Integrative genomic and transcriptomic analysis of *Xanthomonas oryzae* pv. *oryzae* pathotype IV, V, and IX in China reveals rice defense-responsive genes

Zecong Li[1,3], Shujuan Shen[1,3], Kuaifei Xia[1,2], Mingyong Zhang[1,2,4*] and Xuan Zeng[1,2*]

Abstract

Bacterial blight of rice is a devastating disease caused by the gram-negative bacteria *Xanthomonas oryzae* pv. *oryzae* (*Xoo*). Chinese *Xoo* strain pathotypes IV, V, and IX are the major virulent *Xoo* strain types in South China sequentially from the 1990s to the present. Here, we report the isolation of GD0201 and GD0202, which belong to pathotypes IV and IX, respectively, and the complete genome sequence and transcriptomic analysis of GD0201 (IV), GD1358 (V), and GD0202 (IX). We found that resistance genes *xa5*, *Xa23*, and *Xa27* confer strong resistance to all three *Xoo* strains, indicating that they are currently good choices for resistance rice breeding. The genome analysis reveals fewer TAL and non-TAL effector coding genes in GD0202 than in the other two strains, potentially contributing to its strong virulence. Transcriptomic analysis of ZH11 inoculated with the three *Xoo* strains strongly suggests that three *Xoo* strains for better infection repress the *ethylene response factor* (*ERF*) gene family members. Furthermore, weighted gene co-expression network analysis (WGCNA) and protein-protein interaction (PPI) analysis revealed 14 hub genes potentially associated with rice response to the three *Xoo* strains. The expression of several hub genes was validated to be induced by all three *Xoo* strains, suggesting its role in bacterial blight disease response to *Xoo* strains. Genomic analysis of the *Xoo* strains belonging to pathotypes IV, V, and IX, identification of effectors and genes related to *Xoo* virulence in rice plants will provide insights into understanding the molecular mechanism underlying rice-*Xoo* inter-action and the gene expression pattern in response to *Xoo* infection.

Keywords Rice, Chinese *Xoo* strains, Complete genome, Transcriptomic analysis, Virulent effectors, WGCNA

Background

Bacterial blight of rice, one of the most devastating rice diseases, is caused by the gram-negative bacterium *Xanthomonas oryzae* pv. *oryzae* (*Xoo*). *Xoo* colonizes and spreads along the vascular tissue of rice leaves, causing severe loss of yield and damage to grain quality (Mew 1987; Nino-Liu et al. 2006). The most economical and environmentally friendly method of controlling bacterial blight of rice is utilizing genetically inheritable resistance (*R*) genes. More than 40 *R* genes for bacterial blight of rice were identified, and 15 were cloned and character-ized, including *Xa1*, *Xa1–2*, *Xa4*, *xa5*, *Xa7*, *xa13*, *Xa14*,

*Correspondence:
Mingyong Zhang
zhangmy@scbg.ac.cn
Xuan Zeng
zengxuan@scbg.ac.cn
[1] Guangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Guangzhou 510650, China
[2] South China National Botanical Garden, Guangzhou 510650, China
[3] University of Chinese Academy of Sciences, Beijing 10049, China
[4] State Key Laboratory of Plant Diversity and Specialty Crops, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, China

Li *et al. Phytopathology Research*   (2024) 6:28

Page 2 of 14

Xa3/Xa26, Xa10, Xa21, Xa23, xa25, Xa27, Xa31(t), and *xa41(t). Xa4*, widely utilized in the 1980s, encodes a cell wall-associated kinase (WAK) protein that confers resistance to *Xoo* strains by enhancing cell wall strength (Hu et al. 2017). *Xa7, Xa10, Xa27*, and *Xa23* are executor *R* genes that can confer bacterial blight resistance once induced by *Xoo* strains (Gu et al. 2005; Tian et al. 2014; Wang et al. 2015; Chen et al. 2021). The wide utilization of the *R* gene in rice breeding leads to the co-evolution of *Xoo* virulence.

In China, *Xoo* strains are divided into nine pathotypes according to the disease response of *Xoo* on five rice cultivars, including Nangeng 15, Java14, JG30, IR26, and Tetep (Fang et al. 1990). In the 1980s and 1990s, the *Xoo* strain of pathotype IV was the most virulent and epidemic *Xoo* strain type in South China. The wide utilization of *Xa4*, which is resistant to *Xoo* strains of pathotype IV while susceptible to that of pathotype V, has led to an increase in pathotype V *Xoo* strains. In the last decade, various *R* genes were utilized for resistance rice breeding, including *Xa3, xa5, Xa7, Xa21*, and *Xa23*, etc. In recent years, the percentage of pathotype V in the fields of South China has been declining while that of pathotype IX is increasing (Zeng et al. 2002). It is urgent to characterize the mechanism underlying the interaction between rice-*Xoo* and develop resistant rice cultivars to secure rice production.

The major virulent effectors in *Xoo* are secreted by the type III secretion system (T3SS) encoded by the *hrp* gene cluster. Type III effectors are mainly divided into two types, including transcription activator-like (TAL) effectors and non-TAL effectors (Furutani et al. 2009). TAL effectors are highly conserved at the nucleotide and amino acid levels and consist of three functional domains, including an N-terminal secretion and translocation domain, a central repeat domain-containing varying numbers of central repeats with 34 amino acids, and a C-terminal domain with nuclear localization signal and transcription activation domain. The 12th and 13th amino acids (the repeat variable di-residues, or RVD) of each repeat can recognize one nucleotide specifically. Accordingly, the repeat number and composition of each repeat's 12th and 13th amino acids determine the TAL effector binding element (EBE) sequence in the promoter of DNA targets (Boch et al. 2009). For example, TAL effectors AvrXa7, PthXo3, TalC, and Tal5 specifically recognize $EBE_{-AvrXa7}$, $EBE_{-PthXo3}$, EBE_{-TalC}, and EBE_{-Tal5b} in the promoter region of bacterial blight susceptible gene *OsSWEET14*, respectively, to induce susceptible response of rice plants (Eom et al. 2019). Meanwhile, several non-TAL effectors have also been identified and characterized to suppress the pattern-triggered immunity (PTI) of rice. For example, XopK of PXO99[A], which harbors

E3-ubiquitin-ligase activity, can ubiquitinate and degrade embryogenic receptor kinase 2 (OsSERK2), thereby inhibiting PTI upstream of the MAPK cascade, and also, XopZ is reported to be essential for the full virulence of PXO99[A] (Song and Yang 2010; Qin et al. 2018).

AP2/ERF (APETALA2/ethylene-responsive element binding factors) transcription factor family is a large family of plant-specific transcription factors involved in plant development and stress response (Riechmann and Meyerowitz 1998; Nakano et al. 2006). According to the number of AP2/ERF domains, the superfamily is divided into four families: AP2, ERF, RAV, and soloist (Sharoni et al. 2011). The ERF family members, which have only one AP2/ERF domain, play an important role in abiotic and biotic stress. In biotic stress, *OsERF123* was identified as a bacterial blight susceptible gene; it was induced by TAL effector TalB from African *Xoo* strain MAI1, leading to enhanced virulence on rice plants (Tran et al. 2018). *OsERF922*, induced by the rice blast fungus, negatively regulates rice blast resistance (Liu et al. 2012). Overexpression of *OsERF922* represses the expression of defense-related genes and makes rice plants more susceptible to blast fungus. *OsERF3* is rapidly induced by the rice striped stem borer pathogen (Lu et al. 2011). It positively affects two (mitogen-activated protein kinases) MAPKs and two WRKY gene expression levels, as well as jasmonate (JA) and salicylate (SA) concentrations. These findings suggest that *ERF* family genes play key roles in biotic stress response, including bacterial blight of rice.

In this study, the genome of three *Xoo* strains, GD0201, GD1358, and GD0202, belonging to Chinese pathotypes IV, V, and IX, were sequenced and assembled. Based on genomic information and effector analysis, we found that GD0202 contains only 10 complete TAL effector coding genes, far fewer than the number of TAL effector coding genes in GD0201 and GD1358. The transcriptomic analysis of ZH11 leaves inoculated with GD0201, GD1358, and GD0202 revealed that 40 *ERF* genes were repressed by all three *Xoo* strains, which strongly suggests that *Xoo* strains facilitate infection by repressing *ERF* gene expression. Through the weighted gene co-expression network analysis (WGCNA) and protein-protein interaction (PPI) analysis, 14 hub genes were identified. The expression of several hub genes was validated to be induced by all three *Xoo* strains with RT-qPCR. These results provide insights into the molecular mechanism underlying the interaction between rice and *Xoo* and the gene expression pattern of *Xoo* infection.

Results

Isolation and classification of GD0201 and GD0202 strain

Xoo strains were isolated from fields in Guangzhou, Guangdong Province, China, in 2017 and 2020. The strain isolated

Li *et al. Phytopathology Research*   (2024) 6:28

Page 3 of 14

in 2017 was designated as GD0201, while that isolated in 2020 was designated as GD0202. The pathotype of GD0201 and GD0202 was determined by inoculating rice cultivars, including Nangeng 15, Java14, JG30, IR26, and Tetep (Fang et al. 1990). According to the classification criteria, GD0201, incompatible with IR26 and compatible with the other four rice cultivars, belongs to pathotype IV (Table 1). GD0202, compatible with all five rice cultivars, was classified as pathotype IX (Table 1). GD1358 is a strain of Chinese pathotype V isolated decades ago (Zeng et al. 2002).

