA\_03552* was predicted as a candidate effector in *R. solani* AG1IA<sup>28</sup> was identified alternative spliced in R-R and the splice variant were obtained for further research on the pathogenic mechanisms affected by alternative splicing (Supplementary Fig. S5).

Alternative splicing likely regulates the virulence and gene expression of pathogenic fungi<sup>44–46</sup>. In previous studies, secreted proteins, CAZymes, cysteine-rich proteins, ABC transporters, and G protein-coupled receptors (GPCRs) have been identified as pathogenesis-associated factors in *R. solani* AG1IA which involved in pathogenicity<sup>28</sup>. Therefore, we performed statistical analysis of the number of alternative splicing events in these five pathogenesis-associated factors which showed an average proportion of 16.41% have been predicted to be alternative spliced (Supplementary Table S5). Intron retention was the most abundant type of alternative splicing of genes encoding pathogenesis-associated factors. The AS events in above mentioned factors might suggest the development of new pathogenic mechanisms during the co-evolution of strains with their hosts.

The genes that underwent alternative splicing only during host infection compared with non-host (vegetative mycelia) were subjected to further analysis. Surprisingly, during rice infection, 573, 552, and 451 genes underwent alternative splicing in R0 vs. R-R, M0 vs. M-R, and S0 vs. S-R, specifically and respectively, whereas 30 genes underwent alternative splicing in all three comparisons (Fig. 5A). Among the 30 common AS genes, three genes (*AG1IA\_06529*, *AG1IA\_05348*, and *AG1IA\_02459*) were annotated in the PHI database. Surprisingly, twelve genes underwent alternative splicing during infection of all three hosts, eight of which were up-regulated (Fig. 5B). Among these eight genes, except for one hypothetical protein (*AG1IA\_01513*), seven were predicted to be related to pathogenicity, including a cytochrome P450-domain-containing protein (*AG1IA\_00285*) and a glycoside hydrolase family 16 protein (*AG1IA\_09291*) (Supplementary Table S6). Therefore, genes that undergo alternative splicing may play critical roles in the regulation of gene expression, the infection process and environmental adaptation by *R. solani* AG1IA. Further research on these candidate genes should seek to determine how their regulation via alternative splicing regulates the virulence of *R. solani* AG1IA.

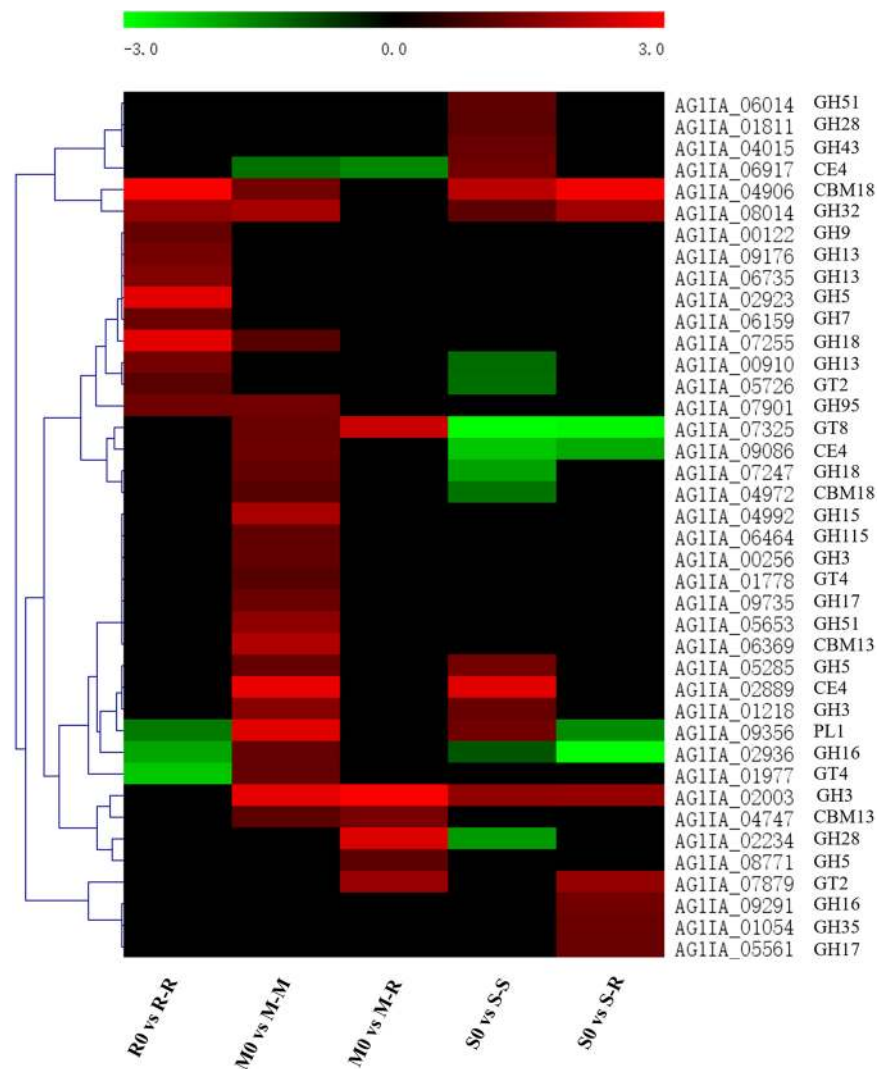
**Secreted proteins and effectors.** We also focused on known pathogenicity-related genes in *R. solani* AG1IA including those encoding secreted proteins, or candidate effectors<sup>28</sup>. Overall, 347 secreted proteins were differentially expressed during the infection of various hosts (Supplementary Fig. S6). Among them, six genes (*AG1IA\_06531*, *AG1IA\_09202*, *AG1IA\_07681*, *AG1IA\_07973*, *AG1IA\_09207* and *AG1IA\_09720*) with two genes (*AG1IA\_09369* and *AG1IA\_10043*) were up-regulated and down-regulated under all infection conditions, respectively. The six genes up-regulated under host infection conditions of different hosts are likely to play an important role in pathogenic mechanisms in the PHI and could be used for further research. Among them, the up-regulation of *AG1IA\_09202* under different infection conditions was confirmed through qRT-PCR (Supplementary Fig. S3). Effectors are known to contribute to disease establishment in the host, disturbing host plant physiology, and most are secreted proteins. In this study, effector P (<http://effectorp.csiro.au/>) was used to predict candidate secreted effectors among secreted proteins<sup>47</sup>. A total of 104 candidate effectors were differentially expressed (Supplementary Table S7), and 64 up-regulated candidate secreted effectors of *R. solani* AG1IA were further analysed (Fig. 6). Among them, most genes showed different expression patterns under different host infections; in contrast, three genes (*AG1IA\_07973*, *AG1IA\_09202* and *AG1IA\_04998*) were significantly up-regulated, but only *AG1IA\_10043* was significantly down-regulated under all infection conditions. Notably, 23, 15, 11, 35 and 19 genes were up-regulated in R-R, M-M, M-R, S-S and S-R, respectively. According to our data, the number of up-regulated candidate effectors was higher in the original hosts than non-original hosts, as particularly evident in rice IA and soybean IA. Among these up-regulated candidate effectors, 13, 5, 4, 18 and 4



**Figure 6.** Heat map of up-regulated candidate secreted effectors during infection of various hosts. Sixty-four up-regulated candidate secreted effectors of *R. solani* AG1IA were analysed by prediction algorithms; a hierarchical clustering statistic method based on Pearson's correlation (correlation  $\geq 98\%$ ) and average linkage was used. Red indicates a significant increase in gene expression, black indicates no change in expression and green indicates a significant decrease in gene expression ( $|\log_2\text{FC}| \geq 1$ ). Colour scale represents  $\log_2$  fold change values from -3 to 3.

genes were differentially and specifically up-regulated in R-R, M-M, M-R, S-S and S-R, respectively. Furthermore, we analysed the function of these 104 candidate effectors; only 27 genes were found to have predictable functions, while 77 candidate effectors had no homology to known effectors in the NCBI database and were annotated as hypothetical proteins in this study. Additionally, 17 genes of the abovementioned 27 genes were up-regulated during host infection (Supplementary Table S8).

