SCIENCE CHINA
Life Sciences

•RESEARCH PAPER•

https://doi.org/10.1007/s11427-020-1917-x

Genome sequencing of the bacterial blight pathogen DY89031 reveals its diverse virulence and origins of *Xanthomonas oryzae* pv. *oryzae* strains

Fudan Chen[1], Bingxiao Yan[2,3], Xiangyu Gong[2,3], Helin Li[4] & Zuhua He[1,3*]

[1]*College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China;*
[2]*University of Chinese Academy of Sciences, Beijing 100049, China;*
[3]*National Key Laboratory of Plant Molecular Genetics, CAS Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai 200032, China;*
[4]*School of Life Sciences, Henan University, Kaifeng 475004, China*

Received November 14, 2020; accepted March 16, 2021; published online April 22, 2021

The bacterial pathogen *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), belonging to *Xanthomonas* sp., causes one of the most destructive vascular diseases in rice worldwide, particularly in Asia and Africa. To better understand *Xoo* pathogenesis, we performed genome sequencing of the Korea race 1 strain DY89031 (J18) and analyzed the phylogenetic tree of 63 *Xoo* strains. We found that the rich diversity of evolutionary features is likely associated with the rice cultivation regions. Further, virulence effector proteins secreted by the type III secretion system (T3SS) of *Xoo* showed pathogenesis divergence. The genome of DY89031 shows a remarkable difference from that of the widely prevailed Philippines race 6 strain PXO99A, which is avirulent to rice *Xa21*, a well-known disease resistance (*R*) gene that can be broken down by DY89031. Interestingly, plant inoculation experiments with the PXO99A transformants expressing the DY89031 genes enabled us to identify additional TAL (transcription activator-like) and non-TAL effectors that may support DY89031-specific virulence. Characterization of DY89031 genome and identification of new effectors will facilitate the investigation of the rice-*Xoo* interaction and new mechanisms involved.

DY89031, genome sequencing, TAL effectors, non-TAL effectors, *Xanthomonas oryzae* pv. *Oryzae*

Citation: Chen, F., Yan, B., Gong, X., Li, H., and He, Z. (2021). Genome sequencing of the bacterial blight pathogen DY89031 reveals its diverse virulence and origins of *Xanthomonas oryzae* pv. *oryzae* strains. Sci China Life Sci 64, https://doi.org/10.1007/s11427-020-1917-x

INTRODUCTION

The bacterial genus *Xanthomonas*, which belongs to the Gram-negative γ-proteobacteria, consists of many important plant pathogens that cause diseases in hundreds of plant species, including several economically important crops such as rice. *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) colonizes in the vascular tissues of rice and causes the bacterial blight disease, which is one of the most destructive diseases in rice resulting in substantial yield loss, with the most serious regions in southeast Asia and west Africa (Mew, 1987; Niño-Liu et al., 2006).

In the past three decades, considerable progresses have been achieved to understand the rice-*Xoo* pathosystem and interaction. To counteract *Xoo* infection, rice plants have evolved innate immunity strategies similar to those in other plant models (Jones and Dangl, 2006), such as adopting the disease resistance (*R*) genes (*Xa*) against *Xoo* (Gao and He, 2013; Deng et al., 2020). Thus far, more than 45 *Xa* loci have been found (Ji et al., 2020; Zhang et al., 2020), including RLK receptor kinases and NLR immune receptors, and a few *Xa* genes have been shown to be capable of conferring broad-

*Corresponding author (email: zhhe@cemps.ac.cn)

spectrum resistance against *Xoo* (Gao and He, 2013; Jiang et al., 2020a; Jiang et al., 2020b; Wang and Chai, 2020; Wang et al., 2020). For example, *Xa21*, the first cloned *Xa* gene, confers resistance to many *Xoo* strains (Song et al., 1995). Breeding application of these *Xa* genes ensures economical and environmentally friendly control of the bacterial blight disease in rice production. However, the deployment of these *Xa* genes renders the co-evolution of *Xoo* virulence. As a consequence, *Xa21* has lost resistance to many Korean strains including the K1 type DY89031 (da Silva et al., 2003), limiting its potential in rice breeding and production (Pruitt et al., 2015).

A plethora of virulence determinants have been identified in *Xoo,* which include extracellular polysaccharides (EPS), lipopolysaccharides (LPS), adhesins, and cell wall degrading enzymes (An et al., 2020; Sinha et al., 2013). Notably, *Xoo* is equipped with at least six types (I–VI, or T1SS–T6SS) of virulence-associated secretion systems with a rich diversity in structures, host targets and functions, ensuring its successful colonization and adaptation to plant hosts (Gerlach and Hensel, 2007; Timilsina et al., 2020). The best-characterized secretion system in *Xoo* is the type III secretion system (T3SS), which functions in the translocation of bacterial virulence proteins/effectors into host cells to hijack host cellular processes and transcriptome, resulting in immune suppression to facilitate pathogen infection. Transcription activator-like (TAL) effectors are well-known virulence factors secreted by the T3SS of *Xoo* (Miller et al., 2011; Pfeilmeier et al., 2016), which are highly conserved and possesse three functional domains: an N-terminal secretion and translocation domain, a central DNA-binding domain and a C-terminal transcription activation domain (Bogdanove et al., 2010). TAL effectors target the promoter motifs of a group of susceptibility (*S*) genes, such as the *SWEET* (sugars will eventually be exported transporters) genes, and regulate their transcriptions once they are internalized into the nuclei of host cells. For instance, the expression of *OsSWEET11*and *OsSWEET13* is induced by the TAL effectors PthXo1 and PthXo2, while that of Os*SWEET14* is induced by AvrXa7, PthXo3, TalC and Tal5 (Yang et al., 2006; Wu et al., 2007; Antony et al., 2010; Chen et al., 2012; Streubel et al., 2013; Zhou et al., 2015). Besides TAL effectors, several non-TAL effectors of *Xoo* have also been identified to suppress the pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) of hosts. For example, XopN, XopJ, XopB and XopS of *X. campestris* pv. *vesicatoria* are capable of suppressing immune responses of plants (Kim et al., 2009; Bartetzko et al., 2009; Schulze et al., 2012). XopK of the *Xoo* strain PXO99A can inhibit immune responses upstream of the mitogen-activated protein kinase (MAPK) cascade, and XopZ of PXO99A is required for its full virulence (Song and Yang, 2010; Qin et al., 2018).

Genome sequencing of *Xoo* greatly facilitates the dissection of pathogenicity and the study of virulence factors involved in host adaptation. To date, several sequenced genomes of *Xoo* strains have been released including the three most representative strains, the Korean strain KACC10331 (also known as KXO85), the Japanese strain MAFF311018 (also known as T7174) and the Philippine strain PXO99A (Lee et al., 2005; Ochiai et al., 2005; Salzberg et al., 2008). PXO99A is one of the most virulent strains of *Xoo* worldwide, but is avirulent to *Xa21* (Pruitt et al., 2015). Interestingly, a highly virulent Korean (K1) strain DY89031 (also known as J18), is capable of overcoming the *Xa21*-mediated resistance (da Silva et al., 2003).

In the present study, in order to identify and deploy new *Xa* gene(s) against the widespread *Xoo* strain DY89031, we sequenced and assembled its complete genome and compared it with other *Xoo* representative genomes to reveal its distinct virulence characteristics, providing a foundation to study the rice-*Xoo* interaction and discover novel *Xa* genes.