The virulence of GD0201, GD1358, and GD0202 was tested on rice cultivars containing various *R* genes (Additional file 1: Table S1). Rice cultivars IRBB5, IRBB8, CBB23, and IRBB27, which harbour *R* gene *xa5*, *Xa8*, *Xa23*, and *Xa27*, respectively, conferred strong resistance to GD0201. IRBB5, CBB23, and IRBB27 also confer strong resistance to both GD1358 and GD0202. These results imply that currently, IRBB5 (*xa5*), CBB23 (*Xa23*), and IRBB27 (*Xa27*), are able to confer strong resistance to *Xoo* in fields of China and are good choices for resistance rice breeding.

To further verify the virulence of GD1358 and GD0202, 206 rice cultivars collected worldwide, which belong to the 3000 Rice Genomes Project, were inoculated with GD1358 and GD0202 in the field in 2018 and 2020, respectively. All rice varieties tested were susceptible to GD0202, while 89.8% of the cultivars were susceptible to GD1358 (Additional file 1: Table S2). This indicates that both GD1358 and GD0202 are highly virulent *Xoo* strains, while the virulence of GD0202 is more potent than that of GD1358.

Sequencing, function annotation, and phylogenetic analysis of GD0201, GD1358, and GD0202 genome

Genomes of the three *Xoo* strains were sequenced with third- and second-generation sequencing to figure

out the difference between the three genomes. The sequencing and assembly information are listed in Additional file 1: Table S3, while the graphical representation of genomes is shown in Fig. 1. The genome size of GD0201 is ~ 5.05 Mb and contains 4866 predicted genes and 19 TAL effector coding genes (*tal*s). GD1358 has a genome size of ~ 5.15 Mb and is predicted to contain 4902 genes and 18 *tal*s. In contrast, GD0202 possesses a smaller genome size of ~ 4.97 Mb with an estimated 5017 genes, surpassing the gene number of both GD0201 and GD1358 (Table 2).

Besides GD0201, GD1358, and GD0202, the genomes of 12 *Xoo* strains originated from different regions were subjected to Average Nucleotide Identity (ANI) analysis. ANIb analysis showed that GD0201, GD1358, and GD0202 had average identities of 99.47, 99.53, and 99.39% to PXO99[A] (Additional file 1: Table S4). ANIm analysis revealed that GD0201, GD1358, and GD0202 exhibited average identities of 99.49, 99.57, and 99.50%, respectively, to PXO99[A] (Additional file 1: Table S4). The similarity between the three strains and Asian *Xoo* strains exceeded 99%, while the similarity between the three strains and African *Xoo* strains fell below 98%. This suggests that GD0201, GD1358, and GD0202 are evolutionarily close to Asian *Xoo* strains. Therefore, GD0201, GD1358, and GD0202 are identified as *Xoo* strains based on ANI analysis.

To investigate the population structure of *Xoo* strains further, we performed phylogenetic analyses of GD0201, GD1358, GD0202, and 83 other *Xoo* strains with their protein sequences (Fig. 2). According to the phylogenetic tree, all *Xoo* strains were divided into two lineages. One lineage is African *Xoo* strains. In the phylogenetic tree, one clade contains only African strains, such as AXO1947, MAI1, and BAI3, indicating that African *Xoo* strains were phylogenetically close to each other. The other lineages predominantly comprise Asian strains. GD1358, a type V strain isolated more than 30 years ago, clusters closely with *Xoo* strains from the Philippines, including PXO86, PXO145, PXO236, PXO83, and PXO211, which were also isolated decades ago (Oliva et al. 2019). While GD0201 and GD0202, isolated within the last 10 years, exhibit close phylogenetic proximity to *Xoo* strains GXO2001 and GXO2006, isolated in Guangxi Province, China, the neighboring province of Guangdong. This suggests that the evolution of *Xoo* strains was influenced by the local environment over time.

Non-TAL and TAL effectors encoded by the GD0201, GD1358, and GD0202

Non-TAL effectors are secreted by the T3SS of *Xoo* and are involved in host immunity suppression (Furutani

Table 1 Pathotype classification of *Xoo* strains according to the disease response of rice

Rice cultivar [b]	*Xoo* strains [a]	
	GD0201	GD0202
Nangeng 15	14.3 ± 5.3 (S)[b]	21.8 ± 4.7 (S)
Java 14	10.2 ± 1.7 (S)	17.9 ± 4.7 (S)
JG30	26.4 ± 3.8 (S)	30.1 ± 4.4 (S)
IR26	1.9 ± 1.1 (R)	12.1 ± 3.8 (S)
Tetep	19.0 ± 4.4 (S)	15.3 ± 4.3 (S)
Pathotype	IV	IX

[a] The unit of the bacterial lesion length is centimeter

[b] Six-week-old rice plants were inoculated with *Xoo* strains. Lesion length and disease phenotype of the inoculated plants were scored at 2 weeks after inoculation. For disease scoring: R, resistant, lesion length ≤ 3.0 cm; MR, moderately resistant, 3.0 cm < lesion length ≤ 6.0 cm; MS, moderately susceptible, 6.0 cm < lesion length ≤ 9.0 cm; S, susceptible, lesion length > 9.0 cm
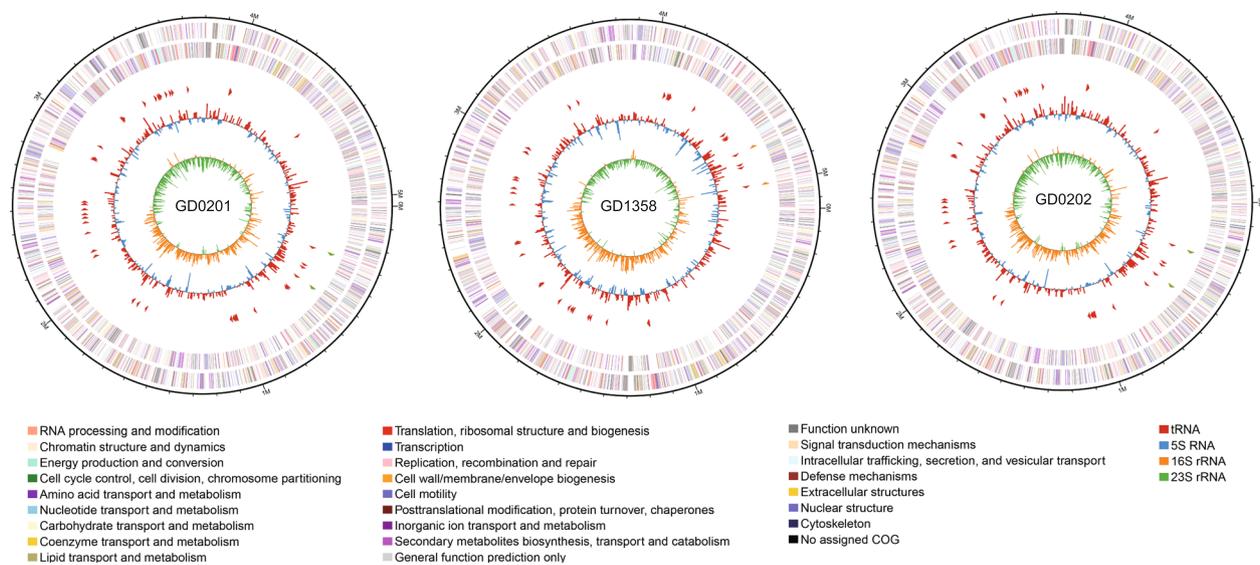
Fig. 1 Graphical representation of GD0201, GD1358, and GD0202 genomes. The outmost circle is the size of the genome; the second and third circles represent coding sequence (CDS) on the sense and antisense strands, respectively; different colors represent CDS with different COG (Clusters of Orthologous Genes) functional classifications; the fourth circle are rRNA and tRNA; the fifth circle shows the GC content. The outward red bars indicate GC content is higher than the whole genome, and the higher the bar, the larger the difference. The inward blue bars indicate GC content is lower than that of the whole genome; the higher the bars, the larger the difference. The innermost circle is the GC skew, calculated by G-C/G + C; when the value is positive, the sense strand tends to transcribe CDS, while when the value is negative, the antisense strand tends to transcribe CDS

Table 2 Summary of three *Xoo* genome genomes

	GD0201	GD1358	GD0202
Length (bp)	5,049,476	5,150,506	4,969,229
GC content (%)	63.69	63.69	63.71
Annotated genes	4866	4902	5017
Transfer RNA	60	54	53
Ribosomal RNA operon	2	2	2
TAL effector genes	19	18	14

et al. 2009; Song and Yang 2010). To identify non-TAL effectors encoded by the three *Xoo* strains, HMM files corresponding to various non-TAL effectors in the National Center for Biotechnology Information (NCBI) protein family models were utilized to search three *Xoo* strain genomes. The HMM accession number, sequence cutoff, and domain cutoff for each non-TAL effector are listed in Additional file 1: Table S5. According to the results, the number of non-TAL effectors in GD0202 is fewer than that of GD0201 and GD1358. Besides the three *Xoo* strains, we have also analyzed the non-TAL effectors in other Chinese *Xoo* strains. We found AvrBs2, XopF, XopK, XopL, XopN, XopV, XopX, XopZ, XopAA, XopAD, XopAE, and XopAF were presented in all the analyzed *Xoo* strains, indicating the importance of these non-TAL effectors for the virulence of *Xoo*

strains. While XopW presents in 8 out of 18 Chinese *Xoo* strains, and XopI only presents in African strain AXO1947. Notably, GD0202 harbors the fewest non-TAL effector coding genes among all the analyzed *Xoo* strains, totaling 12. Further investigations were needed to study whether the number of non-TAL effectors is related to *Xoo* strains' virulence.