**Expression of key hydrolases.** CAZymes, which include glycosyl transferases (GTs), glycoside hydrolases (GHs), polysaccharide lyases (PLs), carbohydrate esterases (CEs) and non-catalytic carbohydrate-binding modules (CBMs), have been a focus of many research studies because of their importance in the penetration and successful infection by fungal pathogens<sup>28</sup>. In this study, 40 putative CAZymes of DEGs, which were up-regulated during the infection of various hosts, were further analysed and are likely to be involved in plant infection or survival (Fig. 7). In total, 29 genes (7 in R-R, 12 in M-M, 3 in M-R, 4 in S-S and 3 in S-R) were specifically up-regulated, and 6 genes (1 in R-R and 5 in S-S) were specifically down-regulated. During the infection of rice, all up-regulated DEGs coding for CAZymes in R-R, M-R and S-R belonged to the GH, CBM and GT families, but

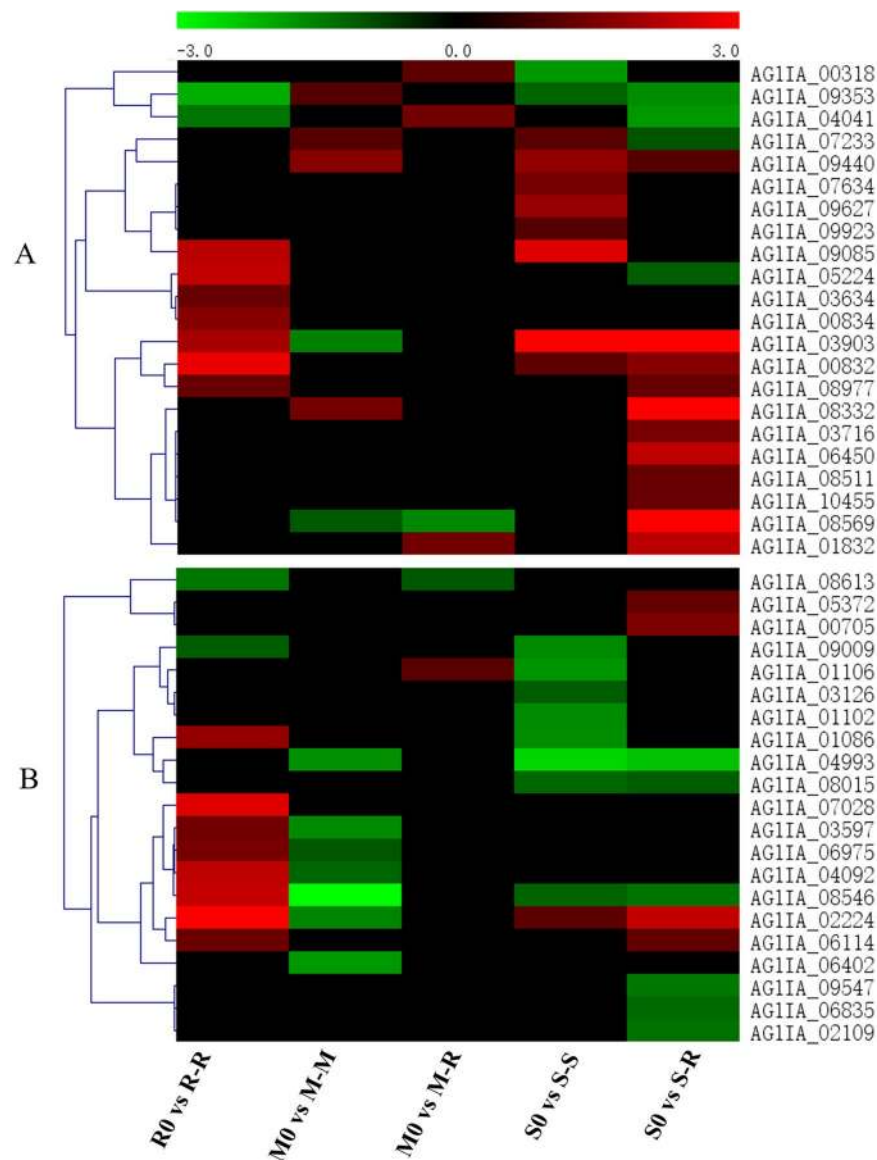


**Figure 7.** CAZyme family genes expressed in rice IA, maize IA and soybean IA during infection of various hosts. In total, 40 up-regulated genes were clustered; of these, 29 genes showed a significant host-specific up-regulated expression pattern during the infection of different hosts. A hierarchical clustering statistic method based on Pearson's correlation (correlation  $\geq 98\%$ ) and average linkage was used. Red indicates a significant increase in gene expression, black indicates no change in expression and green indicates a significant decrease in gene expression ( $|\log_2FC| \geq 1$ ). Colour scale represents  $\log_2$  fold change values from  $-3$  to  $3$ .

no DEGs belonged to the CE families. Interestingly, eleven putative CAZymes in R-R were up-regulated which were more than in M-R and S-R. Twenty-three putative CAZymes were up-regulated during M-M infection, the highest for any conditions and nearly four-fold higher than M-R. Similarly, the number of putative CAZymes up-regulated in maize IA during maize infection was nearly two-fold higher than during rice infection. However, the number of up-regulated CAZymes in S-S was also nearly 1.5-fold that of S-R. From the data above, more up-regulated hydrolase genes participate in the infection of their original host compared to the non-original host. Interestingly, some CAZymes up-regulated in R-R and M-M were significantly down-regulated in S-S, which might indicate different gene regulation patterns for CAZymes in rice IA, maize IA and soybean IA. To better understand the pathogenesis of CAZymes during infection of different hosts, up-regulated genes encoding putative enzymes involved in the degradation of plant cell wall components (mainly GHs and PLs) were further classified (Supplementary Table S9). We conclude that CAZymes that mediate cellulose, hemicellulose and starch could be more necessary during rice infection. Therefore, CAZymes that mediate the degradation of celluloses, hemicellulose and chitin could be likely more required during maize infection, and those that mediate pectin degradation could be more important during soybean infection.