RESULTS

DY89031 is highly virulent to most of the rice accessions

We inoculated 2,347 globally collected rice genetic accessions in the field to test the resistance against DY89031 and PXO99A, and found that DY89031 was virulent to most of the rice accessions with a small fraction (22.7%) exhibiting resistance, whereas fewer varieties were resistant to PXO99A compared to DY89031 (Table 1, Figure 1A and C). *Xa21* confers resistance to a group of *Xoo* strains, including PXO99A, KACC10331 and MAFF311018. However, plants carrying *Xa21* are highly susceptible to DY89031 (Figure 1B and D). Interestingly, those rice accessions, such as Kasalath, TN1 and 9311 that are resistant to DY89031, usually exhibit high susceptibility to PXO99A. Thus, we propose that these DY89031-resistant varieties likely retain race-specific *R* genes against DY89031 but not PXO99A, which may arm with special virulence strategies.

Genome sequencing of DY89031 reveals its unique features

In order to identify the cognate avirulent protein/effector candidates that interact directly or indirectly with the novel resistance protein(s) to DY89031, we sequenced the genome of DY89031 using the third generation sequencing technology. The whole genome was submitted to the NCBI Database (BioProject ID: PRJNA675962). We discovered that the genome of DY89031, with a single circular chromosome, consists of 4.97 Mb with 63.69% GC content and shares a similar size with that of the north Asian strains KACC10331 and MAFF311018, but is ~0.3 Mb smaller than that of PXO99A. The DY89031 genome contains 4,794 predicted

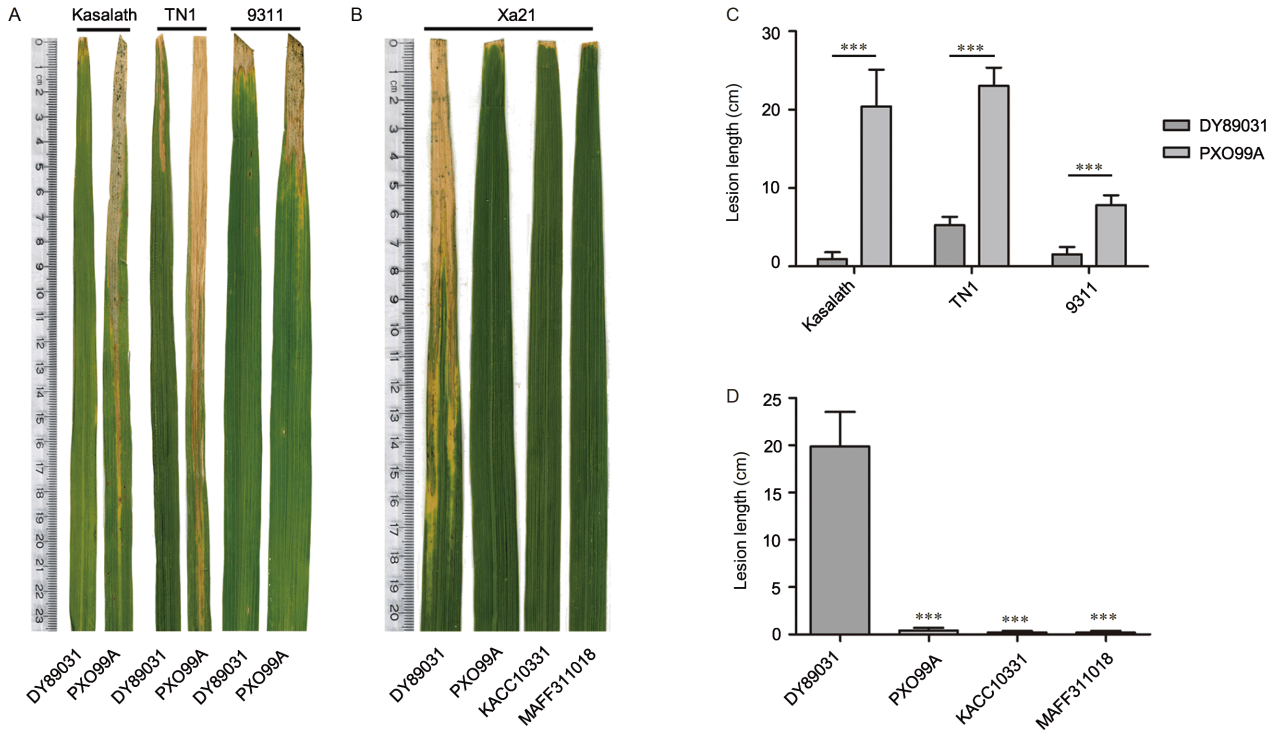

Figure 1 DY89031 displayed a virulence spectrum different from that of PXO99A. A and C, Disease symptoms (A) and lesion lengths (C) of rice variety Kasalath, TN1 and 9311 at 14 dpi inoculated with *Xoo* strain DY89031 and PXO99A. B and D, Disease symptoms (B) and lesion lengths (D) of *Xa21*-carrying plants at 14 dpi inoculated with *Xoo* strain DY89031, PXO99A, KACC10331 and MAFF311018, showing that DY89031 is virulent to *Xa21*-carrying plants. Data shown are means±SD from three biological replicates. Asterisks indicate statistically significant difference in comparison with the control (Student's *t*-test, ***, $P<0.001$).

Table 1 Inoculation results of 2,347 rice accessions with DY89031 and PXO99A[a)]

Strain	Phenotype			
	R	MR	MS	S
DY89031	534	158	260	1,395
PXO99A	27	20	182	2,118

a) R, resistance with 0–3 cm lesion length; MR, moderate resistance with 3–5 cm lesion length; MS, moderate susceptibility with 5–7 cm lesion length; S, susceptibility with>7 cm lesion length.

genes, of which 3,239 (67.6%) genes encode proteins according to COG annotation, and others are predicted to encode hypothetical proteins with unknown functions (Table 2). Therefore, *Xoo* likely diverges as South and North Asia lineages.

A comprehensive frame of the DY89031 genome is presented in Figure 2A, including the predicted TAL genes and the COG function classification of genes. MUMmer (Kurtz et al., 2004) was used to perform global comparison between the DY89031 and PXO99A genomes and analyze the relative positioning and orientation of contigs in the genomes. The result shows that the two genomes share 95% similarity in the matched sequences. We found rich rearrangements and inversions between the DY89031 and PXO99A genomes. The major rearrangement events between DY89031 and PXO99A occurs in the region from 146,726 bp to 4,860,584 bp (Figure 2B, Table S1 in Supporting Information).

Interestingly, we found that the DY89031 genome obtains additional sequences, which contains 148 predicted genes (Table S2 in Supporting Information). Seven vgrG1 proteins (vgrG1-4, 5, 9, 10, 13, 14, 17) are lost in PXO99A, which are essential for a functional T6SS assembly and effector delivery (Cianfanelli et al., 2016). Particularly, VirB5 and VirB6 are two crucial components involved in T4SS which have fundamental effects on disease development of both animal and plant pathogens (Alegria et al., 2005; Souza et al., 2011). These genes are also absent in PXO99A and MAFF311018 but present in the DY89031 and KACC10331 genomes. These data suggest that unique virulence factors may have emerged in the Korean strain but not in PXO99A and MAFF311018.