According to the genome sequence of three *Xoo* strains, TAL effectors form clusters in the genome (Fig. 3). RVDs of TAL effectors from the three *Xoo* strains are listed in Additional file 1: Table S6–S8. Interestingly, GD0202 harbors 14 *tal*s, but only 10 of these genes encode TAL effectors with RVDs; the remaining four genes are pseudogenes encoding truncated TAL effectors lacking the central repeat domain. Furthermore, tal1c$_{-GD0201}$ and tal1c$_{-GD0202}$ are the same as AvrXa23D, and tal8b$_{-GD1358}$ is AvrXa23A. Both AvrXa23D and AvrXa23A are capable of inducing the expression of the *R* gene *Xa23* (Xu et al. 2022). As a result, all three *Xoo* strains are incompatible with CBB23. Also, all three *Xoo* strains harbor an AvrXa27-like effector coding gene in the genome, leading to an incompatible response with IRBB27.

Besides the three *Xoo* strains, we analyzed more than 15 Chinese *Xoo* strains. All TAL effectors were categorized into different classes utilizing annoTALE (Grau et al. 2016) and listed in Additional file 1: Table S9. Consistent with previous studies, TAL effectors in the
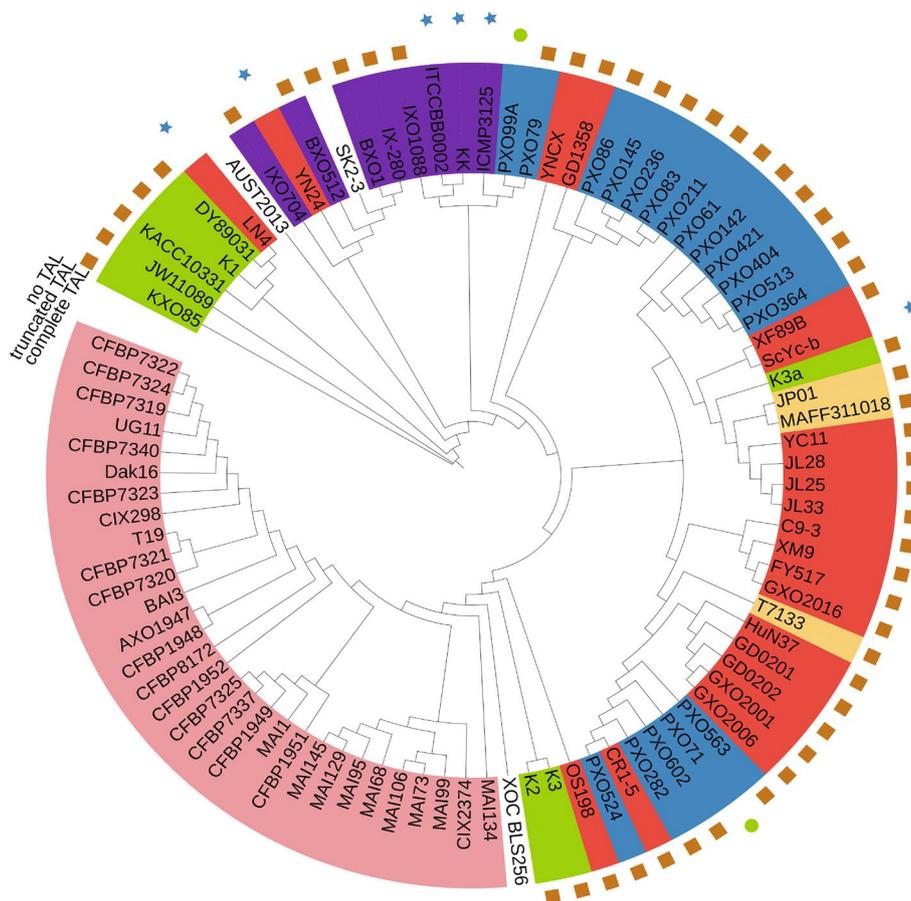
Fig. 2 Phylogenetic analysis of 86 *Xoo* strains from different regions worldwide. African strains, pink; Chinese strains, orange; the Philippines strains, blue; Indian strains, purple; Korean strains, green; Japanese strains, yellow. The brown square represents the strains containing complete TalAG effector. The green circle represents the strains that contain truncated TalAG effector. The blue star represents the strains that do not contain TalAG effector

African *Xoo* strain AXO1947 were classified into distinct classes compared to those of Asian *Xoo* strains. The Philippines strains PXO99[A] and PXO71 harbor TAL effectors classified into classes that are absent from Chinese *Xoo* strains. This indicates that location is an important factor that influences the evolution of *tal*s. Furthermore, the absence of AvrXa23-like effectors in strains C9-3, HuN37, and JL33, as well as the lack of AvrXa27-like effectors in strains JL25, JL28, and ScYc-b, suggests these strains may defeat the resistance of rice cultivars harboring *Xa23* or *Xa27*. Thus, novel *R* genes should be utilized to prevent the outbreak of strains lacking AvrXa23 and AvrXa27. According to the analyzed results, effectors assigned to TalAP present in all the Asian *Xoo* strains analyzed, and those assigned to TalAE and TalAG classes also present extensively in Asian *Xoo* strains. We analyzed 59 Asian strains and found that 51 harbor effectors assigned to TalAG, and 2 strains (PXO99A and PXO71) contain

truncated TalAG, while six *Xoo* strains lack this TAL effector from this class (Fig. 2). This suggests effectors belonging to TalAP, TalAE, and TalAG classes may be important for the virulence of Asian *Xoo* strains.

ERF genes may be repressed by *Xoo* strains to facilitate infection

To identify key genes and pathways regulated by *Xoo* stains to facilitate infection, leaves of Zhonghua11 (ZH11) were inoculated with GD0201, GD1358, and GD0202, respectively, with syringes. Inoculated leaves were harvested 1 hour and 48 hours post-inoculation (hpi) for cDNA library preparation and sequencing. After a quality check, the clean reads were mapped to the Nipponbare reference genome. Over 90% of clean reads were successfully mapped to the Nipponbare genome.

To understand the pattern of gene expression induced by GD0201, GD1358, and GD0202, a pairwise
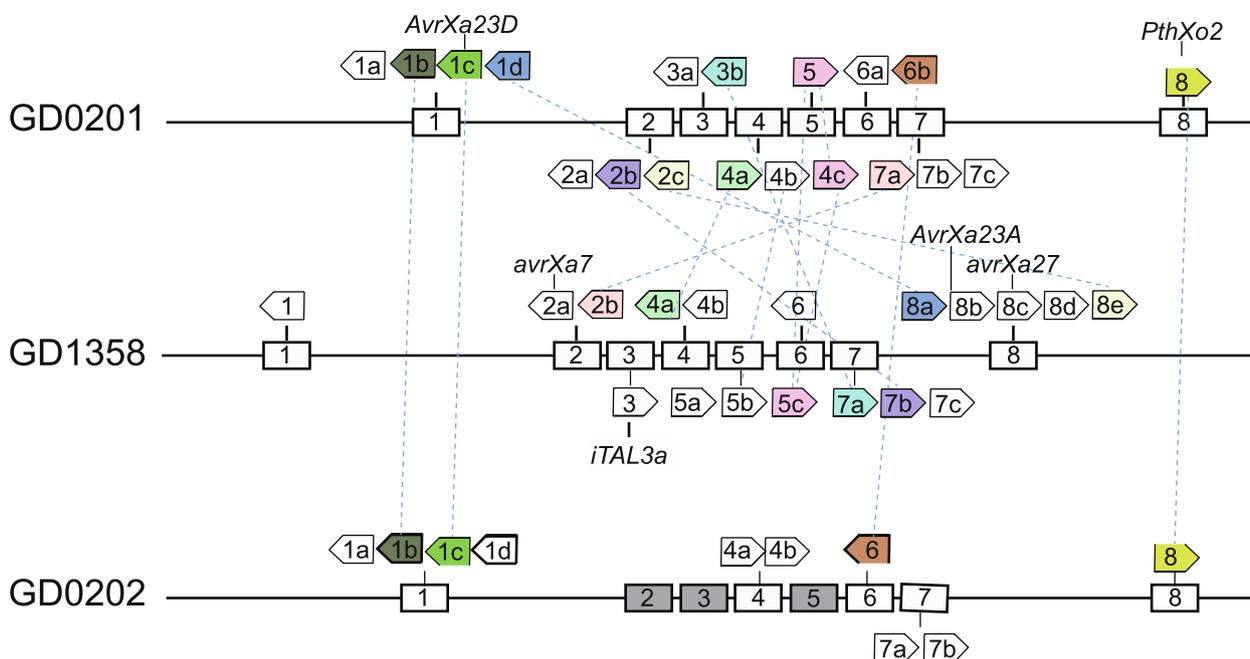
Fig. 3 TAL effectors in *Xoo* strains. Arrangement of *tals* in genomes of three *Xoo* strains. Each box represents one cluster of *tals* in the genome of *Xoo*. One arrow indicates one *tal* gene, and the arrow's orientation represents the *tal* gene's orientation. *Tals* with the same color are identical. White *tals* are strain-specific effectors based on the genome sequence. Pseudogenes have a premature stop codon and do not encode any central repeats, which are indicated in grey boxes

comparison was made between ZH11 leaves inoculated with each strain at 1 hpi vs 48 hpi. A total of 10,239 significantly differential expressed genes (DEGs) (5083 down-regulated and 5156 up-regulated) were identified between ZH11 leaves inoculated with GD0201 1 hpi and 48 hpi. For ZH11 leaves infected with GD1358 1 hpi and 48 hpi, 8662 genes (4930 down-regulated and 3732 up-regulated) were identified. Meanwhile, 9936 genes (5126 down-regulated and 4810 up-regulated) were differently regulated by GD0202. In total, 3422 and 2808 genes were down-regulated and up-regulated by all three *Xoo* strains, respectively (Additional file 2: Figure S1).