**Secondary metabolites.** Fungal pathogens generally produce an array of secondary metabolites, some of which are involved in pathogenesis, and their production is frequently correlated with a specific stage of morphological differentiation<sup>48</sup>. The cytochrome P450 (CYP) monooxygenases of fungi are involved in many essential cellular processes and in primary and secondary metabolic pathways during host infection<sup>49</sup>. Sixty-eight putative

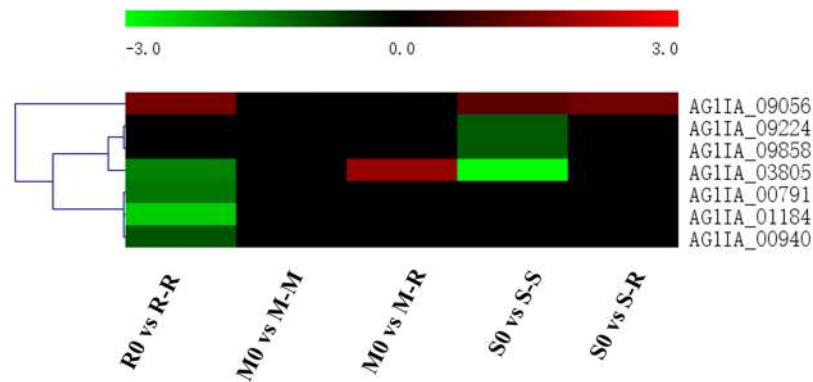




**Figure 8.** Differential expression of secondary metabolite-associated genes in *R. solani* AG1IA during infection of various hosts. (A) Expression profiles of up-regulated cytochrome P450 during infection of various hosts. (B) Expression profiles of differentially expressed candidate ABC transporters during infection of various hosts. A hierarchical clustering statistic method based on Pearson's correlation (correlation  $\geq 98\%$ ) and average linkage was used. Red indicates a significant increase in gene expression, black indicates no change in expression and green indicates a significant decrease in gene expression ( $|\log_2\text{FC}| \geq 1$ ). Colour scale represents  $\log_2$  fold change values from  $-3$  to  $3$ .

CYPs were identified in the *R. solani* AG1IA genome, and are likely to play a role in the metabolism and biotransformation of compounds toxic to the host<sup>28</sup>. Among them, twenty-two CYPs were differentially up-regulated during the infection of different hosts (Fig. 8A). As seen from the heat map, rice IA and soybean IA required a larger number of up-regulated CYPs during infection, whereas maize IA required relatively few. Therefore, CYP genes differentially expressed during infection likely play different roles in the metabolism and biotransformation of compounds toxic to the host.

Transporters export toxic molecules to the surrounding environment<sup>50</sup>. In toxin-producing fungi, transporters mediate the efflux of endogenously produced molecules. Among them, ABC transporters in plant pathogenic fungi are involved in the transport of plant-defense compounds and fungal pathogenicity factors. *R. solani* AG1IA is predicted to harbour 48 ABC transporters<sup>28</sup>. In this study, 21 ABC transporters were differentially expressed during the infection of different hosts (Fig. 8B). More DEGs encoding ABC transporters were up-regulated in R-R compared to M-R and S-R. Surprisingly, no ABC transporters were significantly up-regulated in M-M, but five were specifically down-regulated, similar to S-S. Therefore, our data suggest that ABC transporters could be more important during infection of rice than of maize or soybean.



**Figure 9.** Differential expression of G-protein-coupled receptors (GPCRs) in *R. solani* AG1IA during infection of various hosts. Seven DEGs of G-protein-coupled receptors (GPCRs) were analysed by prediction algorithms using the hierarchical clustering statistic method based on Pearson's correlation (correlation  $\geq$  98%) and average linkage. Red indicates a significant increase in gene expression, black indicates no change in expression and green indicates a significant decrease in gene expression ( $|\log_2\text{FC}| \geq 1$ ). Colour scale represents  $\log_2$  fold change values from  $-3$  to  $3$ .

**Proteins related to the MAPK pathway.** G-protein-coupled receptors (GPCRs) are involved in fungi-plant interactions and in the regulation of morphogenesis, mating, infection, and virulence in plant-pathogenic fungi<sup>51, 52</sup>. Thirteen GPCR-like genes have been identified in *R. solani* AG1IA<sup>28</sup>. In this study, seven predicted GPCR-like DEGs were found during the infection of different hosts (Fig. 9). Interestingly, *AG1IA\_09056*, which was predicted to contain a NCD3G nine-cysteine domain, was found to be up-regulated in R-R, S-S and S-R. *AG1IA\_03805*, which is a predicted homologue of *Homo sapiens* mPR-like GPCRs, was up-regulated in M-R but down-regulated in R-R and S-S. *AG1IA\_00940*, *AG1IA\_01184* and *AG1IA\_00791* were significantly and specifically down-regulated in R-R. *AG1IA\_00940* was predicted as an RGS (Regulator of G-protein Signaling) domain-containing protein which play pivotal roles in upstream regulation of fundamental biological processes in fungi, including vegetative growth, sporulation, mycotoxin/pigment production, pathogenicity, and mating<sup>53</sup>. While *AG1IA\_01184* was predicted to be a *stm1*-related protein which have been verified to participate in apoptosis-like cell death in yeast<sup>54</sup>. These differentially expressed GPCR genes could also reflect the adaptation of *R. solani* AG1IA to different host plants.

## Discussion

Host selection affects the population genetic structure of *R. solani* AG1IA, and the pathogenicity of *R. solani* strains of different groups may differ among host plants<sup>36, 55, 56</sup>. Rice IA, maize IA, and soybean IA were isolated from Xichang (altitude of 1500 m), Chengdu (altitude of 500 m) and Yaan (altitude of 1000 m), respectively. Hyphal fusion is only permitted between genetically similar heterokaryons sharing the same allele specificities at all *het* loci. When two interacting heterokaryons are genetically dissimilar, with different allelic specificities at some or all *het* loci, interaction hyphae cell death prevents dissimilar individuals from anastomosing<sup>57–59</sup>. Different ecotypes of rice IA, maize IA, and soybean IA could not anastomose with each other, indicating that they are genetically different. Cytological and morphological differences further suggested that the long-term co-evolution of strains with their host plants led to the generation of genetic differences. Rice IA, maize IA, and soybean IA can infect rice, but were more aggressive towards their original hosts, consistent with host specialization. Although *R. solani* AG1IA is a soil-borne fungus that can infect a broad range of hosts, each strain has adapted to its host in terms of morphology, physiology, and biochemistry. Moreover, rice is mainly cultivated in regions of high temperature and high humidity, whereas maize and soybean are mainly cultivated in arid regions. The abovementioned strains exist in specific niches in their various hosts due to ecological adaptation and long-term evolutionary selection.

Transcriptomic analyses of the strains during infection of different hosts were performed using Illumina next-generation sequencing. To the best of our knowledge, this is the first study to use comparative transcriptome profiling analysis to explore the pathogenic mechanisms of *R. solani* AG1IA during infection of different hosts. The currently available sequences enhance our understanding of the transcription profiles of *R. solani* AG1IA ecotypes during the infection of different hosts. In fact, the differences between our 25 sequenced strains from rice ecological cultivation areas were very small (unpublished). However, differences in the proportions of reads mapped in rice IA, maize IA and soybean IA indicated that their genomic sequences differ markedly.