In addition, we observed that the PXO99A genome contains sequence (total ~0.3 Mb) that is absent in DY89031

Table 2 Summary of the DY89031 genome and genomes of three other *Xoo* strains

	DY89031	PXO99A	KACC10331	MAFF311018
Length (bp)	4,979,456	5,240,075	4,941,439	4,940,217
GC content (%)	63.7	63.6	63.7	63.7
Predict genes	4,794	5,083	4,637	4,372
Transfer RNA	53	55	54	53
Ribosomal RNA operons	2	2	2	2
TAL gene	13	19	13	17

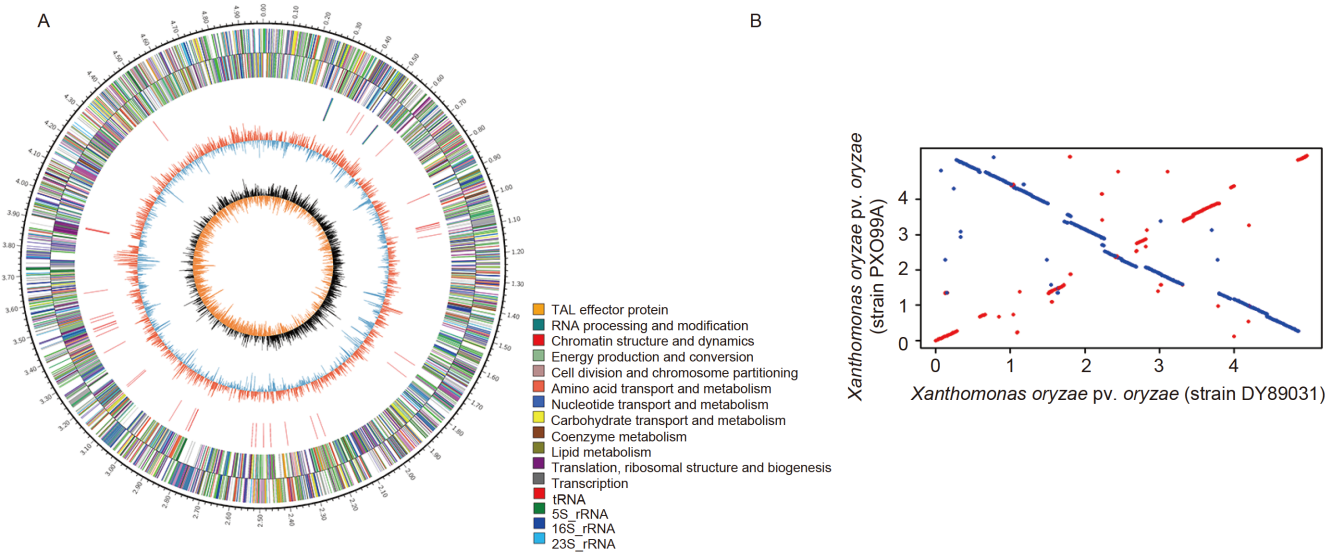

Figure 2 Graphical representation of the DY89031 genome and its rearrangement compared with the PXO99A genome. A, Graphical representation. The Outmost circle is genome size, with 0.1 Mb intervals. The second and the third circle are CDS on the sense and antisense strands. Different colors indicate CDS according to COG functional classification. The fourth circle represents rRNA and tRNA. The fifth circle shows the GC content, with red indicating GC content higher than the average of the whole genome, and blue indicating the opposite. Central circle is the GC skew value. B, Comparison of the DY89031 genome against the PXO99A genome, as determined by the MUMmer tool. The line with slope 1 represents forward matching (red) while the line with slope −1 represents reverse (blue).

(Table 2). We analyzed the large segments (>1 kb) in the 0.3 Mb region and found that these segments encode ~144 potential proteins (Table S3 in Supporting Information) and the majority of them are transposable elements (TEs), which is consistent with the abundant inversions spanning the genome. Besides, there are six type IV secretion proteins Rhs known for the transportation of molecules through membranes (Ahmad et al., 2019). The type II toxin-antitoxin RelE/ParE family is also absent in the DY89031 genome, which functions in diverse physiological processes including antibiotic persistence and genomic stabilization (Fraikin et al., 2020). Together, we propose that these proteins may contribute to different cellular processes of DY89031 as compared with PXO99A.

Phylogenetic analysis implies that the evolution of *Xoo* is influenced by rice cultivation conditions and domestication

To further investigate the population structure of *Xoo* strains, we performed the phylogenetic tree analysis using protein sequences of DY89031 and other 62 *Xoo* strains/isolates available (Table S4 in Supporting Information), which were mainly collected from Asia, Africa, South America and Oceania. The 63 strains could be divided into 2 branches (Figure 3), one mainly consists of Asian strains, while the other only consists of African strains, such as AXO1947, Ug11 and BAI3. This result suggests that the strains from Asia and Africa were diverged in an early stage and their evolution might be influenced by rice domestication and cultivation.

Moreover, five clades are classified in the first branch. Clade I contains three strains from Yunnan, China (such as YN18), and most strains in Clade II are located in tropical regions, including the South America strain (CIAT from Columbia) and South Asia strains, like Philippine strains (such as PXO99A), Indian strains (such as BXO25), Thai strain (SK2-3) and Nepaulese strain (NX0260). Clade III includes strains from north Asia, such as Korean strains (DY89031, KXO85/KACC10331) and Chinese strains

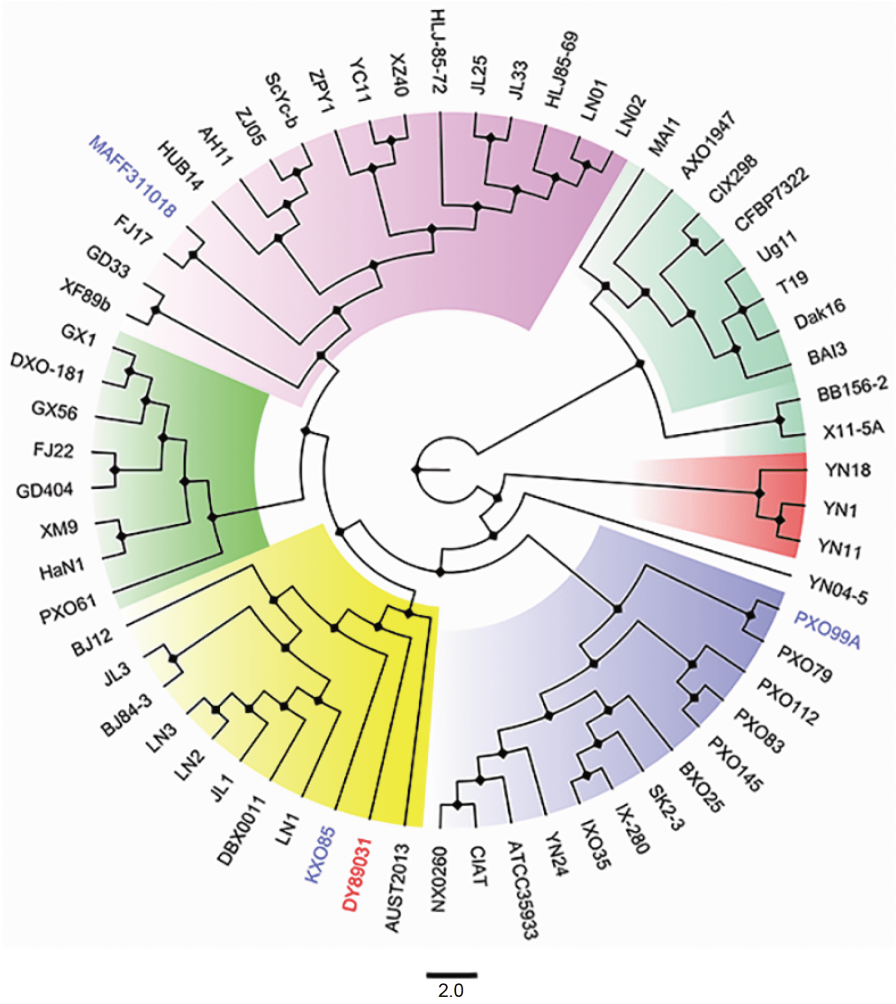

Figure 3 Phylogenetic tree of 63 *Xoo* strains from different regions. The 63 strains were divided into 2 branches. The first branch mainly consists of Asian strains, which could be sub-grouped into five clades indicated by differ colors. Blue represents the south Asian strains including these from Philippines and India; yellow represents north Asian strains from north China and Korea; South China strains are mainly concentrated in the green clades, strains from Japan and other China districts are in purple, and southeast China (Yunnan) strains are in red.