The top 10 most highly up-regulated and down-regulated genes by the three *Xoo* strains were selected, and replicates were deleted. Of the 20 downregulated gene sets, 6 genes were found to encode proteins belonging to the ERF protein family (Fig. 4a). Interestingly, 40 out of 138 *ERF* genes in the rice genome were commonly downregulated by all three *Xoo* strains (Fig. 4b). Only two *ERF* genes were commonly up-regulated. Eight *ERF* genes, including *LOC_Os08g36920*, *LOC_Os08g43200*, *LOC_Os06g03670*, *LOC_Os02g52670*, *LOC_Os02g54050*, *LOC_Os01g66270*, *LOC_Os01g12440*, and *LOC_Os11g06770* were strongly repressed by all three *Xoo* strains (Fig. 4b). Several *ERF* genes were reportedly involved in rice response to pathogens (Lu et al. 2011; Liu et al. 2012). This strongly suggests that the three *Xoo*

strains may facilitate the infection process by repressing *ERF* gene expression.

Modules responsible for biotic and abiotic stress were correlated with *Xoo* infection

WGCNA was performed to construct co-expressed networks and identify co-expression modules. The sample hierarchical clustering analysis suggested by the resulting trees did not identify strong outliers except for one sample treated with water for 2 days, but it had little effect on subsequent results (Additional file 2: Figure S2). The gene expression matrix contains 17,764 genes. After being filtered using the MAD (median absolute deviation) method, the top 15,000 genes were retained.

In this study, the fitting degree of the scale-free topological model was 0.85, and the soft threshold for network construction was selected as 14. The obvious modules were identified, and different colors were used to represent them. A total of 13 modules including MEblack (287 DEGs), MEblue (5387 DEGs), MEbrown (912 DEGs), MEgreen (619 DEGs), MEgreenyellow (90 DEGs), MEgrey (418 DEGs), MEmagenta (123 DEGs), MEpink (273 DEGs), MEpurple (96 DEGs), MEred (448 DEGs), MEtan (36 DEGs), MEturquoise (5639 DEGs), MEyellow (672 DEGs) were obtained (Additional file 2: Figure S3). These modules were positively or negatively correlated with the response to the bacterial blight.
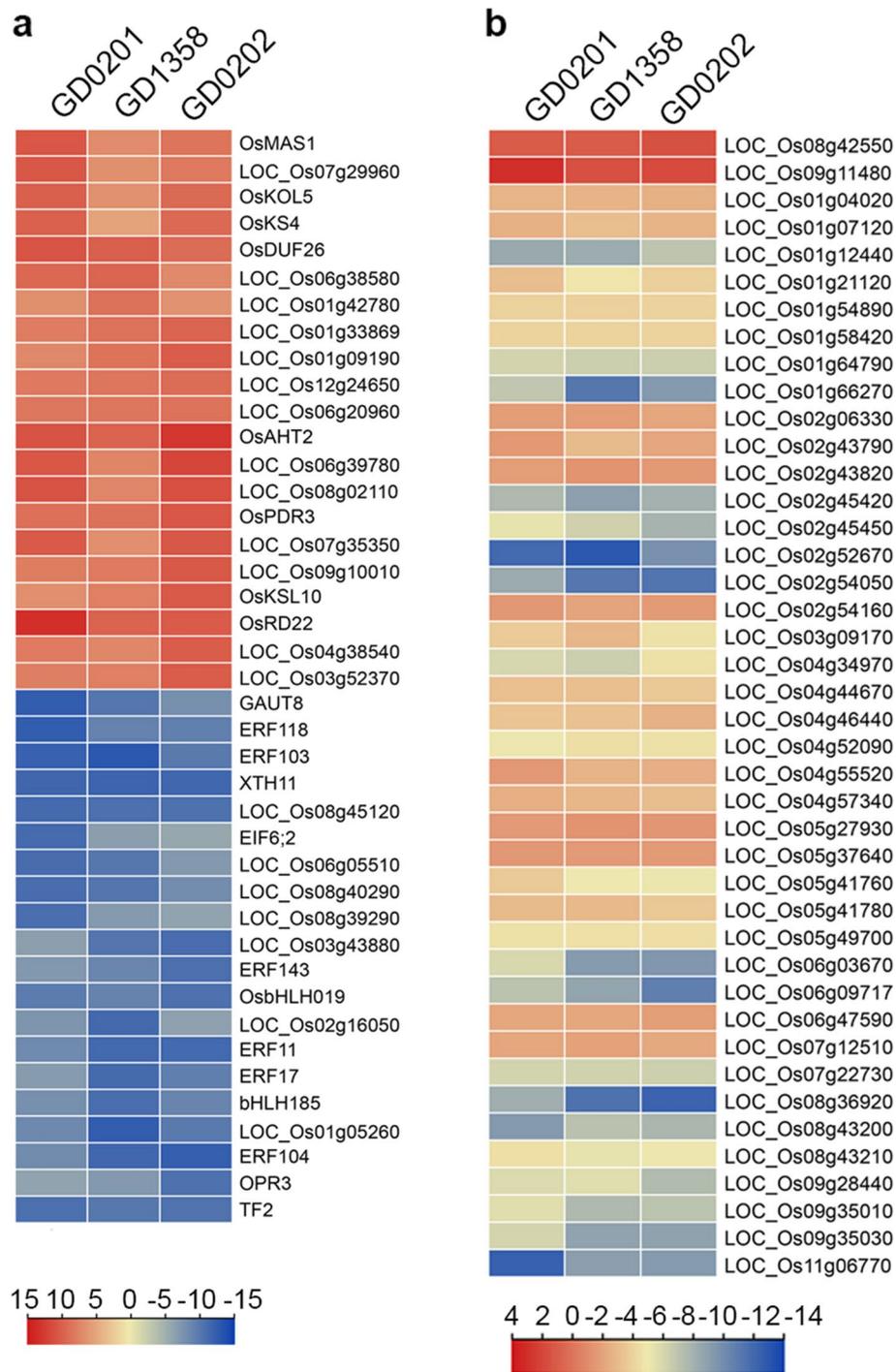
Fig. 4 Heatmap of the top 10 regulated genes and members of the ERF family regulated by three *Xoo* strains after inoculation 48 hours at the tillering stage. **a** The top 10 most highly up-regulated and down-regulated genes are found in three *Xoo* strains. **b** The heatmap of 42 *ERF* genes regulated by three *Xoo* strains

ZH11 inoculated with ddH$_2$O at 1 hpi and 48 hpi and inoculated with three *Xoo* strains at 1 hpi were set as controls, and ZH11 inoculated with three strains at 48 hpi were set as case for gene modules and phenotype correlation analysis. When the absolute values of *cor* >0.5 and *p* <0.01 were set as the strict criterion, several specific

Li *et al. Phytopathology Research*   (2024) 6:28

Page 8 of 14

modules were identified as significantly correlated with rice-*Xoo* interaction. The MEblue (*cor*=0.81, *p*=1.9e-6), MEpurple (*cor*=0.76, *p*=1.5e-5), MEbrown (*cor*=0.58, *p*=3.2e-3) modules showed a high correlation with the case. The MEyellow (*cor*=0.85, *p*=1.1e-7), MEturquoise (*cor*=0.76, *p*=1.7e-5), MEpink (*cor*=0.62, *p*=1.1e-3), and MEblack (*cor*=0.57, *p*=3.5e-3) modules showed high correlation with control (Fig. 5a).

Gene Ontology analysis revealed that genes belonging to the MEblue, MEturquoise, and MEyellow modules are mainly related to metabolic processes. The genes belonging to the MEpruple module were found to be mainly related to stress and defence response. A protein-protein interactions network of expressed genes in the MEpurple module was constructed for hub gene identification. Genes in the MEpurple module were imported into the STRING database for protein interaction analysis and were divided into three clusters using the K-means clustering algorithm. Clustered results were then imported into Cytoscape for analysis. The protein-protein interaction networks of expressed genes in the MEpurple module were present in Fig. 5b. Fourteen genes have been identified as hub genes that are potentially related to rice response to the three *Xoo* strains. According to the trait ontology, most of these hub genes are related to disease response and abiotic tolerance (Additional file 1: Table S10) (Kurata and Oryzabase 2006). Furthermore,

we validated the expression pattern of several hub genes in response to three *Xoo* strains. *Os03g0743500* (*OsCML4*), *Os06g0727400* (*OsRLCK220*), *Os01g0660200* (*PR31*), *Os01g0134700* (*CBP-like*), and *Os04g0611400* (*VSR7*) exhibited induced expression in response to all three *Xoo* strains (Fig. 6). Notably, the fold change of *OsCML4* is highest among the analyzed hub genes, suggesting a crucial role of *OsCML4* played in the biotic stress imposed by GD0201, GD1358, and GD0202. Further investigation is needed to verify the involvement of hub genes in *Xoo* response.