To evaluate the pathogenic mechanisms of different strains of *R. solani* AG1IA during the infection of different hosts, the transcriptomes of the three strains were compared during the infection their original and non-original hosts. DEGs in R-R were significantly involved in acetaldehyde catabolism, the extracellular region and aldehyde dehydrogenase [NAD(P)<sup>+</sup>] activity, unlike the DEGs in M-R and S-R. In particular, aldehyde dehydrogenase is associated with pathogenesis and plays important roles in normal physiological processes and in response to stressors<sup>60</sup>. Pathogenic bacteria overcome the oxidative stress generated by host immune effectors by expressing a variety of aldehyde dehydrogenases<sup>61</sup>. What's more, DEGs in S-R were significantly involved in oxalate metabolism and the peroxisomal matrix, which markedly differs from the roles of DEGs in S-S. The peroxisome is closely

related to the pathogenicity of plant-pathogenic fungi, particularly filamentous fungi, by influencing fat metabolism and melanin synthesis<sup>62,63</sup>. Peroxisome metabolism affects invasion by many plant pathogens<sup>64</sup>. Interestingly, DEGs involved in the peroxisomal matrix were only significantly enriched during the infection of rice (R0 vs. R-R, M0 vs. M-R and S0 vs. S-R). Therefore, we hypothesized that the pathogenic mechanism of *R. solani* AG1IA is closely associated with the host plant.

Splicing in general, and alternative splicing in particular, is important for the regulation of gene expression and disease pathogenesis<sup>65</sup>. RNA-seq has revealed that a high proportion of multiexonic genes (>42.4% in rice<sup>66</sup>, 61% in *Arabidopsis*<sup>67</sup> and 95% in humans<sup>68</sup>) undergo alternative splicing. In contrast, a previous study revealed that an average of 6.4% of fungal genes are affected by this process<sup>46</sup> while the proportion of alternatively spliced (AS) genes in *Trichoderma longibrachiatum*<sup>69</sup> is markedly higher (48.9%) than in *M. grisea* (1.6%)<sup>70</sup>, *Ustilago maydis* (3.6%)<sup>71</sup> and *A. oryzae* (8.55%)<sup>72</sup>, suggesting that alternative splicing events in fungi are far more frequent than has been reported and could play an important role in fungal transcriptome diversity. Several proteins arising from specific AS transcripts have been reported, and genes with alternative splicing events were used to develop a transposon tagging system that was verified in a fungal pathogen<sup>73–75</sup>. Alternative splicing events are developmentally regulated and associated with responses to various environmental conditions<sup>76,77</sup>, the unexpectedly high proportion of AS genes especially pathogenesis-associated genes suggests that alternative splicing could be an important mechanism in the gene regulation of *R. solani* AG1IA. Further investigations are needed to confirm whether *R. solani* AG1IA produces a large number of AS forms to adapt to different host infections which might explain its wide host range. However, previous studies have confirmed the relationship between alternative splicing and fungal pathogenicity. For example, a virulence gene (UrRm75) in *U. maydis* and a ste12-like transcription factor that was essential for growth and pseudohyphal development in many fungi have also been reported to be alternatively spliced<sup>78,79</sup>. The eight genes predicted to be up-regulated during all host infections will be utilized in further research on the role of alternative splicing in the pathogenesis or adaption of *R. solani* AG1IA during different host infections.

Non-toxic plant fungal proteins are predicted to be secreted proteins<sup>80</sup>, namely virulence factors that facilitate colonization during PHIs. These secreted proteins likely constitute the majority of effectors that elicit host responses and mediate the ability of fungi to perceive and respond to the environment<sup>81,82</sup>. Secretion of effector proteins by plant pathogens such as rust fungi<sup>83</sup> and *Magnaporthe oryzae*<sup>84</sup> into host tissues promotes infection by manipulating host processes. There are few reports on the effectors of *Rhizoctonia solani*, although 25 candidate effectors were predicted in *R. solani* AG1IA four years ago<sup>28</sup>. Surprisingly, nearly half of the candidate effectors (44 genes out of 104 genes) were found to be specifically up-regulated under different host infections and 17 have homology protein domain with known effectors. Indeed, effectors may have undergone rapid evolutionary changes, which may account for the varying virulence of the various AGs, ISGs, and ecotypes of *R. solani*, as well as their ability to infect a wide range of plants<sup>28</sup>. For example, *AG1IA\_07075*, which was predicted as a CFEM domain-containing protein, was only up-regulated in M-M and S-S but not during rice infection. CFEMs are domains containing eight cysteines (e.g., Pth11 from *Magnaporthe grisea*) and are proposed to have important roles in fungal pathogenesis<sup>85</sup>. In contrast, *AG1IA\_03245*, which was predicted to be a kinase domain-containing protein, was only up-regulated in R-R; *AG1IA\_08303*, which was predicted as a peroxidase, family 2 domain-containing protein, was only up-regulated in rice IA and soybean IA. Thus, the differentially expressed candidate effectors are likely to be key factors in varying pathogenicity among host infections in *Rhizoctonia solani* AG1IA.

There is a clear relationship between CAZymes and the infection process in fungi<sup>22,86–88</sup>. Many CAZymes are involved in the biosynthesis, modification, binding, and catabolism of carbohydrates. Among them, CAZymes of GHs and PLs are cell wall-degrading enzymes that play important roles in plant biomass decomposition by fungi and bacteria<sup>89</sup>. Pathogens secrete numerous cell wall-degrading enzymes that breach the plant cell wall and use them as sources of nutrients. CAZyme expression markedly differ during the infection of different hosts<sup>28</sup> and some are even specifically up- and down-regulated. GH families (e.g., GH3, GH5, GH7, GH9, and GH17) play important roles in the fungal degradation of celluloses with different enzymatic activity, primarily  $\beta$ -1,3-1,4-glucanase,  $\beta$ -1,4-glucosidase,  $\beta$ -1,3 glucosidase,  $\beta$ -1,4-endoglucanase and cellobiohydrolase in this study. While GH families, such as GH15, GH16, GH35, GH51 and GH95, are involved in hemicellulose degradation through the activity of various enzymes, mainly  $\alpha$ -fucosidase,  $\alpha$ -glucuronidase and  $\alpha$ -arabinofuranosidase as well as  $\beta$ -1,4-galactosidase and GH18, which are involved in chitin degradation, GH13, GH15 and GH18 families are involved in starch degradation<sup>90</sup>. Interestingly, the GH13 member  $\alpha$ -amylase catalyses the hydrolysis of starch and sucrose in crops and has been previously studied<sup>28</sup> but genes in the GH13 family, such as *AG1IA\_09176*, *AG1IA\_06735* and *AG1IA\_00910*, were only specifically up-regulated in R-R in this study. In addition, *AG1IA\_01811*, *AG1IA\_04015* and *AG1IA\_09356* belong to families, such as GH28, GH43 and PL1, which play a role in pectin breakdown and were highly and specifically expressed in S-S compared to S-R, in agreement with the fact that the cell walls of dicots contain higher levels of pectin than monocots<sup>91</sup>. The cell wall composition of dicots and monocots differs, particularly in terms of the proportion of pectin and hemicelluloses, and CAZyme family diversity is correlated with host specificity<sup>92,93</sup>. Thus, our data are consistent with the hypothesis that pathogens of monocots are better adapted for degradation of monocot cell walls while pathogens of dicots are better adapted for degradation of dicot cell walls, reflecting host preferences<sup>90</sup>. To a certain extent, strains of *R. solani* AG1IA can adapt to different hosts and use different mechanisms of infection, depending on the host. These important findings will lay the groundwork for studies of the pathogen-host interaction of *R. solani* and the discovery of a large number of DEGs will enable further studies of the intrinsic differences in the pathogenicity in *R. solani* AG1IA towards different hosts. In turn, this may lead to the development of molecular breeding strategies for selecting species-specific resistant breeding materials in different hosts.