(BJ84-3 and LN1 from Beijing and Liaoning). Clade IV–V mainly consists of other Asian strains. Based on these, we propose that *Xoo* strains were likely evolved differently in different regions. In this scenario, large scale rice cultivation and *Xoo* epidemics in south Asia resulted in the early evolution of *Xoo* in this area. Thus, PXO99A might be a pioneer strain in comparison with the north Asia strains including DY89031 KACC10331 and MAFF311018. Together, phylogenetic analysis suggests that *Xoo* might have coevolved with rice domestication and its resistance gene selection.

Novel TAL effectors encoded by the DY89031 genome

TAL effectors contain multiple 33–35 amino acid long sequence repeats and usually bind to and activate the host *S* gene promoters (Boch and Bonas, 2010; Ji et al., 2016). TAL effectors are highly conserved in amino acid sequences, except two variable amino acids at position 12 and 13 of each repeat (termed "RVD"), which determine the recognition specificity of promoter motifs of host genes (Boch and Bonas, 2010). The predicted DY89031 TAL effectors contain variable RVD sequences and their repeats range from 8.5 to 23.5 (Table 3). Importantly, a number of *Tal* gene clusters and effectors are diversely distributed among the *Xoo* stains.

There are seven Tal loci in DY89031 encoding 12 effectors, with one pseudo gene, *Tal5*, and nine loci encoding 19 effectors in PXO99A, six loci with 13 effectors in KACC10331, and seven loci with 17 effectors in MAFF311018 (Figure 4A). These results suggest that the Tal loci have been subjected to extensive evolutionary selection, and that divergent effectors contribute to different virulence toward different rice target genes. On the other hand, the comparison of the TAL effector sequences from the four strains, DY89031, KACC10331, MAFF311018, and PXO99A revealed that half of the DY89031 TAL effectors showed similarity to those of PXO99A, with only a few

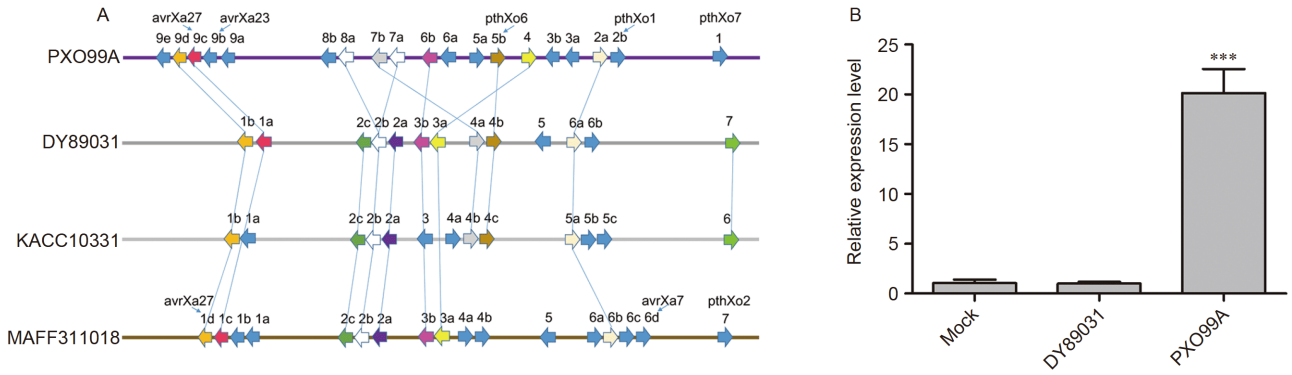

Figure 4 TAL effectors in the *Xoo* genomes. A, Gene orientations are indicated by arrows. TAL genes in the same clusters are numbered. The identical effectors with significant synteny and identical RVD sequences are lined with the same colors. Note that strain-specific effectors were identified based on the genome sequences, which are indicated in blue and not lined. B, Expression of the *S* gene *OsSWEET11* detected by qRT-PCR in TP309 at 24 h postinoculation with DY89031 and PXO99A, which is normalized to *OsActin1* gene. Data shown are means±SD from three biological replicates. Asterisks indicate statistically significant difference in comparison with the mock (Student's *t*-test, ***, $P<0.001$).

Table 3 TAL effectors and RVD sequences in DY89031 and similar genes in 3 other strains[a)]

Gene	Repeat	RVD sequence	PXO99A	KACC-10331	MAFF-311018
Tal1b	12.5	NI-NN-NI-HG-HG-HD-NG-HD-HG-HD-HD-HD-NG	Tal9d	Tal1b	Tal1d
Tal1a	15.5	NI-NN-N*-NG-NS-NN-NN-NN-NI-NN-NI-NG-HD-HD-NI-NG	Tal9c	–	Tal1c
Tal2c	17.5	NI-NS-HD-NG-NS-NN-HD-N*-NN-NN-NI-NN-HD-NG-HD-HD-HD-NG	–	Tal2c	Tal2c
Tal2b	8.5	NI-HG-NI-NI-NI-NN-HD-NS-NN	Tal7a/8a	*Tal2b*	*Tal2b*
Tal2a	23.5	NN-HD-NS-NG-HD-NN-N*-NI-HD-NS-HD-NN-HD-NN-HD-NN-NN-NN-NN-NN-NN-NN-HD-NG	–	*Tal2a*	*Tal2a*
Tal3b	20.5	NI-N*-NI-HG-NI-NI-NS-HD-NN-HD-NS-NG-SS-HD-NI-NI-NN-NI-NN-NI-NG	*Tal6b*	–	*Tal3b*
Tal3a	8.5	NS-HG-NN-NN-NN-NI-NI-NG-HD	*Tal4*	–	*Tal3a*
Tal4a	19.5	NI-HG-NI-NG-HG-HD-NS-NG-HD-NN-NG-HG-NG-HD-HG-HD-HD-NI-NN-NG	*Tal7b*	*Tal4b*	–
Tal4b	22.5	NI-H*-NI-NN-NN-NN-NN-NN-HD-NI-NS-HG-HD-NI-N*-NS-NI-NI-HD-N*-NS-NI-N*	*Tal5b*	*Tal4c*	–
Tal5	5.5	NS-HD-HD-HD-NN-H*	–	–	–
Tal6a	19.5	NI-NG-NN-NG-NK-NG-NI-NN-NI-NN-NI-NN-NS-NG-NS-NN-NI-N*-NS-NG	*Tal2a*	*Tal5a*	*Tal6b*
Tal6b	18.5	HD-HD-HD-NS-N*-N*-HD-HD-NS-NS-NN-NN-NI-NG-NN-NI-N*-NS-N*	–	–	–
Tal7	21.5	NI-HG-NI-NN-NI-NN-HD-HD-HD-HD-NS-NS-HD-NI-NI-HD-NG-HD-HD-HD-NG-NG	–	*Tal6*	–

a) * predicted deletion of the thirteenth codons in the repeats.

differences in SNPs or RVD repeat. In addition, the effectors Tal1b/2b/6a were shared by the four strains (Figure 4A, Table 3), suggesting a same origin.