Discussion

GD1358 is a *Xoo* strain isolated over 30 years ago in Guangdong Province, China (Fang et al. 1990). GD0201 and GD0202 are two *Xoo* strains isolated in Guangdong Province in recent years. Since the 1980s, rice cultivars moderately resistant or resistant to epidemic *Xoo* strains were planted in fields in China to ensure the yield of rice country-wide. As a result, epidemic *Xoo* strains were subjected to selection pressure from resistant rice cultivars. Therefore, the virulence of *Xoo* strains is evolving. From the 1990s to the present, the major virulent *Xoo* strain types in South China have been pathotypes IV, V, and IX sequentially. GD0201, GD1358, and GD0202 are classified as pathotypes IV, V, and IX, respectively. The complete genome assembly of three *Xoo* strains, identification
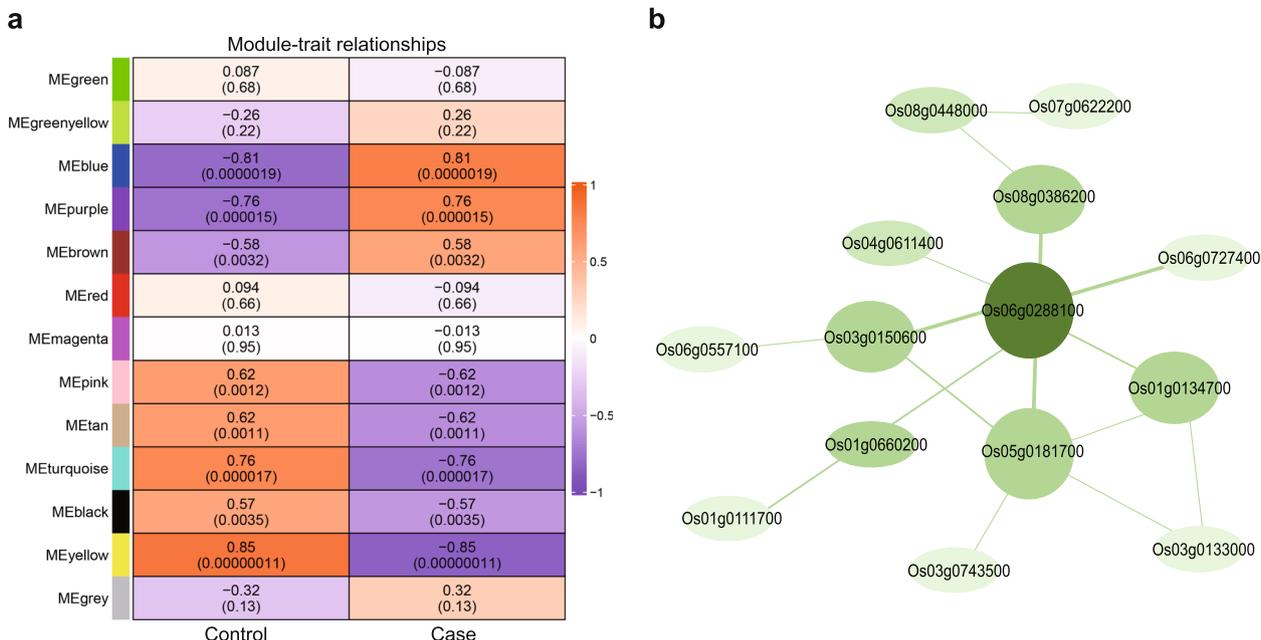

Fig. 5 WGCNA and PPI analysis of ZH11 leaf samples inoculated with three *Xoo* strains. **a** Heatmap of the correlation between Eigengene and disease phenotype of ZH11 inoculated with ddH$_2$O and three *Xoo* strains. **b** The 14 hub genes identified from the interaction network analysis within the purple module. Nodes represent genes. Degree defines the node size. Betweenness centrality defines the color of the node. The combined score defines the thickness of the edge
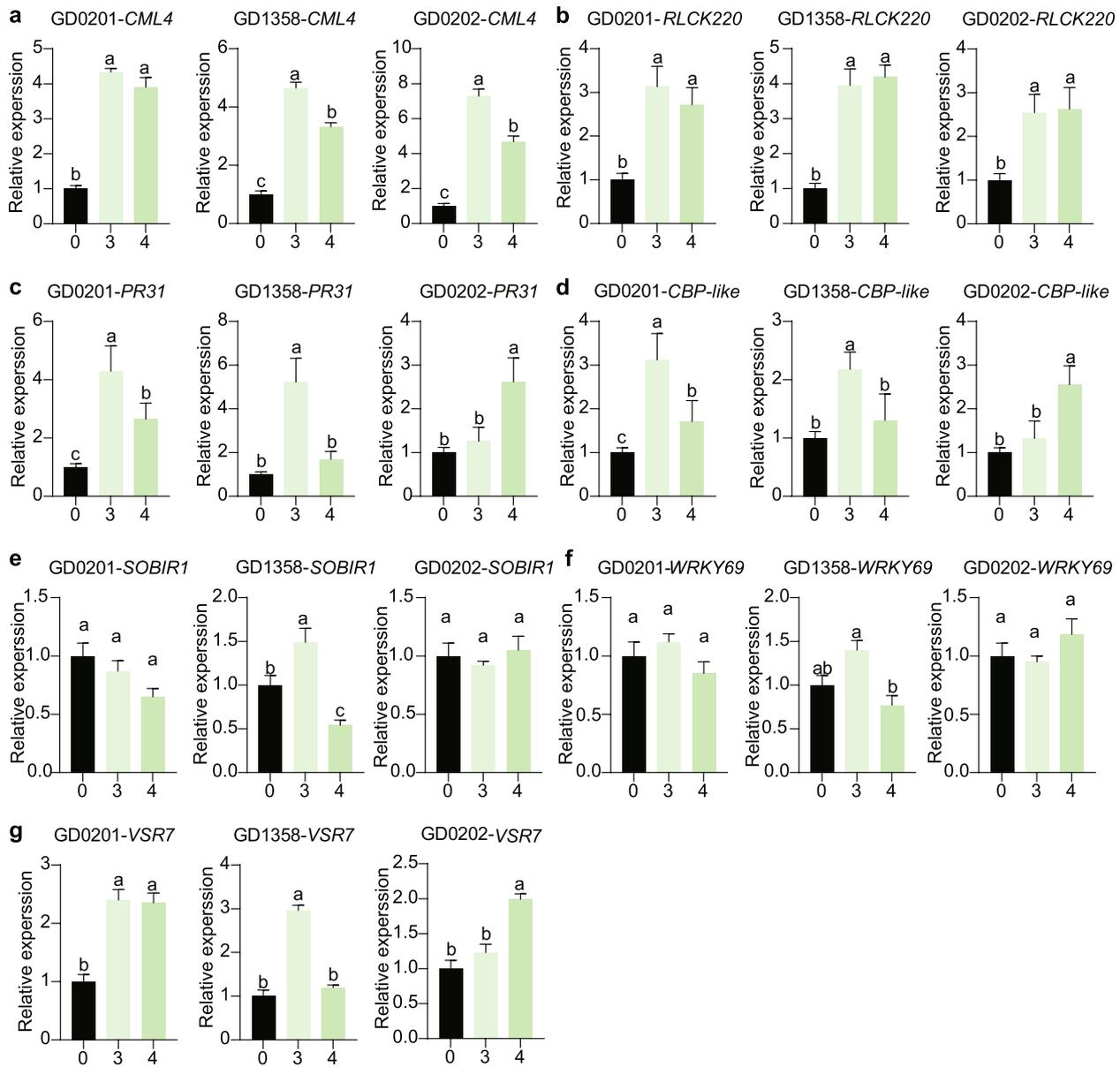
Li *et al. Phytopathology Research*   (2024) 6:28

Page 9 of 14

Fig. 6 The expression level of several hub genes in response to three *Xoo* strains. **a–g** The relative expression level of *Os03g0743500* (*OsCML4*), *Os06g0727400* (*OsRLCK220*), *Os06g0727400* (*OsRLCK220*), *Os01g0134700* (*CBP-like*), *Os04g0611400* (*VSR7*), *Os06g0288100* (*SOBIR1*), and *08 g0386200* (*WRKY69*) in ZH11 at 0, 3, and 4 days post-inoculation (dpi) with the GD0201, GD1358, and GD0202. Bars represent average expression obtained from three independent RNA samples, with standard deviation. The letters above the bars represent the result of Tukey's HSD test. Identical letters correspond to means that are not significantly different from each other ($\alpha = 0.05$). This experiment was repeated three times with similar results

of effectors and transcriptomic analysis of DEGs regulated by three *Xoo* strains are significant for understanding the mechanism underlying *Xoo*-rice interaction and providing insights into bacterial blight-resistant rice breeding.