## Materials and Methods

**Origin, isolation and identification of strains.** Three sequenced strains of *R. solani* AG11A were used in this study. The three strains, rice IA, maize IA, and soybean IA, were randomly isolated from heavily infected rice (*O. sativa* L.), maize (*Z. mays* L.) and soybean (*G. max*) plants in Sichuan province in areas different altitudes (Xichang, Chengdu, and Yaan). The infected samples were thoroughly washed with sterile water, cut into small pieces, surface-sterilized with 1% sodium hypochlorite solution, and plated on agar containing no added nutrients<sup>94</sup>. Target strains were obtained by hyphae tip transplantation, and the purified strains were transferred to a block of PDA (15 g/L, Amresco, Solon, OH, USA) cultured at 28 °C for 24 h and stored at 4 °C. The three ecotypes showed identical ITS1 and ITS4 rDNA sequences, microscopic appearances, and abilities to cause disease in maize and soybean cultivars<sup>95</sup>.

**Anastomosis test.** An anastomosis test was performed according to the method of S. Stepniewska-Jarosz<sup>96</sup>. Cultures of three isolated *R. solani* stains were paired with known *R. solani* AG11A obtained from the Erxun Zhou, South China Agricultural University. The comparisons were performed on microscope slides with a sterile piece of cellophane predipped in a PDA agar medium before use. Three 5-mm disks from young *R. solani* cultures were placed on the slide (the tester in the centre and two unidentified *R. solani* isolates). The slides were placed in a moist chamber and incubated in the dark at 28 °C for 18–20 h until the hyphae from opposite disks overlapped. The overlapping portion was examined under the microscope for hyphal fusion observation.

**Growth and mycelia.** Small pieces of mycelium were removed from the margin of the mycelia and added to fresh solid PDA at 28 °C for 1 week. This was repeated three times. The mycelium diameter was measured every 12 h to assess growth, and changes in mycelia colour were observed.

**Disease index test.** The detached leaf method was used to assess the pathogenicity of rice IA, maize IA, and soybean IA<sup>97</sup>. Samples of cultivar 93–11 (rice), ZhengHong 505 (maize) and NanDou 5 (soybean) of similar sizes and leaf ages were placed on moist filter paper in a porcelain dish after being washed in sterile water. A block of PDA containing 2-day-old mycelia from *R. solani* AG11A was placed in the middle of the leaf surface, covered with plastic wrap, and placed at 28 °C with 78% humidity. The leaves were observed every 12 h from 0–48 h post infection and speckles of disease were counted. Control leaves were inoculated with fresh PDA. Each experiment was repeated in triplicate. ImageJ software (NIH, Bethesda, MD, USA) was used to determine the spot area<sup>98</sup>.

**Sample treatment, total RNA extraction, cDNA library construction, and sequencing.** Rice leaves inoculated with rice IA (R-R), maize IA (M-R) and soybean IA (S-R), maize leaves inoculated with maize IA (M-M), and soybean leaves inoculated with soybean IA (S-S) were cultured in individual petri dishes at 28 °C with 78% humidity. Rice IA (R0), maize IA (M0), and soybean IA (S0) were transferred to PDA at the same time and under the same culture conditions as in the experiment group. When more than half of the leaves exhibited lesions, mycelia were removed from the leaves by scraping<sup>28</sup> and frozen in liquid N<sub>2</sub>. Total RNA was extracted using a Fungal RNA Kit (Omega, Norcross, GA, USA), following the manufacturer's recommendations, and suspended in DEPC-treated water. RNA quantity was assessed with a NanoDrop2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA). After all of the samples were confirmed by sequencing, they were sent to Beijing Biomarker Technologies (Beijing, China) for cDNA library construction and sequencing. Libraries were prepared using NEBNext mRNA Library Prep Master Mix Set for Illumina (NEB, E6110) and NEBNext Multiplex Oligos for Illumina (NEB, E7500) according to manufacturer's instruction and paired-end sequencing was done using Illumina Genome Analyzer II technology (Illumina) to obtain reads of 100 bp length. Original paired-end reads were identified, and low-quality and primer/adaptor-contaminated reads were filtered using NGSQC Toolkit<sup>99</sup>. The high quality reads were used for downstream analysis after quality filtering. Then, transcripts were aligned to scaffold genome sequences using BLAST<sup>100</sup>. The sequencing data generated in this study have been deposited in Sequence Read Archive (SRA) data base and under the accession number SRP106262.

**Alternative splicing analysis and validation.** Alternative splicing events in *R. solani* AG11A genes were identified using transcriptome sequences from the published *R. solani* AG11A genome<sup>29</sup>. The alternative splicing events were constructed by assembling transcripts of the genome (Genome-guided Trinity for Gene Structure Annotation, [http://trinityrnaseq.sourceforge.net/genome\\_guided\\_trinity.html](http://trinityrnaseq.sourceforge.net/genome_guided_trinity.html)). RNA-seq reads were aligned to the genome using GSNAP<sup>101</sup>. The genome-based transcripts were assembled using Trinity<sup>102</sup>. With GMAP and PASA<sup>103</sup>, AS events in genes were identified. RT-PCR was used for alternative splicing validation and cDNA was used as a template to perform RT-PCR. Primers of *AG11A\_03552* are listed (Supplementary Table S10).

**Gene ontology enrichment and pathway analysis.** The GO terms or metabolic pathways significantly enriched in DEGs were identified by performing functional enrichment analyses. The Nr BLAST results were imported into the Blast2 GO software for annotation with GO terms. This analysis involved mapping of all of annotated genes to GO terms in the database and enumeration of the genes associated with each term. Then a Perl script was used to plot the GO functional classifications for unigenes with a GO term hit to visualize the distribution of gene functions. The obtained annotation was enriched and refined using Top GO (R package) with the “elim” method and Kolmogorov-Smirnov test. GO terms with a corrected *q* value of <0.05 were considered significantly enriched in DEGs.

**Differential expression analysis.** In this study, gene expression (false discovery rate [FDR ≤ 0.01]) was assessed using EdgeR<sup>104</sup> software. The FDR was used as a key indicator for screening significantly differentially expressed genes. The transcripts with log<sub>2</sub> fold change ≥ 1 (up-regulated genes) and ≤ (−1) (down-regulated genes) with FDR cut off ≤ 0.05 were considered as significantly differentially expressed genes.



**Prediction, annotation, and analysis of secreted proteins.** Secretomes were predicted following two rules: the gene contains a signal peptide at the N-terminus, as predicted by at least two types of software, and the gene lacks a transmembrane helix as predicted by at least one type of software. Prediction software included SignalP 4.0 (<http://www.cbs.dtu.dk/services/SignalP>)<sup>105</sup>, TargetP 1.1 (<http://www.cbs.dtu.dk/services/TargetP>)<sup>106</sup>, Phobius (<http://phobius.sbc.su.se>), PrediSi (<http://www.predisi.de/index.html>) and TMHMM 2.0 (<http://www.cbs.dtu.dk/services/TMHMM>). Extracellular proteins were predicted using Wolfpsort<sup>107</sup>. Secreted proteins with a maximum of 200 amino acids and at least 4% cysteine residues were considered cysteine-rich proteins. Effector P (<http://effectorp.csiro.au/>) was also used to predicted genes which show probability for being secreted effectors. Multi Experiment Viewer\_ENREF\_110 software was used for making heat map.