Rich diversity is found in all of the four strains (Figure 4A). The Tal2 locus in PXO99A encodes PthXo1, which has been functionally identified as an essential virulence factor responsible for the induction of the S gene *OsSWEET11,* a wild type allele of resistance-conferring *xa13* (Eom et al., 2019). However, PthXo1 (Tal2b in PXO99A) is lost in the north Asia strains. We found that the transcript level of *OsSWEET11* is significantly lower in DY89031 compared to that in PXO99A (Figure 4B), suggesting that DY89031 loses the virulence targeting *OsSWEET11*. This result also suggests that the Tal2 locus might be ancient and essential for *Xoo* virulence on the rice host. Tal1a of DY89031 is a homolog of AvrXa27, and the N-terminal truncated protein is similar to Tal2b and Tal4a, but whether the truncated TAL effector is still active requires further investigation. In addition, DY89031 contains a Tal6b, which is distinct from other strains, it also contains an additional Tal7 shared by another Korean strain KACC10331 (Tal6). All three north Asia strains (DY89031, KACC10331 and MAFF311018) contain the *Tal2* locus which encodes 3 TAL effectors, Tal2a, Tal2b and Tal2c (Figure 4A), suggesting that these strains may obtain new virulence features for infection. It would be worthy to further identify the targets of Tal2, Tal6 and Tal7 in the host by extensive biochemical and molecular experiments.

Non-TAL effectors encoded by the DY89031 genome

Except TAL effectors, non-TAL effectors are also secreted by the T3SS in *Xoo* (Furutani et al., 2009), which are involved in the suppression of the host immunity (Qin et al.,

Table 4 Non-TAL effectors in DY89031, MAFF311018, KACC10331 and PXO99A

Effector class	DY89031	PXO99A	KACC10331	MAFF311018
XopW	J18_00141	PXO_03356	–[a)]	XOO0037
XopF	J18_00076	PXO_03413	XOO0074	XOO0103
AvrBs2	J18_00017	PXO_03330	XOO0168	XOO0148
XopN	J18_00357	PXO_02760	XOO0343	XOO0315
XopY	J18_01663	PXO_04866	–	XOO1488
XopK	J18_01851	PXO_01625	XOO1768	XOO1669
XopT	J18_02455	–	XOO4824	XOO2210
XopZ	J18_02670	PXO_06125/PXO_01041	XOO2543	XOO2402
XopAA	J18_03185	PXO_00234	XOO3022	XOO2875
XopU	J18_03187	PXO_00236	–	XOO2877
XopAB	J18_03512	–	–	XOO3150
XopP	J18_03600	PXO_02107	XOO3426	XOO3222
XopV	J18_04241	PXO_04172	XOO4033	XOO3803
XopX	J18_04500	PXO_03702	XOO4287	XOO4042
XopR	J18_04602	PXO_03819	XOO4391	XOO4134
XopQ	J18_04681	PXO_03901	XOO4466	XOO4208
XopC	J18_03598	PXO_02108	XOO3424	XOO3221
XopA	J18_00098	PXO_03392	XOO0095	XOO0081
XopAD	J18_04614	PXO_03833	XOO4401	XOO4145
XopAE	J18_00069	PXO_03420	XOO0065	XOO0110
XopL	J18_01846	PXO_01620	XOO1762	XOO1662
HpaA	J18_00082	PXO_03408	XOO0079	XOO0097
XopM	J18_00100	PXO_03390	–	–
XopAY	J18_03625/03626	PXO_02081/02079	–	–
XopAV	J18_03626	PXO_02079	–	–
XopAZ	J18_01980	PXO_01757	–	–
XopAU	J18_03627	PXO_02078	–	–

a) These with dishes indicate non-annotated homologs.

2018). Up to now, only a few non-TAL effectors, mainly Xop proteins, have been functionally identified in *Xoo* (Jiang et al., 2020a; Jiang et al., 2020b; Kim et al., 2009). For instance, XopQ, XopX and XopZ are reported to function as suppressors of LipA-induced innate immunity in rice (Sinha et al., 2013). Analysis of the DY89031 genome reveals that at least 27 Xop-like proteins are encoded therein (Table 4). In contrast to the TAL effectors, non-TAL effectors appear to be more conserved, suggesting their conserved evolution and functionality in the rice-*Xoo* interaction. Interestingly, XopT is likely absent in PXO99A. Whether XopT contributes to its compatible phenotype observed in the worldwide collection of rice germplasm (Table 1) needs further investigation.

Ectopic expression reveals potential function of the DY89031 effectors

To identify the differences of virulence factors between DY89031 and PXO99A, we compared the genome-wide effector genes and mainly focused on effector-like genes with different sequences. Totally, we focused on 27 genes that encode proteins predicted to be secreted by T3SS and could be expressed in the transformants of PXO99A (Tables S5 and S6 in Supporting Information). The PXO99A transformants expressing the DY89031 effector genes were individually used to inoculate Kasalath, a PXO99A-susceptible and DY89031-resistant variety. Interestingly, some PXO99A transformants showed decreased virulence, such as 29# and 15#, which express a copper resistance protein and a hypothetical T3SS protein, respectively (Figure 5B and C, Table 5). The remaining PXO99A transformants exhibited a virulence phenotype identical to PXO99A. Therefore, we propose that the T3SS effectors with sequence differences between PXO99A and DY89031 likely contribute to virulence differentiation. Our results may also explain the lower virulence of DY89031 in comparison with PXO99A (Figure 1A and C).