Our study found bacterial blight *R* genes, including *xa5*, *Xa23*, and *Xa27*, confer strong resistance to pathotypes IV, V, and IX in China. Currently, these *R* genes are

widely utilized in rice breeding in China. However, virulent *Xoo* strains that are able to overcome the resistance conferred by these *R* genes are emerging. For example, *Xoo* strain JNXO and LYG50 overcome the resistance of *xa5*; strain C9-3 and AH28 can overcome the resistance of *Xa23* (Chen et al. 2019; Chen et al. 2022). This suggests the resistance conferred by bacterial blight *R* genes is not permanent; they can be overcome by the *Xoo* strains that

Li *et al. Phytopathology Research*   (2024) 6:28

Page 10 of 14

lose certain avirulent effectors or gain novel effectors. This suggests that more efforts should be made to understand the molecular mechanism underlying the virulence of *Xoo* strains.

Based on the complete genome sequence of the three *Xoo* strains we studied, we found there are fewer complete *tal*s in GD0202 compared to GD0201, GD1358, and most Asian *Xoo* strains. Many of the *tal*s in GD0202 encode pseudogenes, meaning they encode truncated TAL effectors. Loss of avirulent *tal*s may lead to strong virulence of GD0202. Meanwhile, TAL effectors belong to TalAP, TalAE, and TalAG classes and are present extensively in most Asian *Xoo* strains. This also suggests that TalAP, TalAE, and TalAG class effectors are important for the virulence of Asian *Xoo* strains, and the effector-binding elements of TalAP, TalAE, and TalAG are good choices for generating broad-spectrum resistant rice plants.

In the transcriptomic analysis of ZH11 leaves inoculated with three *Xoo* strains, forty members of the ERF protein family genes were repressed by all three *Xoo* strains. This strongly suggests that *Xoo* strains may facilitate the infection process by repressing the expression of ERF protein family genes. Several members of the ERF protein family gene have been reported to be involved in plant immunity. For example, *OsERF123* is a bacterial blight *S* gene induced by TAL effector TalB from African *Xoo* strain MAI1 (Tran et al. 2018). In addition, rice blast pathogen *Magnaportha oryzae* can induce the expression of *OsERF83*, which positively regulates rice resistance to rice blast (Tezuka et al. 2019). Ectopic expression of *OsAP2/ERF152* is able to induce the expression of SA, JA, and ET-responsive genes and enhance rice resistance to bacterial and fungal infection (Pillai et al. 2020). These results indicate the ERF protein family positively regulates rice resistance to bacteria and fungi. This strongly suggests that GD0201, GD1358, and GD0202 may facilitate its infection process by repressing the expression of ERF protein family genes. We will further explore the function of the ERF protein family in bacterial blight response.

According to WGCNA analysis, fourteen hub genes were identified to be potentially related to response to *Xoo* strains. According to the gene Trait Ontology (Kurata and Oryzabase 2006), most of the hub genes were related to plant disease response or abiotic stresses. *Os03g0743500*, highly induced by three *Xoo* strains, is a calmodulin-like gene (*OsCML4*) that functions in regulating different abiotic stresses, such as drought and salt stress (Yin et al. 2015; Asif et al. 2022). Overexpression of *OsCML4* significantly improves rice growth performance and enhances survival rates under drought conditions. This is achieved by efficiently clearing reactive oxygen species (ROS) and inducing other stress-related genes in an abscisic acid (ABA)-independent manner, imparting drought resistance (Yin et al. 2015). This mechanism may also enhance the resistance of rice to bacterial blight. Further investigation into the specific functions and regulatory mechanisms of *OsCML4* in response to *Xoo* infection could provide valuable insights into the molecular basis of rice resistance to this pathogen. Other hub genes, such as *PR31* and *OsRLCK220*, which are involved in disease response, were also induced by three *Xoo* strains, suggesting their potential role in bacterial blight. Interestingly, hub genes *Os06g0288100* and *08g0386200* were not induced at 3 and 4 dpi. *Os06g0288100* encodes an LRR receptor kinase (OsSOBIR1) and participated in rice PTI (pattern-triggered immunity) response and antiviral defense (Zhang et al. 2021). *08g0386200* encodes a WRKY transcription factor 69 that has been reported to respond to *M. oryzae* infection (Li et al. 2021). It is possible that *Os06g0288100* and *08g0386200* were induced at the early stage, and the expression of *Os06g0288100* and *08g0386200* were declined at 3 and 4 dpi. Alternatively, it is possible that these two genes may not be directly responsible for the response to *Xoo* strains. The function of all hub genes in rice response to *Xoo* strains needs to be further verified.

Conclusions

This study identified three highly pathogenic Chinese bacterial blight strains, GD0201, GD1358, and GD0202. By genome sequencing and assembly, their phylogenetic relationship and effectors were analyzed. *xa5*, *Xa27*, and *Xa23* confer strong resistance to all three *Xoo* strains. The *ERF* and 14 hub genes were potentially involved in rice response to three *Xoo* strains with transcriptomic analysis.

Methods

Isolation of *Xoo* strains

Xoo-infected rice leaves were isolated from infected rice plants in the field. Briefly, infected rice leaves were sterilized by immersing them in 75% ethanol for 1 minute and washed three times using sterile distilled water. Subsequently, infected leaves were homogenized with mortar and pestle. The homogenate and its diluent were spread on peptone sucrose agar (PSA) plates (10 g/L peptone, 10 g/L sucrose, 1 g/L glutamic acid, 16 g/L agar, and pH 7.0) and incubated at 28°C for 2 d until single colonies were formed. Single *Xoo* colonies were isolated from solid plates and stored under −80°C for further virulence verification. GD0201 and GD0202 were isolated in 2017 and 2020, respectively, from Guangdong, China. GD1358 is kindly provided by Dr. Yin Zhongchao from Temasek Life Sciences Laboratory.

Xoo strain inoculation and disease assays

Xoo strain inoculation was carried out using the leaf-clipping method. Briefly, *Xoo* strains were cultivated on PSA plates for 2 d at 28℃. *Xoo* strains were collected and resuspended in sterile distilled water at an optical density of 0.5 at 600 nm. *Xoo* strains were applied to rice leaves by cutting about 5 cm from the leaf tips using scissors dipped in the inoculum. At least 3 individual rice plants of one cultivar were inoculated with one *Xoo* strain. Lesion length was measured 14 days after inoculation with rulers. For RNA-seq analysis, *Xoo* strains were infiltrated into rice leaves with syringes. Leaves samples were harvested 0, 1 or 48 h after infiltration.

Plants and growth conditions

In 2020 and 2021, rice plants were grown in paddy fields at the South China Botanical Garden in Guangzhou, China. Six-week-old rice plants were utilized for *Xoo* inoculation and disease scoring.

DNA extraction and genome sequencing

Genomic DNA of *Xoo* strains was extracted with the Bacteria DNA kit (OMEGA; Cat: #D3494–01) according to the manufacturer's instructions. DNA quality and quantity were determined using the TBS-380 fluorometer (Turner BioSystems Inc., Sunnyvale, CA). High qualified DNA sample ($OD_{260/280} = 1.8 \sim 2.0$, $> 6 \mu g$) is utilized to construct fragment library.

For Illumina pair-end sequencing of each strain, at least 3 µg high-qualified DNA was utilized for sequencing library construction. Paired-end libraries with insert sizes of ~400 bp were prepared following Illumina's standard genomic DNA library preparation procedure. Purified genomic DNA is sheared into smaller fragments with a desired size by Covaris, and blunt ends are generated by using T4 DNA polymerase, adapters are ligated to the ends of the DNA fragments. The desired fragments can be purified through gel-electrophoresis, then selectively enriched and amplified by PCR. Finally, the qualified Illumina pair-end library will be used for Illumina NovaSeq 6000 sequencing.

For Pacific Biosciences sequencing, an aliquot of 8 µg DNA was spun in a Covairs g-TUBE (Covaris, MA) at 6000 rpm for 60 s using an Eppendorf 5424 centrifuge. DNA fragments were then purified, end-repaired and ligated with SMR Tbell sequencing adapters. Sequencing libraries were purified three times using 0.45× volumes of Agencourt AMPure XPbeads (Beckman Coulter Genomics, MA). 20 kb insert whole genome shotgun libraries were generated and sequenced on a Pacific Biosciences RS instrument using standard methods at Biozeron (Shanghai, China).

Genome assembly and annotation

The raw reads were trimmed and quality controlled by Trimmomatic (v0.36) (Bolger et al. 2014). ABySS was first used for GD0201 and GD1358 genome assembly with multi-Kmer parameters (Jackman et al. 2017). Canu was subsequently utilized to assemble PacBio-corrected long reads (Koren et al. 2017). Finally, GapCloser software was applied to fill the remaining local internal gaps and correct single-base polymorphism for final assembly results (Xu et al. 2020). Unicycler (0.4.8) (Wick et al. 2017) was used for GD0202 genome assembly.

For the prokaryotic organism, we used the ab initio prediction method to obtain gene models for *Xoo* strain. Gene models were identified using GeneMark (Besemer et al. 2001). Then all gene models were blastp against the non-redundant (NR in NCBI) database, SwissPort (http://uniprot.org), Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/), and Database of Clusters of Orthologous Genes (COG) (http://www.ncbi.nlm.nih.gov/COG) for functional annotation. In addition, tRNAs were identified using the tRNAscan-SE (v1.23) (Chan et al. 2021), and rRNAs were determined using RNAmmer (v1.2) (Lagesen et al. 2007). The genome diagram was developed by the Circos (V0.69) software (Krzywinski et al. 2009). TAL effector coding sequences were analyzed using ANNOTALE 1.4 (Grau et al. 2016). Non-TAL effectors were searched with HMM files corresponding to different non-TAL effectors from NCBI Protein Family Models. The HMM accession number and corresponding sequence and domain cutoff of each non-TAL effector were listed in Additional file 1: Table S5.