**Candidate gene expression analysis via qRT-PCR.** A total of twelve candidate genes belonging to different categories from RNA-seq were used for gene expression analysis via qPCR. First-strand cDNA was synthesized from total RNA using a Transcript First Strand cDNA Synthesis Kit (Roche, Upper Bavaria Germany). qPCR was performed using the SoAdvanced Universal SYBR Green Supermix (Bio-Rad, Hercules CA, USA). The qPCR reactions were performed in a final volume of 10  $\mu$ L containing 5  $\mu$ L of 2  $\times$  MasterMix, 0.5  $\mu$ L of 100  $\mu$ M of each primer, 1  $\mu$ L of ddH<sub>2</sub>O and 3  $\mu$ L of cDNA. The reactions occurred at 95 °C for 30 sec, followed by 40 cycles of 95 °C for 10 sec, 58 °C for 30 sec, and melting curve analysis from 65 °C to 95 °C at 0.5 °C increments. Primers for qPCR were designed based on our predicted gene sequences (Supplementary Table S11). Primers for 18 S were used for internal controls, and all of the qPCR primers were tested with RT-PCR before use. Fold changes were determined by the  $2^{-\Delta\Delta C_t}$  method based on three technical replicates per sample. All qPCR reactions were repeated at least three times.

## References

- Sneh, B., Burpee, L. & Ogoshi, A. *Identification of Rhizoctonia species* (APS press, 1991).
- PARMETER JUNIOR, J., Sherwood, R. & Platt, W. Anastomosis grouping among isolates of Thanatephorus cucumeris. *Phytopathology* **59**, 1270–1278 (1969).
- Ogoshi, A. Ecology and pathogenicity of anastomosis and intraspecific groups of Rhizoctonia solani Kuhn. *Annual review of phytopathology* **25**, 125–143 (1987).
- Yang, G. & Li, C. *General description of Rhizoctonia species complex* (INTECH Open Access Publisher, 2012).
- Yang, G., Conner, R., Chen, Y., Chen, J. & Wang, Y. Frequency and pathogenicity distribution of Rhizoctonia spp. causing sheath blight on rice and banded leaf disease on maize in Yunnan, China. *Journal of Plant Pathology*, 387–392 (2008).
- Ciampi, M. *et al.* Genetic structure of populations of Rhizoctonia solani anastomosis group-1 IA from soybean in Brazil. *Phytopathology* **98**, 932–941 (2008).
- Taheri, P., Gnanamanickam, S. & Höfte, M. Characterization, genetic structure, and pathogenicity of Rhizoctonia spp. associated with rice sheath diseases in India. *Phytopathology* **97**, 373–383 (2007).
- Sridevi, G., Parameswari, C., Sabapathi, N., Raghupathy, V. & Veluthambi, K. Combined expression of chitinase and  $\beta$ -1, 3-glucanase genes in indica rice (*Oryza sativa* L.) enhances resistance against Rhizoctonia solani. *Plant science* **175**, 283–290 (2008).
- Prasad, B. & Eizenga, G. Rice sheath blight disease resistance identified in *Oryza* spp. accessions. *Plant Disease* **92**, 1503–1509 (2008).
- Kunihiro, Y. *et al.* [QTL analysis of sheath blight resistance in rice (*Oryza sativa* L.)]. *Yi chuan xue bao=Acta genetica Sinica* **29**, 50–55 (2002).
- Qingxian, L. & Zhenjun, X. Studies on the biological characters of Rhizoctonia isolated from several crops in Jiangsu. *Journal of Yangzhou University (Agriculture and Life Sciences Edition)* **3** (1988).
- Zou, J. *et al.* Mapping quantitative trait loci controlling sheath blight resistance in two rice cultivars (*Oryza sativa* L.). *Theoretical and Applied Genetics* **101**, 569–573 (2000).
- Lehtonen, M., Ahvenniemi, P., Wilson, P., German-Kinnari, M. & Valkonen, J. Biological diversity of Rhizoctonia solani (AG-3) in a northern potato-cultivation environment in Finland. *Plant Pathology* **57**, 141–151 (2008).
- Wei, Y. *et al.* Haplotype variation and phylogeography of Rhizoctonia solani AG1-IA strains based on rDNA5, 8S-ITS and  $\beta$ -actin gene sequence analyses. *Mycological progress* **13**, 247–255 (2014).
- Das, S. *et al.* Genetic variability and pathogenicity of Rhizoctonia solani associated with black scurf of potato in New Zealand. *Plant Pathology* **63**, 651–666 (2014).
- Zhou, Q., Hwang, S., Fu, H., Strelkov, S. & Gossen, B. Genetic variation of Rhizoctonia solani isolates from canola in Alberta, Canada. *Canadian Journal of Plant Science* **94**, 671–681 (2014).
- Wang, L., Liu, L. M., Wang, Z. G. & Huang, S. W. Genetic Structure and Aggressiveness of Rhizoctonia solani AG1-IA, the Cause of Sheath Blight of Rice in Southern China. *Journal of Phytopathology* **161**, 753–762 (2013).
- Qi, Y., Liu, Y. & Rong, W. [RNA-Seq and its applications: a new technology for transcriptomics]. *Yi chuan = Hereditas/Zhongguo yi chuan xue hui bian ji* **33**, 1191–1202 (2011).
- Oh, Y. *et al.* Transcriptome analysis reveals new insight into appressorium formation and function in the rice blast fungus Magnaporthe oryzae. *Genome Biol* **9**, R85 (2008).
- Skibbe, D. S., Doehlemann, G., Fernandes, J. & Walbot, V. Maize tumors caused by Ustilago maydis require organ-specific genes in host and pathogen. *Science* **328**, 89–92 (2010).
- Gao, L., Tu, Z. J., Millett, B. P. & Bradeen, J. M. Insights into organ-specific pathogen defense responses in plants: RNA-seq analysis of potato tuber-Phytophthora infestans interactions (2013).
- Martinez, D. *et al.* Genome, transcriptome, and secretome analysis of wood decay fungus Postia placenta supports unique mechanisms of lignocellulose conversion. *Proceedings of the National Academy of Sciences* **106**, 1954–1959 (2009).
- O’Connell, R. J. *et al.* Lifestyle transitions in plant pathogenic Colletotrichum fungi deciphered by genome and transcriptome analyses. *Nature genetics* **44**, 1060–1065 (2012).
- Teixeira, P. J. P. L. *et al.* High-resolution transcript profiling of the atypical biotrophic interaction between Theobroma cacao and the fungal pathogen Moniliophthora perniciosa. *The Plant Cell Online* **26**, 4245–4269 (2014).
- Schmidtke, C. *et al.* Genome-wide transcriptome analysis of the plant pathogen Xanthomonas identifies sRNAs with putative virulence functions. *Nucleic acids research*, gkr904 (2011).
- ul Rehman, R., Hakeem, K. R., Tahir, I., Padder, B. A. & Teli, M. A. In *Crop Improvement* 269–293 (Springer, 2013).
- Losada, L. *et al.* Mobile elements and mitochondrial genome expansion in the soil fungus and potato pathogen Rhizoctonia solani AG-3. *FEMS microbiology letters* **352**, 165–173 (2014).
- Zheng, A. *et al.* The evolution and pathogenic mechanisms of the rice sheath blight pathogen. *Nature communications* **4**, 1424 (2013).