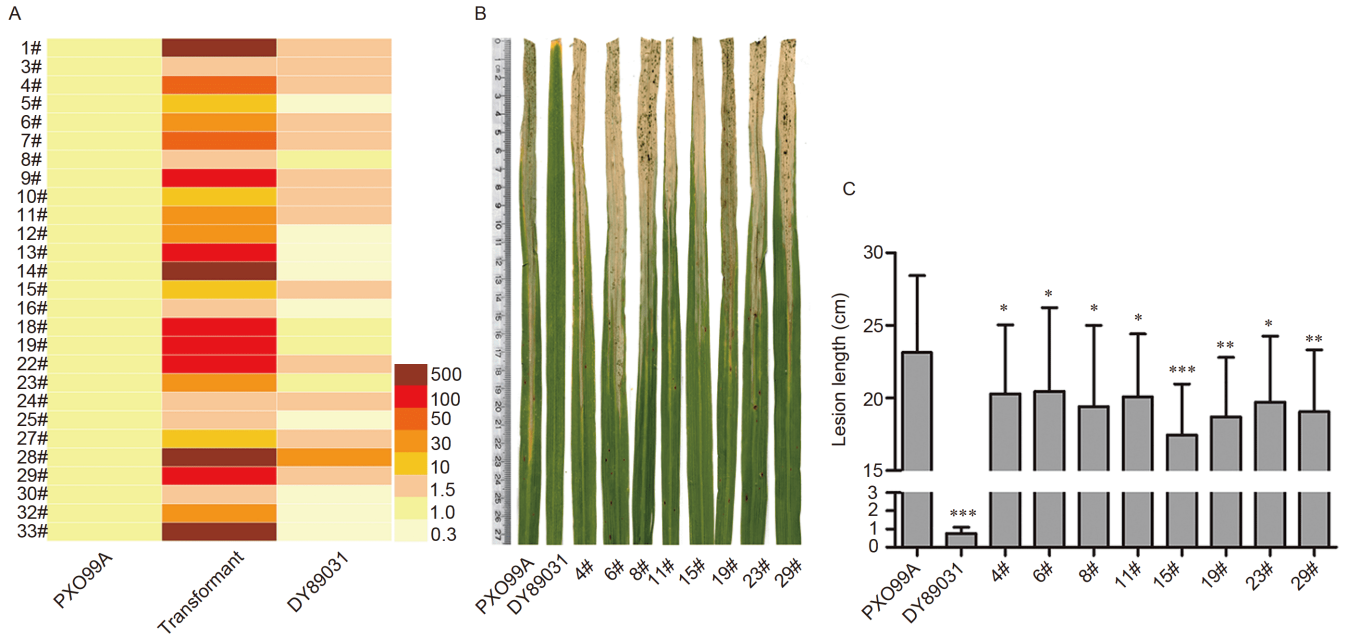

Figure 5 Inoculation analyses of the DY89031 effectors in PXO99A transformants. A, Expression of 27 transferred effector genes in PXO99A transformants, which is detected by qRT-PCR. Transcript levels of the effector genes in the transformants were compared with DY89031 and wild type PXO99A. Note that wild type PXO99A does not contain the same genes, only showing basal expression of endogenous homologues. 16S rRNA gene was used as an internal control to normalize expression levels. Data shown are means±SD from three technical replicates. The primers were listed in Table S6 in Supporting Information. B and C, Disease symptoms (B) and lesion lengths (C) of Kasalath at 14 dpi with the wild type PXO99A, DY89031 and transformants. Note that most of the PXO99A transformants expressing the individual DY89031 effectors decreased virulence in Kasalath that is resistant to DY89031. Asterisks indicate statistically significant difference in comparison with the control (Student's *t*-test, *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$).

Table 5 Information on DY89031 effector genes showing partial avirulence phenotypes in transformed PXO99A

Transformant	Gene	Amino acid	Annotation
4#	J18_04715	161	Putative esterase
6#	J18_04820	200	GTP cyclohydrolase
8#	J18_04831	85	Hypothetical T3SS protein
11#	J18_03349	358	Hypothetical T3SS protein
15#	J18_00005	278	Hypothetical T3SS protein
19#	J18_00554	366	Hypothetical T3SS protein
23#	J18_02024	105	Hypothetical T3SS protein
29#	J18_00787	602	Copper resistance protein A

DISCUSSION

Xanthomonas oryzae pv. *oryzae* is an obligate vascular pathogen in rice, which causes leaf blight by infecting rice plants from stomata and wounds on leaf blade and then multiplying in vascular bundles. To defeat this pathogen, rice has evolved diverse strategies of immune response including expressing different types of resistance genes, dominant or recessive (Li et al., 2020). However, due to the rapid changes/evolution of pathogen virulence, rice production has been threatened by this disease worldwide, particularly in flooding regions. In this study, we sequenced the genome of a north Asia *Xoo* strain, DY89031, and compared it with other representative *Xoo* genomes comprehensively. We found rich diversities in genome size, structure, and particularly in effector genes, which shape the pathogenicity of the strains. Accordingly, we show that *Xoo* might have evolved with diverse origins.

Xoo strains secrete specific TAL effectors, which function as transcription factors to activate host susceptibility genes. Our genome analysis revealed differences in the number of TAL loci and proteins among the strains compared. The genetic relationship of 63 strains revealed the possible geographical and environmental adaptions. Interestingly, *Xoo* strains collected from Asia where cultivated rice has evolved and diseases prevailed with plentiful rainfall harbor more TAL effectors than the Africa strains which only contain 8 TAL effectors (Gonzalez et al., 2007). These results not only support the notion that the long co-evolution with the rice host has greatly increased pathogenicity of the bacterium, but

also suggest a direct manifestation of environmental adaptation to pathogenic virulence and host resistance with natural selection.

Up to 80% sequenced rice accessions are susceptible to DY89031, indicating that many rice varieties have kept the *R* gene(s) against the strain. In contrast to the south Asia strain PXO99A, the efficient *R* gene identified so far to this strain is *Xa21*, which was originally derived from wild rice (Song et al., 1995). However, DY89031 can overcome the *Xa21*-mediated resistance. Whether the occurrence of new TAL effectors such as Tal2a, Tal2c and Tal7 may reflect new cognate *R* genes specific to DY89031 remains unknown. Our experiments on the PXO99A transformants expressing the DY89031 effectors suggested that some of the effectors may function in ETS (effector-triggered susceptibility). We propose that the Korean strains including DY89031 might have evolved to possess special virulence, which is probably associated with the domestication history of cultivated rice.

Exploring the rice-*Xoo* pathosystem and analyzing the pathogenic factors of *Xoo* will facilitate further understanding of the co-evolution of rice-*Xoo*, as well as the developing of strategies to explore new disease resistance genes for rice breeding. Currently, more than 45 *Xa* genes/loci have been discovered (Ji et al., 2020), and several of them have been used in rice breeding and production. However, single *Xa* gene-mediated resistance is easy to be broken down in agriculture. Pyramiding different *R* genes provides a practical approach for durable resistance. Among the reported *Xa* genes, approximately 25% were screened from wild rice, such as *Xa21*, *Xa23*, *Xa27*, *Xa29*, *Xa32*, *Xa35*, *Xa36*, *Xa38*, *xa41*, and *Xa45* (Bhasin et al., 2011; Hutin et al., 2015; Ji et al., 2020). Thus, wild rice is a precious treasure to identify novel *R* genes against the bacterial blight disease. Nevertheless, our finding of DY89031-resistant germplasm will facilitate discovery of novel *R* genes against this pathogen.

MATERIALS AND METHODS

Plant material and bacterial growth

Rice varieties used in this study included Kasalath, TN1, 9311, and *Xa21* transgenic line 106 in TP309 background. A total of 2,347 rice germplasm resources were obtained from 3,000 rice genomes project (Wang et al., 2018). Leaves of 8-week-old plants were inoculated with the *Xoo* strains, DY89031, PXO99A (wild type and transformants), KACC103311 and MAFF311018.

Genome sequencing and assembly

Genomic DNA of *Xoo* strain DY89031 was extracted by DNeasy Blood Tissue Kit (Qiagen, German) and used for sequencing, which was performed using PacBio® RS II System (Pacific Biosciences, USA) by Shanghai Biotechnology Corporation. The data provided by PacBio was qualified by the index of mean read length and quality, and assembled by SPAdes-3.5.0. Prokka software combined with Swiss-prot library profile was employed in gene prediction. The predicted genes were annotated with the GO function by blast2GO algorithm, and the COG, KEGG and ARDB annotations were obtained using RPS-BLAST, KOBAS and BLASTP tools by comparing the predicted genes with the CDD/COG, KO and ARDB databases, respectively. The tRNA and rRNA data are predicted by the tRNAscan-SE-1.23 and RNAmmer1.2 software. The DY89031 genome diagram was drawn by the Circosv0.69 software. The TAL effector coding sequences were scanned using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and Software ANNOTALE 1.4.1. The DY89031 genome information was applied to draw the rearrangement file by Mummer plot. The whole genome sequence of DY89031 was deposited in the NCBI BioProject Database (BioProject ID: PRJNA675962).

Virulence assays

Xoo strains were cultivated for 48–72 h, washed and resuspended in sterile distilled water at a concentration of A_{600}=1.0 for inoculation on rice. Plant leaves were inoculated with strains by the leaf-clipping method (Yang et al., 2008). Disease evaluation was performed 14 days after inoculation (dpi) by measuring the lesion length. All *Xoo* strains were cultivated in PSA medium (10 g L^{-1} peptone, 10 g L^{-1} sucrose, 1 g L^{-1} glutamine, 15 g L^{-1} agar) with antibiotics at 28°C.