RNA extraction, library construction, and sequencing

For mRNA isolation, leaves were ground to powder with liquid nitrogen and 100 mg of powder was utilized for RNA extraction. The total mRNA was isolated with Eastep Super RNA isolation kit (Promega, Cat.# LS1040). The yield and purity of RNA were assessed by recording absorbance at 260 and 280 nm on Nanodrop 8000 Spectrophotometer (Thermo Scientific). Library construction and sequencing were performed by Novogene (Beijing, China).

Gene expression analyses

A 4 cm leaf section was entirely infiltrated and collected at 0, 3, and 4 days post-inoculation (dpi) of three strains, respectively. Each individual sample contained three independent infiltration areas, and three independent replicate samples were processed. After the grinding of the sample, total RNA was extracted from plant leaves using the Eastep Super total RNA isolation kit

(Promega, Cat. # LS1040). 1μg total RNA was reverse transcribed into cDNA using HiScript II 1st Strand cDNA Synthesis Kit (Vazyme, Cat. R212–01). All gene expression studies were performed by RT-qPCR in a LightCycler (ROCHE), using ChamQ SYBR Color qPCR Master Mix (Vazyme, Cat. Q421–02). The average transcript levels of each gene were detected with gene-specific primers (Additional file 1: Table S11). The expression of the *OsActin* gene was used to standardize the RNA sample (Hu et al. 2017).

Reads mapping and annotations

The genome and gene information of Nipponbare (*Oryza sativa* L. subsp. *japonica*) was utilized as a reference and downloaded from Ensembl Plants (Bolser et al. 2017). High-quality (QV > 25) transcriptome libraries from all 24 samples were mapped individually to the Nipponbare genome using hisat2 (v2.2.1) with default parameters (Kim et al. 2019). Transcript abundance and differential gene expression were estimated using the subread and DESeq2 along with FPKM (Fragments Per Kilobase of transcript per Million mapped reads) from reference-guided mapping (Love et al. 2014). Significance tests for differential expression were based on a modified exact test. FDR < 0.05 and $|Log_2FC| > 1$ criteria were set to retrieve the significantly differential expressed genes (DEGs) between different time points. The ggplot2, a package of R, was used to produce the heat maps of differentially expressed genes.

Construction of WGCNA co-expression network

The WGCNA co-expression network was constructed using WGCNA shiny in TBtools (Langfelder and Horvath 2008; Chen et al. 2020). Inputting the gene counts format as the expression matrix. Pearson's coefficient was used to calculate the correlation between genes to determine the co-expression similarities between the two genes. An appropriate adjacency matrix weight parameter β value was selected to satisfy the scale-free network distribution precondition, and the Topological Overlap Matrix (TOM) was constructed to cluster and segment the modules. The correlation between each network module and the sample phenotype were analyzed, to select modules related to the time points and different samples.

Identification and analysis of vital modules and key genes

The significance of the different time points and different rice sample modules were compared; the correlation of module-trait value $cor \geq 0.5$, $p < 0.05$, indicating a statistically significant difference and vital module. The module

eigengene (ME) value can be obtained by analyzing the correlation between gene expression and the corresponding module eigengene. The final value of the correlation coefficient is gene significance (GS). GS reflects the correlation between gene expression and phenotype data. The higher the ME and GS, the more relevant the gene is to the research phenotype.

Phylogenetic tree and ANI analysis

Besides GD0201, GD1358, and GD0202, genome sequences of *Xoo* stains were retrieved from the NCBI database. PhyloPhlAn3.0 (v3.0.67) was utilized to construct phylogenetic trees with all the protein sequences of each strain based on Multiple Sequence Alignment (MSA) results according to maximum likelihood method (Asnicar et al. 2020). iTOL (https://itol.embl.de/) was used to visualize the phylogenetic tree (Letunic and Bork 2021). The average nucleotide identity (ANI) values among 13 genome sequences, including GD0201, GD1358, GD0202, and other *Xoo* strains were calculated using the JSpeciesWS online service (https://www.ribocon.com/jspeciesws.html) (Richter et al. 2015).

Function enrichment and visualization of key genes

Key genes were subjected to the Gene Ontology (GO) database and the Kyoto Encyclopedia of Genes and Genomes (KEGG) for function and pathway enrichment analysis (Ashburner et al. 2000; Kanehisa and Goto 2000; Consortium et al. 2023); $p < 0.05$ was the significant threshold. Cluster and visualization of key genes were constructed with the STRING database and analyzed network plugins of Cytoscape (3.9.1) (Shannon et al. 2003).

Abbreviations

ANI	Average nucleotide identity
AP2/ERF	Ethylene response factor gene family
COG	Clusters of orthologous groups of proteins
DEGs	Differential expressed genes
dpi	Days post-inoculation
EBE	Effector-binding elements
FDR	False discovery rate
GS	Gene significance
KEGG	Kyoto encyclopedia of genes and genomes
MAD	Median absolute deviation
ME	Module eigengene
non-TAL effector	Non-transcription activator-like effector
PTI	Pattern-triggered immunity
PPI	Protein-protein interaction
RT-qPCR	Real-time quantitative reverse transcription polymerase chain reaction
RVDs	Repeat variable diresidues
T3SS	Type III secretion system
TAL effector	Transcription activator-like effector
TOM	Topological overlap matrix
WAK	Wall-associated kinase gene family
WGCNA	Weighted gene co-expression network analysis
Xoo	*Xanthomonas oryzae* pv. *oryzae*

Li *et al. Phytopathology Research*   (2024) 6:28

Page 13 of 14

Supplementary Information

Additional file 1: Table S1. Disease response of different rice cultivars to *Xoo* strains GD0201, GD1358, and GD0202. **Table S2.** Disease phenotype of 206 rice cultivars inoculated with GD1358 and GD0202. **Table S3.** Sequencing and assembly information of three *Xoo* genomes and predicted protein percentage against various protein databases. **Table S4.** ANI analysis between strains GD0201, GD1358, GD0202, and other *Xanthomonas* strains. **Table S5.** Non-TAL effector coding genes in various *Xoo* strains. **Table S6.** RVDs of TAL effectors in GD0201. **Table S7.** RVDs of TAL effectors in GD1358. **Table S8.** RVDs of TAL effectors in GD0202. **Table S9.** Classification of TAL effectors in different *Xoo* strains according to RVDs. **Table S10.** List of 14 hub genes. **Table S11.** Primers used in the RT-qPCR analyses.

Additional file 2: Figure S1. Venn diagram showing the number of DEGs. **Figure S2.** Hierarchical clustering analysis of ZH11 samples inoculated with three *Xoo* strains. **Figure S3.** Data preparation for WGCNA analysis of ZH11 leaf samples inoculated with three *Xoo* strains and ddH$_2$O.

Availability of data and materials
The raw sequencing data generated from this study have been deposited in the Genome Sequence Archive in the National Genomics Data Center (https://ngdc.cncb.ac.cn/). The complete genome sequences of GD0201, GD1358, and GD0202 were deposited under the accession number: PRJCA017491, and the raw data of rice RNA-seq data was deposited under the accession number: PRJCA017482.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

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

References

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. Nat Genet. 2000;25(1):25–9. https://doi.org/10.1038/75556.

Asif S, Kim EG, Jang YH, Jan R, Kim N, Asaf S, et al. Identification of the *OsCML4* gene in rice related to salt stress using QTL analysis. Plants (Basel). 2022;11(19):2467. https://doi.org/10.3390/plants11192467.

Asnicar F, Thomas AM, Beghini F, Mengoni C, Manara S, Manghi P, et al. Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. Nat Commun. 2020;11(1):2500. https://doi.org/10.1038/s41467-020-16366-7.

Besemer J, Lomsadze A, Borodovsky M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. 2001;29(12):2607–18. https://doi.org/10.1093/nar/29.12.2607.

Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009;326(5959):1509–12. https://doi.org/10.1126/science.1178811.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20. https://doi.org/10.1093/bioinformatics/btu170.

Bolser DM, Staines DM, Perry E, Kersey PJ. Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomic data. Methods Mol Biol. 2017;1533:1–31. https://doi.org/10.1007/978-1-4939-6658-5_1.

Chan PP, Lin BY, Mak AJ, Lowe TM. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res. 2021;49(16):9077–96. https://doi.org/10.1093/nar/gkab688.

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13(8):1194–202. https://doi.org/10.1016/j.molp.2020.06.009.

Chen G, Xu Z, Yang Y, Zou L, Zhu B. Classification of pathotypes of Chinese *Xanthomonas oryzae* pv. *oryzae* and resistance breeding strategies for bacterail blight. J Shanghai Jiaotong Univ (Agricul Sci). 2019;37:67–73. https://doi.org/10.3969/J.ISSN.1671-9964.2019.01.012. (in Chinese)

Chen X, Li Q, Wang J, Zhang Y, Tang L, Huang S, et al. Genome resource of a hypervirulent strain C9-3 of *Xanthomonas oryzae* pv. *oryzae* causing bacterial blight of rice. Plant Dis. 2022;106(2):741–4. https://doi.org/10.1094/PDIS-07-21-1565-A.

Chen X, Liu P, Mei L, He X, Chen L, Liu H, et al. *Xa7*, a new executor *R* gene that confers durable and broad-spectrum resistance to bacterial blight disease in rice. Plant Commun. 2021;2(3):100143. https://doi.org/10.1016/j.xplc.2021.100143.