29. 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. (2014).
30. Wibberg, D. *et al.* Establishment and interpretation of the genome sequence of the phytopathogenic fungus *Rhizoctonia solani* AG1-IB isolate 7/3/14. *Journal of biotechnology* **167**, 142–155 (2013).
31. Wibberg, D. *et al.* Draft genome sequence of the sugar beet pathogen *Rhizoctonia solani* AG2-2IIIB strain BBA69670. *Journal of biotechnology* **222**, 11–12 (2016).
32. Wibberg, D. *et al.* Development of a *Rhizoctonia solani* AG1-IB Specific Gene Model Enables Comparative Genome Analyses between Phytopathogenic *R. solani* AG1-IA, AG1-IB, AG3 and AG8 Isolates. *PLoS one* **10**, e0144769 (2015).
33. Anderson, J. P. *et al.* Proteomic analysis of *Rhizoctonia solani* identifies infection-specific, redox associated proteins and insight into adaptation to different plant hosts. *Molecular & Cellular Proteomics* **15**, 1188–1203 (2016).
34. Zhu, C. *et al.* De novo transcriptome analysis of *Rhizoctonia solani* AG1 IA strain early invasion in *Zoysia japonica* root. *Frontiers in microbiology* **7** (2016).
35. Venu, R. *et al.* RL-SAGE and microarray analysis of the rice transcriptome after *Rhizoctonia solani* infection. *Molecular Genetics and Genomics* **278**, 421–431 (2007).
36. ChenJiaoZi, W., XiaoXia, H., Mei, Y. & ErXun, Z. Cross-pathogenicity of *Rhizoctonia* spp. from rice, maize and wheat on these three crops. *Journal of South China Agricultural University* **36**, 82–86 (2015).
37. Xiaoxia, H., Qiqi, C. & Zhengkai, P. Comparison of differences in biological characteristics of *Rhizoctonia* spp isolated from three kinds of crops. *Journal of Huazhong Agricultural University* (2012).
38. YI, R.-h, ZHU, X.-r & ZHOU, E.-x On the evaluation of different methods by artificially inoculating with rice sheath blight fungus (*Rhizoctonia solani*)[J]. *Journal of Guangzhou University (Natural Science Edition)* **3**, 005 (2003).
39. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
40. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology* **28**, 511–515 (2010).
41. Dutton, M. V. & Evans, C. S. Oxalate production by fungi: its role in pathogenicity and ecology in the soil environment. *Canadian journal of microbiology* **42**, 881–895 (1996).
42. Li, L. *et al.* MoPEX19, which is essential for maintenance of peroxisomal structure and woronin bodies, is required for metabolism and development in the rice blast fungus. *PLoS one* **9**, e85252 (2014).
43. Bhabhra, R. & Askew, D. Thermotolerance and virulence of *Aspergillus fumigatus*: role of the fungal nucleolus. *Medical mycology* **43**, 87–93 (2005).
44. Rodríguez-Kessler, M. *et al.* Isolation of UmRrm75, a gene involved in dimorphism and virulence of *Ustilago maydis*. *Microbiological research* **167**, 270–282 (2012).
45. Loftus, B. J. *et al.* The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. *Science* **307**, 1321–1324 (2005).
46. Grützmann, K. *et al.* Fungal alternative splicing is associated with multicellular complexity and virulence: a genome-wide multi-species study. *DNA research* **21**, 27–39 (2014).
47. Sperschneider, J. *et al.* EffectorP: predicting fungal effector proteins from secretomes using machine learning. *New Phytologist* **210**, 743–761 (2016).
48. Vogel, J. Unique aspects of the grass cell wall. *Current opinion in plant biology* **11**, 301–307 (2008).
49. Keller, N. P., Turner, G. & Bennett, J. W. Fungal secondary metabolism—from biochemistry to genomics. *Nature Reviews Microbiology* **3**, 937–947 (2005).
50. Slot, J. C. & Rokas, A. Multiple GAL pathway gene clusters evolved independently and by different mechanisms in fungi. *Proceedings of the National Academy of Sciences* **107**, 10136–10141 (2010).
51. Li, L., Wright, S., Krystofova, S., Park, G. & Borkovich, K. Heterotrimeric G protein signaling in filamentous fungi. *Annual review of microbiology* **61**, 423–452 (2006).
52. Schumacher, J. In *Botrytis—the Fungus, the Pathogen and its Management in Agricultural Systems* 247–267 (Springer, 2016).
53. Wang, Y. *et al.* Characterizations and functions of regulator of G protein signaling (RGS) in fungi. *Applied microbiology and biotechnology* **97**, 7977 (2013).
54. Ligr, M. *et al.* The proteasomal substrate Stm1 participates in apoptosis-like cell death in yeast. *Molecular biology of the cell* **12**, 2422–2432 (2001).
55. Carling, D., Kuninaga, S. & Brainard, K. Hyphal anastomosis reactions, rDNA-internal transcribed spacer sequences, and virulence levels among subsets of *Rhizoctonia solani* anastomosis group-2 (AG-2) and AG-BI. *Phytopathology* **92**, 43–50 (2002).
56. Yi, R., Liang, C., Zhu, X. & Zhou, E. Genetic diversity and virulence variation of rice sheath blight strains (*Rhizoctonia solani* AG-1 IA) from Guangdong Province. *Journal of Tropical and Subtropical Botany* **10**, 161–170 (2001).
57. Rayner, A. D. The challenge of the individualistic mycelium. *Mycologia*, 48–71 (1991).
58. Worrall, J. J. Somatic incompatibility in basidiomycetes. *Mycologia*, 24–36 (1997).
59. Glass, N. L. & Kaneko, I. Fatal attraction: nonself recognition and heterokaryon incompatibility in filamentous fungi. *Eukaryotic Cell* **2**, 1–8 (2003).
60. Singh, S. *et al.* Aldehyde dehydrogenases in cellular responses to oxidative/electrophilic stress. *Free Radical Biology and Medicine* **56**, 89–101 (2013).
61. Mydel, P. *et al.* Roles of the host oxidative immune response and bacterial antioxidant rubrerythrin during *Porphyromonas gingivalis* infection. *PLoS Pathog* **2**, e76 (2006).
62. Kimura, A., Takano, Y., Furusawa, I. & Okuno, T. Peroxisomal metabolic function is required for appressorium-mediated plant infection by *Colletotrichum lagenarium*. *The Plant Cell* **13**, 1945–1957 (2001).
63. Ramos-Pamplona, M. & Naqvi, N. I. Host invasion during rice–blast disease requires carnitine-dependent transport of peroxisomal acetyl-CoA. *Molecular microbiology* **61**, 61–75 (2006).
64. Wang, Z.-Y., Soanes, D. M., Kershaw, M. J. & Talbot, N. J. Functional analysis of lipid metabolism in *Magnaporthe grisea* reveals a requirement for peroxisomal fatty acid  $\beta$ -oxidation during appressorium-mediated plant infection. *Molecular plant-microbe interactions* **20**, 475–491 (2007).
65. Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476 (2008).
66. Zhang, G. *et al.* Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. *Genome research* **20**, 646–654 (2010).
67. Marquez, Y., Brown, J. W., Simpson, C., Barta, A. & Kalyna, M. Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. *Genome research* **22**, 1184–1195 (2012).
68. Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature genetics* **40**, 1413–1415 (2008).
69. Xie, B.-B. *et al.* Deep RNA sequencing reveals a high frequency of alternative splicing events in the fungus *Trichoderma longibrachiatum*. *BMC genomics* **16**, 1 (2015).
70. Ebbole, D. J. *et al.* Gene discovery and gene expression in the rice blast fungus, *Magnaporthe grisea*: analysis of expressed sequence tags. *Molecular plant-microbe interactions* **17**, 1337–1347 (2004).
71. Ho, E. C., Cahill, M. J. & Saville, B. J. Gene discovery and transcript analyses in the corn smut pathogen *Ustilago maydis*: expressed sequence tag and genome sequence comparison. *BMC genomics* **8**, 1 (2007).