Quantitative real-time PCR

RNAs were extracted using TRIzol reagent (Invitrogen, USA) from rice leaves inoculated. Total RNAs (0.5 μg) were used for cDNA synthesis with ReverTra Ace qPCR RT Master Mix with gDNA remover (TOYOBO, Japan). Quantitative real-time PCR (qRT-PCR) was conducted using the SYBR Premix Ex Taq kit (TaKaRa Biotechnology, Japan) on CFX96 Real-time PCR Instrument (Bio-Rad, USA) following the manufacturer's instructions. The rice gene *OsActin1* was used as a control to normalize expression levels, and the result was repeated with three technical and biological repeats. Primers used to amplify *OsSWEET11* and *OsActin1* were listed as follows: OsActin1-F, 5′-TGTATGCCAGTGGTCGTACCA-3′; OsActin1-R, 5′-CCAGCAAGGTCGAGACGAA-3′; OsSWEET11-F, 5′-AGGAGGTTTCTTGTCCATGG-3′; and OsSWEET11-R, 5′-CCTCGAGTTGGTCTTCACCA-3′.

Phylogenetic tree

The genome sequences of 62 *Xoo* strains were retrieved from NCBI database (https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/529/), and OrthoFinder2.2.7 was served to construct phylogenetic trees based on Multiple Sequence Alignment (MSA) results according to maximum likelihood method. RAxML-NG was used to infer the evolutionary tree and Figtree was used to visualize the phylogenetic tree. Different clades were colored.

Plasmid construction and *Xoo* transformation

The sequences of candidate genes were amplified by PCR and then cloned into vector pUFR034 digested by *Eco*RI and *Bam*HI. The resulting plasmids were first amplified in *E. coli* and sequenced, and then transformed into competent recipient cells of PXO99A.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *This work was supported by Chinese Academy of Sciences (XDB27040201) and the National Natural Science Foundation of China (3181101746). We would like to thank Prof. Jianlong Xu (Institute of Crop Sciences, CAAS) for providing rice germplasm resources, Prof. Gongyou Chen (Shanghai Jiao Tong University) for providing Xoo strains, Prof. Yiwen Deng for rice disease evaluation, Dr. Zeling Xu (South China Agricultural University) for helping with the genome sequencing analysis.*

References

Ahmad, S., Ranaghan, K.E., and Azam, S.S. (2019). Combating tigecycline resistant *Acinetobacter baumannii*: A leap forward towards multi-epitope based vaccine discovery. Eur J Pharm Sci 132, 1–17.

Alegria, M.C., Souza, D.P., Andrade, M.O., Docena, C., Khater, L., Ramos, C.H.I., da Silva, A.C.R., and Farah, C.S. (2005). Identification of new protein-protein interactions involving the products of the chromosome- and plasmid-encoded type IV secretion loci of the phytopathogen *Xanthomonas axonopodis* pv. *citri*. J Bacteriol 187, 2315–2325.

An, S.Q., Potnis, N., Dow, M., Vorhölter, F.J., He, Y.Q., Becker, A., Teper, D., Li, Y., Wang, N., Bleris, L., et al. (2020). Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen *Xanthomonas*. FEMS Microbiol Rev 44, 1–32.

Antony, G., Zhou, J., Huang, S., Li, T., Liu, B., White, F., and Yang, B. (2010). Rice *xa13* recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene *Os-11N3*. Plant Cell 22, 3864–3876.

Bartetzko, V., Sonnewald, S., Vogel, F., Hartner, K., Stadler, R., Hammes, U.Z., and Börnke, F. (2009). The *Xanthomonas campestris* pv. *vesicatoria* type III effector protein XopJ inhibits protein secretion: evidence for interference with cell wall-associated defense responses. MPMI 22, 655–664.

Bhasin, H., Bhatia, D., Raghuvanshi, S., Lore, J.S., Sahi, G.K., Kaur, B., Vikal, Y., and Singh, K. (2011). New PCR-based sequence-tagged site marker for bacterial blight resistance gene *Xa38* of rice. Mol Breeding 30, 607–611.

Boch, J., and Bonas, U. (2010). *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol 48, 419–436.

Bogdanove, A.J., Schornack, S., and Lahaye, T. (2010). TAL effectors: finding plant genes for disease and defense. Curr Opin Plant Biol 13, 394–401.

Chen, L.Q., Qu, X.Q., Hou, B.H., Sosso, D., Osorio, S., Fernie, A.R., and Frommer, W.B. (2012). Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. Science 335, 207–211.

Cianfanelli, F.R., Alcoforado Diniz, J., Guo, M., De Cesare, V., Trost, M., and Coulthurst, S.J. (2016). VgrG and PAAR proteins define distinct versions of a functional type VI secretion system. PLoS Pathog 12, e1005735.

da Silva, F.G., Shen, Y., Dardick, C., Burdman, S., Yadav, R.C., de Leon, A. L., and Ronald, P.C. (2003). Bacterial genes involved in type I secretion and sulfation are required to elicit the rice *Xa21*-mediated innate immune response. MPMI 17, 593–601.

Eom, J.S., Luo, D., Atienza-Grande, G., Yang, J., Ji, C., Thi Luu, V., Huguet-Tapia, J.C., Char, S.N., Liu, B., Nguyen, H., et al. (2019). Diagnostic kit for rice blight resistance. Nat Biotechnol 37, 1372–1379.

Deng, Y., Ning, Y., Yang, D.L., Zhai, K., Wang, G.L., and He, Z. (2020). Molecular Basis of disease resistance and perspectives on breeding strategies for resistance improvement in crops. Mol Plant 13, 1402–1419.

Fraikin, N., Goormaghtigh, F., and Van Melderen, L. (2020). Type II toxin-antitoxin systems: evolution and revolutions. J Bacteriol 202, e00763.

Furutani, A., Takaoka, M., Sanada, H., Noguchi, Y., Oku, T., Tsuno, K., Ochiai, H., and Tsuge, S. (2009). Identification of novel type III secretion effectors in *Xanthomonas oryzae* pv. *oryzae*. MPMI 22, 96–106.

Gao, M.J., and He, Z.H. (2013). Studies on innate immunity in rice (in Chinese). Sci Sin Vitae 43, 1016–1029.

Gerlach, R.G., and Hensel, M. (2007). Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. Int J Med Microbiol 297, 401–415.

Gonzalez, C., Szurek, B., Manceau, C., Mathieu, T., Séré, Y., and Verdier, V. (2007). Molecular and pathotypic characterization of new *Xanthomonas oryzae* strains from West Africa. MPMI 20, 534–546.

Hutin, M., Sabot, F., Ghesquière, A., Koebnik, R., and Szurek, B. (2015). A knowledge-based molecular screen uncovers a broad-spectrum *OsSWEET14* resistance allele to bacterial blight from wild rice. Plant J 84, 694–703.

Ji, C., Ji, Z., Liu, B., Cheng, H., Liu, H., Liu, S., Yang, B., and Chen, G. (2020). *Xa1* allelic *R* genes activate rice blight resistance suppressed by interfering TAL effectors. Plant Commun 1, 100087.

Ji, Z., Ji, C., Liu, B., Zou, L., Chen, G., and Yang, B. (2016). Interfering TAL effectors of *Xanthomonas oryzae* neutralize *R*-gene-mediated plant disease resistance. Nat Commun 7, 13435.