Consortium TGO, Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, et al. The gene ontology knowledgebase in 2023. Genetics. 2023;224(1):1–14. https://doi.org/10.1093/genetics/iyad031.

Eom JS, Luo DP, Atienza-Grande G, Yang J, Ji CH, Luu VT, et al. Diagnostic kit for rice blight resistance. Nat Biotechnol. 2019;37(11):1372–9. https://doi.org/10.1038/s41587-019-0268-y.

Fang C, Xu Z, Guo C, Yin S, Wu S, Xu X, et al. Studies on pathotypes of *Xanthomonas campestris* pv. *oryzae* in China. Acta Phytopathol Sin. 1990;20(2):81–8. http://zwblxb.magtech.com.cn/CN/Y1990/V20/I2/81. (in Chinese)

Furutani A, Takaoka M, Sanada H, Noguchi Y, Oku T, Tsuno K, et al. Identification of novel type III secretion effectors in *Xanthomonas oryzae* pv. *oryzae*. Mol Plant-Microbe Interact. 2009;22(1):96–106. https://doi.org/10.1094/MPMI-22-1-0096.

Grau J, Reschke M, Erkes A, Streubel J, Morgan RD, Wilson GG, et al. AnnoTALE: bioinformatics tools for identification, annotation and nomenclature of TALEs from *Xanthomonas* genomic sequences. Sci Rep. 2016;6:21077. https://doi.org/10.1038/srep21077.

Gu K, Yang B, Tian D, Wu L, Wang D, Sreekala C, et al. *R* gene expression induced by a type-III effector triggers disease resistance in rice. Nature. 2005;435(7045):1122–5. https://doi.org/10.1038/nature03630.

Hu K, Cao J, Zhang J, Xia F, Ke Y, Zhang H, et al. Improvement of multiple agronomic traits by a disease resistance gene via cell wall reinforcement. Nat Plants. 2017;3:17009. https://doi.org/10.1038/nplants.2017.9.

Jackman SD, Vandervalk BP, Mohamadi H, Chu J, Yeo S, Hammond SA, et al. ABySS 2.0: resource-efficient assembly of large genomes using a bloom filter. Genome Res. 2017;27(5):768–77. https://doi.org/10.1101/gr.214346.116.

Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27–30. https://doi.org/10.1093/nar/28.1.27.

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–15. https://doi.org/10.1038/s41587-019-0201-4.

Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res. 2017;27(5):722–36. https://doi.org/10.1101/gr.215087.116.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. Genome Res. 2009;19(9):1639–45. https://doi.org/10.1101/gr.092759.109.

Kurata N, Oryzabase YY. An integrated biological and genome information database for rice. Plant Physiol. 2006;140(1):12–7. https://doi.org/10.1104/pp.105.063008.

Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 2007;35(9):3100–8. https://doi.org/10.1093/nar/gkm160.

Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinf. 2008;9:559. https://doi.org/10.1186/1471-2105-9-559.

Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293–W6. https://doi.org/10.1093/nar/gkab301.

Li Y, Liao S, Mei P, Pan Y, Zhang Y, Zheng X, et al. *OsWRKY93* dually functions between leaf senescence and in response to biotic stress in rice. Front Plant Sci. 2021;12:643011. https://doi.org/10.3389/fpls.2021.643011.

Liu D, Chen X, Liu J, Ye J, Guo Z. The rice ERF transcription factor *OsERF922* negatively regulates resistance to *Magnaporthe oryzae* and salt tolerance. J Exp Bot. 2012;63(10):3899–911. https://doi.org/10.1093/jxb/ers079.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. https://doi.org/10.1186/s13059-014-0550-8.

Lu J, Ju H, Zhou G, Zhu C, Erb M, Wang X, et al. An EAR-motif-containing ERF transcription factor affects herbivore-induced signaling, defense and resistance in rice. Plant J. 2011;68(4):583–96. https://doi.org/10.1111/j.1365-313X.2011.04709.x.

Mew TW. Current status and future prospects of research on bacterial blight of rice. Annu Rev Phytopathol. 1987;25(1):359–82. https://doi.org/10.1146/annurev.py.25.090187.002043.

Nakano T, Suzuki K, Fujimura T, Shinshi H. Genome-wide analysis of the ERF gene family in Arabidopsis and rice. Plant Physiol. 2006;140(2):411–32. https://doi.org/10.1104/pp.105.073783.

Nino-Liu DO, Ronald PC, Bogdanove AJ. *Xanthomonas oryzae* pathovars: model pathogens of a model crop. Mol Plant Pathol. 2006;7(5):303–24. https://doi.org/10.1111/j.1364-3703.2006.00344.x.

Oliva R, Ji CH, Atienza-Grande G, Huguet-Tapia JC, Perez-Quintero A, Li T, et al. Broad-spectrum resistance to bacterial blight in rice using genome editing. Nat Biotechnol. 2019;37(11):1344–50. https://doi.org/10.1038/s41587-019-0267-z.

Pillai SE, Kumar C, Dasgupta M, Kumar BK, Vungarala S, Patel HK, et al. Ectopic expression of a cell-wall-degrading enzyme-induced *OsAP2/ERF152* leads to resistance against bacterial and fungal infection in *Arabidopsis*. Phytopathology. 2020;110(4):726–33. https://doi.org/10.1094/phyto-10-19-0395-r.

Qin J, Zhou X, Sun L, Wang K, Yang F, Liao H, et al. The *Xanthomonas* effector XopK harbours E3 ubiquitin-ligase activity that is required for virulence. New Phytol. 2018;220(1):219–31. https://doi.org/10.1111/nph.15287.

Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics. 2015;32(6):929–31. https://doi.org/10.1093/bioinformatics/btv681.

Riechmann JL, Meyerowitz EM. The AP2/EREBP family of plant transcription factors. Biol Chem. 1998;379(6):633–46. https://doi.org/10.1515/bchm.1998.379.6.633.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504. https://doi.org/10.1101/gr.1239303.

Sharoni AM, Nuruzzaman M, Satoh K, Shimizu T, Kondoh H, Sasaya T, et al. Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. Plant Cell Physiol. 2011;52(2):344–60. https://doi.org/10.1093/pcp/pcq196.

Song C, Yang B. Mutagenesis of 18 type III effectors reveals virulence function of XopZ$_{PXO99}$ in *Xanthomonas oryzae* pv. *oryzae*. Mol Plant-Microbe Interact. 2010;23(7):893–902. https://doi.org/10.1094/mpmi-23-7-0893.

Tezuka D, Kawamata A, Kato H, Saburi W, Mori H, Imai R. The rice ethylene response factor OsERF83 positively regulates disease resistance to *Magnaporthe oryzae*. Plant Physiol Biochem. 2019;135:263–71. https://doi.org/10.1016/j.plaphy.2018.12.017.

Tian D, Wang J, Zeng X, Gu K, Qiu C, Yang X, et al. The rice TAL effector–dependent resistance protein XA10 triggers cell death and calcium depletion in the endoplasmic reticulum. Plant Cell. 2014;26(1):497–515. https://doi.org/10.1105/tpc.113.119255.

Tran TT, Pérez-Quintero AL, Wonni I, Carpenter SCD, Yu Y, Wang L, et al. Functional analysis of African *Xanthomonas oryzae* pv. *oryzae* TALomes reveals a new susceptibility gene in bacterial leaf blight of rice. PLoS Pathog. 2018;14(6):e1007092. https://doi.org/10.1371/journal.ppat.1007092.

Wang C, Zhang X, Fan Y, Gao Y, Zhu Q, Zheng C, et al. XA23 is an executor R protein and confers broad-spectrum disease resistance in rice. Mol Plant. 2015;8(2):290–302. https://doi.org/10.1016/j.molp.2014.10.010.

Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017;13(6):e1005595. https://doi.org/10.1371/journal.pcbi.1005595.

Xu M, Guo L, Gu S, Wang O, Zhang R, Peters BA, et al. TGS-GapCloser: a fast and accurate gap closer for large genomes with low coverage of error-prone long reads. GigaScience. 2020;9(9):giaa094. https://doi.org/10.1093/gigascience/giaa094.

Xu Z, Xu X, Wang Y, Liu L, Li Y, Yang Y, et al. A varied AvrXa23-like TALE enables the bacterial blight pathogen to avoid being trapped by *Xa23* resistance gene in rice. J Adv Res. 2022;42:263–72. https://doi.org/10.1016/j.jare.2022.01.007.

Yin XM, Huang LF, Zhang X, Wang ML, Xu GY, Xia XJ. *OsCML4* improves drought tolerance through scavenging of reactive oxygen species in rice. J Plant Biol. 2015;58:68–73. https://doi.org/10.1007/s12374-014-0349-x.

Zeng L, Huang S, Wu S. Resistance of IRBB21 (*Xa21*) to five races of bacterial blight in Guangdong. Acta Phytophylacica Sin. 2002;29(2):97–100. https://doi.org/10.13802/j.cnki.zwbhxb.2002.02.001. (in Chinese)

Zhang H, Chen C, Li L, Tan X, Wei Z, Li Y, et al. A rice LRR receptor-like protein associates with its adaptor kinase OsSOBIR1 to mediate plant immunity against viral infection. Plant Biotechnol J. 2021;19(11):2319–32. https://doi.org/10.1111/pbi.13663.