72. Wang, B. *et al.* Survey of the transcriptome of *Aspergillus oryzae* via massively parallel mRNA sequencing. *Nucleic acids research* **38**, 5075–5087 (2010).
73. Zhao, C., Waalwijk, C., de Wit, P. J., Tang, D. & van der Lee, T. RNA-Seq analysis reveals new gene models and alternative splicing in the fungal pathogen *Fusarium graminearum*. *BMC genomics* **14**, 21 (2013).
74. Freitag, J., Ast, J. & Böcker, M. Cryptic peroxisomal targeting via alternative splicing and stop codon read-through in fungi. *Nature* **485**, 522–525 (2012).
75. Kempken, F. & Kück, U. restless, an active Ac-like transposon from the fungus *Tolypocladium inflatum*: structure, expression, and alternative RNA splicing. *Molecular and cellular biology* **16**, 6563–6572 (1996).
76. Zhao, C., Waalwijk, C., de Wit, P. J., Tang, D. & van der Lee, T. RNA-Seq analysis reveals new gene models and alternative splicing in the fungal pathogen *Fusarium graminearum*. *BMC genomics* **14**, 1 (2013).
77. Kalsotra, A. & Cooper, T. A. Functional consequences of developmentally regulated alternative splicing. *Nature Reviews Genetics* **12**, 715–729 (2011).
78. Rodriguez-Kessler, M. *et al.* Isolation of UmRm75, a gene involved in dimorphism and virulence of *Ustilago maydis*. *Microbiol Res* **167**, 270–282, doi:10.1016/j.micres.2011.10.007 (2012).
79. W S Hoi, J. & Dumas, B. Ste12 and Ste12-like proteins, fungal transcription factors regulating development and pathogenicity. *Eukaryot Cell* **9**, 480–485, doi:10.1128/EC.00333-09 (2010).
80. Kamoun, S. Groovy times: filamentous pathogen effectors revealed. *Current opinion in plant biology* **10**, 358–365 (2007).
81. de Jonge, R., Bolton, M. D. & Thomma, B. P. How filamentous pathogens co-opt plants: the ins and outs of fungal effectors. *Current opinion in plant biology* **14**, 400–406 (2011).
82. Dean, R. A. *et al.* The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* **434**, 980–986 (2005).
83. Petre, B., Joly, D. L. & Duplessis, S. Effector proteins of rust fungi. *Frontiers in plant science* **5**, 416 (2014).
84. Liu, W. *et al.* Recent progress in understanding PAMP-and effector-triggered immunity against the rice blast fungus *Magnaporthe oryzae*. *Molecular plant* **6**, 605–620 (2013).
85. Kulkarni, R. D., Kelkar, H. S. & Dean, R. A. An eight-cysteine-containing CFEM domain unique to a group of fungal membrane proteins. *Trends in biochemical sciences* **28**, 118–121 (2003).
86. Charaoui-Boukerzaza, S. & Hugouvieux-Cotte-Pattat, N. A family 3 glycosyl hydrolase of *Dickeya dadantii* 3937 is involved in the cleavage of aromatic glucosides. *Microbiology* **159**, 2395–2404 (2013).
87. Lombard, V., Ramulu, H. G., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic acids research* **42**, D490–D495 (2014).
88. Espagne, E. *et al.* The genome sequence of the model ascomycete fungus *Podospora anserina*. *Genome Biol* **9**, R77 (2008).
89. Amnuaykanjanasin, A. & Daub, M. E. The ABC transporter ATR1 is necessary for efflux of the toxin cercosporin in the fungus *Cercospora nicotianae*. *Fungal genetics and biology: FG & B* **46**, 146–158, doi:10.1016/j.fgb.2008.11.007 (2009).
90. Cantarel, B. L. *et al.* The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic acids research* **37**, D233–D238 (2009).
91. Boraston, A. B., Bolam, D. N., Gilbert, H. J. & Davies, G. J. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochemical Journal* **382**, 769–781 (2004).
92. Lagaert, S., Belien, T. & Volckaert, G. Plant cell walls: Protecting the barrier from degradation by microbial enzymes. *Seminars in cell & developmental biology* **20**, 1064–1073, doi:10.1016/j.semcdb.2009.05.008 (2009).
93. B C King1, K. D. W., Nicholas V Nenni2,5, Larry P Walker3, Gary C Bergstrom1, Donna M Gibson4\*. Arsenal of plant cell wall degrading enzymes reflects host preference among plant pathogenic fungi. *Biotechnology for Biofuels* **4** (2011).
94. Sinclair, J. *Rhizoctonia solani*: special methods of study. *Rhizoctonia solani: Biology and Pathology*, 199–217 (1970).
95. Ghosh, S., Gupta, S. K. & Jha, G. Identification and functional analysis of AG1-IA specific genes of *Rhizoctonia solani*. *Current genetics* **60**, 327–341 (2014).
96. S. S. pniewska-J, M. M. k. a. F. O. A. Tn *Forest pathology* Vol. 36 97–109 (2006).
97. Sayler, R. J. & Yang, Y. Detection and quantification of *Rhizoctonia solani* AG-1 IA, the rice sheath blight pathogen, in rice using real-time PCR. *Plant Disease* **91**, 1663–1668 (2007).
98. Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with ImageJ. *Biophotonics international* **11**, 36–42 (2004).
99. Patel, R. K. & Jain, M. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS one* **7**, e30619 (2012).
100. Kent, W. J. BLAT—the BLAST-like alignment tool. *Genome research* **12**, 656–664 (2002).
101. Wu, T. D. & Watanabe, C. K. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**, 1859–1875 (2005).
102. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology* **29**, 644–652 (2011).
103. Haas, B. J. *et al.* Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* **461**, 393–398 (2009).
104. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140, doi:10.1093/bioinformatics/btp616 (2010).
105. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods* **8**, 785–786 (2011).
106. Emanuelsson, O., Nielsen, H., Brunak, S. & von Heijne, G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of molecular biology* **300**, 1005–1016 (2000).
107. Horton, P. *et al.* WoLF PSORT: protein localization predictor. *Nucleic acids research* **35**, W585–W587 (2007).

## Acknowledgements

This work was supported by the National Natural Science Foundation (No. 31400130) and the National 973 Project (No. 2014CB160304). We thank all of the participants in the workshop for their valuable contributions. We acknowledge Beijing Biomarker Technologies for performing RNA-seq of *R. solani* AG1IA strains.

## Author Contributions

Y.X. completed the main work of this article, analysis of the data, and drafting of the manuscript. B.F. and A.Z. managed the project. J.H. and B.F. prepared the R.N.A. samples. Y.X., M.Z., D.Z. and L.P. performed the disease index test and gene expression validation. J.Z., S.L., L.W. provided useful suggestions; Y.X. carried out bioinformatic analyses; Y.L., J.Z., P.L. and A.Z. revised the manuscript. All authors read and approved the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-10804-1

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017