Jiang, G., Liu, D., Yin, D., Zhou, Z., Shi, Y., Li, C., Zhu, L., and Zhai, W. (2020a). A rice *NBS-ARC* gene conferring quantitative resistance to bacterial blight is regulated by a pathogen effector-inducible miRNA. Mol Plant 13, 1752–1767.

Jiang, N., Yan, J., Liang, Y., Shi, Y., He, Z., Wu, Y., Zeng, Q., Liu, X., and Peng, J. (2020b). Resistance genes and their interactions with bacterial blight/leaf streak pathogens (*Xanthomonas oryzae*) in rice (*Oryza sativa* L.)—an updated review. Rice 13, 3.

Jones, J.D.G., and Dangl, J.L. (2006). The plant immune system. Nature 444, 323–329.

Kim, J.G., Li, X., Roden, J.A., Taylor, K.W., Aakre, C.D., Su, B., Lalonde, S., Kirik, A., Chen, Y., Baranage, G., et al. (2009). *Xanthomonas* T3S effector XopN suppresses PAMP-triggered immunity and interacts with a tomato atypical receptor-like kinase and TFT1. Plant Cell 21, 1305–1323.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. (2004). Versatile and open software for comparing large genomes. Genome Biol 5, R12.

Lee, B.M., Park, Y.J., Park, D.S., Kang, H.W., Kim, J.G., Song, E.S., Park, I.C., Yoon, U.H., Hahn, J.H., Koo, B.S., et al. (2005). The genome sequence of *Xanthomonas oryzae pathovar oryzae* KACC10331, the bacterial blight pathogen of rice. Nucleic Acids Res 33, 577–586.

Li, W., Deng, Y., Ning, Y., He, Z., and Wang, G.L. (2020). Exploiting broad-spectrum disease resistance in crops: From molecular dissection

to breeding. Annu Rev Plant Biol 71, 575–603.
Mew, T.W. (1987). Current status and future prospects of research on bacterial blight of rice. Annu Rev Phytopathol 25, 359–382.
Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., et al. (2011). A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29, 143–148.
Niño-Liu, D.O., Ronald, P.C., and Bogdanove, A.J. (2006). *Xanthomonas oryzae pathovars*: model pathogens of a model crop. Mol Plant Pathol 7, 303–324.
Ochiai, H., Inoue, Y., Takeya, M., Sasaki, A., and Kaku, H. (2005). Genome sequence of *Xanthomonas oryzae* pv. *oryzae* suggests contribution of large numbers of effector genes and insertion sequences to its race diversity. JARQ 39, 275–287.
Pfeilmeier, S., Caly, D.L., and Malone, J.G. (2016). Bacterial pathogenesis of plants: future challenges from a microbial perspective. Mol Plant Pathol 17, 1298–1313.
Pruitt, R.N., Schwessinger, B., Joe, A., Thomas, N., Liu, F., Albert, M., Robinson, M.R., Chan, L.J.G., Luu, D.D., Chen, H., et al. (2015). The rice immune receptor XA21 recognizes a tyrosine-sulfated protein from a Gram-negative bacterium. Sci Adv 1, e1500245.
Qin, J., Zhou, X., Sun, L., Wang, K., Yang, F., Liao, H., Rong, W., Yin, J., Chen, H., Chen, X., et al. (2018). The *Xanthomonas* effector XopK harbours E3 ubiquitin-ligase activity that is required for virulence. New Phytol 220, 219–231.
Salzberg, S.L., Sommer, D.D., Schatz, M.C., Phillippy, A.M., Rabinowicz, P.D., Tsuge, S., Furutani, A., Ochiai, H., Delcher, A.L., Kelley, D., et al. (2008). Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99A. BMC Genomics 9, 204.
Schulze, S., Kay, S., Büttner, D., Egler, M., Eschen-Lippold, L., Hause, G., Krüger, A., Lee, J., Müller, O., Scheel, D., et al. (2012). Analysis of new type III effectors from *Xanthomonas* uncovers XopB and XopS as suppressors of plant immunity. New Phytol 195, 894–911.
Sinha, D., Gupta, M.K., Patel, H.K., Ranjan, A., and Sonti, R.V. (2013). Cell wall degrading enzyme induced rice innate immune responses are suppressed by the type 3 secretion system effectors XopN, XopQ, XopX and XopZ of *Xanthomonas oryzae* pv. *oryzae*. PLoS ONE 8, e75867.
Song, C., and Yang, B. (2010). Mutagenesis of 18 type III effectors reveals virulence function of $XopZ_{PXO99}$ in *Xanthomonas oryzae* pv. *oryzae*. MPMI 23, 893–902.
Song, W.Y., Wang, G.L., Chen, L.L., Kim, H.S., Pi, L.Y., Holsten, T., Gardner, J., Wang, B., Zhai, W.X., Zhu, L.H., et al. (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science 270, 1804–1806.
Souza, D.P., Andrade, M.O., Alvarez-Martinez, C.E., Arantes, G.M., Farah, C.S., and Salinas, R.K. (2011). A component of the *Xanthomonadaceae* type IV secretion system combines a VirB7 motif with a N0 domain found in outer membrane transport proteins. PLoS Pathog 7, e1002031.
Streubel, J., Pesce, C., Hutin, M., Koebnik, R., Boch, J., and Szurek, B. (2013). Five phylogenetically close rice *SWEET* genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. New Phytol 200, 808–819.
Timilsina, S., Potnis, N., Newberry, E.A., Liyanapathiranage, P., Iruegas-Bocardo, F., White, F.F., Goss, E.M., and Jones, J.B. (2020). *Xanthomonas* diversity, virulence and plant-pathogen interactions. Nat Rev Microbiol 18, 415–427.
Wang, J., and Chai, J. (2020). Molecular actions of NLR immune receptors in plants and animals. Sci China Life Sci 63, 1303–1316.
Wang, J., Chern, M., and Chen, X. (2020). Structural dynamics of a plant NLR resistosome: transition from autoinhibition to activation. Sci China Life Sci 63, 617–619.
Wang, W., Mauleon, R., Hu, Z., Chebotarov, D., Tai, S., Wu, Z., Li, M., Zheng, T., Fuentes, R.R., Zhang, F., et al. (2018). Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature 557, 43–49.
Wu, X.M., Li, Y.R., Zou, L.F., and Chen, G.Y. (2007). Gene-for-gene relationships between rice and diverse avrBs3/pthA avirulence genes in *Xanthomonas oryzae* pv. *oryzae*. Plant Pathol 56, 26–34.
Yang, B., Sugio, A., and White, F.F. (2006). *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. Proc Natl Acad Sci USA 103, 10503–10508.
Yang, D.L., Li, Q., Deng, Y.W., Lou, Y.G., Wang, M.Y., Zhou, G.X., Zhang, Y.Y., and He, Z.H. (2008). Altered disease development in the *eui* mutants and *Eui* overexpressors indicates that gibberellins negatively regulate rice basal disease resistance. Mol Plant 1, 528–537.
Zhang, B., Zhang, H., Li, F., Ouyang, Y., Yuan, M., Li, X., Xiao, J., and Wang, S. (2020). Multiple alleles encoding atypical NLRs with unique central tandem repeats in rice confer resistance to *Xanthomonas oryzae* pv. *oryzae*. Plant Commun 1, 100088.
Zhou, J., Peng, Z., Long, J., Sosso, D., Liu, B., Eom, J.S., Huang, S., Liu, S., Vera Cruz, C., Frommer, W.B., et al. (2015). Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. Plant J 82, 632–643.

SUPPORTING INFORMATION

The supporting information is available online at https://doi.org/10.1007/s11427-020-1917-x. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.